



Norwegian University
of Life Sciences

Master's Thesis 2020 30 ECTS
Faculty of Science and Technology

Establishment of Confocal Microscopy Method for Detecting Double-Strand Breaks (DSBs) in DNA after Radiation

Tina Sørvik

Environmental Physics and Renewable Energy

Acknowledgements

I want to thank my supervisors Cecilia Marie Futsæther, Nina F. Edin, Joe A. Sandvik and Ingunn Hanson, for their professional guidance and great insights. I have been very fortunate for having such enthusiastic supervisors.

Thanks also to Stefan Schrunner for helping with the R-coding and for inspirational conversations. Thanks to Delmon Arous for his contribution and helpfulness.

I would like to thank Hilde S. Skeie for the great collaboration and for being both an excellent partner in the laboratory and friend. Thanks also to Isak B. Lande for the support and excellent conversations. Finally, I would like to thank everyone at the Biophysics and Medical Physics group at UiO for being both very welcoming and inspirational.

Abstract

Cancer is classified as a pandemic, causing more than 9 million deaths per year. Cancer treatment requires insights into cellular responses to radiation and repair mechanisms of damage. It is much interested in quantifying damage extent over time following radiation because the reduction in damage can be interpreted as a repair rate. Of particular interest is the double-strand breaks (DSB) on DNA because it is a severe form of damage that could potentially cause cell death.

The aim of this study was to develop a method for detecting the number of double-strand breaks (DSB) induced in cells following radiation and characterise the DSB's spatial distribution using 3D reconstructed images obtained by confocal microscopy.

Six experiments were conducted on cancerous cells to test the developed method and evaluate the damage extent following irradiation. Methodological development consisted of both determining the protocol for sample preparation and image analysis of the acquired images.

The final protocol for sample preparation used cells attached to the microscope slide and keeping their position through the entire experiment from irradiation to imaging. The benefit of this method is that the damage extent and repair can be view in light of the cells environment. Environmental conditions such as cell density might impact the cellular response.

The measurements of damage extent were consistent with results obtained by using flow cytometry which detects fluorescence intensity. The accuracy of DSB detection is limited to doses of 2 Gy and lower, as the DSB density becomes too high to distinguish at larger irradiation doses. The number of DSBs detected was, however, lower than the theoretically expected value. More investigation into the method's parameters is required for better a DSB count.

Sammendrag

Kreft er klassifisert som en pandemi, som forårsaker mer enn 9 millioner dødsfall per år. Kreftbehandling krever innsikt i cellulære reaksjoner på stråling og reparasjonsmekanisme for skade. Det er en stor interesse i å kvantifisere skadeomfang over tid etter stråling fordi reduksjonen i skade kan tolkes som en reparasjonshastighet. Av spesiell interesse er dobbeltstrengbruddene (DSB) på DNA fordi det er en alvorlig form for skade som potensielt kan forårsake celledød.

Målet med denne studien var å utvikle en metode for å oppdage antall dobbeltstrengede brudd (DSB) induisert i celler etter stråling og karakterisere DSBs romlige distribusjon ved bruk av 3D-rekonstruerte bilder oppnådd ved konfokal mikroskopi.

Seks eksperimenter ble utført på kreftceller for å teste den utviklede metoden og evaluere skadeomfanget etter bestråling. Metodologisk utvikling besto av både å bestemme protokollen for prøveforberedelse, og bildeanalyse av de anskaffede bildene.

Den endelige protokollen for prøveforberedelse brukte celler festet til mikroskopets glass og holdt sin posisjon gjennom hele eksperimentet fra bestråling til avbildning. Fordelen med denne metoden er at skadeomfanget og reparasjonen kan sees i lys av cellemiljøet. Miljøforhold som celletetthet kan påvirke cellens respons.

Målingene av skadeomfang var i samsvar med resultater oppnådd ved bruk av *flow cytometry* som detekterer fluorescensintensitet. Nøyaktigheten av DSB-deteksjon er begrenset til doser på 2 Gy og lavere, ettersom DSB-tettheten blir for høy til å skille ved større bestrålingsdoser. Antallet detekterte DSBs var imidlertid lavere enn den teoretisk forventede verdien. Mer undersøkelse av metodens parametere er nødvendig for en bedre DSB-telling.

List of Abbreviations

A	Adenine
ATCC	American Type Culture Collection
BER	Base Excision Repair
BSA	Bovine Serum Albumin
C	Cytosine
CI	Confidence interval
Cdk	Cyclin-dependent kinases
CP	Charged Particles
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DNEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DSB	Double-Strand Break
FCCF	Flow Cytometry Core Facility
FIJI	Fiji Is Just ImageJ
FITC	Fluorescein Isothiocyanate
G	Guanine
HRR	Homologous Recombination Repair
IP	Ionization Potential
LAF	Laminar Flow Cabinet
LD	Lethal Damage
LET	Linear Energy Transfer
NA	Numerical Aperture
NHEJ	Nonhomologous End-Joining
OPTICS	Ordering Points To Identify the Clustering Structure
PCC	Pearson Correlation Coefficient
PI	Propodium Iodide
PLD	Potentially Lethal Damage
PMT	Photomultiplier Tube
PAB	Primary Antibody
RT	Radiation Therapy
RBE	Relative Biological Effectiveness
SAB	Secondary Antibody
SLD	Sublethal Damage
S-phase	Synthesis phase
SSB	Single-Strand Break
SSD	Source Surface Distance
T	Thymine
UiO	University of Oslo
3D-OC	3D-Oject Counter

Contents

1	INTRODUCTION	1
2	THEORY	3
2.1	Cell biology	3
2.1.1	DNA	3
2.1.2	Cell cycle and control system	3
2.2	Radiation physics	5
2.2.1	Radiation interaction with matter	5
2.2.2	Linear Energy Transfer	7
2.2.3	Relative biological effectiveness	8
2.2.4	X-ray production	8
2.3	Radiobiology	9
2.3.1	Direct and indirect effect	9
2.3.2	Radiation damage	9
2.3.3	Damage repair	10
2.3.4	Radiosensitivity	11
2.4	Fluorescence imaging	12
2.4.1	The histone protein H2AX	13
2.4.2	Confocal microscopy	14
2.4.3	Flow cytometry	14
2.5	Analysis of confocal microscopy images	15
2.5.1	Object detection	15
2.5.2	Image characterization using Moran's I	15
2.5.3	OPTICS Clustering	16
3	METHOD	19
3.1	Experimental process and setup	19
3.2	Cell line	20
3.3	Cell line cultivation	20
3.3.1	Chemicals Equipment and aseptic technique	20
3.3.1.1	Chemicals	20
3.3.1.2	Equipment	20
3.3.1.3	Aseptic technique	21
3.3.2	Cell cultivation procedure	21
3.3.3	Seeding	21
3.4	Cell treatment	21
3.4.1	X-ray irradiation	21
3.4.1.1	Set up	22
3.4.1.2	Irradiation procedure	22

3.4.1.3	Dosimetry	23
3.4.2	Incubation	24
3.5	Staining protocols	24
3.5.1	H2AX Assay using cells in suspension	24
3.5.1.1	Equipment	24
3.5.1.2	Chemicals	24
3.5.1.3	Protocol	25
3.5.1.3.1	Protocol alterations	25
3.5.2	H2AX assay using adherent cells	25
3.5.2.1	Equipment	25
3.5.2.2	Chemicals	25
3.5.2.3	Protocol	26
3.5.2.4	Tests conducted to determine the protocol	26
3.5.2.4.1	Cell seeding amount	26
3.5.2.4.2	Fixation and permeabilization chemicals	26
3.5.2.4.3	Antibodies	26
3.6	Confocal imaging	27
3.7	Image processing and analysis	28
3.7.1	Image pre-processing	28
3.7.2	3D object counter	29
3.7.3	Spatial distribution evaluation	30
3.7.3.1	Moran's I	30
3.7.3.2	OPTICS clustering	31
3.8	Comparison of survival data and damage extent determined by the method	31
4	RESULTS	33
4.1	Developing the DSB quantification method	33
4.1.1	Development of suspension cell assay	33
4.1.2	Confocal microscopy image acquisition	34
4.1.3	Additional experiments to determine assay	35
4.1.4	Cell sampling significance	35
4.1.5	Duration of methodological processes	35
4.2	Experimental results from adherent cell assay (experiment A1)	36
4.2.1	Dose response	36
4.2.2	Characterization of γ -H2AX fluorescence and cell nuclei	44
4.2.3	LQ-model performance: P(VoIM/VoIN) plotted against survival data	47
4.2.4	Evaluation of Moran's I	48
4.2.5	Evaluation of OPTICS clustering	49
4.3	Results from suspension cell assay experiment S1	51
4.4	Summary of Result using Flow cytometry	52
5	DISCUSSION	55
5.1	Methodological considerations and inherent uncertainty	55

5.1.1	Cellular variations	55
5.1.2	Assay	55
5.1.3	Image acquisition using confocal microscopy	56
5.1.4	Image processing	57
5.1.5	Method evaluation	58
5.2	Performance of the method (DSB detection)	58
5.3	γ-H2AX experiment results	59
5.4	γ-H2AX detection using flow cytometry VS confocal microscopy	60
5.5	DSB's compared to the LQ-model and survival data	61
5.6	Method compared to FocAn	61
5.7	Future considerations	61
6	CONCLUSION	62
APPENDIX		I
A	ASSAY PROTOCOLS	I
A.1	Protocol for suspension cell assay	I
A.2	Protocol for Adherent cell assay	II
A.3	Protocol for fixation and permeabilization testing	II
A.4	Protocol for antibody testing	III
B	PROGRAMMING SCRIPTS	III
B.1	Programme for determining cell confluency	III
B.2	Programme for determining Moran's I and z-value	XXVII
B.3	Programme for determining clusters	XXIX
B.3.1	Programme for determining clusters by OPTICS	XXIX
B.3.2	Programme for determining cluster by k-means	XXXI
B.3.2.1	Gap Statistics function	XXXIV
C	ADDITIONAL RESULTS	XXXV
C.1	PI incubation test	XXXV
C.2	Antibody tests	XXXV
		XXXVI

C.3	Fixation and permeabilization test	XXXVII
C.3	Cell seeding confluence	XXXVIII
C.4	X-ray dose rate at SSD 40	XXXVIII

1 Introduction

In 2018 the *World Health Organization* reported approximately 9.6 million deaths worldwide due to cancer, making the disease classified as a pandemic [1]. Cancer is characterised by uncontrolled cell division, called proliferation. A key distinction of cancer cells is thus that they mutate to continue proliferation indefinitely and one of the cells central roles is to maintain genomic integrity. Ever since the crucial role of DNA in biology was discovered, there has been a strong interest in developing exploratory assays for DNA damage.

As *Deoxyribonucleic acid* (DNA) damage is a threat to genomic stability, cells have repair mechanisms to combat any such changes. DNA is coiled around histone protein H2AX and is therefore present throughout the genome. *Double-strand break* (DSB) denotes a breakage in both strands of the double-stranded DNA helix and can cause a fatal outcome if left unrepaired. One of the first responses to DSB damage is the phosphorylation of H2AX [2-4]. Phosphorylation means that a phosphate group (PO_4^-) is added to the protein molecule. Phosphorylated H2AX is denoted g-H2AX and is considered as an important step in signalling and initiation of the repair process. The phosphorylation of H2AX starts on the site of DSB but expands to the surrounding area [5-8]. The phosphorylated region can be visualised microscopically by using a fluorescent antibody specific for g-H2AX[4, 9-11]. The number of g-H2AX foci reflects the number of DSB within the cell [9] and can be used to assess the extent of DNA DSB damage. The technique is highly sensitive and detects effects of radiation doses as small as 0.003 Gy [8, 10]. Furthermore, the DSB foci develop quickly and are detected as soon as 30 minutes after irradiation. Creation of g-H2AX foci visualises the detection of DSBs, while reduction in the number of foci as a function of time reflects DNA DSB repair kinetics.

Cancer is frequently treated with radiation therapy (RT), both for palliative and curative effects[12]. A substantial drawback to radiation treatment is that the healthy tissue is also irradiated. The most common radiation used in clinical treatment is high energy X-rays. The maximum dose deposited by X-rays occurs near the entrance surface. Lately, there has been a great interest in proton radiation for clinical use. Accelerated protons deposit their energy in a smaller and more defined range where the maximum dose occurs at the end of the track. Due to its recent entrance in clinical practice [13], more research on proton irradiation and its effects are needed.

Many cellular investigations rely on quantitative analysis, such as colony count and protein detection. In radiation biology, foci analysis reveals insight into DNA damage and repair kinetics. Knowledge of genotoxicity induced by ionizing radiation is essential for clinical

application, both for increasing impact on the desired region and minimising the effects elsewhere.

Currently, there are different ways to detect and count foci from images obtained by fluorescence microscopy [11]. The traditional method consists of manual counting which is both time-consuming and highly susceptible for human error. Another strategy is using analysis software like MetaMorph (Molecular Devices Inc., USA), but many rely heavily on time-consuming and error-prone manual labour or are highly specialised and expensive. Furthermore, the availability of programs considering the 3D nature of the cell nucleus is very limited. Epifluorescence microscopes, such as zero resolution flow cytometers provide a time-efficient method for estimating overall fluorescence level, but cannot distinguish foci [14].

This thesis aims to produce a method for foci detection using confocal microscopy and reconstructed 3D images for better documentation and spatial evaluation. Flow cytometry was used simultaneously to confocal imaging to detect foci for comparison of the methods. The foci detection was automated to decrease bias, and batch processing was applied to images for time-saving purposes. The method provides a procedure for dealing with intensity variations so that interexperimental data can be compared. Furthermore, all software and programs are free, accessible, and adaptable. The method was developed and tested on data obtained through six experiments. The experimental goal was to gain greater insight into the surviving mechanisms of A549 cells by measuring damage extent at different time points after x-ray irradiation. The spatial distribution of DSB foci was also quantified using Moran's I and clustering algorithms and analysed to see if it could reveal additional insight into the radiation impact.

This thesis starts by explaining the principles and theory needed to understand the methodological procedure and the following results, in chapter 2 (Theory). The theory covered are cell biology, radiation physics and radiobiology, followed by principles of fluorescence imaging and lastly the theoretical foundation for the analytical tool used. In chapter 3 (Method), the assays for sample preparation are explained along with details of image acquisition procedure and analysis. The results of methodological development and experimental findings are presented in chapter 4 (Results). In chapter 6 (discussion), the experimental results are discussed, and the method is evaluated in contrast to alternatives. Lastly, in chapter 6 (conclusion), the findings are summarised.

2 Theory

2.1 Cell biology

2.1.1 DNA

DNA holds all genetic information of the organism. This large molecule consists of two complementary strands organised in a helical formation. From strand to strand, the DNA helix structure is approximately 2 nm broad [15]. Alternating sugar and phosphate groups serve as the strand foundation, which is bound to bases. There are four different bases arranged in unique sequences along the strand specifying all genes. The bases are divided into pyrimidine and purines based on ring structure, single- and double-ring, respectively. Pyrimidine includes complimentary bases *thymine* (T) and *cytosine* (C). Purines consist of complimentary *adenine* (A) and *guanine* (G). DNA coils around histone proteins into a compact formation. Five families of histone proteins serve as building blocks for organising DNA, where four of them are core proteins. Amongst core histone proteins, H2A plays an essential role in initiating connectivity, giving rise to the structure and characteristic shape of chromosomes. Compacted DNA takes the form of a chromosome, as illustrated in figure 1.

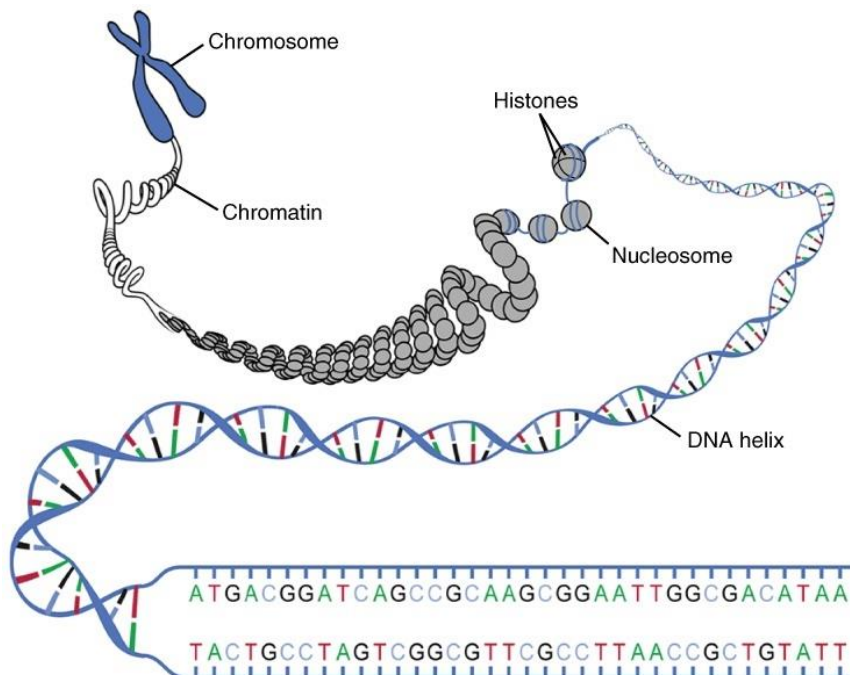


Figure 1 Organisation of the genome. Genes are specified by base sequence. DNA strands are coiled up and structured by the help of histones into chromosome structures [16].

2.1.2 Cell cycle and control system

Cell proliferation sustains life in a multicellular organism by producing new cells to replace old or damaged ones [15]. The cell cycle takes care of proliferation and denotes the series of steps

a cell undergoes in order to duplicate itself. The objective is to duplicate genetic material accurately and segregate all cell constituents evenly between the two daughter cells. Figure 2 illustrates the process with its central division between *interphase* and *mitosis* along with further subdivisions. Interphase prepares the cell for division by duplicating organelles and produce protein building blocks, while mitosis divides the chromosomes and cytokinesis divides the cells. The essential task of duplicating DNA in the chromosome occurs in the *synthesis phase* (S-phase) of the interphase. A new DNA strand is synthesised using antecedent DNA strand as a template. The process strives to be faultless to avoid introducing errors to progeny. The replication process is reliable, only making mistakes every 10 billion base-pairs [17]. Gap phases before and after the S-phase takes care of other preparatory tasks such as cell growth and monitoring of external and internal conditions. Environment evaluation is especially crucial during the *G1* phase. During unfavourable division conditions, the cell might enter resting-phase, also denoted *G0* taking, no further action before a situational change occurs. If the environment is favourable for division, the cell will start proliferation. The first task is organising daughter nuclei by distributing chromosome copies evenly between them. Cytokinesis completes mitosis by segregating the cell into two identical progenies. Duration of the mammalian cell cycle is approximately 24 hours, where the S-phase consume 8-10 hours of it [18].

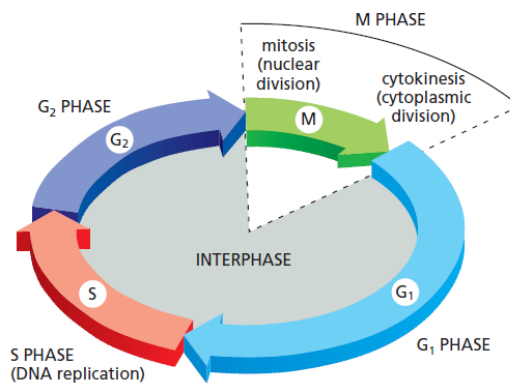


Figure 2 Schematic of cell cycle phases. Interphase is the preparatory phase before mitosis and cell division [15].

The control system orchestrates the initiation of all events in the cell cycle [15]. Commencement of proliferation depends on the environmental conditions, and whether there is a need for new cells. The main actors of control are various types of *cyclin-dependent kinases* (Cdk). Cdk are activated by cyclins which undergo a cycle of synthesis and degradation through the cell cycle. Activated Cdks drive the process forward. Erroneous continuation of proliferation can have detrimental consequences for the organisms. Therefore, the activation of Cdks requires multiple steps. Checkpoints after vital phases evaluate if the conditions are adequate for progression. The control system regulates its action according to extra- and intracellular signals that reveal environmental and cellular conditions, respectively [15].

2.2 Radiation physics

Radiation takes the form of a particle or waves able to transport energy without any transporting medium. There are different types of radiation with distinct characteristics [19]. Radiation is categorized by its ability to ionize. Ionizing radiation transmits enough energy to counteract bonding energy between electrons and the nucleus and thereby ejects an electron from the atom. Generally speaking, ionizing radiation's quantum energy exceeds the ionization potential (IP) of an absorber. Non-ionizing radiation holds less energy per photon than the ionization potential of the absorber. Non-ionizing radiation has insufficient energy to ionize, but can excite electrons. Within ionizing radiation, there is a further distinction between direct and indirect ionization [19]. The first group consist of charged particles such as protons while the latter constitute neutrons and electromagnetic radiation such as X-rays. Directly ionizing radiation deposits energy through coulomb interactions between charged particles and orbital electrons of the target atom. Direct ionizing radiation deposits energy through a one-step process, while indirect ionizing radiation first releases a charged particle in the absorber which then goes on to deposit energy in a directly ionizing manner.

X-rays and gamma rays are photons with identical properties only distinguishable by their origin. X-rays are produced from electrons either undergoing de-excitation or deceleration within an electrostatic field. The processes give rise to characteristic X-rays and bremsstrahlung radiation, respectively [19]. Gamma rays are equal in nature to X-rays, but originate from an atomic nucleus or an annihilation reaction. Both forms of radiation have *energy* (E), *frequency* (f) and *linear momentum* (p) related by the equations 1 and 2 below, where h is Planck's constant and c the speed of light.

$$E = hf \quad (1)$$

$$E = pc \quad (2)$$

2.2.1 Radiation interaction with matter

There are three main mechanisms in which an X-ray photon can interact with matter, namely *photoelectric absorption*, *Compton scattering* and *pair production* [19]. When X-rays interact with matter, three possible events occur depending on radiation energy and absorbing material. Through the photon trajectory, it might not interact with the matter at all. It might collide with a particle across its path and deposit some or all of its energy. The impact will lead to a change of course denoted as scattering.

Photoelectric absorption denominates the process of photon departing all its energy to an electron and ionizing it. The energy transmitted releases the electron and provides the photoelectron with kinetic energy. The probability of this phenomenon occurring depends on the incident photon energy and the atomic number of the absorbing material. There is a higher frequency of photoelectric absorption at higher atomic number and on the lower energy range, approximately at 0-0.5 MeV [19].

For clinical X-rays, Compton scattering is of particular importance. The process consists of a photon colliding with an electron and depositing a fraction of its energy and momentum. The photon continues with lower energy (higher wavelength) and results in new Compton processes until finally ending with photoelectric effect. The transfer of kinetic energy is sufficient to ionize the atom. The probability of this interaction lessens with higher photon energy and increases with the number of available electron targets. Biological tissue consists mostly of water and other elements with a low atomic number. For these materials, Compton scattering is the dominating interaction mechanism in the energy range of 100 keV - 30 MeV [19], which overlaps the energy range applied in medical use.

X-rays (or gamma) create an electron and positron pair when passing close to an atomic nucleus if the photon energy exceeds 1.022 MeV [19]. The process is called pair production and requires high photon energy. During the near nucleus trajectory, the photon interacts with the electromagnetic field resulting in a pair creation that conserves both energy and momentum. Shortly after creation, the product particles will combine through annihilation and produce two new gamma photons. The likelihood of pair production events increases with increasing atomic number of the nucleus as that results in a larger electromagnetic field. Pair production is the dominating mechanism for high photon energies.

All forms of radiation utilize charged particles (CP) in interaction with the material at some point. CPs are encompassed by their electric field and loses energy by the interaction of this field with charged electrons or nuclei in the matter [20]. During the trajectory of heavy CP both nuclei and electron in target material feel the coulomb forces, but linear momentum transfers almost exclusively to the lighter electron through the impulse exerted by the CP [20]. A collision with a nucleus is only likely at high initial energy for the charged particle and plays a small role in the general energy loss. The energy loss per interaction is relatively small, resulting in many depositions before the charged particle slows down. The influence in terms of deflection is negligible so that the charged particle's trajectory is approximately a straight line. Berthe-Block formula describes the energy loss per unit length [20]. A key feature of energy loss of charge particles is that fluctuations in rate averages out, resulting in a definite probability of radiation range which depends on energy, mass, charge and target medium. The energy loss rate increases with decreasing energy, resulting in higher ionization density as the particle slows down toward the end of its range. The maximum energy deposition rate at the final distance is a distinct characteristic of charged particle interaction named Bragg peak, illustrated by the red curve in figure 3.

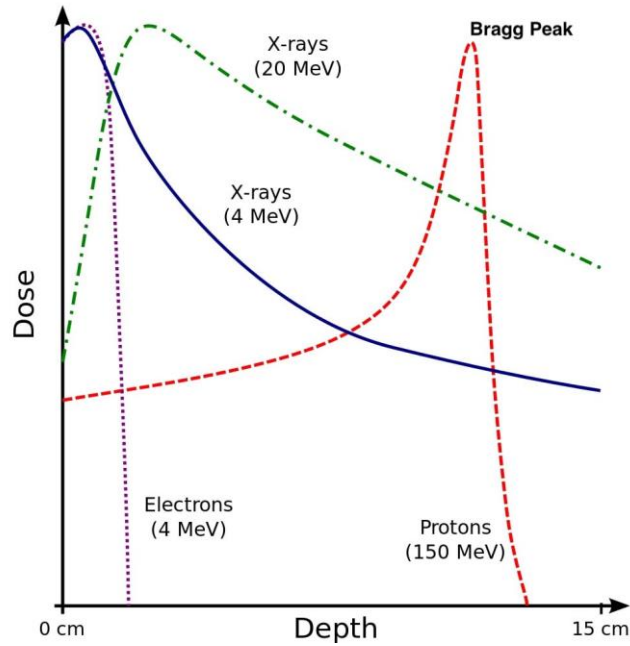


Figure 3 Dose deposition as a function of distance for different radiation forms. The red stipulated curve displays the Bragg Curve of protons [21].

The light mass of electron alters the manifestation of the interaction mechanisms, and since X-rays produce freed electrons these differences are described below. At a given energy, electrons have a much higher speed than heavy charged particles—greater speed results in less time for electric interaction and hence a lower energy loss rate. Without collision, the electron retains its energy longer, enabling it to penetrate deeper. During an electron-electron collision, the incident electron deposits a greater fraction of its energy than a heavy charged particle. More often than not, the electron reaching the final range is not the incident electron. Electrons are subjected to substantial variance in energy deposition and deflection, resulting in a less defined range. Deflection degree coupled with high speed induces emission of Bremsstrahlung radiation at a higher rate for electrons than other heavier charged particles [20]. Energy loss through collision is consistent for electrons across energy range and material. However, energy loss through radiation depends on the atomic number of the target material and increases substantially with the energy of the electron [20].

2.2.2 Linear Energy Transfer

Linear Energy Transfer (LET) expresses the average energy deposited per unit distance [22]. Equation 3 gives the LET expression,

$$LET = \frac{dE}{dl} \quad (3)$$

where dE is average energy in units keV and dl is distance in μm . Though energy deposition varies along the radiation track, an energy or track averaged LET is a useful measure of radiation quality [22].

2.2.3 Relative biological effectiveness

Relative biological effectiveness (RBE) is a means to compare two radiation forms by their biological impact [22]. The comparison consists of the ratio between a test-radiation dose and a reference-radiation dose that results in equivalent biological outcome. The RBE formula is provided in equation 4.

$$RBE = \frac{D_{ref}}{D_r} \quad (4)$$

Here D_r is the dose of radiation desired for comparison with the reference radiation with dose D_{ref} .

2.2.4 X-ray production

X-ray tube produces X-rays by accelerating electrons towards metal and inducing Bremsstrahlung radiation [23]. The cathode releases electrons and an ample potential difference between the cathode and the anode accelerates and directs the electrons into a beam that bombards the anode as illustrated in figure 4. In addition to bremsstrahlung production, the incoming electrons can induce ionizations or excitation of inner electrons in the target material resulting in replacement by outer electrons accompanied by photon emission [24]. Collision energy deposition creates excess heat which must be removed from the anode by a cooling system.

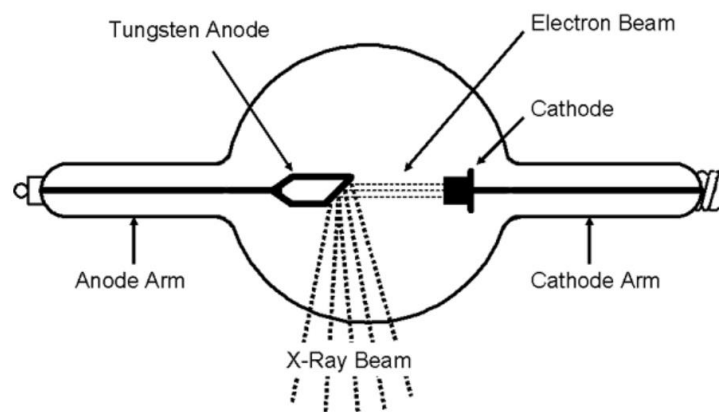


Figure 4 Basic X-ray tube setup [25].

The photons emitted by electron deexcitation result in characteristic X-rays in the beam spectrum, which disrupt the desired homogeneous energy distribution [23]. An example of a characteristic X-rays spectrum is depicted in figure 5. To decrease the characteristic X-ray peaks, filters of aluminium and copper are used, which even out the beam spectrum, but also lower the intensity.

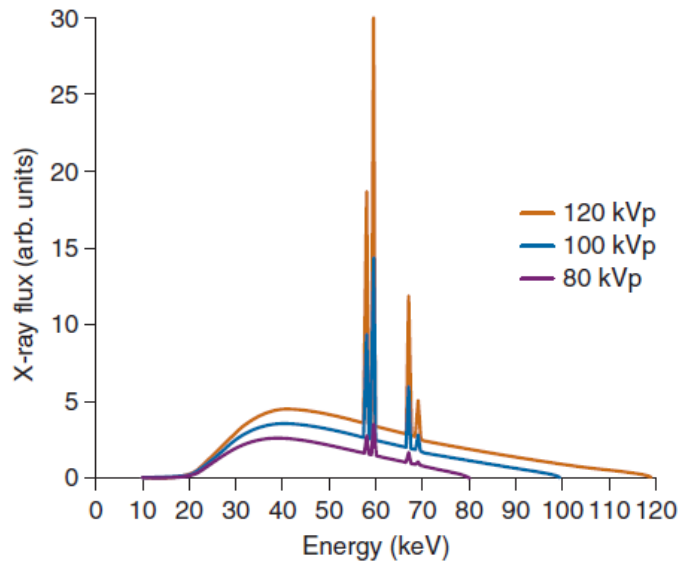


Figure 5 X-ray spectra for a tungsten anode at three different voltages. Characteristic X-rays produce the spikes in the plot. These are filtered out by aluminium and copper filters. Bremsstrahlung produce the continuous spectrum. A legend of applied voltage is placed on the right hand side, the unit kVp stands for “peak voltage in kilovolts” denoting the maximum energy of x-rays with these settings [26].

2.3 Radiobiology

2.3.1 Direct and indirect effect

The critical target in terms of biological effect is DNA, including the 3-dimensional structure as well as base sequence and genetic information. *Direct effects* denote the process of radiation energy deposition directly to an atom of the DNA structure [22]. The impact on DNA starts a chain of events resulting in biological change. Other cell constituents are less vital, but still impose a treat to DNA when irradiated if they transform into a free radical. A *free radical* is an atom or molecule with an unpaired orbital electron making it chemically very reactive. Water radicals are of particular importance because the cells consist of 80 % water [22]. *Indirect effect* names the process of a radiation-induced free radical which diffuses toward DNA and deposits its energy there [22]. DNA damage caused by direct effects predominantly occurs with high LET radiation, while the indirect action occurs more with sparse radiation such as X-rays.

2.3.2 Radiation damage

Radiation energy deposition follows the interaction track within the target medium resulting in a local impact rather than a homogeneous distribution. Blobs and spurs denominate the extent of regional impact [22]. A *spur* contains on average three ion pairs within a diameter of 4 nm and energy of up till 100 eV. In comparison, a *blob* has roughly 12 ion pairs within a 7 nm

diameter at energy upwards to 500 eV. Both have a circumference large enough to encompass the DNA double helix and will cause several injuries to DNA if the regions overlap [22]. Add to that the indirect action of diffused free radicals and the scope of DNA harm is considerable. X-rays produce 95% spurs, while protons depositing energy at the back of the Bragg peak (high LET) produce blobs to a much higher degree [22]. The difference in deposition events accounts for some of the qualitative differences in radiation impact due to the varying complexity of lesions and the subsequent repair possibilities.

There are three damage types in which the DNA double helix can suffer, namely *base damage*, *single-strand break (SSB)* and *DSB*. The cell has an assortment of repair pathways to apply in case of injury; the selection depends on damage type and cell cycle stage [22].

Loss of one or more bases or injury to one-sided sugar-phosphate structures give rise to SSB [22]. SSB repairs correctly given that the corresponding region and complementary bases on the other strand are intact and can be used as a template. Immutable loss of DNA fragment and the base sequence information is the real treat of genomic stability. DSB arise from one particle breaking opposite strands or two independent SSB close enough in time and space to constitute a DSB. DSB cleave the chromosome leading to *chromosome aberrations* [22]. An assortment of possible aberrations is presented in figure 6. Interaction of two DSB may ensue mutation, cell death or carcinogenesis.

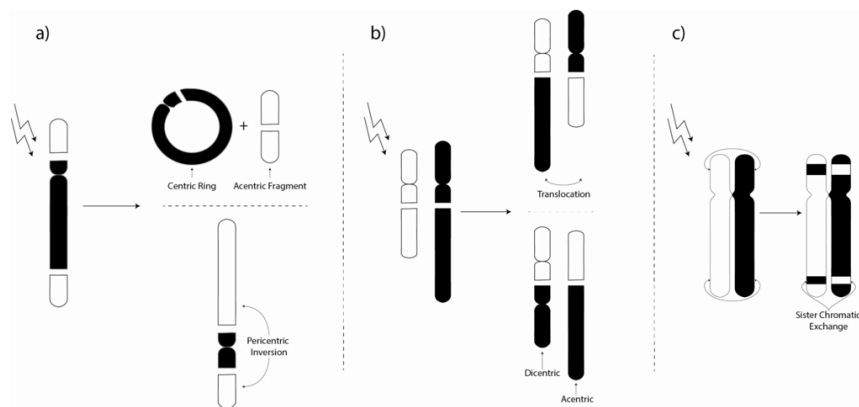


Figure 6 Examples of chromosome and chromatid aberrations. (a) intra-chromosomal aberrations, (b) inter-chromosomal aberrations, and (c) sister-chromatid exchange [27].

2.3.3 Damage repair

DSB cause a more serious treat than SSB due to reparation possibilities. *Base excision repair (BER)* fix depurination and deamination of bases by replacing them [22]. The high accuracy of this repair mechanism generates rapidly restored strands. The more complex nature of DSB makes them harder to reconstruct. Two essential repair pathways take place in restoration; *Homologous recombination repair (HRR)* and *Nonhomologous End-Joining (NHEJ)* [22]. HRR repairs perfectly by utilizing the sister chromatic as a template. Application of HRR mechanism

depends on sister chromatid availability, and thus it appears predominantly in the S/G2 phase. NHEJ mediate broken strand ends to recombine, resulting in an imperfect repair. NHEJ occurs mostly in the G0 and G1 phase, where the template from another chromatid is not accessible, but can also be active in the S/G2 phase. The success of DSB repair depends on repair pathway choice, which is influenced by cell cycle stage and other unknown factors.

In addition to repair by HRR and NHEJ, DSB can suffer two alternative destinies. The ends at strand cleavage are sticky due to unpaired bases, enabling it to reconnect with other broken and exposed DNA ends [22]. The ends can re-join with other ends caused by DSB and produce deformed chromosome after the next mitosis. This will likely put an end to proliferation and even cell death if the damage is severe. Alternatively, the ends may fail to re-join and a chromosome part is lost after mitosis.

Damages are classified according to the degree of lethality and repairing potential [22]. *Sublethal damages* (SLD) include SSB and base damages, which are not lethal if its quickly repaired. However, sublethal damages close in time and space may create *potentially lethal damage* (PLD). Cell cycle arrest can prevent the lethal outcome of PLD by providing more time for repair before mitosis. *Lethal damage* (LD) overwhelms repairing capacity of the cell and lead to cell death.

2.3.4 Radiosensitivity

Radiosensitivity describes the susceptibility of organisms and biological materials to radiation. Terasima and Tolmach studied how radioresistance varies across cell lifecycle. The experimental setup consisted of irradiating cell cultures and documenting survival fraction. Detecting cell age influence on survival required a synchronised cell culture. During mitosis, the attachment of the cells to the surface is weaker than during interphase [28, 29]. This phenomenon facilitates mitotic cell removal by use of a shaking system. The age-response curve produced by Terasima and Tolmach is represented in figure 7.

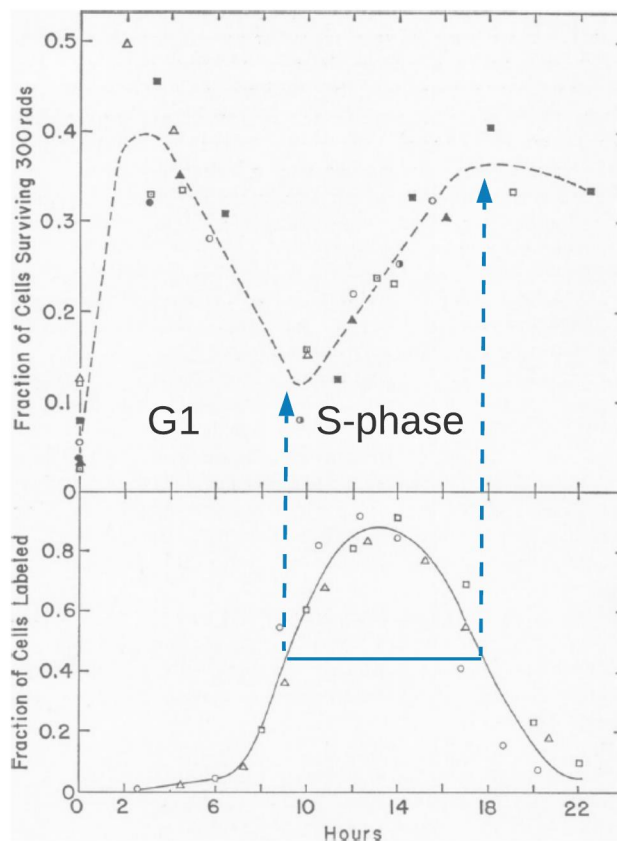


Figure 7 Age-response curve demonstrates cell survival post-irradiation as a function of cell cycle progression. This age-response curve is for HeLa-S3 cells irradiated with 3 Gy X-ray radiation at different times after mitotic cell selection (upper panel) and fraction of labelled cells denoting cell cycle phase (lower panel). Illustration is modified from [29].

The cells demonstrate prodigious radioresistance during the beginning of G1 phase compared to early S-phase. DSB at DNA synthesising inception leads to lethal asymmetric chromosome aberrations such as *dicentric*, *acentric* and *chromosome ring* illustrated in [figure X](#). The fatal outcome of DSB highlights the importance of repair prior to S-phase initiation. At the beginning of the G1 phase, there is enough time for repairs, and hence the cell is more resistant to radiation damages. S-phase progression increases radioresistance in accordance with DNA synthesising completion. At the end of the S-phase, all the chromosomes appear in duplicate, facilitating HRR perfect replication resulting in improved radioresistance. In mitosis, repair opportunity is minimal, making mitotic cells exceedingly radiosensitive compared to any other interphase stage.

2.4 Fluorescence imaging

Fluorescence imaging revolutionized biotechnology by providing high contrast images of cellular material. Fluorescence image acquisition depicts microscopic structures and opens up for quantification through image processing. Fluorescent labelling allows for examination of a specific compound. Fluorophores improve both sensitivity and specificity by easing target

detection and increasing signal-to-noise ratio [30]. A commonly used labelling probe is antibodies due to their selective binding. Labelling which highlights the desired substance coupled with high-resolution documentation enables in-depth studies of biological material.

Fluorescence imaging achieves exceptional contrast by a focused collection of emitted light from fluorophores. A fluorophore becomes excited by absorbed light and emits lower wavelength light when de-excited. Both incident and returning light have a wavelength within a specific range. Filters are used for wavelength selection, ensuring that only the intended light passes. An excitation light source, usually a mercury lamp or laser, passes through an excitation filter suited for the absorption of the specific fluorophore in use. Returned light passes through an emission filter differentiating weak and intense light. At the heart of the microscope's working principle is directing, focusing and collection of light. Different lenses cooperate in orchestrating the light, including *objective-*, *collector-* and *condensing lenses* as with any microscope [31]. The *numerical aperture* (NA) designates accepted angles for subsuming and emission and is an essential lens characteristic. *Dichroic mirrors* segregate light by letting some light pass through while reflecting others, depending on their wavelength. Finally, photodetectors convert light into a digital image, which can be examined as is, or further enhanced by image processing.

A few features of fluorescence microscopy require intricate consideration when planning the setup and procedure. Each fluorophore studied necessitates a separate imaging channel for optimized parameter tuning according to its nature. Use of multiple fluorophores gives the risk of *bleedthrough*, where the channel detects an undesired fluorophore along with the designated one [32]. Filters optimized for a specific fluorophore reduce the risk of bleedthrough by channelling out unwanted wavelengths. Additionally, choosing fluorophores with disparate emission spectra ease the filtering task. *Photobleaching* denotes permanent degradation of the fluorescence ability cause by photochemical alterations [33]. This process naturally occurs during excitation light exposure and manifests itself by intensity reduction. Preventative actions against photobleaching include reducing laser intensity and exposure duration as well as minimizing access of oxygen scavengers by proper sealing of the specimen. Light can also induce photochemical damage to the biological tissue [33]. Example of such damages includes protein denaturation and loss of enzymic activity. Knowledge of biological impact from light is essential both for acquisition and evaluation of results.

2.4.1 The histone protein H2AX

H2AX plays a vital role by marking damaged sites in DNA and initiating responsive action [34]. H2AX is a member of the H2A histone family which forms chromatin and is therefore present ubiquitously in the genome. A DSB leads to activation and recruitment of *ataxia-telangiectasia mutated* ATM and other kinases to phosphorylate H2AX at Serine-139 and (called γ -H2AX) [2]. γ -H2AX kickstarts the downstream signalling pathway of repair, and γ -H2AX foci are used as a marker of DSBs [2].

Preliminary description of Serine-139 specific antibodies by Rogakou in 1998 opened up for γ -H2AX and hence DSB detection [34]. Antibodies are proteins with a structure adapted to bind to suitable targets, antigens [15]. Onto the primary antibody attached to γ -H2AX, a secondary antibody with fluorescence properties is attached, allowing for detection with applications such as flow cytometry and confocal microscopy.

2.4.2 Confocal microscopy

Confocal microscopy can reconstruct a virtual plane section several micrometres downwards into the specimen where conventional microscopes would require physical segmentation [30]. The strength of confocal configuration is the focus to a single spatial point both for illumination and information collection of returned emitted light. Figure 8 illustrate the beam path. Focusing of the excitation beam through a microscope objective facilitates precise regional impact with a diameter as small as $0.5 \mu m$. The microscope objective also narrows the returning emitted light. Constricted sample illumination reduces out of focus light and hence reduces image noise. In contrast to a traditional microscope, the returning light is projected before viewing. Scattered light obscures the image by contributing to a pixel which does not reflect the light's origin. A small pinhole aperture allows only linear trajectory light to pass through for detection. Located behind the pinhole is a *photomultiplier tube* (PMT) which detects and multiplies the input. The image is built a pixel at the time by scanning the corresponding focal point in the specimen. The focal point is moved across the selected area of the specimen, scanning all spots within the frame and building the image pixel by pixel. Furthermore, the focal point can be varied in-depth, allowing virtual-sectioning and consecutive 3D reconstruction.

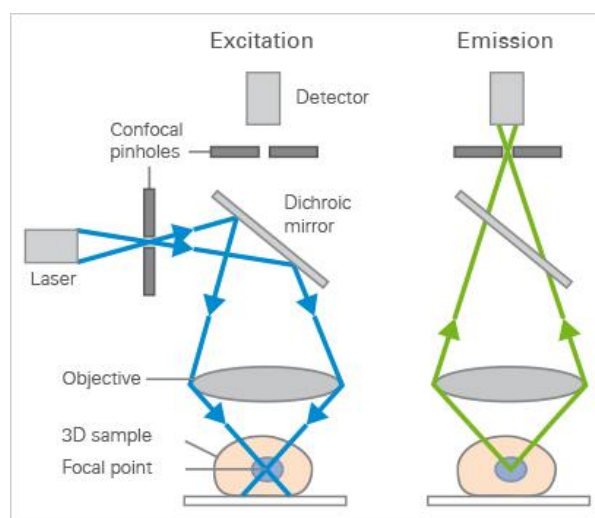


Figure 8 Diagram of beam path following excitation (left) and emission (right) for a confocal microscope [35].

2.4.3 Flow cytometry

Instead of image creation, flow cytometry records physical information of particles as they flow in a stream [36]. The system requires suspended particles so that any tissue must be dissolved into a solution beforehand. The fluid system guides particles in a stream that

intersect with the trajectory path of beams. Beams consisting of different lasers illuminate the particles resulting in excitation light as well as scattering. Surrounding the fluid stream are optical filters, which direct the output signal to designated detectors for recording. Flow cytometry records information about particle size, granularity, complexity and of course fluorescence intensity [36]. The information is analysed to present relative statistics about the whole population within the fluid. The method is accurate and time-efficient, but does not provide a detailed insight into the internal conditions of the cell as confocal microscopy does.

2.5 Analysis of confocal microscopy images

This section will explain the theoretical foundation of applied processing and analysis techniques to acquired images by confocal microscopy. The goal is to quantify DSBs and obtain spatial information of radiation-induced DSB in cells.

2.5.1 Object detection

The *Fiji (Fiji Is Just ImageJ)* [37] plugin *3D-object counter (3D-OC)*[38, 39] can be used to extract DSB location within an image along with other physical properties such as volume. Segmentation by threshold differentiate pixels into object pixels and background must be determined and configured manually. As this process does not consider noise, the monomodal nature of these images makes the threshold-setting a critical decision. The outcome of 3D-OC is very sensitive to the threshold setting, even a small threshold alteration will influence the output directly.

The 3D-OC runs connectivity analysis two times over the image stack. A connectivity analysis is the process of detecting all object pixels (having pixel value above the threshold value) and collecting connecting pixels into a single object. Each object is given a number as label. For each pixel in an object, the adjacent pixels are checked if they classify as object-pixel (having pixel value above threshold). If the neighbouring pixel classify as object pixels they are included in the original object and given the same label. Adjacent pixels include 4 pixels on the current image and 9 pixels at the same location in the previous image in the image-stack. After the entire image-stack is analysed for objects, the connectivity analysis is conducted again to combine objects that are diminutive connected but marked as two independent objects during the first connectivity analysis.

2.5.2 Image characterization using Moran's I

The Moran's I algorithm developed by statistician Patrick A. P. Moran measures spatial autocorrelation [40], i.e. quantifies the similarity of observations at different locations. Moran's I is not commonly applied outside the field of geography.

The tool assesses values $x(i,j)$ within the image x at the location denoted by i and j as well as their spatial relation. \bar{x} is the mean of image x . The weight matrix W indicates proximity between regions and scales, where $w_{ij} (\geq 0)$ is the entry corresponding to image value x_{ij} . A

common strategy is weighting neighbours with 1 and all other pixels with 0. Neighbourhood size varies depending on the problem at hand. Equation 5 presents the mathematical formulation of Moran's I, where N denotes the number of elements (pixels) within the image.

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (5)$$

Normalization causes the resulting Moran's I to be a number ranging between -1 and 1, in theory at least. Some studies indicate that the maximum value depends on the weight matrix and its dimensions [41]. Nevertheless, negative values specify clustering, positive dispersion and approximately zero a random pattern as illustrated in figure 9.

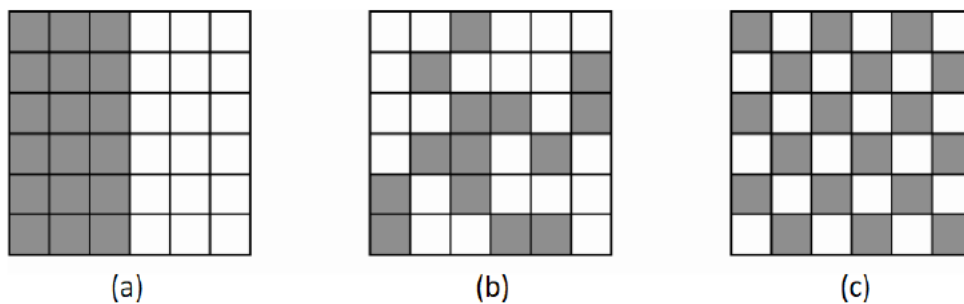


Figure 9 Pattern characteristics indicated by Moran's I. (a) displays complete clustering corresponding to a Moran's value of -1. The pattern in (b) is random and would be indicated by a Moran value close to 0. (c) has a perfect dispersed pattern which would result in a Moran value of -1 [42].

2.5.3 OPTICS Clustering

Clustering algorithms aim at revealing the internal structure of a dataset [43]. Decomposing data into meaningful subclasses sharing some characteristics is a challenging undertaking. There are many viable options of clustering algorithms, *Density-based spatial clustering of applications with noise* (DBSCAN) being one of them. The different algorithms offer alternative approaches, but all require parameter input. Dependence on parameter input raises some issues; they are hard to determine, and the choice substantially influences the outcome. Furthermore, real-world data are complex often with a skewed trait which cannot be accurately decomposed using a single global parameter. *Ordering Points To Identify the Clustering Structure* (OPTICS) [43, 44] omits input parameter complications by ordering the data to represent density-based clustering.

OPTICS is an augmented extension of DBSCAN, so DBSCAN is a natural starting point for explaining the theoretical foundation of this new algorithm. The defining parameters DBSCAN operates with is the maximum radius of the neighbourhood (ϵ) and the minimum number of points within the given neighbourhood ($MinPts$). These parameters define core points and border points. A *core point* satisfies the condition of having more than $MinPts$ number of points within neighbourhood ϵ of itself, while a *border point* has fewer than $MinPts$ number

of points within neighbourhood ϵ of itself but is still within ϵ of a core point. Noise is any point that does not classify as either a core- or border point. The liability of DBSCAN is that a highly dense cluster can completely absorb a less-dense cluster for a given *MinPts*. Figure 10 illustrates the problem. One set of parametric input can reveal either cluster A, B and C, or cluster C1, C2 and C3.

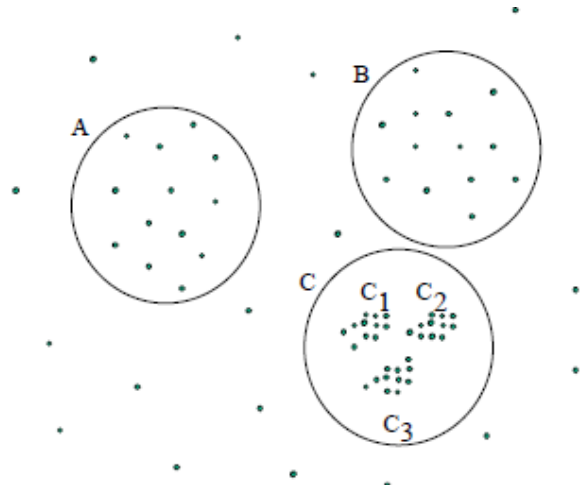


Figure 10 DBSCAN would detect either the groups A, B and C, or C1, C2, C3 depending on paramteric input [43].

OPTICS resolve this issue by sequential density clustering starting with higher-density groupings. The algorithm stores cluster-ordering information by two parameters, namely core distance and reachability distance. Their mathematical expression is presented below.

Core distance of an object p : Let ϵ be the smallest radius so that the resulting neighbourhood $N_\epsilon(p)$ has at least *MinPts* number of objects, where *MinPts* is a natural number. $card(N)$ denotes the cardinality of set N .

$$Core_distance_{\epsilon, MinPts}(p) = \begin{cases} \text{undefined if } card(N_\epsilon(p)) < MinPts \\ MinPts - distance(p), \text{ otherwise} \end{cases}$$

Reachability distance of an object p from core object o :

$$Reachability_distance_{\epsilon, MinPts}(p, o) = \begin{cases} \text{undefined if } o \neq \text{core object} \\ \max(core - distance(o), distance(o, p)), \text{ otherwise} \end{cases}$$

The core distance of an object p is the shortest ϵ so that the corresponding neighbourhood contains at least *MinPts* objects. Reachability distance from object p to the core object o is the

minimum radius making p density-reachable from o . Plotted reachability distance reveals intrinsic grouping as displayed in figure 11. Members of clusters have low reachability to the nearest neighbour. This trait is apparent in the plot by deep valleys. The more profound valleys signal denser clusters.

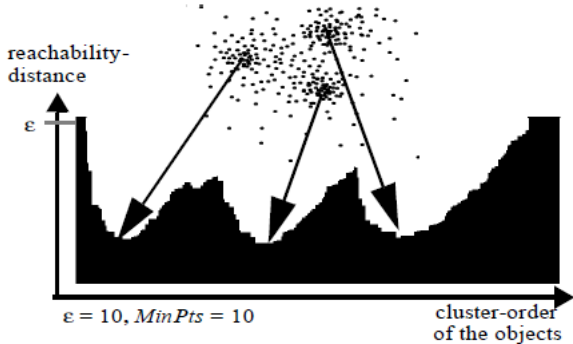


Figure 11 Reachability plot reveal group density by valley depth [43].

OPTICS resolves DBSCAN issues by cluster-ordering the database according to density. The information output is equivalent to density-based clustering algorithms for a wide spectrum of parameter settings [43]. The algorithm allows for both automatic and interactive analysis and can detect hierarchically nested clustering structures.

3 Method

3.1 Experimental process and setup

The main goal of these experiments was to establish a method for quantifying the amount of DSB and their distribution within the cell. Furthermore, the impact of radiation dose and time-development of DSB foci were examined. All experiments followed one of two assay pathways and all analysis paths are outlined in figure 12. All experiments underwent four stages, namely preparations, treatment, image acquisition and analysis. This chapter provides in depth details of all experiment steps.

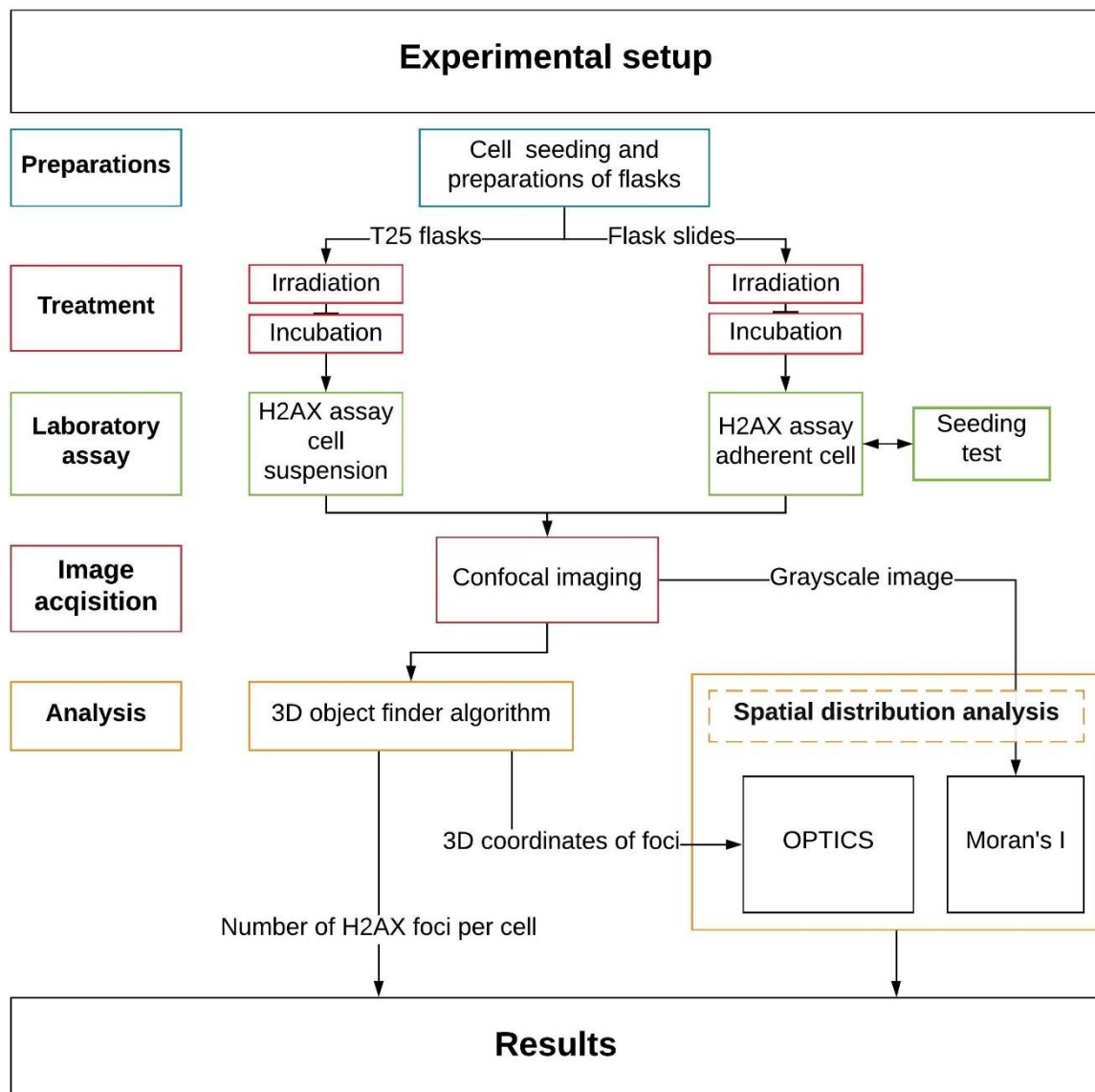


Figure 12 Flowchart outlining the methodology processes. Experiments were conducted using either H2AX cell suspension assays or H2AX adherent cell assays. The other methodology pathways were principally alike for both assays, except for some adjustments made along the way. All alterations and adjustments will be presented in this chapter. Illustration by author.

3.2 Cell line

Natural shortening of DNA telomers prohibit cells from proliferating indefinitely. Mutations in cancer cells allow them to bypass this aging effect and immortalize them in terms of cell division [45]. Immortalized condition can be achieved either naturally or by manipulation, either way it allows consistent experimentation with the same cell type through subculturing.

The experiments were conducted with cells from the A549 cell line obtained from American Type Culture Collection (ATCC). The cell line stems from human lung carcinoma that is derived from a 58-year-old Caucasian male [46].

3.3 Cell line cultivation

3.3.1 Chemicals Equipment and aseptic technique

The following materials were used throughout the laboratory work. Any additional equipment and chemicals applied during experiment will be noted under the designated section.

3.3.1.1 Chemicals

Sterile filtered DMEM F-12 medium (Lonza, Belgium) primed with 10% fetal bovine serum (Euroclone, Devon, UK) and 1% penicillin/streptomycin (Eu-roclone, Devon, UK) was used to provide cells with necessary nutrition and protection against infections. Trypsinisation detach cells from the flask bottom by cleaving the amino acids that anchored the cells to the surface [47]. The trypsin used was Trypsin-Versene Mixture (Lonza, USA). PBS Phosphate-buffered saline (Lonza, Belgium) was used to purify cell colonies from other chemicals such as excess medium. Equipment and the workspace was disinfected with 75% ethanol Antibac overflatedesinfeksjon (Antibac, Norway) and LAF-benck with additional Rely+On Virkon (DuPont, United Kingdom). Cell medium, PBS and trypsin were used throughout all experiments.

3.3.1.2 Equipment

Cell colony were subcultured and seeded into 25 cm² sterile vent/close cap flasks (Thermo Fisher Scientific Nunc A/S, Denmark) denoted T25 in figure 12. Protocols were administered using disposable plastic pipettes (Sarstedt, Germany) with electrical handles (pipetus-akku Hirschmann Laborgeräte, Germany) or rubber bubbles. Cell counting was done by transferring a solution sample by 20-200 µL Pipet-Lite XLS Single Channel Manual Pipette (RAININ, USA) to a Bürker chamber (KOVA, USA) and evaluated under an optical microscope or Norma XS (iPrasene, France). A 10x magnification microscope (Nikon TMS, Japan) was used to examine cell condition and progress of single cell suspension as well as cell counting by Bürker chambers. The cells were stored in incubators for optimal living conditions holding 37 °C, 80% humidity and 5% CO₂. Two Steri-Cult 200 CO₂ incubators (Forma Scientific, USA), and one Thermo Forma Series II, Water Jacketed CO₂ Incubator (Forma Scientific, USA) were used.

3.3.1.3 Aseptic technique

In order to prevent contamination, all samples were handled within a Laminar flow cabinet (LAF-bench), either Class 100 Laminar Air Flow cabinet (Gelaire, Australia) or VB 2040 Laminar Air Flow cabinet (Odd A. Simonen, Norway). Prior to use, the LAF-bench and all equipment was sterilized with ethanol. The LAF-bench was sterilized again after use.

3.3.2 Cell cultivation procedure

The cell line A549 has adherent culturing properties, meaning that cells stick to the container bottom. Given optimal environment the cells will proliferate until the entire surface is covered. The level of cell coverage is referred to as confluency. High confluence induces contact inhibition which will bring cell growth to a standstill as eventuated [48]. Two times per week a fraction of the cells were transferred to another contained in order to preserve undisturbed proliferation. Every Monday and Friday the population was reduced and subcultured according to proliferation rate. Perpetuating healthy cells also requires replacing old medium with new. Medium was refreshed every Wednesday and under each subculturing. Both tasks were conducted by the laboratory engineer at the Biophysics and Medical Physics Cell Laboratory at the Department of Physics, University of Oslo (UiO).

In the sterilized environment of a disinfected LAF-bench, new flasks were prepared by labelling and adding 5 ml of fresh medium. Flasks containing the cell culture were transferred to the LAF-bench and rinsed twice with 1.5 ml of trypsin. Depending on the temperature and freshness of the trypsin, the cells were then incubated from 1-5 minutes. The flask was then lightly smacked against the counter to help loosen the cells and was then examined in the microscope to verify the intended effect. Medium was added to the cells and gently aspirated to obtain an even solution with dispersed cells. The amount of medium corresponded to degree of dilution. For the cell line A549, dilution calculations were based on a population size factor of 2.3 per day. Lastly, the required number of cells were transferred to a new flask and incubated.

3.3.3 Seeding

Cell cultures needed for experiments were ordered from the cell laboratory and prepared 1-3 days prior to the experiment. The procedure was alike the one for cell cultivation, except that the degree of dilution was adjusted according to pending radiation dose and incubation time before harvest.

3.4 Cell treatment

3.4.1 X-ray irradiation

Irradiation was performed with an X-ray unit PANTAK PMC 1000 (Pantak, USA).

3.4.1.1 Set up

The X-ray unit is connected to a water bath which supplies the chamber shelf with warm water in order to keep the environment at steady 37 °C during irradiation. The irradiation chamber shelf is a moveable plate allowing to adjust the height placement within the chamber and thus changing the dose rate. A schematic overview of the X-ray unit is displayed in figure 13. The irradiation beam was filtered through 1.52 mm aluminium and 0.70 mm copper. The dose rate at 60 cm Source Surface Distance (SSD) was calculated by Anne-Marit Rykkeliid following the principles of IAEA TRS-398 and the papers of Waldeland [49] and Rosser [50]. The settings used throughout were 220 kV and 10 mA.

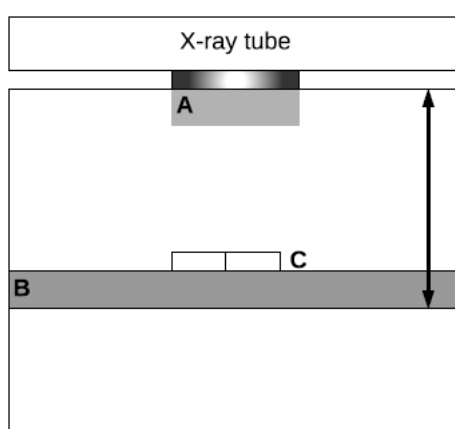


Figure 13 Schematic illustration of the X-ray machine used throughout this work. **A** denotes the exit of the radiation beam. **B** is the adjustable plate (symbolised by the arrow) which is heated to keep the cells at 37 °C during irradiation. The plate height used is 40 cm from source (**A**) (SSD 40) and 60 cm from source (SSD 60). **C** is the flask placeholder centered in beam. Illustration by author.

3.4.1.2 Irradiation procedure

Prior to irradiation, the cells had been prepared and incubated with the appropriate amount of medium. Preparations include seeding and marking flasks with replica identification and treatment to be given. On the day of irradiation, the flask lids were closed and transported to Rotgen lab in a disinfected styrofoam box. The flasks were placed in a holder and centred in the middle of the X-ray chamber for most even dose distribution. The holder was uniquely designed by Efim Brondz, University of Oslo, to hold the flask used and is made of PMMA-material. Samples were irradiated in batches of four and immediately returned to the incubator after treatment. The control samples were brought along with one of the batches to mimic the treatment's environment conditions. The duration of irradiation was calculated according to dose rate.

3.4.1.3 Dosimetry

The dose rate at 40 cm SSD was determined by measuring the total dose delivered after 30 seconds at 220 kV and 10 mA. The measurements were taken at three locations within a T25 flask at each position in the placeholder. Figure 14 shows all four flask positions and locations within the flasks where dose was measured. The inner, middle, and outer region were all measured to test if the distribution was approximately homogeneous.

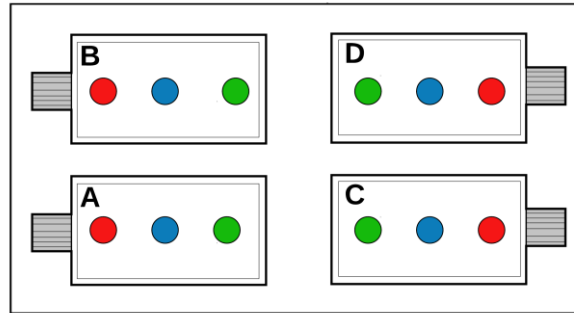


Figure 14 The figure illustrates the T25 flasks' position in the placeholder. Dosimetry was measuring on each place and each position within the flask denoted by coloured circles. The inner, middle, and outer region correspond to green, blue and red circle respectively. Illustration by author.

Ionizing radiation induces a current in an ionization chamber. The current reveals the dose with accurate calibration. IBA FC65-G ionization chamber (PTW-Freiburg, Germany) was used with a standard imaging electrometer MAX-4000 (Standard Imaging, USA) to measure the current. The dose to water is given in equation 6 [23, 49, 50].

$$D_{water} = M_u N_k K_u p_u \left(\frac{\mu_{en}}{\rho} \right)_{water,air} \quad (6)$$

Where D_{water} is the dose in water in Gy. M_u is the chamber reading measured in nC multiplied by correction factor for ambient conditions K_{tp} . M_u is given in equation 7. N_k is the calibration factor for standard ambient conditions. Factor K_u accounts for spectral distribution changes due to medium transfer from air to water. The $\left(\frac{\mu_{en}}{\rho} \right)_{water,air}$ ratio is the mass energy absorption coefficient between water and air averaged over photon spectrum at 5 cm water depth. p_u is the perturbation factor. All known factors included in calculations are listed in table 1.

$$M_u = M * \frac{273.2 + T * p_0}{273.2 + T_0 * p} = M * k_{Tp} \quad (7)$$

The correction factor K_{Tp} was determined using ambient conditions for calibration, $T_0 = 20.0$ °C and pressure $P_0 = 1013$ hPa, and ambient conditions during dosimetry $T = 30.2$ °C and pressure $P = 1022$ hPa.

Table 1 Known dosimetry factors included in equation 6.

$N_{k,air}$	K_u	$\left(\frac{\overline{\mu_{en}}}{\rho}\right)_{water,air}$	p_u	K_{Tp}
43.77±0.39	≈1	1.07	1.023±0.001	1.03

3.4.2 Incubation

After irradiation treatment the cells were kept in an incubator for a designated amount of time. The incubator holds a temperature of 37 °C and keeps the CO₂ level at 5.0 %. Duration of incubation time varied from 0.5 h to 72 h with 24 h intervals. During this time there were no medium changes regardless of duration.

3.5 Staining protocols

3.5.1 H2AX Assay using cells in suspension

3.5.1.1 Equipment

After treatment, the cell cultures was trypsinized and transferred to a 10 ml tube (Sarstedt, Australia) containing medium. Protocols where administered using disposable plastic pipettes (Sarstedt, Germany) with electrical handles (pipetus-akku Hirschmann Laborgeräte, Germany) or rubber bubbles. The cells were centrifuged using a Rotofix 32A (Hettich, Germany), mega star 600 (VWR, Germany) and mega star 600R (VWR, Germany). The samples were centrifuged at 1000 *rpm* for 4 *minutes*. However, the centrifugal force was regulated both up and down depending on the cell pallet hardness. After centrifugation, the supernatant was removed using Vacusafe (Integra, Switzerland) and vortexed using IKA MS3 digital orbital shaker (VWR, Germany). The final solution was filtered through a 03-50/37 nylon mesh (Sefar AG, Switzerland) onto microscope slides (VWR, Germany) and covered with 18x18 *mm* cover glasses (VWR, Germany). Generic nail polish was used to seal the cover glasses.

3.5.1.2 Chemicals

In addition to cell medium, PBS and trypsin, the following chemicals were used the suspension cell assay: PI Propodium iodide (Sigma, USA), BSA Bovine serum albumin (Sigma, USA), and Methanol (Sigma-Aldrich, Germany). The primary antibody used in this assay was Anti-phospho-Histone H2A.X (ser139) clone JBW301 (EMO Millipore Corp, USA) while the secondary antibody used was Polyclonal Rabbit Anti-Mouse Immunoglobulins/FITC (Dako, Denmark). Before mounting the cover slip on a microscope slide a drop of mounting medium Vectashield H-1000 (Vector, Norway) was applied.

3.5.1.3 Protocol

The protocol used was derived from another protocol provided by *Flow Cytometry Core Facility (FCCF), Department of Core Facilities, Institute for Cancer Research, The Norwegian Radium Hospital*. The protocol's procedure consisted of a series of chemical treatments with washing and centrifugation in between. The adapted protocol can be found in appendix A.1. The cell samples were fixated using methanol which dehydrates the cells. The fixation procedure includes addition of 400 ml of PBS prior to methanol. The PBS was added to ensure a slow and steady increase of methanol concentration. If the methanol concentration rises too quickly, the permeabilization occurs too fast and breaks the membrane leaving the cell to burst.

3.5.1.3.1 Protocol alterations

Post assay a drop of cell sample was placed on the microscope slide and left to dry. This method produced regions of high cell density making them difficult to distinguish in processing. To improve cell distribution on the microscope slide, the cell drop was filtered during transfer. PI signals were inconsistent between experiments. Different alterations and tests were run to locate the source of low PI signal. Different PI concentrations ranging from 1:10000 to 1:4000 were tested. Furthermore, duration of incubation with PI solution and duration of drying the sample were tested to see if they would explain the weak signal.

3.5.2 H2AX assay using adherent cells

This experimental setup was configured to include all cells seeded that underwent treatment in the final analysis. Suspension cell assay included trypsinization and post-centrifugation supernatant removal. Both these processes entailed cell loss. The adherent assay seeds cells to microscope slide prior to any treatment. The goal was to get a comprehensive look at how cells are affected by looking at all cells, not just a remaining cell-subset withstanding the procedure. A secondary motive for this experimental setup was to achieve a better cell distribution on the microscope slide and exclude difficulties with cell overlap in the z-direction. Such effects will ease the imaging process and make the image more useful and informative.

3.5.2.1 Equipment

Prior to treatment, the cell cultures were trypsinized and subcultured in SlideFlask (Thermo Fisher Scientific Nunc A/S, Denmark) containing medium. The dilution factor was determined using a cell counter Norma XS (iPrasene, France). Protocols were administered using disposable plastic pipettes (Sarstedt, Germany) with electrical handles (pipetus-akku Hirschmann Laborgeräte, Germany) or rubber bubbles. The washing solution was swirled using Rotamix 120 (Heidolph, Germany) and subsequently removed using Vacusafe (Integra, Switzerland). The SlideFlask cap was removed to mount 18x18 mm cover glass (VWR, Germany). Generic nail polish was used to seal the cover glasses.

3.5.2.2 Chemicals

The chemicals used for H2AX assay with adherent cells is equal to the assay with suspension cells. See section 3.5.1.2 for specification of chemicals.

3.5.2.3 Protocol

This protocol follows the same procedure as suspension cell assay (section 3.5.1.3) but was adapted to suit flask slides which can be seen in figure 15. All centrifugation steps were replaced by gentle swirling. Furthermore, the incubation times with chemical solutions were adjusted according to increased contact area of the adherent cells. To avoid mechanical stress to the cells, all liquids were administered to the flask slide cap and subsequently gently swirled across the cells. Washing after methanol fixation was upscaled to overcome increased surface tension caused by methanol's dehydration effect. Meticulous washing at this stage was important for PAB solution coverage. When the surface tension prohibited PAB and/or SAB solution from complete coverage of cells, the slide was lightly tapped from below.

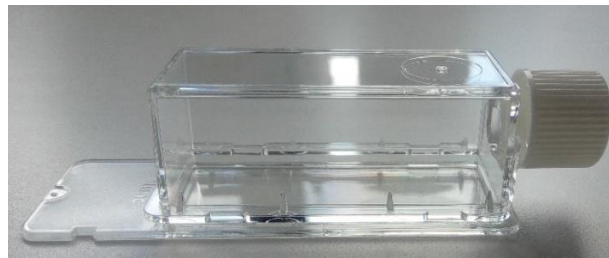


Figure 15 Image of flask slides used in the adherent cell assay. Illustration by author.

3.5.2.4 Tests conducted to determine the protocol

3.5.2.4.1 Cell seeding amount

A seeding test was performed to determine sufficient cell seeding on flask slides given the treatment dose and the longest incubation time. For each dose investigated, three flask slides were seeded with 150 000, 175 000 and 200 000 cells. Cell count was done using a Bürker chamber (KOVA, USA). After irradiation all samples were incubated for 72 hours and then stained with Methylene Blue Hydrate (Sigma, USA). The microscope slides were imaged using a Perfection V850 Pro flatbed scanner (Epson, Indonesia). To determine the confluency, the images were processed and evaluated using several MATLAB programs written by Delmon Arous, UiO. The programs are attached in appendix at B.1.

3.5.2.4.2 Fixation and permeabilization chemicals

The protocol was tested using methanol for fixation with varying freezer incubation time to see what produced the best results. Furthermore, formaldehyde (Sigma-Aldrich, Germany) was tested in combination with methanol. The full protocol can be found in appendix A.2.

3.5.2.4.3 Antibodies

The protocol was tested with two other antibody combinations and evaluated for antibody affinity and sensitivity. The conjugated antibody Anti-phospho-Histone H2A.X (Ser139) Antibody, clone JBW301, FITC conjugate SDS (Sigma-Aldrich, USA) was tested along with the combination of PAB Anti-phospho-Histone H2A.X (ser139) clone JBW301 (EMO Millipore Corp, USA) with SAB Goat Anti-Mouse IgG H&L (FITC) (ab6785) (abcam). A range of concentrations

for each antibody combination was tested. The protocol used for testing antibodies is found in appendix A.4.

3.6 Confocal imaging

A subset of cells from each sample were imaged sectionally with an increment of 0.5 μm to reproduce a 3D construction of the cell. Image acquisition was performed with a Leica TCS SP2 (Leica, Heidelberg) confocal scanner. Originally two replicates of each experiment treatment were prepared for imaging, but due to cost and imaging-duration it was reduced to only one specimen. In total 88 samples and 2704 cells were imaged throughout this project.

For the entirety of this work a 100x oil objective with NA of 1.4 was used. Each image was composed of two channels, one for DSB stained with H2AX fluorescent antibodies and another for DNA stained with PI. The channels were images in sequence to minimize bleedthrough. The images were acquired at speed 400 Hz, with a line average of 4. Table 2 displays the confocal settings used throughout all experiments. Electronic zoom of 2 or 3 was used.

Table 2 Confocal microscope settings used throughout all experiments.

Channel	1	2
Laser wavelength	488 nm	543 nm
Intensity	35 %	60 %
Detector range	497-550	560-630
Detector gain	550 %	550 %
Pinhole size	182.12 μm	182.12 μm

Frame for acquisition was chosen following a set pattern to avoid imaging the same region twice. In addition, the distance between two frames was large enough to avoid photobleaching effects from previous scanning. The imaging frames were selected to maximize the number of cells within the frame to reduce the duration of both image acquisition and analysis. An effort was made to exclude cells overlapping in the z-direction, but this was not always feasible due to cell population morphology. Cells overlapping but separate in the z-direction can be segmented and utilized but prolong the analysis duration. Cells overlapping with each other could not be utilized in further analysis as proper separation was impossible. The adherent cell assay produced samples with better cell distribution than the suspension cell assay. The spread of cells created more regions suitable for imaging, and hence locations could be sampled at random and not selected. On average each sample required five frames to image enough cell nuclei. Therefore, each sample was imaged in each corner and centre of cover glass. Furthermore, an overview image of each selected frame was acquired with software zoom 1. This image was taken to examine the environment and neighbouring cells of those cells that were analysed.

Originally 50 cells were sampled from each specimen, but this was decreased to 30 cells due to extensive duration of both image acquisition and processing. Furthermore, the software zoom was downregulated to 2 from 3 to include more cells within each imaging frame. Different pinhole sizes were tested to check if size could compensate for some information loss caused by a lower hardware zoom.

3.7 Image processing and analysis

Confocal imaging provided two images, one of each colour channel, for each section taken in the z-direction. The red colour channel originated from PI staining, meaning that the channel's signal marked the presence of cell DNA. The cell nucleus produced a strong signal with PI staining, while other cells parts produced a vague signal due to some fragmented DNA throughout the cell cytoplasm. Therefore, the red signal can be utilized to locate the cell nucleus and determine its circumference within a given section. The green channel originates from γ -H2AX fluorescence antibody staining, so its signal denotes DNA damage in the form of DSB. The presence of a DSB in a stained cell presents itself as a γ -H2AX signal focus. Processing and analysis of these images were done using Fiji along with programs created in R [51] and Python [52].

3.7.1 Image pre-processing

All pre-processing was done using Fiji. The image sequences of each channel were converted into stacks and then united by merging the colour channels into a single image. The physical dimensions were saved in the stacks to enable measurements in the images. Signals from the red channel, originating from PI, were used to locate and segment out single nuclei. Occasionally the PI signal in the red channel was insufficient to use in nucleus detection. In such cases, the red channel underwent *Li* thresholding [53, 54] prior to merging of the stacks.

The segmentation process consisted of marking a region and cropping the stack to the smallest rectangle covering the marked region. Marking was performed either by using a pre-defined shape and selecting placement and size or by drawing the shape containing the cell nucleus. The segmented region was fixed throughout the stack, and therefore drawn according to the largest nucleus section. Since the rectangular shaped stack crop covered an approximately round nucleus, the pixels in stack corners originated from the nucleus environment. To reduce non-nucleus signal everything outside of the segmented area was blackened. Since cell nucleus size varied though the z-depth and the segmented region was fixed, some parts of the image still originated from the nucleus' environment. Finally, the stack was separated into two stacks containing its respective channel and each channel was used for further analysis. See figure 16 for illustration of the pre-processing steps.

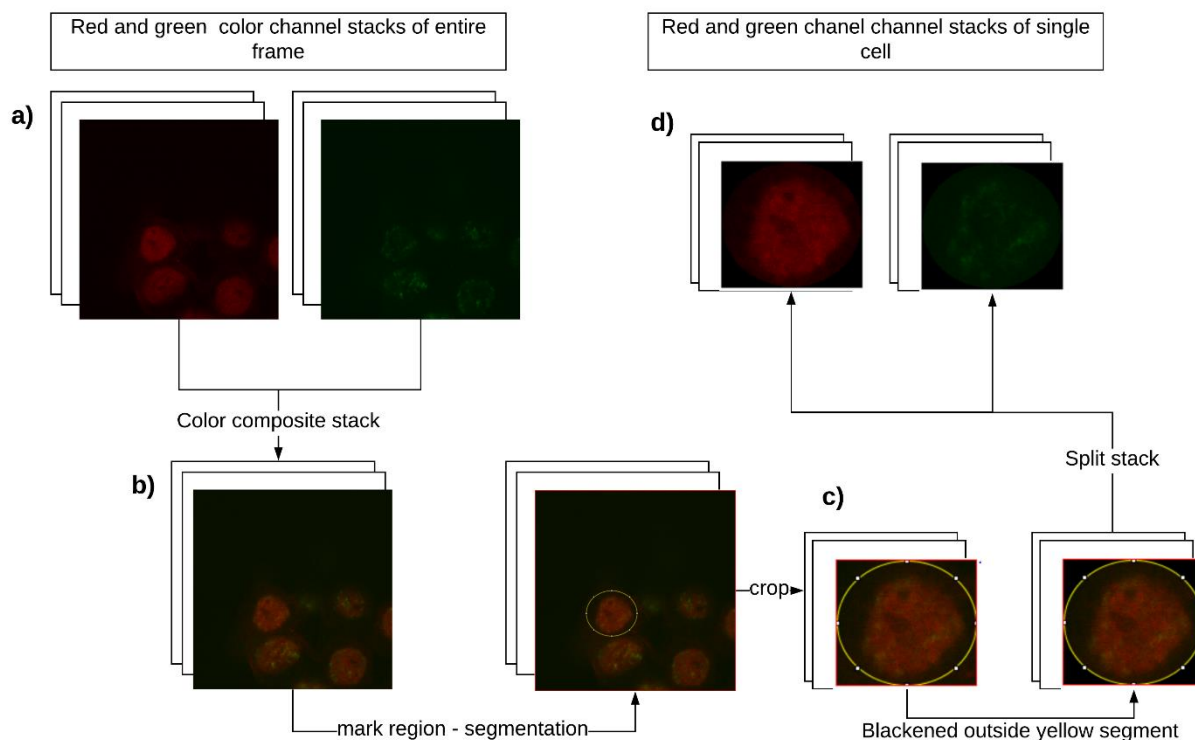


Figure 16 Pre-processing steps applied to images. The image set from a single sample was collected into two stacks, one for each colour channel (a). The stacks were merged by combining the colour channels and segmented by cropping according to selecting region (b). Pixels outside of the marked region were set to zero (c). The resulting cropped and blackened stack was split into two separate stacks containing the respective colour channel (d). Illustration by author.

3.7.2 3D object counter

The Fiji plugin *3D-OC* was used to quantify DBS foci and record physical properties of them, such as position coordinates and volume. The macro input was stacks of green colour channel images from pre-processing. For time saving purpose the *3D-OC* specifications were recorded in a Fiji macro and then utilized in batch processing for all samples within a specimen. Originally a threshold of 90 was used throughout. This threshold was determined by manually counting foci and testing which threshold yielded the equivalent result for multiple samples within the first experiment. The intensity varied between experiments, so this method did not provide a good foundation for comparison. Therefore, the stack was re-processed using the *Stack Meancenter5* plugin authored by Knut Kvaal, Norwegian University of Life Sciences.

The image was converted to 32-bit float for enhanced calculations. Each image was centred by determining its average pixel-value and subtracting it. Subsequently each pixel value was divided by the image standard deviation. The global minimum and maximum of the entire stack was determined. The global minimum was subtracted from each image in the stack. The result was an 8-bit stack with minimum value zero. The resulting image utilized the whole range of histogram values, as illustrated in figure 17.

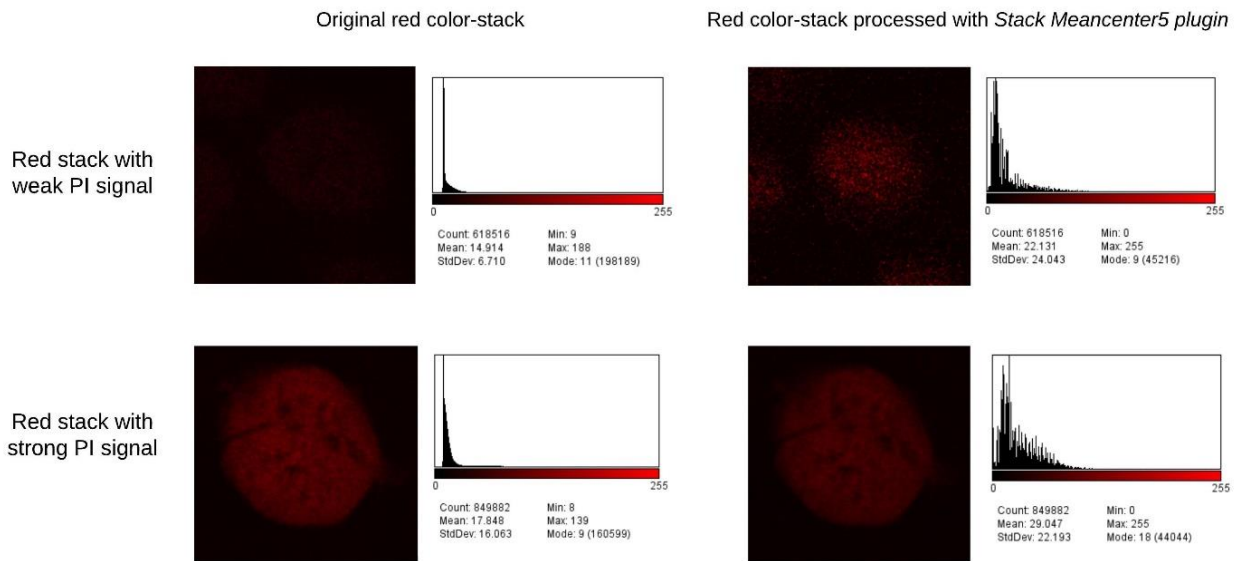


Figure 17 Stack-Image and -histogram before (left) and after (right) processing with Fiji plugin “Stack Meancenter5” authored by Knut Kvaal. The image-stack with weak PI signal is enhanced while the stack with strong PI signal is dimmed after plugin processing. The histograms reveal that maximum pixel intensity prior to processing with plugin was 188 and 139, while after it was 255. The plugin alters the image to utilize the entire pixel-range and shift distribution towards centrum of range. Illustration by author.

Originally only the number of marks was recorded with 3D-OC, but the analysis was changed to include the intensity of each focus and focus volume. These properties were extracted from the green colour channel. The red colour channel was used to determine nucleus volume using 3D-OC when the PI signal was sufficient.

3.7.3 Spatial distribution evaluation

3.7.3.1 Moran’s I

The R-programme used to calculate Moran’s I was written by Stefan Schrunner, University of Oslo, and can be found in appendix at B.2. The program input was green colour channel stacks converted into greyscale in Fiji. At its final version the programme assesses both location and pixel values within each image to determine the Moran’s coefficient. The Moran’s I is calculated under the null hypothesis of spatial randomness. The programme was expanded by author to include a significance test, to assess whether the Moran’s I was significantly clustered or dispersed. The significance test was conducted under normality assumption and assuming independent sampling. The test was carried out by comparing the z-value to the critical value of ± 1.96 at 95 % confidence interval (CI).

Originally the script input was coordinates of objects (foci) detected by 3D-OC along with outer boundaries enclosing the cell. This previous method yielded a substantial information loss by

only considering binary information, i.e. there was a DSB at a given location or not. Furthermore, the scale of a cell nucleus is extremely small from a macroscopic perspective and consequently there was little distinction of foci-coordinates. To overcome difficulties introduced by subtle differences in coordinates, cell downscaling was added to the code. Downscaling consist of divided the cell nucleus boundaries into subspaces by a grid. The value count and location of each box served as foundation for further calculations. The principle is illustrated in figure 18.

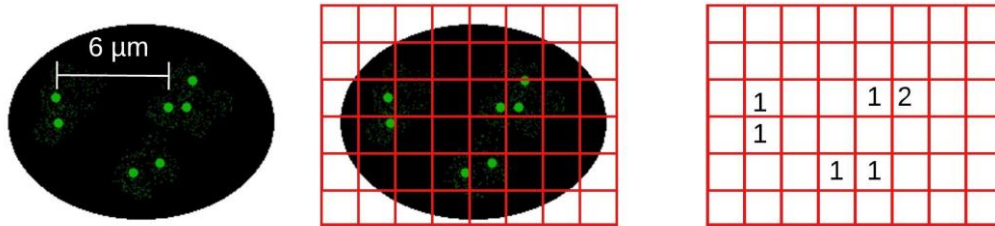


Figure 18 Illustration of downscaling principle used in original Moran's I calculations. The small scale of a cell produced little difference between coordinates. To solve the problem, the cell was sectioned into regions by adding a virtual grid on top of it. The new coordinates correspond to the centre of each grid-rectangle, and the value count within each rectangle was the new value to be used in calculations. Illustration by author.

3.7.3.2 OPTICS clustering

The Python-programme used to evaluate OPTICS of each cell nucleus can be found in appendix at B.3.1 The programme input was coordinates of each foci detected by 3D-OC. Included in the script is code for producing plots representing the groupings detected.

Originally clustering was determined by using K-means and gap-statistics. The programme determining k-means clustering can be found in appendix B.3.2, and the function authored by Mikael Vejdemo-Johansson [55] in appendix B.3.3.

3.8 Comparison of survival data and damage extent determined by the method

The Linear Quadratic (LQ) model describes the probability of survival for a cell after radiation. The LQ model is given in equation 8:

$$S = e^{(-\alpha D - \beta D^2)} \quad (8)$$

Where S is the probability of survival, D is the given dose and parameters α and β describes the cell's radiosensitivity. Furthermore, the term $\alpha D - \beta D^2$ is a measure of radiation damage. Survival fraction (SF) is the ratio of seeded cells to the number of colonies formed after irradiation and incubation for approximately ten days. The percentage of accumulated DSB volume relative to the nucleus volume (P) is a measure of damage extent determined by the developed method. Assuming that the LQ-model describes SF and that the term $\alpha D - \beta D^2$ is proportional to the measure of damage extent P , equation 8 can be written as equation 9:

$$\ln(SF) = -kP \quad (9)$$

Where k is some constant. The constant k was determined by plotting SF and P data for same doses and performing linear regression.

4 Results

Two objectives have driven the work of this thesis, namely constructing a method for investigating radiation impact and evaluating the results on the cell line A549. Results from using confocal imaging and flow cytometry were compared for investigating the previously mentioned two objectives. Table 3 presents an overview of experiments conducted through the work of this thesis. Three experiments were conducted in collaboration with Hilde Solesvik Skeie who performed the flow cytometry analysis. The early stage experiments were used to test and evaluate the methodological procedure. Only experiments with unique findings are presented in this chapter.

Table 3 Overview of experiments conducted throughout this project. Two protocols were used to produce samples for investigation, namely suspension cell assay (S) and adherent cell assay (A). Experiments marked with ‘*’ were done in collaboration with Hilde S. Skeie who evaluated the results using flow cytometry. The column “Time points [H]” denotes the incubation time of samples. In addition to the irradiation doses specified by column “Dose [Gy]”, each experiment had a control sample which was not irradiated.

Experiment	Irradiation	Dose [Gy]	Time points [H]	Cells sampled	Comments
S1	X-ray	0.3, 2	0.5, 48, 72	50	
S2 *	X-ray	2, 5	24, 48, 72	30	CLSM and Flow Cytometry used
S3 *	X-ray	2, 5, 10, 12	0.5, 24	30	CLSM and Flow Cytometry used
S4 *	X-ray	2, 5, 10, 12	0.5, 24, 48, 72	30	CLSM and Flow Cytometry used
A1	X-ray	2, 5, 10, 13	0.5, 24, 48, 72	30	
A2	X-ray	2, 5, 10, 13	0.5, 24, 48, 72	30	

4.1 Developing the DSB quantification method

This section describes and explains changes and development of the methodological procedure. It is sectioned according to the different phases of the experimental procedure. Alterations of the procedure were made based on experimental outcome and literature findings.

4.1.1 Development of suspension cell assay

Centrifugal rotation speed was tested to determine the relative centrifugal force required to create a cell pellet. Too hard centrifugation subjected the cells to unnecessary mechanical stress. Too weak centrifuging could lead to a loosely bound cell pellet, making it more difficult to remove all the supernatant and increasing the risk of cell loss. 200 G proved to be the best level of force in terms of adequate cell pellet formation. During testing several samples were rendered unusable due to insufficient cell pellet formation. A poor cell pellet and supernatant removal made the sample watery, which in turn increased the drying time and oxygen exposure prior to sealing the microscope slide with cover glass and nail polish. In some cases, the cells appeared dry, but examination in microscope revealed that they had not attached to the glass and were floating in vectashield oil, rendering the sample unusable.

Some issues arose for the suspension assay method. Firstly, the size of the cell suspension drop affected the drying and oxygen exposure time. As a result, a smaller drop was preferred. Secondly, the cell density at the centre of the slide was substantially higher than at the periphery. Cell density and cell stacking in the z-direction is of tremendous importance for both imaging and processing. High cell density makes the segmentation process difficult and occasionally impossible because cells cannot be distinguished leading to multiple cells contributing to a single segmentation region. Cell overlapping can take place in both the vertical and the horizontal plane. Cells that are in line but distinctly separated in the z-direction can be adequately segmented, but the segmentation requires stack slitting which is a time-consuming task. Figure 19 is an image sequence from an experiment using suspension cell assay (experiment S2). The image frame is fixed in the horizontal plane while image depth is increasing for each subsequent image. The most centred (marked "A") cell gradually became more covered by other cells as the sequence progresses in the z-direction. At the site of cell overlap (mark "B"), both cells contributed to the pixel values and cannot be distinguished.

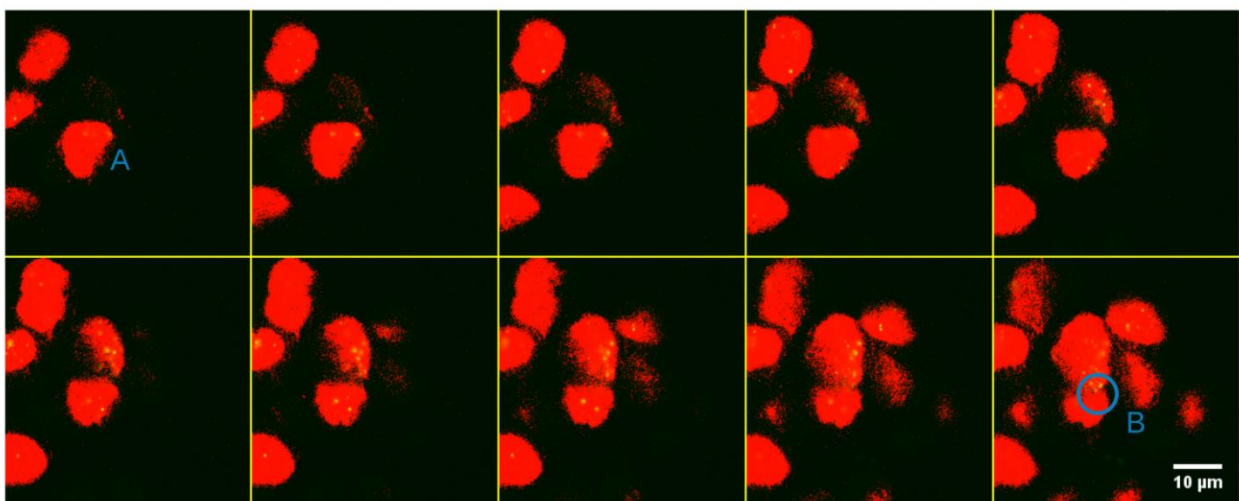


Figure 19 Example of overlap in the z-direction using the suspension cell assay. The most centred cell marked "A" was covered by other cells as the image sequence progresses in the z-direction. When cells overlap, both are contributing to the pixel values and can therefore not be distinguished. Region marked "B" display an area of overlap. The sample from experiment S2 with treatment 5 Gy and examined after 48 h.

Several strategies were tried to improve the cell distribution of the suspension assay. Of the cell sample amounts tested, a cell sample of 15 μ l and filtering the aliquot during transition to the microscope slide, gave the highest cell density with the lowest degree of cell overlap. A smaller sample amount lessened the cell stacking and decreased drying time. Filtering help with cell spread as well as removing cell duplets. However, the best result in terms of both cell distribution and drying time is adherent assay method.

4.1.2 Confocal microscopy image acquisition

Test of zoom and pinhole influence for confocal imaging acquisition DSB detection indicates that image resolution is more important with increasing number of DSB marks. Figure 20 displays plots with results of DSB detection analysis using four different image stacks of the same five cells. The image stacks were a combination of zoom two and three with pinhole 125 and 180 AU. Cells with few DSB marks has little variation in analysis across both zoom and pinhole size. Cells with higher number of DSB marks display larger difference between pinhole sizes at the same zoom and between zooms.

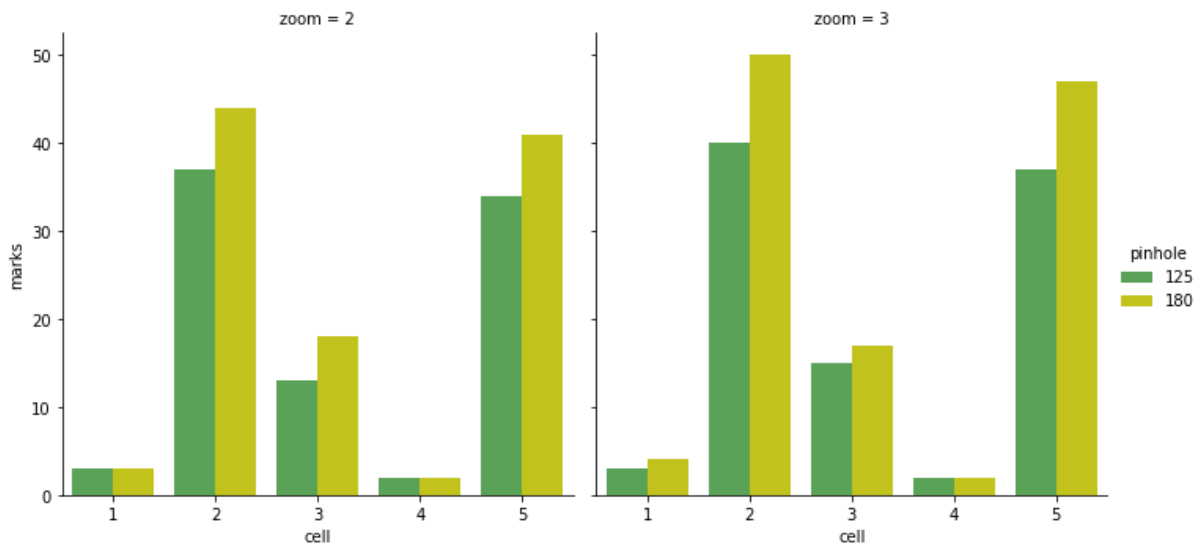


Figure 20 Plot shows differences in DSB detection according to zoom and pinhole size values used during confocal imaging. Five cells are imaged using two- and three times zoom and pinholes of 125 and 180 AU. The plot displays larger differences between both pinholes and zooms in cells with more marks. Cells with few DSB marks, 10 and below, display little difference between the confocal parameter setting.

4.1.3 Additional experiments to determine assay

PI signal intensity varied both between samples of a single experiment and within a single sample. The result of these tests can be found in appendix C.1.

Different antibodies and combinations were tested for affinity and specificity. All experiments were conducted with the optimal antibody set but results of the other antibodies performance can be found in appendix C.2.

Results from cell seeding test for adherent cell assay can be found in appendix C.3.

Results from dose determination in x-ray tube at SSD 40 can be found in appendix C.4.

4.1.4 Cell sampling significance

The number of imaged cells were reduced from 50 cells to 30 cells per sample due to long duration image acquisition. In the adherent cell assay 300 000 cells were seeded per sample. Imaging of 30 cells and confidence level of 90 % results in a margin of error at 15 %. A margin of error at 5 % would require 384 cells to be imaged per sample.

4.1.5 Duration of methodological processes

Table 4 lists the approximate time duration for methodological procedure that differs in duration between the two developing methods, suspension cell assay and adherent cell assay.

Table 4 Approximate time duration of the methodological processes that differs in duration between the two methods, suspension cell assay and adherent cell assay. The processing duration varies between samples

	Suspension cell assay	Adherent cell assay
Sample preparation in laboratory – assay	5 hours for all samples	2,5 hours for all samples
Image acquisition using confocal microscopy	90 min per sample	35 min per sample
Cell segmentation and image processing	90 min per sample	20 min per sample

4.2 Experimental results from adherent cell assay (experiment A1)

4.2.1 Dose response

The adherent cell experiments are used to investigate different parameters for quantifying γ -H2AX fluorescence. The parameters are evaluated across different doses to quantify the radiation impact. Furthermore, the time development of γ -H2AX fluorescence for each dose is examined to study the repair mechanism. The two experiments based on adherent cell assay were performed consecutively. On the final day of the first adherent experiment, a sample displayed unambiguous signs of fungal infection (sample irradiated with 10 Gy and incubated for 72 hours). All samples for the subsequent experiment (experiment A2, table 3) were prepared and stored in the contaminated incubator, and hence exposed to fungal spores. The experiment with contaminated samples was conducted as planned for its methodological value but is not assessed for damage response due to lack of reliability.

Literature suggest that the number of γ -H2AX foci reflects the number of DSB within the nucleus [9]. Figure 21 (a) presents the median DSB marks detected per sample in experiment A1 from table 3. It was expected that the number of γ -H2AX foci would decrease over time in accordance with repair completion on non-lethal damages. Figure 21 (a) shows an unexpected tendency of a higher number of DSB for the larger doses (8 and 10 Gy) after 24 h and 48 h incubation time, compared to 0.5 h. Samples irradiated with lower doses (2 Gy and 5 Gy) has a smaller amount of DSB marks after 48h and 72 h compared to 0.5 h. Figure 21 (b) shows the accumulated intensity signal of marks per cell nucleus as a function of incubation time, while figure 21 (c) display the accumulated mark volume to nucleus volume percentage as a function of incubation time. Both plots 21(b) and 21(c) in figure 22 display the expected tendency of higher γ -H2AX expression at 0.5 hours incubation time and decreasing substantially after that. There are some fluctuations in values from 24 to 72 hours, like 8 Gy intensity value in 21 (b) increase from 24 to 48 and settles somewhere in between at 72 hours. However, the 95% confidence interval for all values at 24 to 48 h are overlapping.

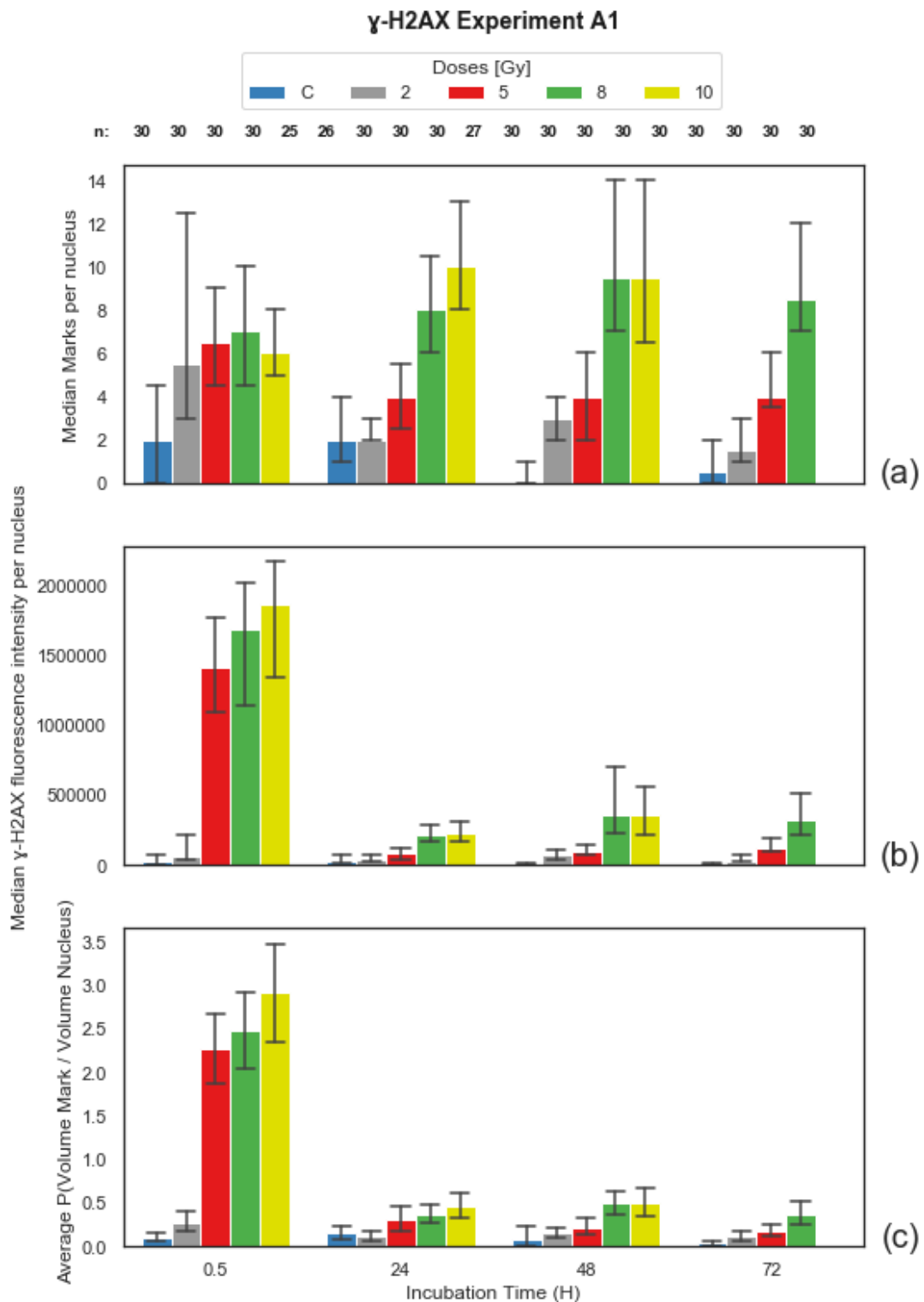


Figure 21 Bar plot of DSB markers within each sample in experiment A1 from table 3. (a) Median number of DSB marks per cell for each sample. (b) Median intensity per nucleus for each sample. (c) Average percentage of accumulated volume of the DSB marks relative to nucleus volume. Each plot has a 95 % CI denoted by error bars. "n" above plot (a) specifies how many cells were processed per sample.

Figure 22 displays plots of DSB marks per cell as a function of incubation time, where each subplot represents a dose and the marker size indicates accumulated volume of DSB marks within the corresponding cell nucleus. The control (dose=0) tends to steadily increase

accumulated DSB mark volume with increasing DSB detected. At 48 *hours* control there is an extreme cell sample with substantially larger DSB marker volume. The extreme cell can be view in figure 25 mark 2. At higher doses (5-10 Gy) and 0.5 *hours* incubation time the average DSB marker volume is generally higher, even at fewer DSB detected. This suggest that the density of DSB is too high to be distinguished.

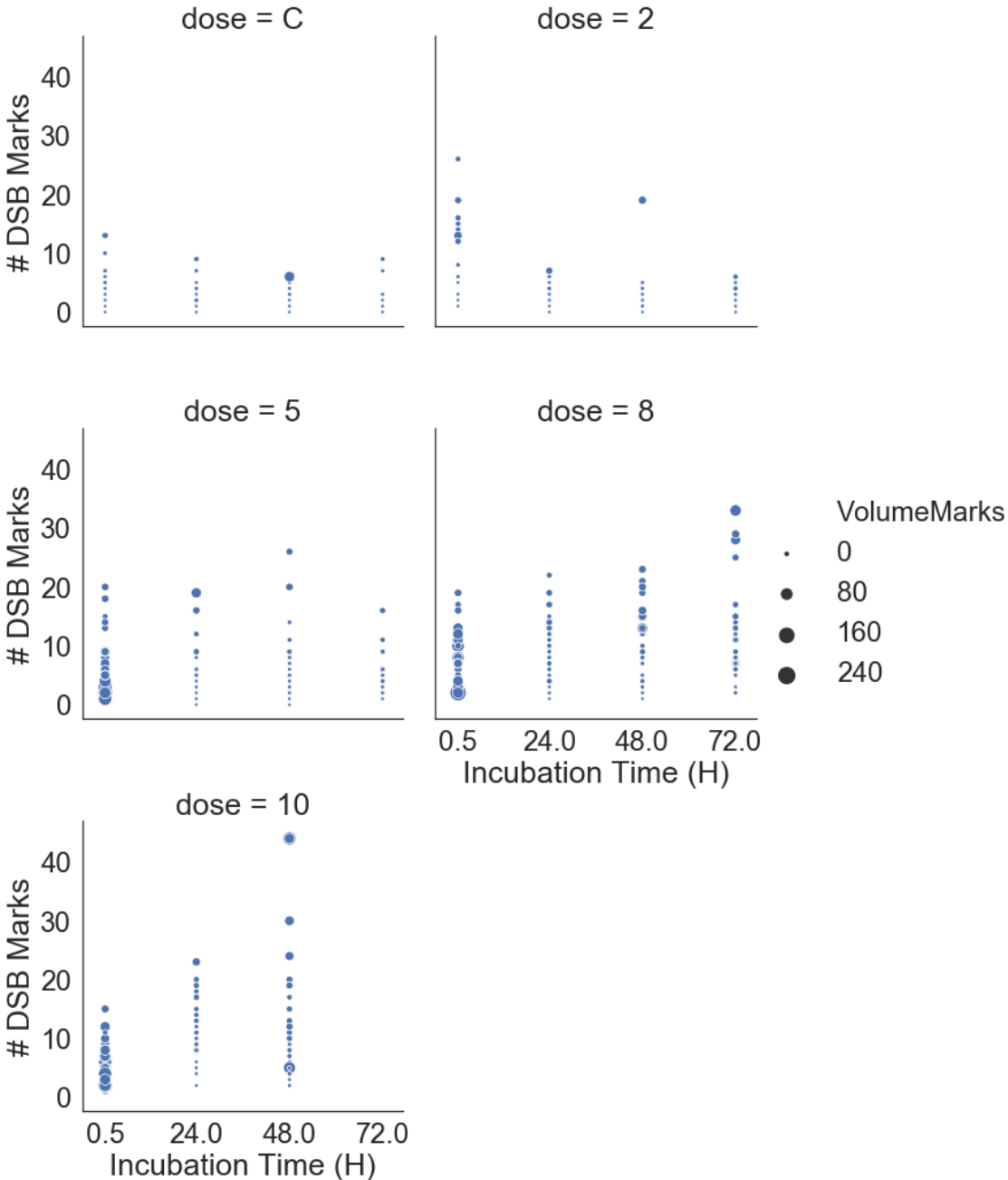


Figure 22 Number of DSB marks per cell plotted for each sample in experiment using adherent cell assay (experiment A1 in table 3). The sizes of each mark represent the accumulated volume of marks within the respective cell nucleus. The doses are in G and C denoted control which is not irradiated (0 Gy).

Figure 23 displays the distribution of the number of marks per nucleus, for the different incubation times (experiment A1 in Table 3). The box plot in figure 24 shows that all distributions are positively skewed, indicating that a normal distribution does not accurately describe the underlying population. The samples treated with higher irradiation doses (5-10 Gy) consistently have a wider distribution than samples treated with lower doses (0 and 2 Gy).

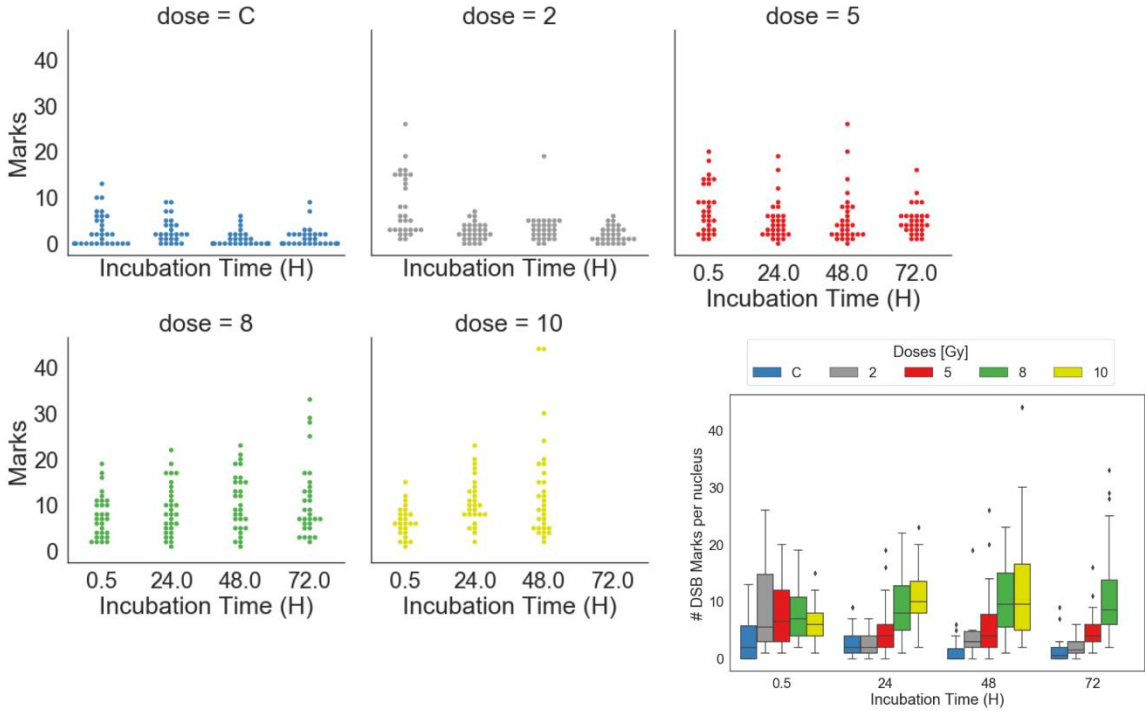


Figure 23 The number of DSB marks detected plotted as a function of incubation time. Each subplot presents the data for a treatment dose, whereas the doses have unit Gray (Gy) and C stands for “control” which is not irradiated. All samples of the experiment using adherent cell assay are presented (experiment A1 Table 3). The box plot displays the distribution of each sample processed in the experiment; all samples have a positive skew, and all cells outside 1.5 interquartile range are classified as an outlier.

Figure 24 present the same information as figure 23, but for the parameter of percentage DSB marker volume relative to nucleus volume. Again, all distributions are positively skewed but less skewed than the distributions for the number of DSB marks. Furthermore, most distributions of volume marks to volume nucleus percentage has a distinct and narrow peak. Interestingly, the peak is considerably wider for samples harvested at 0.5 h incubation time.

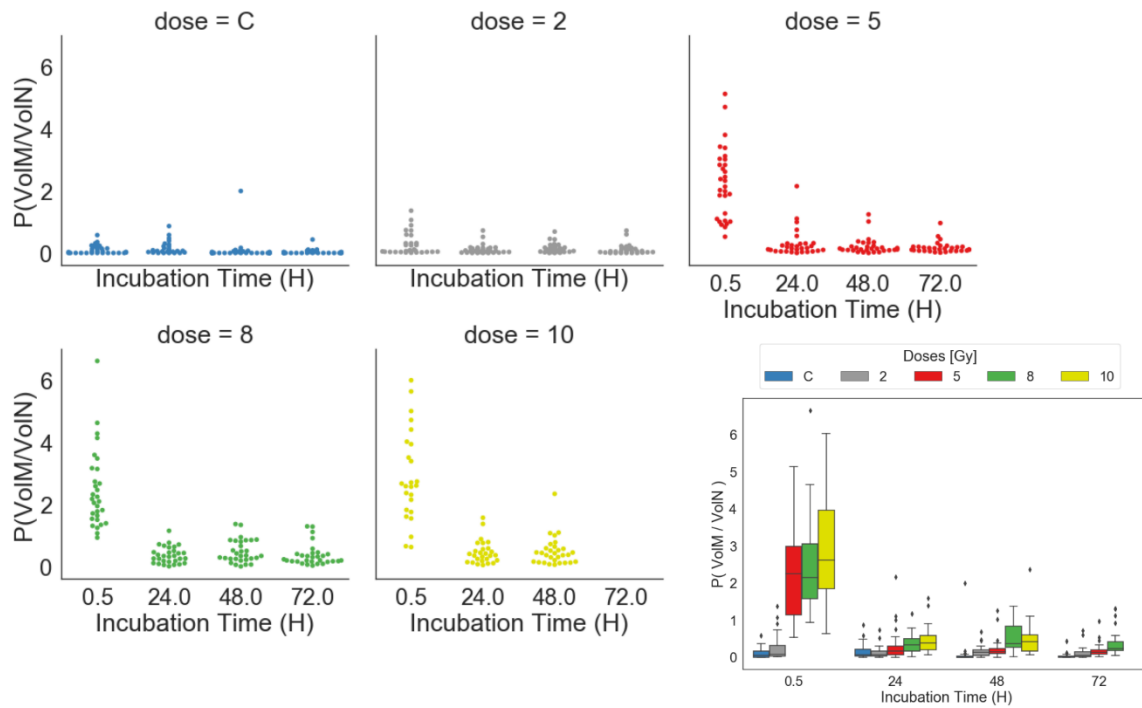


Figure 14 The percentage of DSB mark volume relative to nucleus volume plotted as a function of incubation time. Each subplot presents the data for a treatment dose, whereas the doses have unit Gray (Gy) and C stands for “control” which is not irradiated. All samples of experiment using adherent cell assay are presented (experiment A1 Table 3). The box plot displays the distribution of each sample processed in the experiment; all samples have a positive skew, and all cells outside 1.5 interquartile range are classified as an outlier.

Figure 25 shows a composite-image (both red and green channels merged) acquired with a zoom of two for each sample processed in experiment A1. The columns present incubation time, while the rows present a treatment dose, starting with control (0 Gy) and increasing up to 10 Gy.

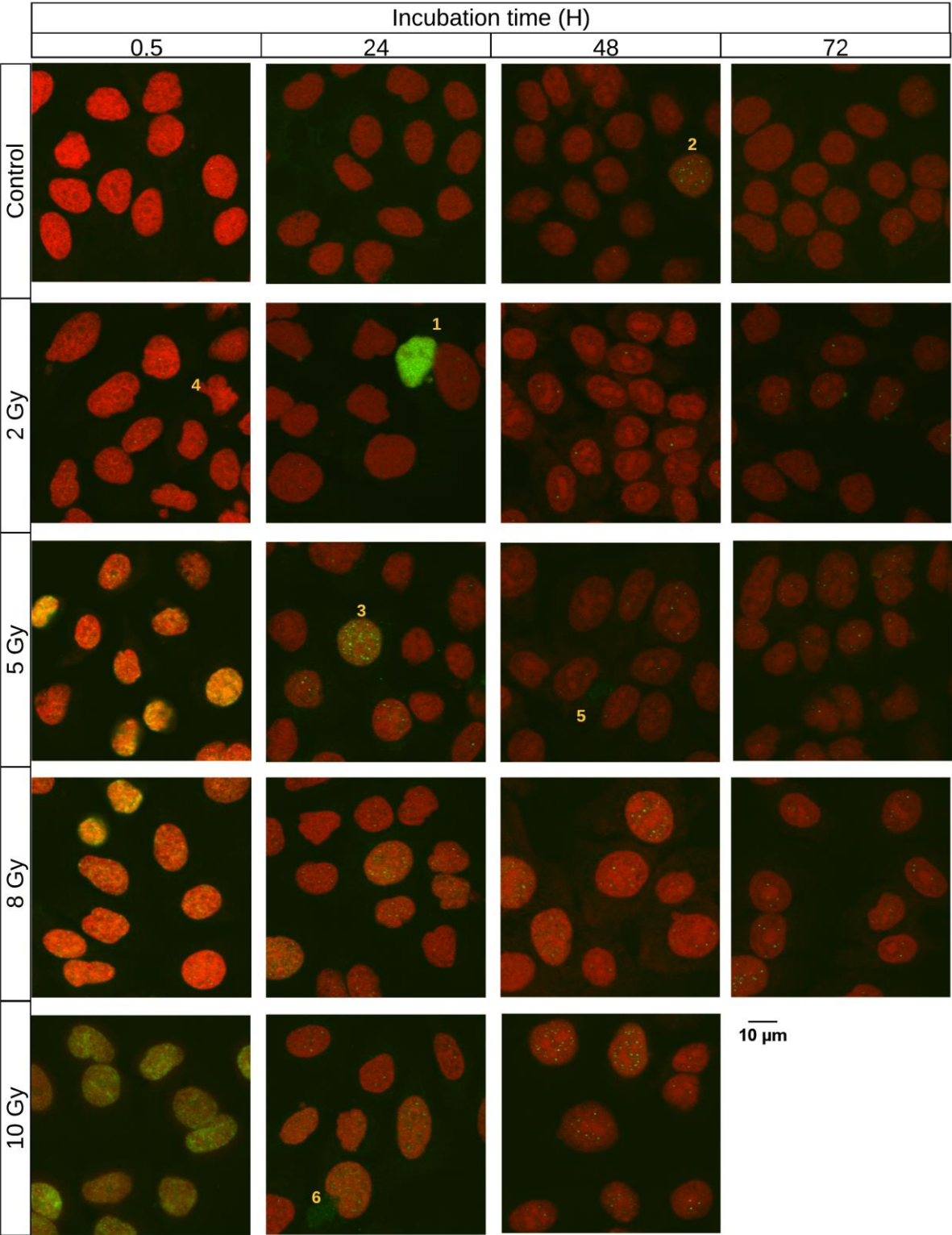


Figure 25 RBG images from each sample processed in experiment A1. The images were acquired using confocal microscopy with a software zoom of 2.

By visual inspection, it's apparent that the radiation impact increases with dose. This is seen as more green colour in the images, presenting itself as focal spots or general glow of the cell. The image in column 0.5 hours and 5 Gy is a good example of how γ -H2AX fluorescence makes the cell nucleus glow. Figure 26 below displays composite-image 0.5 hours and 5 Gy along with its contributing colour channels. The nuclei are generally more dispersed as the dose increases. More dispersed cells make for easier segmentation as cells are clearly distinguished.

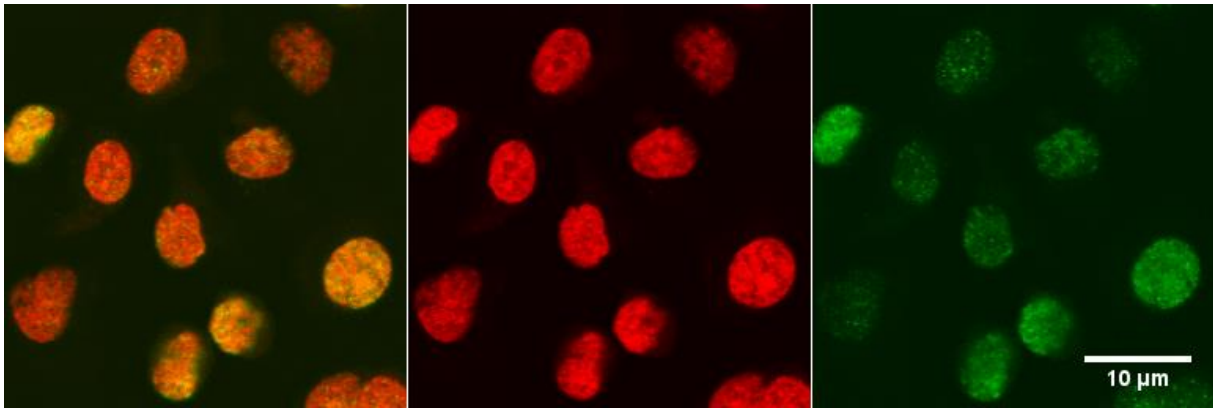


Figure 26 Image from figure 25 (0.5 h, 5 Gy). From left to right, the channels are composite (both red and green channel), red channel originating from PI staining, and green channel originating from H2AX fluorescence staining.

At figure 25 mark 1 is an extreme outlier cell apparent by its bright green colour. Cells illuminating brightly of γ -H2AX fluorescence was not uncommon and was present in all samples processed. Outliers like this were not included in processing due to their extreme nature. This extreme nucleus was in mitosis as the two nuclei were still in the process of separation. The other nucleus (to the right of mark 1) had far less γ -H2AX fluorescence than its counterpart, but was still the most extreme nucleus processed in terms of the number of DSB marks, DSB accumulated volume and intensity. A closer look at cells in mitosis can be found in figure 30 at the end of this section.

Figure 25 (2) and (3) mark cell nuclei that stand out from their neighbouring cells in terms of γ -H2AX fluorescence. Both nuclei have maximum values within their group (sample 48 h control and 24 h, 5 Gy, respectively) in terms of number of DSB marks and accumulated volume and intensity. Figure 25 (4) marks a misshaped cell. The nuclear envelope is clearly damaged. Alike cells were found throughout both experiments with adherent cell culture (experiment A1 and A2 in Table 3).

There are variations of PI signal intensity both between samples (as seen in figure 25) and within a single sample. The PI signal intensity tends to be similar within a region of the sample, but vary between different locations. Figure 27 below displays images of different locations of the same sample to illustrate PI variations. Two locations are imaged using a zoom of both one and two at each location.

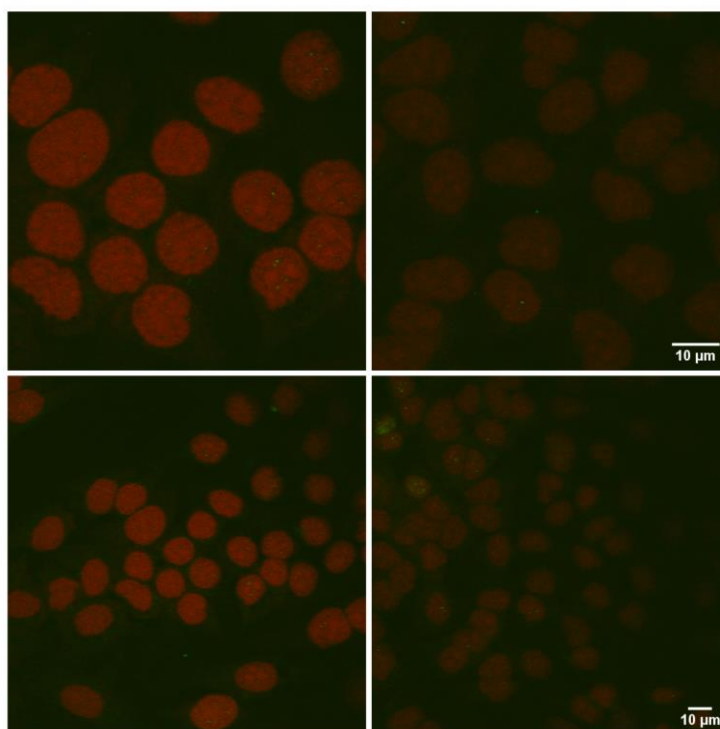


Figure 27 Images of a control sample harvested after 72 h incubation. Two regions of the sample are presented, one in each column. The images in the first row are acquired at a zoom of 1, while images in the second row are acquired using a zoom of 2. The images show how PI signal intensity varies within a single sample but is consistent within a broader region. The scale bars denote 10 μm each, and is valid for both images per row. The sample imaged originates from adherent cell assay (experiment A1 Table 3).

PI stains nuclear acids by intercalating between DNA bases. PI stains the nucleus and other cell constituents containing DNA. RNA is commonly found in the cell cytoplasm and can, therefore, be visualised by using PI staining. Figure 28 shows image 24 h, 8 Gy (figure 25) at a larger zoom along with a duplicate image, which has enhanced brightness for better clarity. The nuclei are encompassed by a PI cloud, most likely to be the cell cytoplasm. PI shades are usually presented throughout the image-stack at some depth, but not throughout all images in the stack. The lack of consistent PI staining of cell cytoplasm makes it challenging to distinguish between binuclear cells and two individual cells.

The most centred cell in figure 28 has green spots on the outside of the nucleus, on the upper left side. It is not uncommon to see γ -H2AX fluorescence outside the nuclei as both cytoplasm and mitochondria contain DNA. In this case, it is concentrated close to the nucleus, and most probably belongs to the corresponding cell cytoplasm. There are also multiple examples of γ -H2AX fluorescence clouds in the samples, like figure 25 (5-6). At 25 (6) there is a clear breach on nucleus membrane which might explain that fluorescence is leaking out. Figure 25 (5) there are no sign of membrane breach, but the membrane might still be compromised and leaking.

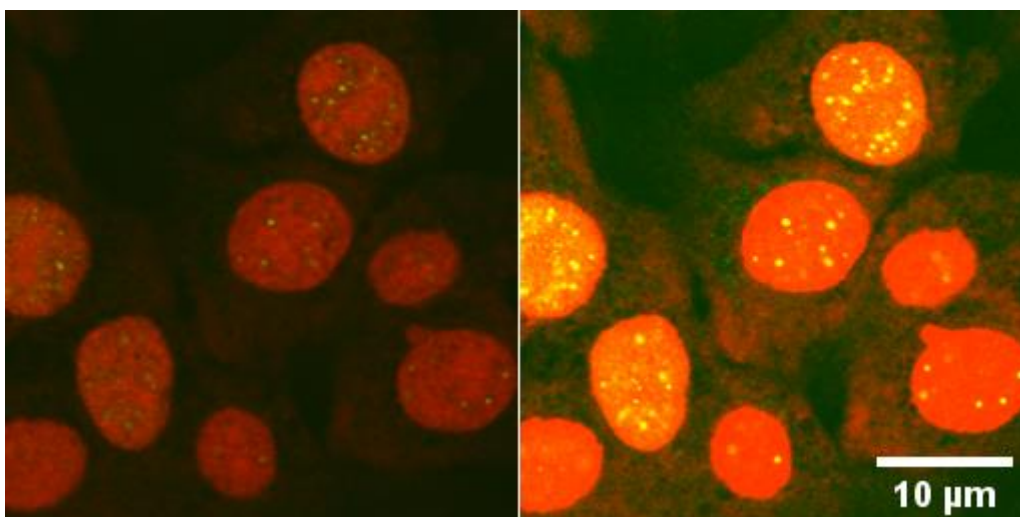


Figure 28 Composite image from figure 25, 24 h, 8 Gy enlarged for better view. The image to the left has enhanced contrast to ease PI staining of cell cytoplasm and γ -H2AX fluorescence outside of the nucleus.

There are regions with higher density PI staining within the nucleus. Regions with a strong PI signal are shown in figure 28, and even more prominent in figure 25, 48 h, 2 Gy. Compacted DNA likely cause the regions of extra strong PI signal intensity. Interestingly, there are few to none DSB mark in these regions.

4.2.2 Characterization of γ -H2AX fluorescence and cell nuclei

Parameters used to characterize γ -H2AX fluorescence and cell nuclei in adherent cell assay are listed in table 5.

Table 5 Parameters used to characterize cell nuclei and γ -H2AX fluorescence in adherent cell assay.

Marks	Number of DSB marks detected by 3D-OC analysis
Volume marks	Accumulated volume of all marks within the nucleus evaluated
IntensityMarks	Accumulated intensity of all DSB marks within the nucleus
IntensityPI	Intensity of the PI-channel of nucleus
P(VoIM/VoIN)	Percentage of accumulated mark volume to accumulated nucleus volume

To explore the relationships between these parameters, they were plotted pairwise against each other in figure 29. The parameters are listed along the x- and y-axis. A univariate kernel density estimate of each parameter is shown along the diagonal. On the upper triangular grid (above the diagonal), the scatter plot is shown, whereas the lower triangular shows a bivariate kernel density plot. The plots on the diagonal and upper triangle are plotted for each dose with a unique colour specified by the legend bar. In the scatter plots, the Pearson correlation coefficient (PCC) was calculated to indicate any potential correlations in the data. The PCC value for each dose is presented in pair plots by " ρ ", where the colour denoted which the dose. The Pearson value ranges from -1 to 1, where -1 is a perfect negative linear relationship between the parameter and +1 is a perfect positive linear relationship. A negative linear relationship means that one parameter decreases as the other parameter increases, while a

positive linear relationship means that both increases or decreases simultaneously. A value at 0 means that there is no linear relationship between the parameters.

The univariate kernel density plots display a narrow peak for parameters "VolumeMarks", "IntensityMarks" and "P(VoIM/VoIN)". The peak arises from control data, and all other doses are relatively small compared to the control. The univariate kernel density plots for "Marks", "VolumeNucleus" and "IntensityPI" generally had broader peaks indicating more variance in the parameters within each group.

The parameter "IntensityMarks" was calculated by summing the intensity of each DSB mark voxel and is therefore correlated to the DSB mark volume. The parameter "P(VoIM/VoIN)" was calculated on "VolumeMarks" relative to "VolumeNucleus". The "P(VoIM/VoIN)" is positively correlated for both "VolumeMarks" and "IntensityMarks", but interestingly not correlated to "VolumeNucleus". "IntensityPI" was calculated using the same principle as "IntensityMarks" - the intensity of each nucleus voxel was added up for the entire volume of the nucleus. This is reflected in the correlation between "Plintensity" and "VolumeNucleus".

The correlation between "Marks" and other parameters differs according to dose. Control samples and cells irradiated with 2 Gy has higher PCC values for "VolumeMarks" ($\rho=0.5$ and $\rho=0.72$, respectively), "IntensityMarks" ($\rho=0.5$ and $\rho=0.72$, respectively) and "P(VoIM/VoIN)" ($\rho=0.61$ and $\rho=0.72$, respectively) than higher dose irradiated samples. Cells irradiated with 8 Gy and 10 Gy has higher PCC values for parameters "VolumeNucleus" ($\rho=0.49$ and $\rho=0.72$, respectively) and "IntensityPI" ($\rho=0.46$ and $\rho=0.72$, respectively) than lower dose irradiated samples.

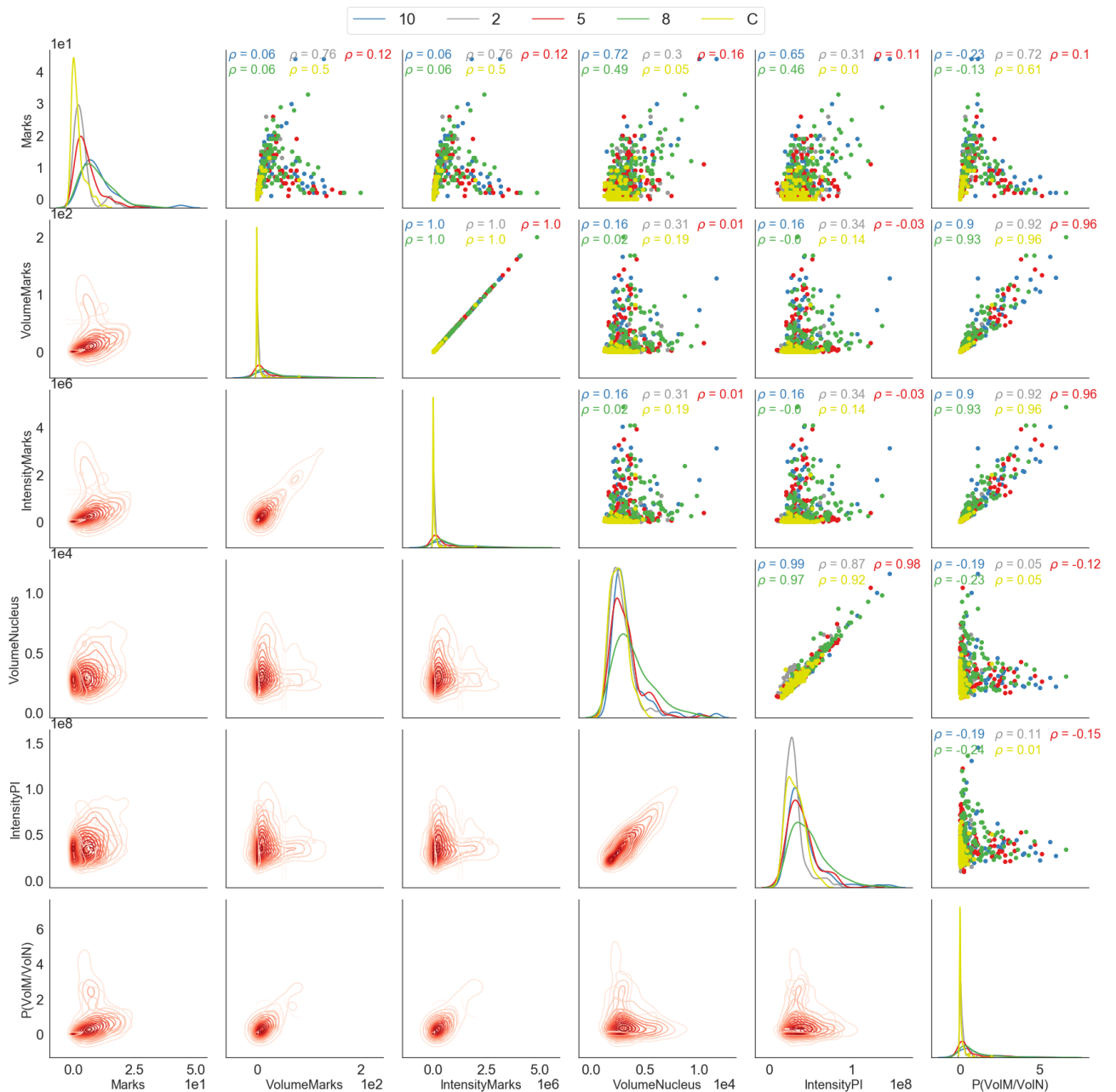


Figure 29 All parameters characterising the cell nuclei, and γ -H2AX fluorescence are plotted against each other. Along the diagonal are univariate kernel density plots of the corresponding parameters separated by colour into subgroups of treatment doses. Above the diagonal are scatterplots for all parameters against each. Pearson's correlation between two parameters are calculated for each treatment dose. The Pearson's value denoted " ρ " ranges from -1 to 1, where 0 indicates no association between the variables. A negative value indicates a negative association, meaning that one variable will increase when the other decreases. A positive value means that both variables increase and decrease together. More variation in data from the best fit line results in a Pearson value tending toward 0. Below the diagonal are the bivariate kernel density estimation (KDE) plot for estimating the probability density function of random variables using a Gaussian kernel and 'Scott' bandwidth. All data stems from adherent cell assay (experiment A1 table 3). The parameters evaluated are listed and explained in table 5. The legend bar denotes colour coding for each dose of unit Gray [Gy] where C is the control samples which are not irradiated.

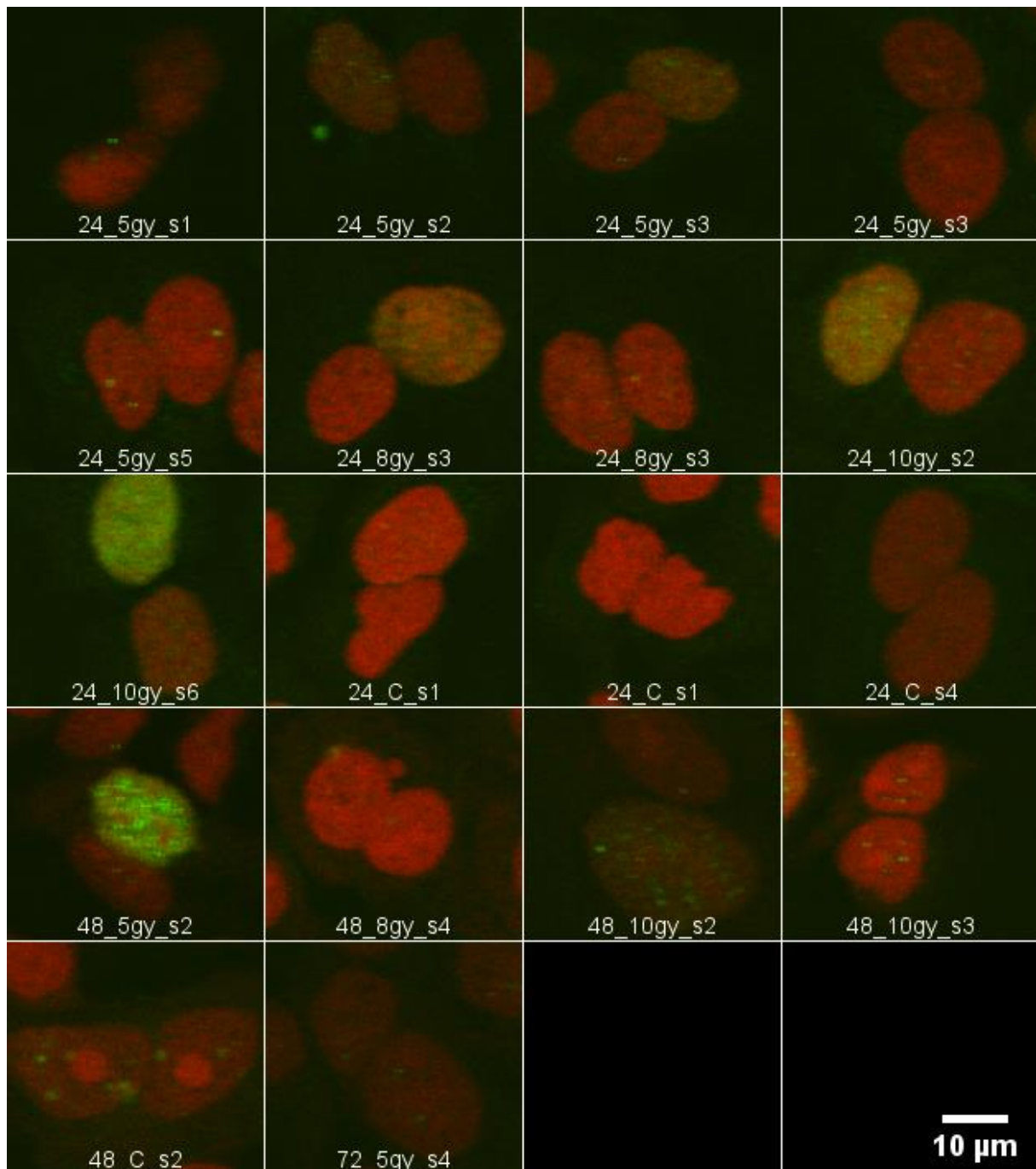


Figure 30 Cells and nuclei still in mitosis (or recently finished mitosis) at the point of fixation. The images are retrieved from overview images acquired at zoom 1. These images were meant to give information about processed cells environment and were not imaged in z-direction to minimise any photobleaching effects. The name on each image specifies which treatment they received (“incubation time”_“dose”) and which stack they originate from. Several of the cell pairs that was irradiated under mitosis has uneven damage extent between the two daughter nuclei.

4.2.3 LQ-model performance: $P(\text{VolM}/\text{VolN})$ plotted against survival data

The percentage of DSB marks volume relative to nucleus volume was plotted against survival data for each dose and evaluated for each incubation time in figure 31. All doses of a single incubation time point are given a colour, with the regression line expression and regression score in the upper right corner with the same colour. The average volume mark to volume nucleus percentage of samples incubated for 24 h, 48 h and 72 h create a straight line with

regression score above 0.92. The 0.5 h incubated samples deviate most from the regression line, with increasing distance from the line with higher doses. The large standard deviation of 0.5 h incubated doses was expected as figure 25 shows a low peakedness for 0.5 h incubation times.

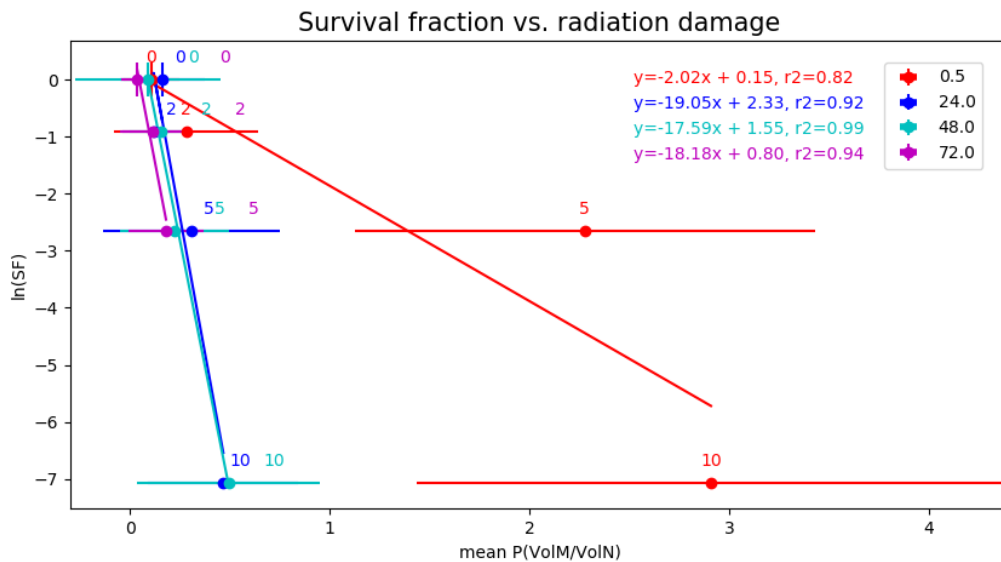


Figure 31 Survival fraction after irradiation are plotted against percentage of DSB marks volume relative to nucleus volume for the same doses. The survival fraction is corrected for multiplicity. The regression line expression and regression score is plotted in the upper right corner where the colour designates which incubation time it represents. The DSB marker and nucleus data was collected using the adherent cell assay (experiment A1 table 3).

4.2.4 Evaluation of Moran's I

Figure 32 displays the median Moran's I within each sample of experiment using adherent cell assay (experiment A1, table 3). All values are positive ranging from 0.05 to 0.45.

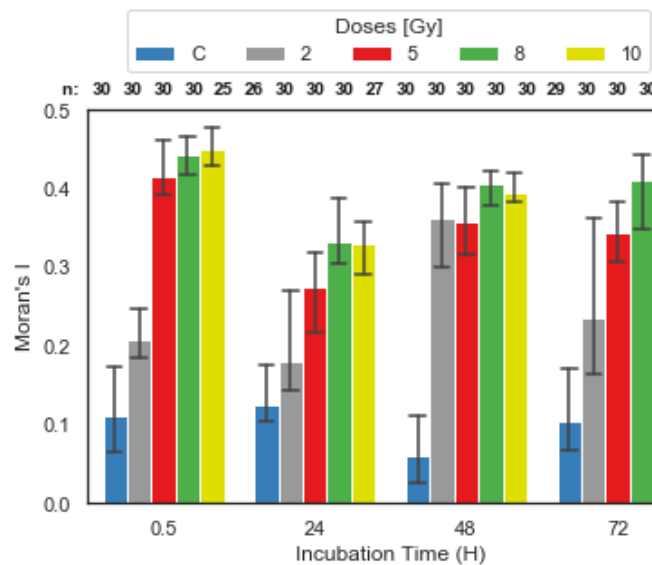


Figure 32 The median Moran's I estimation per sample using adherent cell assay (experiment A1 table 3). All samples have a positive Moran's I value which indicates clustering of the DSB foci. The image stacks were pre-processed with "Stack Meancenter5" plugin. Each plot has a 95 % CI denoted by error bars. "n" above plot (a) specifies how many image stacks that could be evaluated by the Moran's I.

The distribution of all the values of Moran's I and its respective calculated z-values are shown in the two histograms to the left in figure 33. In the rightmost histogram, the respective calculated p-values under the normality assumption are shown. Cells with a p-value below 0.05 means that the null hypothesis of spatial randomness is rejected at the 95 % CI, suggesting that the cells are significantly clustered. 99.6 % of the cells had a p-value below 0.05.

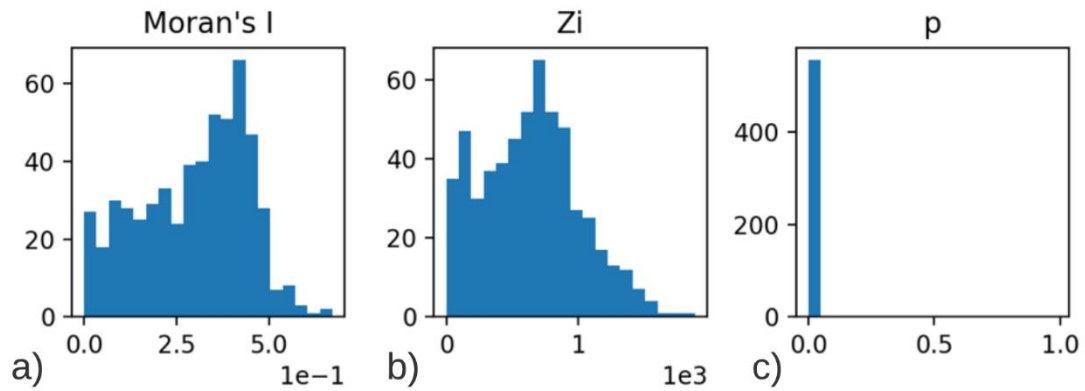


Figure 33 Histograms, with bin size of 20, of the Moran's I (a), the z-value (b) and p-value (c) calculated under normality assumption.

4.2.5 Evaluation of OPTICS clustering

OPTICS values are based on the coordinates of the DSB marks provided by 3D-OC and are therefore based of the same information as dose response results in chapter 4.2.1. A value of one means that there is no clustering of DSB marks in the dataset of a cell nucleus. Any nucleus with less than three DSB marks is automatically set to be not clustered (one cluster) which corresponds to a value of 1 from OPTICS. Figure 34 displays the average clusters for each sample in experiment A1. Only one group, 48 h, 10 Gy, had a clustering average above 1 at 95 % CI.

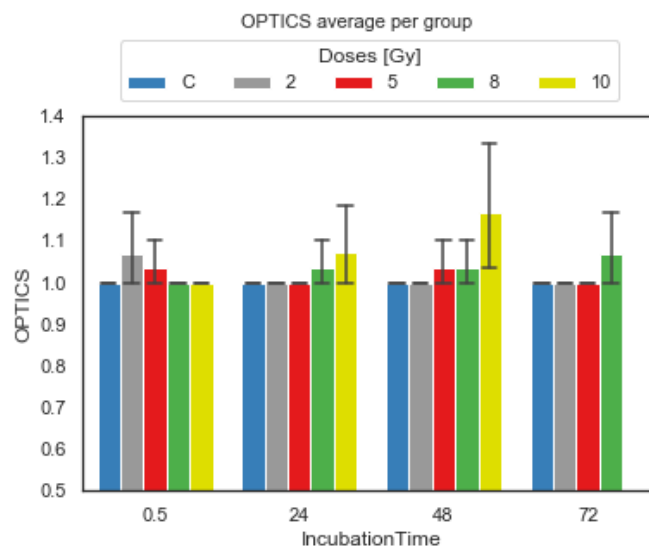


Figure 34 Averaged OPTICS value per group in experiment A1. The estimated values are encompassed by a 95 % CI. Only one group ,48 h, 8 Gy, have a cluster average above 1 with 95 % probability of being true.

Figure 35 displays plots of the number of clusters as a function of number of DSB marks figure 35(a) and accumulated intensity of marks figure 35(b). In figure 35(a), it can be seen that separate clusters appear when the number of marks exceeds around 10 to 20, and thereafter, the occurrence of clusters relative to only one cluster seem to increase. Whereas in 35(b), increasing the variable IntensityMarks further from between 1,000,000 to 2,000,000 does not seem to increase the proportion of more than one clusters compare to only one, rather there might be instead an interval for this variable where separate clusters appears more often.

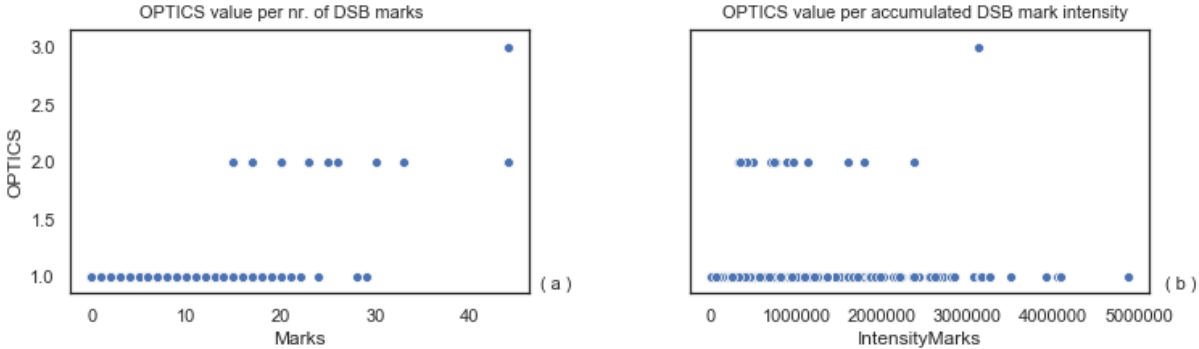


Figure 35 Plot of number of clusters determined by the OPTICS algorithm as a function of number of DSB (a) and DSB intensity (b).

Figure 36, the 3D-plot shows three examples of OPTICS clusters determined by the developed method. Each marker colour determines a cluster, and the shade of the colour represents the depth from our point of view.

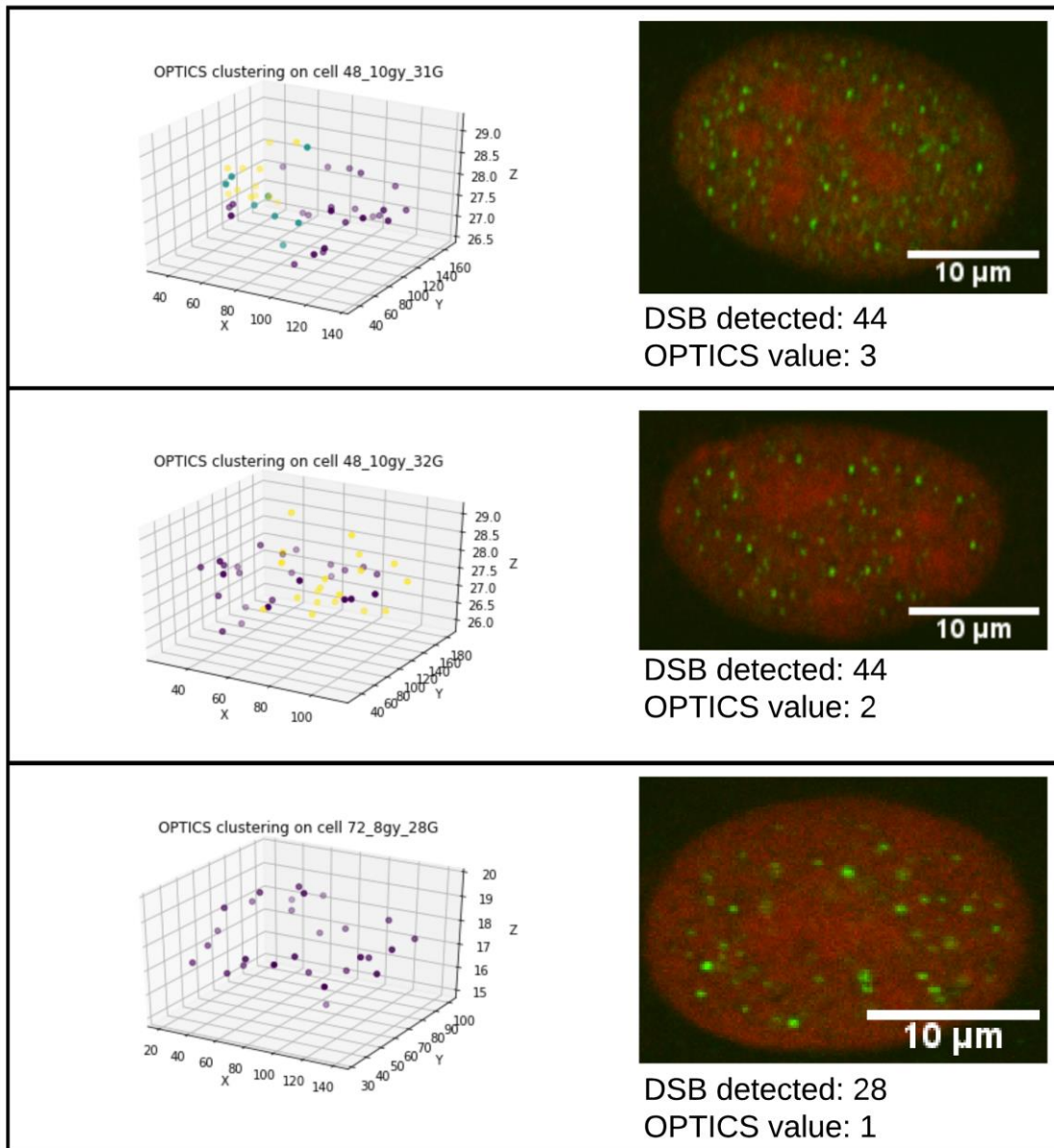


Figure 36 Each row displays the clusters determined by OPTICS along with the cell nucleus evaluated. DSB detected is the number of individual DSB marks determined by 3D-OC analysis. Each cluster determined by OPTICS analysis are given a unique colour. Increasing colour saturation indicates a location toward the front, while dimmer coloured indicates a located at the back. The 3D-plots reflect how DNA dense regions (where PI stain is more prominent) do not have any DSB marks. The title of the 3D-plots specify which cell (experiment A1 table 3) the data originates from ("incubation time_"dose_"cell_id").

4.3 Results from suspension cell assay experiment S1

In experiment S1 cells were irradiated with doses of 0.3 and 2 Gy. 50 cells were images per sample in experiment S1, whereas experiment A1 imaged 30 cells. Additionally, the samples were images using a zoom of three and thus had a greater image resolution. The samples incubated for 24 h contained much water due to insufficient pellet after centrifugation and were therefore unusable. Figure 37 displays suspension cell assay results (experiment S1 table 3) of DSB marks parameters per treatment group as a function of time. Figure 37(a) shows the median DSB marks per treatment group, while figure 37(b) shows the median of accumulated

γ -H2AX fluorescence per cell nucleus. Figure 37 reveals the similar trends between the number of DSB marks and accumulated intensity. The higher resolution images produced with a zoom of three made DSB foci detection easier to distinguish, compared to images acquired with a zoom of two. Furthermore, lower doses gave less density of DSBs which also eased the detection.

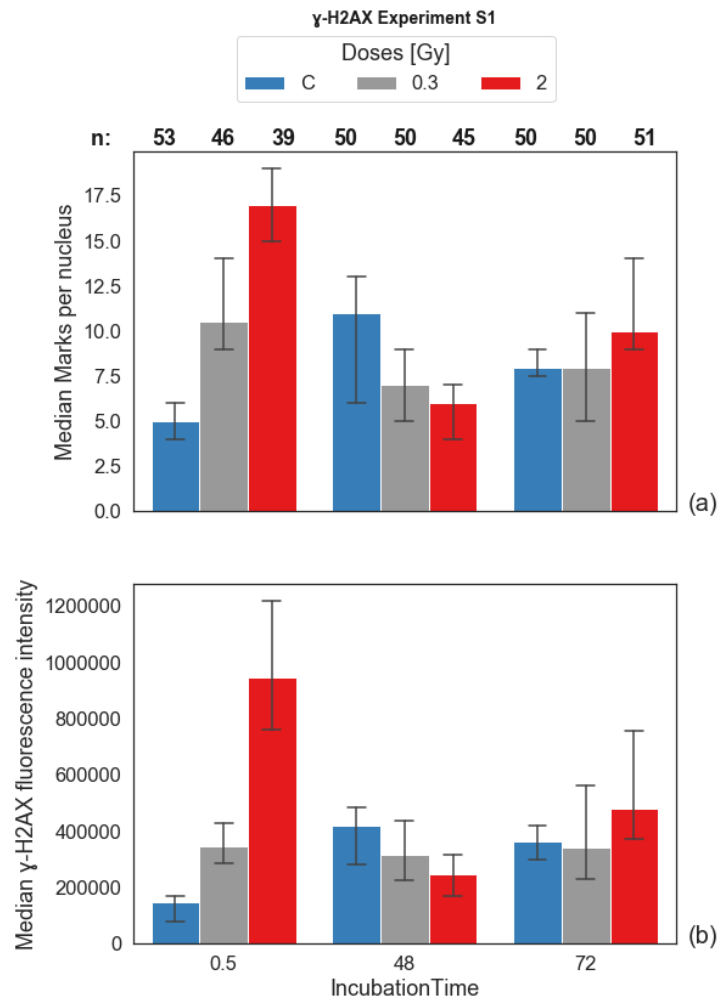


Figure 37 Plots presents results using suspension cell assay (experiment S1 table 3) with a zoom of three. (a) show median marks per nucleus as a function of dose and incubation time, while (b) shows the accumulated intensity of marks.

4.4 Summary of Result using Flow cytometry

Flow cytometry enables γ -H2AX fluorescence evaluation of a much larger cell sample population that was is possible with confocal microscopy. Furthermore, the PI intensity signal is utilised to detect the average percentage of cells in different cell cycle phases.

Figure 35 shows averaged results from six experiments using flow cytometry to analyse the γ -H2AX fluorescence intensity. The experiments are conducted either solely by Hilde Solesvik Skeie or in collaboration with the author. The results display a clear increase of median γ -H2AX fluorescence with increasing dose at 0.5 h harvest. At timepoints 24, 48 and 72 h, 2 Gy and 5 Gy intensity value stabilizes at 1.5 (figure 38 (a) and 38 (b)) while 9 Gy and 12 Gy the intensity

value stabilized at ca. 2 (figure 38 (c) and 38 (d)). At 0.5 h, G1 gave the strongest signal and became more prominent with increasing dose. The difference in PI signal largest at 0.5 h and decreasing with increasing incubation time.

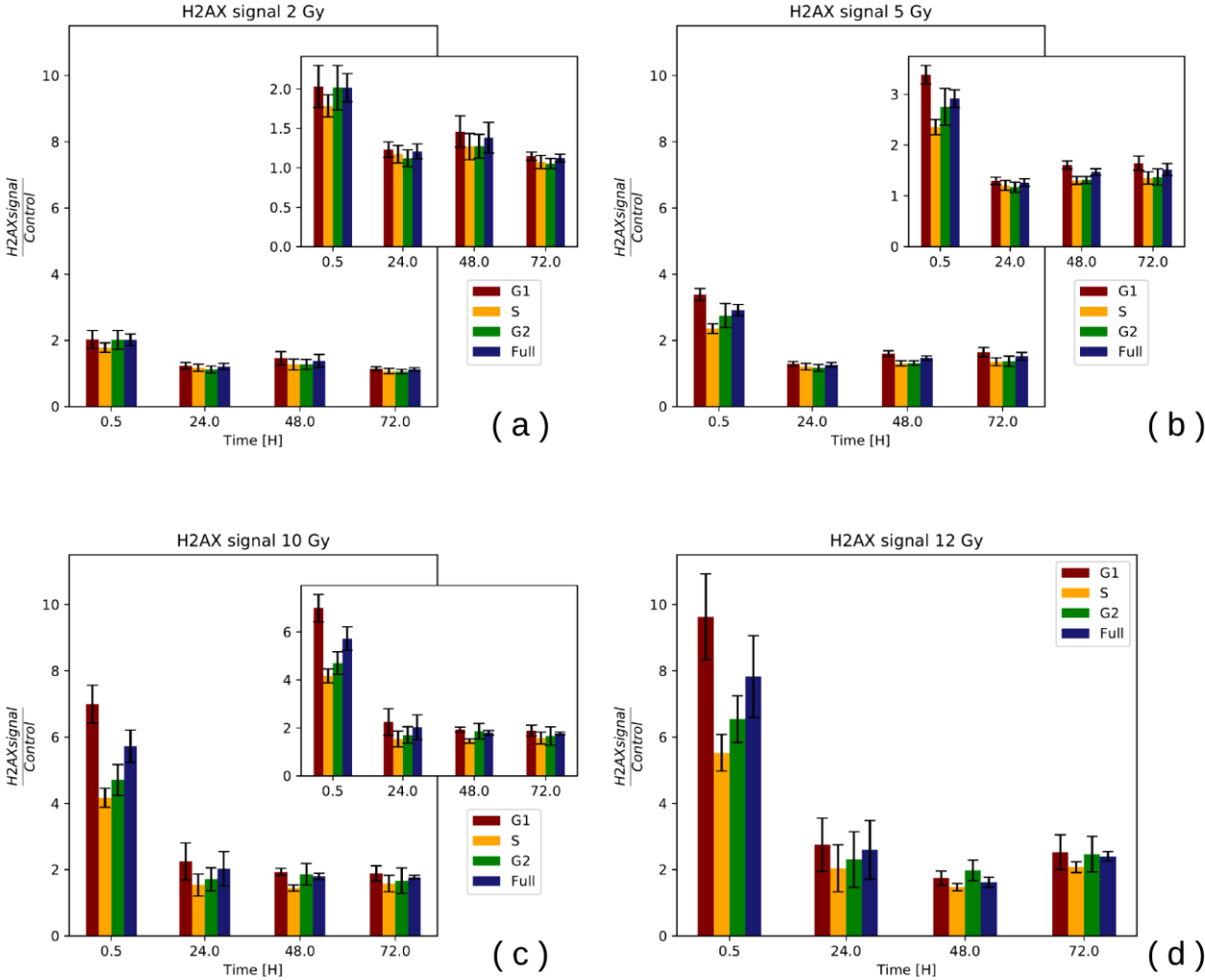


Figure 38 The plots show averaged results from γ -H2AX assay resulting from six experiments in total. The treatment given is x-ray irradiation of 2,5,9 and 12 Gy with Incubation time of 0.5, 24, 48 and 72 h. The Flow cytometry analysis was conducted by Hilde Solesvik Skeie. The signal for a treatment dose and incubation time is further divided into cell cycle phases, represented by bars.

Figure 39 displays the average amount of cells in each phase in percentage for each treatment group. The majority of cells irradiated with 2 Gy was in G1 phase at 0.5 h, with a larger percentage of G1 and lesser percentage of cells S-phase in 24 h. At 48 h and 72 h, the percentages of cells in both G1 and S-phase remain approximately the same as in 24 h. 5 Gy irradiated cells also start with the majority of cells in the G1 phase at 0.5 h and increase in the amount in G1 and G2 in 24 h, while decreasing in the percentage of s-phase cells. The amount

in G1 continues to increase steadily throughout the time points, while G2 percentage declines. 9 Gy irradiated cells display the same tendency with the majority in G1 phase but decreased in percentage in 24 h time point for both G1 and S-phase. An increasing amount of cells in G2 arrest is seen toward 24 h before it declined again. 12 Gy irradiated cells had a slight increase of cells in G2 arrest from 48 to 72 h. Similar for all doses is the approximate stabilisation at a low percentage in S-phase from 24 to 72 h incubation time.

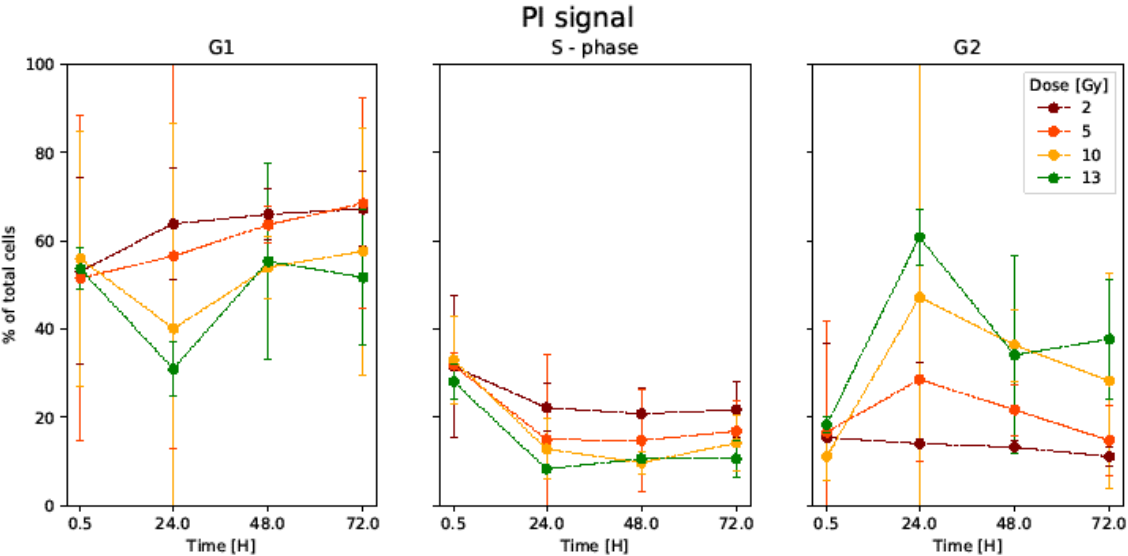


Figure 39 Percentage of cells in different cell cycle phases as a function of harvest time after irradiation. % G1 cells increase from 0.5 to 24 h for lower doses (2 and 5 Gy) and stabilized after the 24 h time point. Larger doses of 10 and 12 Gy decrease the % of G1 cells towards 24 h, before increasing again. Percentage of S-phase cells showed a similar trend for all doses with decreasing amount from 0.5-24h time interval and stabilizing after that. G2 and G2 arrest is complementary of the G1 percentage plot, with high doses (5,10 and 12 Gy) increasing initially before declining again. Cells irradiated with 2 Gy had a steady decrease in percentage of arrested cells.

5 Discussion

5.1 Methodological considerations and inherent uncertainty

5.1.1 Cellular variations

Living cells present substantial uncertainty in the experiment. It is immensely challenging to replicate the exact conditions from week to week. The cells have a varying growth rate resulting in different cell numbers for experiments. Furthermore, the cells will be in different cell cycle phases, which also introduce some uncertainty. DSB susceptibility and repair response vary depending on cell cycle phase [56].

Substantial efforts were made start the assay at the exact same time after irradiation for all experiments. However, this was not always possible due to the number of samples processed together. Consequently, there are some variations in the 0.5 h incubation time, and some samples got more time to initiate repair.

Environmental changes may cause cells to go into cell cycle arrest. Such changes have implications for the experimental results as non-cycling cells have more time for damage repair. Therefore, the uttermost care was taken to minimize any environmental variation during preparation and irradiation of cells. The irradiation chamber was preheated to keep the cells at 37°C during irradiation. Furthermore, the time spent outside of incubator and irradiation chamber was minimized. The cells were irradiated submerged in medium to help maintain both temperature and pH concentration. Medium removal inflicts a rise in pH levels which can lead to cell death if sustained over a longer time. Research conducted at UiO showed decreased survival rate for cells irradiated without medium [57].

The incubators used for these experiments are routinely cleaned and sterilized. Nevertheless, fungus infection reappears in the incubator and damages the biological material. Samples with a clear sign of fungus were discarded and excluded from experiments. The remaining samples have been used to avoid discarding entire experiments and the following cost and time waste. Although none of the processed cells has displayed any sign of infection, there is no way of ensuring that they are not infected. All samples of the second adherent cell assay experiment (experiment A2, table 3) were compromised due to a fungal infection. Each sample of this experiment had a large proportion of cell with broken nuclear envelope and was therefore excluded from analysis. Furthermore, there is no documentation available on how a fungus infection might change radiation damage response in the cell. Once the fungus has invaded an incubator, the spores may remain in the air, presenting an everlasting uncertainty of cell condition and potential reblossoming.

5.1.2 Assay

The suspension cell assay requires multiple stages of aspiration and resuspension along with incubation of different dying solution, which all inevitably introduce some variance between experiments. Every effort was made to ensure equal treatment, but human error will still

produce some uncertainty. The adherent cell assay is substantially less comprehensive and less prone to human-errors.

Antibody and other reagents undergo quality degradation over time. The experiments were conducted over a long period, so the degradation effect will inevitably affect the results, but how much it influences is uncertain. Furthermore, the antibodies used were used up and replaced. Albeit the replacement was from the same manufacturer, some variations in production may occur.

The mechanical stress of suspension cell assay removes some of the necrotic cells and debris. This process creates a more homogenous cell population and does not accurately represent the entire sample. This phenomenon is an important consideration for results evaluated using flow cytometry and suspension cell assay. The adherent cell assay was more gentle on the cells, but might still have caused mechanical stress to the cells.

5.1.3 Image acquisition using confocal microscopy

Confocal microscopy can image structures in great detail with an axial resolution of 400-800 *nm* and a lateral resolution of 200 *nm* [58, 59]. The diameter of γ -H2AX foci is larger than the resolution limit of confocal microscopy with a diameter of approximately 0.5-1 μ m [60] making confocal microscopy appropriate for γ -H2AX foci detection.

A small subset of cells from each sample is imaged using confocal microscopy. The imaged sample-population is small due to imaging duration, and the small quanta produce uncertainty in population estimation. Furthermore, there is considerable bias in cell selection for imaging. The bias is more substantial in suspension assay because many regions of the sample are not suited for imaging due to cell overlap and high cell density prohibiting proper segmentation. Thus, the process involves searching for suitable regions for imaging in the microscopy. Furthermore, the extended duration of image acquisition makes regions with high cell density more desirable, as more nuclei can be imaged within one frame and thus saving some time. The adherent cell assay produces substantially less bias when selecting cells. The cells are evenly distributed through the slide in comparison to the slide produced suspension cell assay. The even cell distribution makes almost the entire microscope slide suitable for imaging and thus enables random sampling. The physical location on the microscope slide was selected following a predefined pattern. The only cause for discarding the random image frame was bright illuminating impurities.

The number of cells processed is not consistent throughout all samples within an experiment. Each experiment had a target number of cells to be imaged and evaluated for γ -H2AX fluorescence. The target number of cells plus some additional cells were imaged for good measure. Occasionally some imaged cells could not be processed due to a slight shift of the microscope stand which relocated the sample and causing one or more cells only to be partially covered by the image frame. Furthermore, the starting- and endpoint-depth of the image stack deviated from the depths configures. These deviations were mostly discovered and corrected for, but some were, not resulting in faulty image stacks that did not image the

whole cell in the z-direction. The most substantial source of cell loss during imaging arose from the suspension cell assay. The level of cell stacking in z-direction was not always detected, rendering cells un-segmentable. Imaging of stacked cells could be avoided by meticulous examination throughout the stack, but this is not desirable as it also induced photobleaching of the fluorophores. An equal number of imaged cells in all samples could have been obtained by reducing the number of cells evaluated to the limiting sample (having fewest cells imaged). Downscaling of other samples was decided against as the original number of cells to be imaged was small.

The confocal parameters were set to avoid pixel saturation yet produce a strong enough signal for detection. The determination method was a subjective assessment based on a try and fail methodology. There may exist a better combination of parameters, but it's not easily quantified as the same regions can't be imaged twice without photobleaching affecting the results. Five cells were images four times each with a combination of two zooms and two pinhole sizes to get an indication on confocal parameters influence of the number of DSB detected by 3D-OC. Due to the small number of cells sampled, the test is statistically insignificant. Furthermore, the reliability of the results is questionable as there were no clear distinction between confocal parameters effect and photobleaching influence.

Fluorophore imaging is usually accompanied by photobleaching and this is an important phenomenon to consider when configuring the method. Photobleaching affects subsequent images in z-stacks by lowering the saturation threshold [ref 36 art]. Photobleaching is especially prominent in large z-stacks and their long acquisition duration. Efforts were made to minimize the z-stack size and pixel dwelling to reduce photobleaching, but the effects are inevitable and will affect the results.

The accumulated pixel intensity of each DSB foci was recorded as a measure of damage extent. An absolute measure of image intensity is generally not an ideal parameter as the intensity depends on many factors in sample production. The factors mat be any human-error during the assay or degradation of reagents. The confocal setting was consistent within each experiment evaluating intensity, but any hardware issues could still affect the recorded signal. The reproducibility of absolute intensity is very poor but was included for better comparison between flow cytometry which measures intensity. Furthermore, the DSB intensity was intended for revealing trends in intensity and not evaluated for its exact value.

5.1.4 Image processing

Cell segmentation was a labour-intensive task, and even though every care was taken to ensure proper segmentation of single cells, human error may occur. The risk of multiple cells contributing to a single segmented image was much greater for experiments based on the suspension cell assay, as the cell density and stacking made single-cell segmentation elusive.

The 3D-OC utilises threshold segmentation to distinguish between DSB foci and background signal. The histograms of cell images are monomodal, which makes threshold-setting

problematic as there is no clear distinction between background and foreground pixels. Furthermore, the histogram centre differs between images so that a single threshold value was not suited for all images, and in fact, rendered many samples unprocessed as all intensity value was below the threshold. The "Stack Meancentre5" plugin improved the issue by both centring and broadening the histograms.

5.1.5 Method evaluation

The DSB foci detection was evaluated to trends of other parameters characterising the DSB foci and DSB intensity trends obtained by flow cytometry. The other parameters consisted of accumulated DSB foci intensity (ADI) per cell nucleus, accumulated DSB foci volume (ADV) and percentage of volume marks relative to nucleus volume (*P*). These parameters were chosen for comparison to the number of DSB because they don't rely on foci distinction. The evaluation goal was to confirm the same trends across parameters and detect limitation of DSB foci detection.

The DSB foci detection was compared to manually counting during the early stages of method development to determine 3D-OC configurations. The testing included a limited number of cells from the first experiments. Ideally, the method should be evaluated to manually counting on a more substantial subset of cells and performed by several individuals to minimise bias.

5.2 Performance of the method (DSB detection)

Automated DSB foci detection has a limited operating range in terms of treatment dose. The number of DSB foci trends deviate from trends of parameters ADI and *P* at an irradiation dose of 5 Gy and larger (figure 21). The deviations indicate that the density of DSB foci produced by 5 Gy irradiation is too high to be accurately distinguished. Cells irradiated with doses of 8 Gy and 10 Gy were found to have the lowest number of DSB foci count at 0.5 h incubation time (fig 21). These findings are inconsistent with previous experiments showing that the DNA damage extent is most substantial at 0.5 h incubation compared to 24 h incubation due to the onset and duration of repair mechanisms [61]. Samples irradiated with 5 Gy display the expected trends of a maximum number of DSB foci at 0.5 h incubation and decreasing, but the repair rate is more substantial in plots utilising the parameters ADI and *P* (figure 21(b) and 21(c)). The plot of accumulated DSB foci volume as a function of the number of DSB foci (figure 22) also reflect the finding of too high DSB foci density for accurate distinction at irradiation doses of 5 Gy and larger. This is apparent by a negative relationship between accumulated DSB mark volume, and the number of DSB marks detected - cells irradiated with 5 Gy having the most massive accumulated marker volume score combined with the lowest counts of the number of DSB foci. Figure 23 display more consistency of a positive relationship between accumulated volume and the number of DSB foci in samples that are not irradiated or irradiated with a dose of 2 Gy, but some outlier cells are deviating from the trend. Scatterplot and Pearson's Correlation Coefficient (PCC) in figure 30 reflect the positive linear relationship between the number of DSB marks and parameter ADI and *P*. Literature suggest that the

number of DSB induced in a cell per Gray is between 30 and 40 [62]. The number of DSB detected by the developed method is drastically lower, i.e. the median number of DSB in the sample treated with 2 Gy and harvested after 0.5 hours incubation was nearly six (fig 21) when the theoretical value is between 60 and 80. The cause of low DSB detection may be that the threshold determining object pixels is too high.

5.3 γ -H2AX experiment results

The following section will elaborate around the experimental findings that arose from visual inspection of the confocal images.

The red channel has protruding regions with relatively strong PI fluorescence compared to the rest of the nucleus. Few to none DSB foci were detected in these regions. The protruding PI region consists of compacted DNA which has a high density of bindings sites for PI fluorescence. X-ray irradiation is characterised by low LET and hence dispersed energy deposition [24]. The dominating form of energy deposition of low LET irradiation is the indirect effect which is limited the compact DNA as the accessibility for water radicals is reduced. High LET irradiation like protons deposition from the back of the Bragg peak predominantly deposits energy through the direct effect. Irradiation of the cells with protons would most likely cause more DSB in compact DNA.

It was hypothesised that nucleus characteristics such as volume and PI intensity could identify subgroups in the cell sample according to the cell cycle. The desire was to enable examination of damage extent and DSB foci distribution in the different cell phases to get a detailed insight to repair mechanisms. The strategy was to apply the same principles used in flow cytometry to gate for different cell cycle phases by investigating the differences in PI intensity. The PI signal intensity varies in the cell cycle [63]. The most prominent distinction is between G1 and G2, due to doubling of DNA content during S-phase replication [15]. The testing for correlation between PI intensity and DNA damage parameters displayed little consistency across doses. The images had been processed with MC5 plugin, which shifts the histogram centre and consequently information of PI signal intensity is lost. Similarly to PI intensity, it was thought that nucleus volume could indicate cell cycle phase, but this did not display the consistent correlation between damage extent parameters P, ADI and ADV. The correlation might be improved with a larger sample size.

Multiple causes for differences in PI signal intensity has been evaluated. It was hypothesised that the PI solution did not cover all cells equally in the suspension cell assay due cell aggregation. This explanation was rejected when the same trend was observed in the adherent cell assay, where all cells are evenly distributed throughout the microscope slide. In both assays, the microscope slide with samples must dry completely before sealing the sample. One theory was that oxygen exposure affected the PI intensity. This could explain regional PI differences as the duration of water evaporations differs across the surface.

However, the samples exposed to oxygen for 70 minutes had similar PI signal to samples exposed for 30 minutes.

Several cells had broken nuclear envelope. This might be a sign of fungal infection. The control samples of adherent cell assay experiment (experiment A1, table 3) showed varying damage extent across incubation times. This observation may support the suspicion of fungal infection, but might also be caused by other forms of stress, such as mechanical stress or high level of confluency. This observation may support the suspicion of fungal infection, but might also be caused by other forms of stress affecting samples differently, such as mechanical stress or high level of confluency. It is not uncommon for cancerous cells to have pre-existing DSB [64], but this effect should be similar across samples.

Daughter nuclei of cell irradiated during mitosis frequently displayed a substantial difference in damage extent between the two. This insight was discovered by visual inspection of the overview images and was not searched for or evaluated for statistical significance. There seemed to a tendency of either both daughter nuclei having low damage extend or one receiving the majority of the impact. If the cells were in mitosis during irradiation this may be caused by one cells receiving the majority of energy deposition. Another possible cause for the differences is that the radiation impact causes genomic instability which results in DNA DSB after some time.

The OPTICS clustering classified approximately all cells to not be clustered. However, the Moran's I classified approximately all cells to be significantly clustered. The values are not directly comparable as the Moran's value depends on both pixel values, the number of pixels, spatial location and definition of the weight matrix, whereas OPTICS only asses the spatial location of DSB. It is a common strategy to use normality assumption when testing for Moran's I, however there are alternative methods that might produce different results [65, 66]. Moran's I can alternatively be used to rank the cells in terms of clustering and or dispersion. A direct comparison of the Moran's I would be appropriate for images obtained by confocal microscopy given that the dimensions are equal for all images and the weight matrix consistently defined.

5.4 γ -H2AX detection using flow cytometry VS confocal microscopy

Confocal imaging detects the same trends of γ -H2AX fluorescence as zero resolution flow cytometry. The most accurate comparison between the two detection methods is achieved by evaluating the flow cytometry data of cells from the entire cell cycle without not normalising to DNA content. Flow cytometry can easily gate between different cell cycle phases. Confocal microscopy is beneficial regarding insight into the internal of the cell, damage distribution and cell environment. The intracellular distribution of DSBs can only be detected using confocal microscopy, but there are limitations to how large irradiation doses that can be assessed. Flow cytometry is limited to measuring fluorophore intensity, but can in return, evaluate several thousand cells in minutes. In contrast, both image acquisition and -processing is labour-intensive and time-consuming using confocal microscopy. Imaging flow cytometry can however detect the number of γ -H2AX foci by imaging cells from multiple

channels simultaneously [67] and offers an efficient method for DSB foci detection with statistical robustness.

5.5 DSB's compared to the LQ-model and survival data

It was hypothesised that logarithmic surviving fraction (SF) and the number of DSBs would form a linear relationship, assuming SF is a function of DNA damage extent. The hypothesis was investigated by plotting the logarithmic SF against the percentage of DSB marks volume relative to the corresponding volume of the nucleus (figure X). The number of marks per cell was not chosen for evaluation due to the lack of reliability at higher irradiation doses. All incubation times displayed a linear relationship to the survival fraction except the 0.5 h incubation time. The SF evaluates survival after approximately ten days, so the most relevant linear regression lines are the ones for 72 h incubation which does not contradict the hypothesis.

5.6 Method compared to FocAn

FocAn is ImageJ based plugin that analysis foci in 3D using z-stacks obtained by confocal microscopy [68]. The algorithm detects individual nuclei and counts the number of DSB foci based only on the channel with foci-staining, and does not require additional DNA-binding fluorophore staining to detect the nucleus. Application of 3D watershed segmentation enables FocAn to distinguish between overlapping foci. FocAn requires little user input which improves usability. The developed method of this thesis is a lot more labour-intensive compared to FocAn, but has the advantage of adaptability in terms of which parameters to evaluate as well as analysis of spatial distribution.

5.7 Future considerations

The DSB foci detection method would be more useful if the very labour-intensive and time-consuming cell segmentation were automatized. It would be possible to run the 3D-OC analysis on the entire image stack produced by confocal microscope on both red and green colour channel. A simple script could use coordinates of all objects detected and map DSB marks to their corresponding nuclei with all parameters investigated. This alternative method does, however, require a lot of processing power from the computer used for analysis.

Mitotic selection of seeded cells would ease the cell cycle phase determination by producing an approximate synchronous cell population. The number of cells imaged is too small to accurately describe the variation in an asynchronous cell population.

More investigation into the threshold settings applied for segmentation is needed to improve DSB detection by the 3D-OC. The number of DSB is consistently too low compared to the theoretical expected value given the radiation doses.

6 Conclusion

A method for quantifying damage extent and spatial distribution of DNA DSBs was successfully made. The experimental procedure for H2AX cell assay was developed over the course of this project in order to optimise the DNA DSB detection and environmental information of the cell. The adherent cell assay is most suited for this purpose, as cell location is maintained through the entire procedure, whereas in the suspension cell assay, all environmental information is lost. Furthermore, there is considerable cell loss in the suspension cell assay due to the trypsinisation and centrifugation steps; this results in a homogenous cell population.

Several parameters of damage extent were investigated, such as DSB foci intensity, and the percentage of mark volume relative to nucleus volume. These two parameters showed trends consistent with γ -H2AX fluorescence which was found by flow cytometry. The number of DSB are however too low compared to the theoretically expected value. This is due to a too high threshold setting. A more elaborate investigation into threshold settings and following DSB detection is needed.

Parameters of spatial distribution were assessed. It was found that almost all of the cells in experiment using adherent cells (experiment A1) were clustered when the significance of Moran's I for these cells was tested. This might be a trait of the radiation quality. Comparison of the spatial distribution of DSB for other radiation types and LET values should be investigated, as it might also be a good measure of damage complexity.

References

1. WHO. *Cancer*. 2018; Available from: <https://www.who.int/news-room/fact-sheets/detail/cancer>.
2. Burma, S., et al., *ATM Phosphorylates Histone H2AX in Response to DNA Double-strand Breaks*. *J Biol Chem*, 2001. **276**(45): p. 42462-42467.
3. Chen, H.T., et al., *Response to RAG-mediated VDJ cleavage by NBS1 and gamma-H2AX*. *Science*, 2000. **290**(5498): p. 1962.
4. Rogakou, E.P., et al., *DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139*. *J Biol Chem*, 1998. **273**(10): p. 5858-5868.
5. E.P. Rogakou, D.R.P., A.H. Orr, V.S. Ivanova, W.M. Bonner,, *DNA doublestranded breaks induce histone H2AX phosphorylation on serine 139*. *J. Biol. Chem.*, 1998. **273**: p. 5858-5868.
6. O.A. Sedelnikova, D.R.P., C. Redon, W.M. Bonner, *Histone H2AX in DNA damage and repair*. *Cancer Biol. Ther.*, 2003. **2**: p. 233-235.
7. Rogakou, E.P., et al., *Megabase chromatin domains involved in DNA double-strand breaks in vivo*. *Journal Cell Biology*, 1999. **146**: p. 905-916.
8. W.M. Bonner, C.E.R., J.S. Dickey, A.J. Nakamura, O.A. Sedelnikova, S. Solier, and Y. Pommier, *H2AX and cancer*,. *Nat. Rev. Cancer*, 2008(8): p. 957-967.
9. Sedelnikova, O.A., et al., *Quantitative detection of (125)IdU-induced DNA double-strand breaks with gamma-H2AX antibody*. *Radiat Res*, 2002. **158**(4): p. 486.
10. Loblrich, M., et al., *In vivo formation and repair of DNA double-strand breaks after computed tomography examinations*. *Proc Natl Acad Sci U S A*, 2005. **102**(25): p. 8984-8989.
11. Alesia N. Ivashkevicha, O.A.M., Andrea J. Smitha, Christophe E. Redonb, and R.F.M. William M. Bonnerb, Pavel N. Lobachevskya,, *gamma-H2AX foci as a measure of DNA damage: A computational approach to automatic analysis*. *Mutation Research / Fundamental and Molecular Mechanisms of Mutagenesis*, 2011. **711**(1-2): p. 49-60.
12. Sejpal, S.V., A. Bhate, and W. Small, *Palliative radiation therapy in the management of brain metastases, spinal cord compression, and bone metastases*. *Semin Intervent Radiol*, 2007. **24**(4): p. 363-74.
13. PTCOG. *Particle therapy facilities in clinical operation*. 2020; Available from: <https://www.ptcog.ch/index.php/facilities-in-operation>.
14. Parris, C.N., et al., *Enhanced γ -H2AX DNA damage foci detection using multimagnification and extended depth of field in imaging flow cytometry*. *Cytometry A*, 2015. **87**(8): p. 717-23.
15. Alberts, B., *Molecular biology of the cell*. 5th ed. ed. 2008, New York: Garland Science.
16. J. Gordon Betts, P.D., Eddie Johnson, Jody E. Johnson, Oksana Korol, Dean Kruse, Brandon Poe, James A. Wise, Mark Womble, Kelly A. Young. *Anatomy and physiology*. Available from: <https://opentextbc.ca/anatomyandphysiology/chapter/3-3-the-nucleus-and-dna-replication/>.
17. Bębenek, A. and I. Ziuzia-Graczyk, *Fidelity of DNA replication—a matter of proofreading*. *Curr Genet*, 2018. **64**(5): p. 985-996.
18. Cooper, G.M., *The Cell : a molecular approach*. 2000, ASM Press: Washington, D.C.
19. Maqbool, M., *An Introduction to Medical Physics*. 2017, Cham: Springer: Cham. p. 43-55.
20. Lilley, J.S., *Nuclear physics : principles and applications*. The Manchester physics series. 2001, Chichester: Wiley.
21. *Proton Therapy*. 2015; Available from: <https://www.oncopods.com/proton-therapy.html>.
22. Hall, E.J., Giaccia, A.J., *Radiobiology For The Radiologist*. 7th ed.: LIPPINCOTT WILLIAMS & WILKINS, a WOLTERS KLUWER business.
23. Attix, F.H., *Introduction to radiological physics and radiation dosimetry*. 1986, New York: Wiley.

24. Bushberg, J.T., J. A. Seibert, and Edwin M. Leidholdt., *Essential Physics of Medical Imaging*, 2011: Wolters Kluwer Health.
25. Frame, P. *Coolidge X-ray Tubes*. 2009; Available from: <https://www.orau.org/ptp/collection/xraytubescoolidge/coolidgeinformation.htm>.
26. Simon R. Cherry, J.A.S., Michael E. Phelps, *Physics in Nuclear Medicine*. 4th ed. 2012: Elsevier Inc.
27. Robertson, A., et al., *The cellular and molecular carcinogenic effects of radon exposure: a review*. *Int J Mol Sci*, 2013. **14**(7): p. 14024-14063.
28. Terasima, T. and L.J. Tolmach, *Variations in several responses of HeLa cells to x-irradiation during the division cycle*. *Biophys J*, 1963. **3**(1): p. 11-33.
29. Terasima, T. and L.J. Tolmach, *Changes in X-ray Sensitivity of HeLa Cells during the Division Cycle*. *Nature*, 1961. **190**(4782): p. 1210-1211.
30. Nwaneshiudu, A., et al., *Introduction to Confocal Microscopy*. *Journal of investigative dermatology*, 2012. **132**(12): p. 1-5.
31. *Fluorescence microscopy from principles to biological applications*, ed. U. Kubitscheck. 2013, Weinheim [Germany: Wiley-Blackwell.
32. Jerome, W.G. and R.L. Price, *Basic Confocal Microscopy*. 2018, Springer International Publishing : Imprint: Springer: Cham.
33. Prasad, P. and B. Masters, *Introduction to Biophotonics*. *Journal of biomedical optics*, 2005. **10**: p. 39901.
34. Yen-Nien, H., et al., *Development of an Automated γ -H2AX Immunocytochemistry Assay*. *Radiat Res*, 2009. **171**(3): p. 360-367.
35. *Microscopy Techniques and Culture Surfaces: Confocal microscopy*. Available from: <https://ibidi.com/content/216-confocal-microscopy>.
36. Weaver, J.L., *Introduction to Flow Cytometry*. *Methods*, 2000. **21**(3): p. 199-201.
37. Rasband, W.S., *ImageJ*. 1997-2018, U. S. National Institutes of Health: Bethesda, Maryland, USA.
38. Bolte, S. and F.P. CordeliÈRes, *A guided tour into subcellular colocalization analysis in light microscopy*. *J Microsc*, 2006. **224**(3): p. 213-232.
39. Cordelières, F.P. *3D object counter*. 2012; Available from: https://imagejdocu.tudor.lu/doku.php?id=plugin:analysis:3d_object_counter:start.
40. Moran, P.A.P., *NOTES ON CONTINUOUS STOCHASTIC PHENOMENA*. *Biometrika*, 1950. **37**(1-2): p. 17-23.
41. Carrijo, T.B. and A.R. da Silva, *Modified Moran's I for Small Samples: Modified Moran's I*. *Geographical analysis*, 2017. **49**(4): p. 451-467.
42. Radil, S.M. *Spatializing social networks: Making space for theory in spatial analysis*. 2011.
43. Ankerst, M., et al., *OPTICS: ordering points to identify the clustering structure*. *SIGMOD Rec.*, 1999. **28**(2): p. 49–60.
44. Schubert, E. and M. Gertz. *Improving the Cluster Structure Extracted from OPTICS Plots*. in *LWDA*. 2018.
45. Margaret, A.K.K. and A. Margaret, *Introduction to the Cellular and Molecular Biology of Cancer*. 2005, GB: Oup Oxford: GB.
46. ATCC. Available from: https://www.lgcstandards-atcc.org/products/all/CCL-185.aspx?geo_country=no.
47. Rawlings, N.D., et al., *Handbook of Proteolytic Enzymes*. 2013, GB: Academic Press: GB.
48. Stoker, M.G.P. and H. Rubin, *Density Dependent Inhibition of Cell Growth in Culture*. *Nature*, 1967. **215**(5097): p. 171-172.
49. WALDELAND, E., HOLE, E. O., SAGSTUEN, E., MALINEN, E., *The energy dependence of lithium formate and alanine EPR dosimeters for medium energy x rays*. *Med Phys*, 2010(37): p. 3569-75.
50. Rosser, K.E., *Measurement of absorbed dose to water for medium energy X-rays*. 1996, UCL (University College London).

51. (2014)., R.C.T., *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria.
52. Van Rossum, G., & Drake, F. L., *Python 3 Reference Manual*. 2009, CA: CreateSpace.: Scotts Valley.
53. Li, C.H. and C.K. Lee, *Minimum cross entropy thresholding*. Pattern recognition, 1993. **26**(4): p. 617-625.
54. Li, C.H. and P.K.S. Tam, *An iterative algorithm for minimum cross entropy thresholding*. Pattern recognition letters, 1998. **19**(8): p. 771-776.
55. Vejdemo-Johansson, M. *Python implementation of the Gap Statistics from Tibshirani, Walther, Hastie to determine the inherent number of clusters in a dataset with k-means clustering*. 2013; Available from: <https://gist.github.com/teruo41/882c8bd7f411bf0adbdd217040bcd3e3>.
56. Shibata, A. and P. Jeggo, *DNA Double-strand Break Repair in a Cellular Context*. Clinical Oncology, 2014. **26**.
57. Rykkelid, A.M., *Method for in vitro cell irradiation with low energy protons*. 2017.
58. Heintzmann, R. and G. Ficz, *Breaking the resolution limit in light microscopy*. Brief Funct Genomic Proteomic, 2006. **5**(4): p. 289-301.
59. Wilson, T., *Resolution and optical sectioning in the confocal microscope*. J Microsc, 2011. **244**(2): p. 113-121.
60. Sisario, D., et al., *Nanostructure of DNA repair foci revealed by superresolution microscopy*. The FASEB Journal, 2018. **32**(12): p. 6469-6477.
61. Beyreuther, E., et al., *DNA double-strand break signalling: X-ray energy dependence of residual co-localised foci of γ -H2AX and 53BP1*. International Journal of Radiation Biology, 2009. **85**(11): p. 1042-1050.
62. Ward, J.F., *DNA Damage Produced by Ionizing Radiation in Mammalian Cells: Identities, Mechanisms of Formation, and Reparability*, in *Progress in Nucleic Acid Research and Molecular Biology*, W.E. Cohn and K. Moldave, Editors. 1988, Academic Press. p. 95-125.
63. Mori, R., et al., *Estimation of the radiation-induced DNA double-strand breaks number by considering cell cycle and absorbed dose per cell nucleus*. Journal of Radiation Research, 2018. **59**.
64. Vilenchik, M.M. and A.G. Knudson, *Endogenous DNA double-strand breaks: Production, fidelity of repair, and induction of cancer*. Proceedings of the National Academy of Sciences, 2003. **100**(22): p. 12871.
65. Levine, N., *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 4.02)*. 2015: Washington, D.C.
66. Mathur, M., *Spatial autocorrelation analysis in plant population: An overview*. Journal of Applied and Natural Science, 2015. **7**(1): p. 501-513.
67. Mori, R., et al., *Estimation of the radiation-induced DNA double-strand breaks number by considering cell cycle and absorbed dose per cell nucleus*. Journal of radiation research, 2018. **59**(3): p. 253-260.
68. Memmel, S., et al., *FocAn: automated 3D analysis of DNA repair foci in image stacks acquired by confocal fluorescence microscopy*. BMC Bioinformatics, 2020. **21**(1): p. 27.

Appendix

A Assay protocols

Table 6 displays an overview of reagents used in the protocols. The dilution factor specifies concentration used, or the highest concentration tested.

Tabell 6 Overview of reagents used in protocols

Protocol name	Full name	Stock concentration	Dilution
Primary antibody	Anti-phospho-Histone H2A.X (Ser139), clone JBW301	1 mg / ml	1 : 500
Secondary antibody	Polyclonal Rabbit Anti-Mouse Immunoglobulins/FITC (Dako, F0232)	2.7 mg / ml	1 : 30
PI	Propodium iodide (Sigma, USA)	1 mg / ml	1 : 5000
Conjugated antibody	Anti-phospho-Histone H2A.X (Ser139) Antibody, clone JBW301, FITC conjugate SDS	100 µg / ml	1 : 25
Secondary antibody	Goat Anti-Mouse IgG H&L (FITC) (ab6785) (abcam)	2 mg/ml	1 : 1000

A.1 Protocol for suspension cell assay

1. Remove medium and wash with approximately 4 ml PBS
2. Add 3 ml Trypsin and place in incubator for approximately 3 minutes
3. Gently pipette with a 2 ml pipette and balloon to ensure maximum number of single cells
4. Transfer to tube containing 5 ml of cold medium
5. Spin down the samples at 200 G for 4 minutes, remove supernatant and gently vortex
6. Resuspend in 3 ml PBS
7. Spin down the samples at 200 G for 4 minutes, remove supernatant and gently vortex
8. Resuspend pellet in 400 µl PBS, and add 2 ml -20° 100 % methanol drop by drop while gently vortexing
9. Store at -20° for 1 hour
10. Add 5 ml PBS to each sample
11. Spin down the samples at 200 G for 4 minutes, remove supernatant and gently vortex
12. Add 50 µl PBS with 1 % BSA suspension containing the primary antibody

13. Vortex and incubate for 30 *minutes* in the dark and with room temperature
14. Add 3 *ml* PBS to each sample
15. Spin down the samples at 200 *G* for 4 *minutes*, remove supernatant and gently vortex
16. Add 5 *ml* PBS to each sample
17. Spin down the samples at 200 *G* for 4 *minutes*, remove supernatant and gently vortex
18. Resuspend in 100 μ l PBS with 1 % BSA suspension containing the secondary antibody
19. Vortex and incubate for 30 *minutes* in the dark with room temperature
20. Add 3 *ml* PBS to each sample
21. Spin down the samples at 200 *G* for 4 *minutes*, remove supernatant and gently vortex
22. Resuspend the cells in 200 μ l PBS containing PI
23. Vortex and incubate for 30 *minutes* in the dark with room temperature
24. Transfer 15 μ l of cell sample through filter to the microscope slide
25. Let the sample dry completely in the dark with room temperature before proceeding
26. Add approximately 10 μ l of VectaShield mounting medium to the sample, place cover slip on top and seal with nail polish around edges

A.2 Protocol for Adherent cell assay

1. Remove medium and wash twice with 1,2 *ml* cold PBS
2. Add 400 μ l PBS, and 2 *ml* -20° 100 % methanol drop by drop while gently swirling
3. Store at -20° for 5 *minutes*
4. Remove methanol and wash with 1,2 *ml* PBS
5. Add 1,2 *ml* PBS and swirl for 4 *minutes*
6. Wash with 1,2 *ml* PBS
7. Add in 500 μ l PBS with 1% BSA containing the primary antibody
8. Swirl and incubate for 30 *minutes* in the dark with room temperature
9. Wash with 1,2 *ml* PBS
10. Add 1,2 *ml* PBS and leave to swirl for 4 *minutes*
11. Add in 500 μ l PBS w/ 1% BSA containing the secondary antibody
12. Swirl and incubate for 30 *minutes* in the dark with room temperature
13. Wash with 1,2 *ml* PBS
14. Add 1,2 *ml* PBS and leave to swirl for 4 *minutes*
15. Resuspend in 500 μ l PBS containing PI
16. Swirl and incubate for 15 *minutes* in room temperature in the dark with room temperature
17. Wash with 1,2 *ml* PBS
18. Add 1,2 *ml* PBS and leave to swirl for 4 *minutes*
19. Break off-cover glass and gently rinse with Milli-Q water
20. Let the sample dry completely in the dark with room temperature before proceeding
21. Add approximately 10 μ l of VectaShield mounting medium to the sample, place cover slip on top and seal with nail polish around edges

A.3 Protocol for fixation and permeabilization testing

Methanol:

1. 1,2 ml ice cold methanol and incubate for 5 min in freezer
2. Remove methanol

Formaldehyde:

1. Fixate with 1.2 mL 4% Formaldehyde in PBS and incubate in the dark for 15 *minutes*
2. Wash twice with 1.2 ml PBS
3. Permeabilize with 1.2 ml -20° 100 % methanol and incubate on ice in the dark for 5 *minutes*

Assay:

1. Remove medium and wash twice with 2 ml cold PBS and remove excess PBS
2. **Fixation and permeabilization**
3. Wash twice with 1.2 ml PBS
4. Add in 2 ml PBS with 1% BSA containing the conjugated antibody
5. Swirl and incubate for 30 minutes in the dark with room temperature
6. Wash twice with 2.5 ml PBS
7. Resuspend in 500 µl PBS containing PI
8. Swirl and incubate for 15 *minutes* in room temperature in the dark
9. Wash twice with 1.2 ml PBS
10. Break off-cover glass and gently rinse with Milli-Q water
11. Let the sample dry completely in the dark with room temperature before proceeding
12. Add approximately 10 µl of VectaShield mounting medium to the sample, place cover slip on top and seal with nail polish around edges

A.4 Protocol for antibody testing

Protocol in appendix section A.3 was used to test different antibodies and antibody combination using methanol for permeabilization.

B Programming Scripts

B.1 Programme for determining cell confluency

CellSegmentation

```
function [areastats_colony, redstats_colony, greenstats_colony, ...  
    bluestats_colony, graystats_colony, pca1stats_colony, ...  
    pca2stats_colony, confluency] = CellSegmentation(sourcepath, ...  
    destpath, sourcepath_dish, pixelsize)
```

```
[~, filename, ext] = fileparts(sourcepath);
```

```

%% If a valid screen size is returned (MATLAB was run without -nodisplay)
if usejava('jvm') && feature('ShowFigureWindows')
    % Start progress bar
    progress = waitbar(0, [[filename ext] ': Reading colony image'], ...
        'Name', 'Segmenting colonies...', ...
        'CreateCancelBtn', 'setappdata(gcf,'canceling',1)');
    setappdata(progress, 'canceling', 0);
    pause(1)
end

%% Read RGB image with color channels
[img_in, ~] = imread(sourcepath);
img_original = img_in;

% Extract color channel and normalize between [0,1]
redchannel = double(img_in(:, :, 1));
greenchannel = double(img_in(:, :, 2));
bluechannel = double(img_in(:, :, 3));

img_red = (redchannel - min(min(redchannel))) ./ ...
    abs(max(max(redchannel)) - min(min(redchannel)));
img_green = (greenchannel - min(min(greenchannel))) ./ ...
    abs(max(max(greenchannel)) - min(min(greenchannel)));
img_blue = (bluechannel - min(min(bluechannel))) ./ ...
    abs(max(max(bluechannel)) - min(min(bluechannel)));

% Create spatial referencing object associated with the image to use it to
% set the x- and y-axes limits in the world coordinate system
sizex = size(img_in, 2);
sizey = size(img_in, 1);
xmax = sizex * pixelsize;
ymax = sizey * pixelsize;

```

```

RI = imref2d(size(img_in));
RI.XWorldLimits = [0 xmax];
RI.YWorldLimits = [0 ymax];

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')
    delete(progress)
    delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
    return
end
if exist('progress', 'var') && ishandle(progress)
    waitbar(0.1, progress, [[filename ext] ...
        ': Grayscale conversion and normalization'], ...
        'Name', 'Segmenting colonies...');
    pause(1)
end

%% Convert to grayscale
img_in = double(rgb2gray(img_in));

%% Normalization between 0-1
[img_norm, img_norm2] = normalizationMinMax(img_in);
img_gray = img_norm;

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')
    delete(progress)
    delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
    return
end
if exist('progress', 'var') && ishandle(progress)

```



```

waitbar(0.2, progress, [[filename ext] ...
    ': Cell dish border extraction'], ...
    'Name', 'Segmenting colonies...');
end

%% Attain only the areas within the cell dish/flask
bw_border = getCellContainerBorder(sourcepath_dish, destpath, filename, img_norm);
img_norm = bw_border .* img_norm;

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')
    delete(progress)
    delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
    return
end
if exist('progress', 'var') && ishandle(progress)
    waitbar(0.3, progress, [[filename ext] ...
        ': Principal component analysis'], ...
        'Name', 'Segmenting colonies...');
end

%% Principal component analysis (PCA)
[~, ~, ~, img_pca1, img_pca2, img_pca3] = runPCA(double(img_original));

img_pca = img_pca2;
if isempty(img_pca2) && ~isempty(img_pca1)
    img_pca = img_pca1;
end

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')

```

```

delete(progress)
delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
return
end
if exist('progress', 'var') && ishandle(progress)
    waitbar(0.35, progress, [[filename ext] ...
        ': Filtration and histogram equalization'], ...
        'Name', 'Segmenting colonies...');
    pause(1)
end

%% Guassian smoothing and removal of outliers

% Gaussian filtering
img_norm1 = imgaussfilt(img_norm, [3 3]);
% Opening-by-Reconstruction
img_e = imerode(img_norm1, strel('disk', 6));
img_obr = imreconstruct(img_e, img_norm1);
% Opening-Closing-by-Reconstruction
img_obrd = imdilate(img_obr, strel('disk', 6));
img_obrcbr = imreconstruct(imcomplement(img_obrd), imcomplement(img_obr));
img_norm1 = imcomplement(img_obrcbr);
img_norm1(img_norm1 > 0.8) = 0.8;

%% Background correction
if mean(mean(img_norm2)) > 0.45
    img_norm = imtophat(img_norm1, strel('disk', 90));
else
    img_norm = imbothat(img_norm1, strel('disk', 90));
end

%% Perform CLAHE (histogram equalization)
img_norm = (adapthisteq(img_norm) - min(min(adapthisteq(img_norm)))) ./ ...

```

```

abs(max(max(adapthisteq(img_norm))) - min(min(adapthisteq(img_norm))));

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')
    delete(progress)
    delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
    return
end
if exist('progress', 'var') && ishandle(progress)
    waitbar(0.4, progress, [[filename ext] ...
        ': BLOB extraction by K-means'], ...
        'Name', 'Segmenting colonies...');
end

%% Edge detection by iterative Otsu segmentation
area = [80 100000000000];
bw_BLOBs = extractBLOBs(img_pca.*bw_border, bw_border, area(1));
bw_seg = bw_BLOBs;

%% Estimate staining density
area_original = sum(bw_border(:));
area_staining = sum(bw_BLOBs(:));
confluency = area_staining/area_original; % cell density

%% Plot
figure(); imshow(img_original)
hold on
visboundaries(bw_seg, 'Color', 'red', 'LineWidth', 1)
errorbar(0.1*size(bw_seg,2), 0.95*size(bw_seg,1), round(1/(pixelsize)), ...
    'horizontal', 'k.', 'LineWidth', 1.5, 'CapSize', 6);
text(0.1*size(bw_seg,2), 0.95*size(bw_seg,1), {'2 mm'}, ...
    'HorizontalAlignment', 'center', 'VerticalAlignment', 'top', 'FontSize', 12);

```

```

hold off
set(gca, 'FontSize', 14)
saveas(h, fullfile(destpath, sprintf('%s-Seg.fig', filename)))

% Colony mean intensity feature
areastats_colony = regionprops(bw_seg, img_norm, 'MeanIntensity');
all_meanintensities = [areastats_colony.MeanIntensity];

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')
    delete(progress)
    delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
    return
end
if exist('progress', 'var') && ishandle(progress)
    waitbar(0.9, progress, [[filename ext] ...
        ': Colony intensity estimation'], ...
        'Name', 'Segmenting colonies...');
    pause(1)
end

%% Colony feature estimations; area, RGB, gray, PCA1 and PCA2
areastats_colony = regionprops(bw_seg, 'Area', 'Centroid', 'PixelIdxList');
redstats_colony = regionprops(bw_seg, img_red, 'PixelValues');
greenstats_colony = regionprops(bw_seg, img_green, 'PixelValues');
bluestats_colony = regionprops(bw_seg, img_blue, 'PixelValues');
graystats_colony = regionprops(bw_seg, img_gray, 'PixelValues');
pca1stats_colony = regionprops(bw_seg, img_pca1, 'PixelValues');
pca2stats_colony = regionprops(bw_seg, img_pca2, 'PixelValues');
for i = 1:length(areastats_colony)
    redstats_colony(i).MedianIntensity = ...
        median(redstats_colony(i).PixelValues);

```

```

greenstats_colony(i).MedianIntensity = ...
    median(greenstats_colony(i).PixelValues);
bluestats_colony(i).MedianIntensity = ...
    median(bluestats_colony(i).PixelValues);
graystats_colony(i).MedianIntensity = ...
    median(graystats_colony(i).PixelValues);
pca1stats_colony(i).MedianIntensity = ...
    median(pca1stats_colony(i).PixelValues);
pca2stats_colony(i).MedianIntensity = ...
    median(pca2stats_colony(i).PixelValues);

redstats_colony(i).SD = std(redstats_colony(i).PixelValues);
greenstats_colony(i).SD = std(greenstats_colony(i).PixelValues);
bluestats_colony(i).SD = std(bluestats_colony(i).PixelValues);
graystats_colony(i).SD = std(graystats_colony(i).PixelValues);
pca1stats_colony(i).SD = std(pca1stats_colony(i).PixelValues);
pca2stats_colony(i).SD = std(pca2stats_colony(i).PixelValues);
end
all_areas = [areastats_colony.Area] .* pixelsize^2;
mean_area = mean(all_areas);
SD_area = std(all_areas);

if isfield(redstats_colony, 'PixelValues')
    redstats_colony = rmfield(redstats_colony, 'PixelValues');
end
if isfield(greenstats_colony, 'PixelValues')
    greenstats_colony = rmfield(greenstats_colony, 'PixelValues');
end
if isfield(bluestats_colony, 'PixelValues')
    bluestats_colony = rmfield(bluestats_colony, 'PixelValues');
end
if isfield(graystats_colony, 'PixelValues')
    graystats_colony = rmfield(graystats_colony, 'PixelValues');
end

```

```

end
if isfield(pca1stats_colony, 'PixelValues')
    pca1stats_colony = rmfield(pca1stats_colony, 'PixelValues');
end
if isfield(pca2stats_colony, 'PixelValues')
    pca2stats_colony = rmfield(pca2stats_colony, 'PixelValues');
end

% Save final segmented colony mask
writematrix(bw_seg, fullfile(destpath, sprintf('%s-SegMask.csv', filename)))

%% Update progress bar
if exist('progress', 'var') && ishandle(progress) && ...
    getappdata(progress, 'canceling')
    delete(progress)
    delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
    return
end
if exist('progress', 'var') && ishandle(progress)
    waitbar(1.0, progress, 'Segmentation completed', ...
        'Name', 'Segmenting colonies...');
    pause(1)
end

%% Close waitbar
if exist('progress', 'var') && ishandle(progress)
    close(progress);
end

%% Delete progress handle if it exists
if exist('progress', 'var') && ishandle(progress), delete(progress); end

end

```

extractBLOBs

```
function [bw] = extractBLOBs(img, border, area_min)

%% Normalization between 0-1 and Gaussian filtering
img = (img - min(min(img))) ./ ...
    abs(max(max(img)) - min(min(img)));
img = imgaussfilt(img, [3 3]);
img = (adapthisteq(img) - min(min(adapthisteq(img)))) ./ ...
    abs(max(max(adapthisteq(img))) - min(min(adapthisteq(img))));

%% Threshold by means of K-means to extract conglomerate colony BLOBs
bw = runKmeans(img);
bw = bw .* border;

% Clean up
bw = imclearborder(bw, 8);
bw = bwmorph(bw, 'bridge', Inf);
bw = bwmorph(bw, 'clean', Inf);
bw = bwpropfilt(bw, 'Area', [area_min/2 2*10000000]);
bw = imfill(bw, 'holes');
bw = bwpropfilt(bw, 'Area', [area_min/2 2*10000000]);
bw = imfill(bw, 'holes');
bw = imdilate(bw, ones(2));
% Remove disk-shaped structures with radius less than 3 pixels
bw = imopen(bw, strel('disk', 3));
% Remove objects containing fewer than a given pixels
bw = bwareaopen(bw, (area_min/2) - 10); % increase to remove more "smudges"
bw = bwpropfilt(bw, 'Area', [area_min/2 2*10000000]);

end
```

getCellContainerBorder

```
function [bw] = getCellContainerBorder(sourcepath, destpath, filename, img_fixed)
```

```
% Read empty cell dish/flask image
```

```
[img, ~] = imread(sourcepath);
```

```
% Normalization between 0-1
```

```
[img_norm, ~] = normalizationMinMax(double(rgb2gray(img)));
```

```
%% Estimate geometric transformation that aligns the two 2D images
```

```
[optimizer, metric] = imregconfig('multimodal');
```

```
tform = imregtform(img_norm, img_fixed, 'rigid', optimizer, metric);
```

```
img_norm = imwarp(img_norm, tform, 'OutputView', imref2d(size(img_fixed)));
```

```
%% Cell dish/flask extraction
```

```
% Preprocessing of the image
```

```
img_norm = imgaussfilt(img_norm, [2 2]);
```

```
img_norm = adapthisteq(img_norm, 'Distribution', 'exponential');
```

```
if ~isempty(size(img_norm, 3)) && size(img_norm,3) == 3
```

```
    img_norm = double(mean(img_norm,3));
```

```
elseif ~isempty(size(img_norm,3)) && size(img_norm,3) == 2
```

```
    img_norm = double(mean(img_norm,2));
```

```
else
```

```
    img_norm = double(img_norm);
```

```
end
```

```
% Threshold detection using most occurring value in histogram
```

```
[counts, x] = hist(img_norm(:), 255);
```

```
delta = 0.25;
```



```

bw1 = img_norm <= x(find(counts == max(max(counts)))) + delta;
bw2 = img_norm >= x(find(counts == max(max(counts)))) - delta;
bw = bw1 & bw2;

```

```

% Clean up

```

```

bw(1:20,:) = 0; % bw(:,1) = 0;
bw(end-20:end,:) = 0; % bw(:,end) = 0;
bw = imclearborder(bw, 8);
bw = imdilate(bw, ones(3));
bw = bwmorph(bw, 'bridge', Inf);
bw = bwareaopen(bw, 16000);
bw = imfill(bw, 'holes');
bw = bwmorph(bw, 'clean', Inf);
bw = imopen(bw, strel('disk', 4));
bw = imclose(bw, strel('disk', 500)); % 200
bw = imfill(bw, 'holes');

```

```

while nnz(bw) < 0.5*size(bw,1)*size(bw,2) && delta > 0

```

```

    delta = delta - 0.025;

```

```

    bw1 = img_norm <= x(find(counts == max(max(counts)))) + delta;
    bw2 = img_norm >= x(find(counts == max(max(counts)))) - delta;
    bw = bw1 & bw2;

```

```

% Clean up

```

```

bw(1:20,:) = 0; % bw(:,1) = 0;
bw(end-20:end,:) = 0; % bw(:,end) = 0;
bw = imclearborder(bw, 8);
bw = imdilate(bw, ones(3));
bw = bwmorph(bw, 'bridge', Inf);
bw = bwareaopen(bw, 16000);
bw = imfill(bw, 'holes');

```

```
bw = bwmorph(bw, 'clean', Inf);
bw = imopen(bw, strel('disk', 4));
bw = imclose(bw, strel('disk', 500)); % 200
bw = imfill(bw, 'holes');

end

writematrix(bw, fullfile(destpath, sprintf('%s-Template.csv', filename)))

%% Clear temporary variables
clear img img_norm optimizer metric tform counts x bw1 bw2

end
```

normalizationMinMax

```
function [img_norm, img_norm2] = normalizationMinMax(img)

img_norm2 = (img - min(min(img))) ./ abs(max(max(img)) - min(min(img)));

if mean(mean(img_norm2)) > 0.45
    img = imcomplement(double(img(:,:,1)));
    img_norm = (img - min(min(img))) ./ abs(max(max(img)) - min(min(img)));
    img_norm(img_norm > 0.8) = 0.8;
else
    img_norm = img_norm2;
end

end
```

RunCellSegmentation

```
clc
clear all
close all
delete(findall(0, 'type', 'figure', 'tag', 'TMWWaitbar'))
warning('off','all')

%% Source and destination directory of the CFU images
sourcepath_img_colonies = uigetdir('C:\', 'Choose assay image folder');
if isequal(sourcepath_img_colonies, 0)
    disp('User selected Cancel')
    return;
end

destpath_img_colonies = uigetdir('C:\', 'Choose destination folder');
if isequal(destpath_img_colonies, 0)
    disp('User selected Cancel')
    return;
end

[file, path, indx] = uigetfile('C:\', 'Choose cell container (template) image');
if isequal(file, 0)
    disp('User selected Cancel')
    return;
end

sourcepath_img_cellcontainer = fullfile(path, file);

%% Pixel size
pixelsize = 21.1667 * 10-3; % mm/pixel

%% Loop through source directories
folderList = getAllFolders(sourcepath_img_colonies);
```

```

for i = 1:length(folderList)

    [path1, foldername1, ~] = fileparts(folderList{i});
    [path2, foldername2, ~] = fileparts(path1);
    destpath = fullfile(destpath_img_colonies, foldername2);
    if ~exist(destpath, 'dir')
        mkdir(destpath)
    end
    destpath = fullfile(destpath, foldername1);
    if ~exist(destpath, 'dir')
        mkdir(destpath)
    end

    fileList = getAllFiles(folderList{i});
    % fileList_CFUcounts = getAllFiles(sourcepath_txt_CFUcounts{i});

    %n_colonies_exp = [];
    filename_vec = {};
    n_colony_seg = [];
    mean_area = []; SD_area = [];
    mean_red = []; SD_red = [];
    mean_green = []; SD_green = [];
    mean_blue = []; SD_blue = [];
    mean_gray = []; SD_gray = [];
    mean_pca1 = []; SD_pca1 = [];
    mean_pca2 = []; SD_pca2 = [];
    confluency_vec = [];
    pixelindexlist = {};

    for j = 1:length(fileList) % 2

        [~, filename, ~] = fileparts(fileList{j});

```

filename

tic

```
[areastats_colony, redstats_colony, greenstats_colony, ...  
    bluestats_colony, graystats_colony, pca1stats_colony, ...  
    pca2stats_colony, confluency] = CellColonySegmentation( ...  
    fileList{j}, destpath, sourcepath_img_cellcontainer, pixelsize);
```

toc

```
colony_xyCentroids = vertcat(areastats_colony.Centroid);  
colony_xCentroids = colony_xyCentroids(:,1);  
colony_yCentroids = colony_xyCentroids(:,2);  
colony_areas = [areastats_colony.Area] .* pixelsize^2;  
colony_MedianRed = [redstats_colony.MedianIntensity];  
colony_MedianGreen = [greenstats_colony.MedianIntensity];  
colony_MedianBlue = [bluestats_colony.MedianIntensity];  
colony_MedianGray = [graystats_colony.MedianIntensity];  
colony_MedianPCA1 = [pca1stats_colony.MedianIntensity];  
colony_MedianPCA2 = [pca2stats_colony.MedianIntensity];  
colony_SDRed = [redstats_colony.SD];  
colony_SDGreen = [greenstats_colony.SD];  
colony_SDBlue = [bluestats_colony.SD];  
colony_SDGray = [graystats_colony.SD];  
colony_SDPKA1 = [pca1stats_colony.SD];  
colony_SDPKA2 = [pca2stats_colony.SD];
```

```
filename_vec{j} = filename;  
n_colony_seg(j) = numel(areastats_colony);  
mean_area(j) = mean(colony_areas);  
SD_area(j) = std(colony_areas);  
mean_red(j) = mean(colony_MedianRed);  
SD_red(j) = std(colony_MedianRed);  
mean_green(j) = mean(colony_MedianGreen);
```

```

SD_green(j) = std(colony_MedianGreen);
mean_blue(j) = mean(colony_MedianBlue);
SD_blue(j) = std(colony_MedianBlue);
mean_gray(j) = mean(colony_MedianGray);
SD_gray(j) = std(colony_MedianGray);
mean_pca1(j) = mean(colony_MedianPCA1);
SD_pca1(j) = std(colony_MedianPCA1);
mean_pca2(j) = mean(colony_MedianPCA2);
SD_pca2(j) = std(colony_MedianPCA2);
confluency_vec(j) = confluency * 100;

% Write to Excel file
header = {'Colony Area (mm2)', 'Centroid x-Coordinate (px)', ...
'Centroid y-Coordinate (px)', 'Median Colony Red (a.u)', ...
'Median Colony Green (a.u)', 'Median Colony Blue (a.u)', ...
'Median Colony Gray (a.u)', 'Median Colony PCA1 (a.u)', ...
'Median Colony PCA2 (a.u)', 'SD Colony Red (a.u)', ...
'SD Colony Green (a.u)', 'SD Colony Blue (a.u)', ...
'SD Colony Gray (a.u)', 'SD Colony PCA1 (a.u)', ...
'SD Colony PCA2 (a.u)'};
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), header, 1, 'A1')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ....
filename)), colony_areas.', 1, 'A2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_xCentroids, 1, 'B2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_yCentroids, 1, 'C2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_MedianRed.', 1, 'D2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_MedianGreen.', 1, 'E2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...

```

```

filename)), colony_MedianBlue.', 1, 'F2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_MedianGray.', 1, 'G2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_MedianPCA1.', 1, 'H2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_MedianPCA2.', 1, 'I2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_SDRed.', 1, 'J2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_SDGreen.', 1, 'K2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_SDBlue.', 1, 'L2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_SDGray.', 1, 'M2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_SDPCA1.', 1, 'N2')
xlswrite(fullfile(destpath, sprintf('%s-ColonyData.xlsx', ...
filename)), colony_SDPCA2.', 1, 'O2')

end

header = {'Filename', 'CFU count', 'Mean size (mm2)', ...
'SD size (mm2)', 'Confluency (%)', 'Mean Red (a.u)', ...
'SD Red (a.u)', 'Mean Green (a.u)', 'SD Green (a.u)', ...
'Mean Blue (a.u)', 'SD Blue (a.u)', 'Mean Gray (a.u)', ...
'SD Gray (a.u)', 'Mean PCA1 (a.u)', 'SD PCA1 (a.u)', ...
'Mean PCA2 (a.u)', 'SD PCA2 (a.u)'};
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), header, 1, 'A1')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), filename_vec.', 1, 'A2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), n_colony_seg.', 1, 'B2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_area.', 1, 'C2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_area.', 1, 'D2')

```



```
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), confluency_vec.', 1, 'E2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_red.', 1, 'F2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_red.', 1, 'G2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_green.', 1, 'H2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_green.', 1, 'I2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_blue.', 1, 'J2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_blue.', 1, 'K2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_gray.', 1, 'L2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_gray.', 1, 'M2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_pca1.', 1, 'N2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_pca1.', 1, 'O2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), mean_pca2.', 1, 'P2')
xlswrite(fullfile(destpath, [foldername2 '.xlsx']), SD_pca2.', 1, 'Q2')

close all

end
```

runKmeans

```
function [bw] = runKmeans(img)

%% Design the feature matrix Z
new_img = zeros(size(img)+2);
new_img(2:end-1,2:end-1) = double(img);

% Image boundary coordinates without first/last row/column
inner_coord = [2 2; size(new_img,1)-1 size(new_img,2)-1];

% 9x2 matrix with 1 row for the relative shift to reach neighbors
[d2, d1] = meshgrid(-1:1, -1:1);
d = [d1(:) d2(:)];

% Cell array to store all 9 shifted images
temp = {};

for i = 1:size(d,1)
    % x-indices of the submatrix when shifted by d(i,1)
    coord_x = (inner_coord(1,1):inner_coord(2,1)) + d(i,1);
    % y-indices of the submatrix when shifted by d(i,2)
    coord_y = (inner_coord(1,2):inner_coord(2,2)) + d(i,2);
    % image matrix resulting from shift by d(i,)
    temp{i} = reshape(new_img(coord_x, coord_y), 1, []);
end

% Column-wise bind all 9 shifted images (as vectors) from the array
Z_feat = vertcat(temp{:}).';

%% K-means

% Number of clusters that will definitely be present
```

```

k = 2; % background/foreground
% Define background-/foreground centroid-cluster vector
start_mat = [repelem(0.2,9); repelem(1,9)];
% The feature matrix M is now size(img,1)*size(img,2)-by-9
[clusterIndexes, ~] = kmeans(Z_feat, k, 'Distance', 'sqeuclidean', ...
    'MaxIter', 1000, 'Start', start_mat); % 'Replicates', 5
img_Kmean = reshape(clusterIndexes, size(img,1), size(img,2));

% Get pixels that are labeled as colonies
[maxValue, indexOfMaxValue] = max(img_Kmean);

if nnz(img_Kmean == min(img_Kmean(:))) > nnz(img_Kmean == max(img_Kmean(:)))
    bw = (img_Kmean == max(img_Kmean(:)));
else
    bw = (img_Kmean == min(img_Kmean(:)));
end

%% Clear temporary variables
clear new_img inner_coord d d1 d2 temp M_feat coord_x coord_y k ...
    start_mat clusterIndexes clusterCenters img_Kmean maxValue indexOfMaxValue

end

```

runPCA

```
function [coeff, score, latent, img_pca1, img_pca2, img_pca3] = runPCA(img)

%% Size and number of color channels
[rows, columns, numberofcolorbands] = size(img);

%% Get an N by 3 array of all the RGB values. Each pixel is one row (RGB)
% Column 1, 2 and three are red, green and blue values, respectively
listOfRGBValues = double(reshape(img, rows * columns, 3));
% listOfRGBValues_ref = double(reshape(img_ref, rows * columns, 3));
X = listOfRGBValues.>';
N = size(listOfRGBValues, 1); % number of pixels

% Get the principal components (eigenvectors of the covariance matrix S
% of the mean-deviation form of the 2D RGB matrix data)
% score is P .* B.' representing the transform into PCA space
% latent are eigenvalues of S
[coeff, score, latent] = pca(listOfRGBValues);
% Take the components and transform the RGB list into a PCA list
transformedImagePixelList = listOfRGBValues * coeff;

% transformedImagePixelList is also an N by 3 matrix of values
% Column 1, 2 and 3 are the values of 1st, 2nd and 3rd principal component
% Extract each column and reshape back into a rectangular image the same size as the original image
img_pca1 = double(reshape(score(:,1), rows, columns));
img_pca2 = double(reshape(score(:,2), rows, columns));
img_pca3 = double(reshape(score(:,3), rows, columns));

% Compute the percentages of the total variance explained (displayed)
% by the principal components. That is, percentage of information contained
% in each principal component image
prct_pca1 = (latent(1)/sum(latent)) * 100;
```

```
prct_pca2 = (latent(2)/sum(latent)) * 100;
prct_pca3 = (latent(3)/sum(latent)) * 100;

%% Clear temporary variables
clear row columns numberofcolorbands listOfRGBValues X N ...
    transformedImagePixelList prct_pca1 prct_pca2 prct_pca3

end
```

B.2 Programme for determining Moran's I and z-value

moran3D_new3.R

```

1 library(abind)
2 library(tensor)
3
4 moran3D_final <- function(img, w, r = 1)
5 {
6   # img ... array with arbitrary number of dimensions, but >= 4 rows per dimension
7   # w ... either a vector with weights for all neighbors, or a scalar (if all neighbors are weighted equally)
8   # r ... size of the neighborhood (must comply with w and dim(img), s.t. length(w) = prod(dim(img) - 2*r) - 1)
9
10  # special case, if weight vector is of length 1 (this means that all neighbors should be weighted equally)
11  if(length(w) == 1){
12    w <- rep(w, (2 * r + 1) ^ length( dim(img) ) - 1) # build vector by repeating the value of w
13  }
14
15
16
17  nd <- length(dim(img)) # number of dimensions
18
19  shift.vec <- expand.grid( # shift vectors to determine neighbors
20    rep(
21      list(c(0, -1:-r, 1:r)),
22      nd
23    )
24  )
25
26  img <- img - mean(img, na.rm = TRUE) # normalize image values
27
28  img.list <- list() # a list of images, corresponding to the shifts in each direction
29
30  for(i in 1:nrow(shift.vec)){
31
32    inds <- rep( # help variable to access the dimensions of an array
33      alist(,)[1],
34      nd
35    )
36
37    for(j in 1:nd){
38
39      inds[j] <- (r + 1 + shift.vec[i,j]) : (dim(img)[j] - r + shift.vec[i,j]) # add current shift to the jth dimension
40    }
41
42    img.list[[i]] <- do.call("c", c(list(img), inds)) # access image values at the position inds
43  }
44
45
46
47  arr <- abind(img.list[-1], along = (nd + 1)) # turn the image list into a (nd+1)-dimensional array
48
49  # calculate the weighted sum of the nominator via a (vector*tensor)-product over the neighborhood array (performance)
50  nom <- tensor(w, arr, alongA = 1, alongB = (nd + 1))
51
52  nom <- nom * img.list[[1]] # elementwise multiplication of the weighted neighbor and the original image pixel
53
54  # number of pixels taken into account for Moran's I (= total number of pixels excl. border region) ***** modified *****
55  n <- sum(!is.na(nom))
56
57  nom <- sum(nom, na.rm = TRUE) # sum over all pixels to calculate nominator
58  denom <- sum(img.list[[1]]^2, na.rm = TRUE) # calculate denominator
59
60  fac <- 1 / sum(w) # calculate normalization factor (n cancels out)
61
62  I <- fac * nom / denom # Moran's I
63
64  EI <- (-1) / (n - 1) # expected value
65
66  S0 <- n * sum(w) # sum of weights ***** n times! *****
67
68  S1 <- 2 * n * sum(w^2, na.rm = TRUE) # outer sum ... n times, inner sum ... 2*w^2 (constant for all i) ***** n times! *****
69
70  S2 <- 4 * n * sum(w, na.rm = TRUE)^2 # outer sum ... n times, inner sum ... 2*sum(w)^2 (constant for all i)
71
72  A <- n * ((n^2 - 3*n + 3)*S1 - n*S2 + 3*S0^2)
73  D <- (sum(img.list[[1]]^4, na.rm = TRUE)) / ((sum(img.list[[1]]^2, na.rm = TRUE))^2)
74  B <- D * ((n^2 - n)*S1 - 2*n*S2 + 6*S0^2)
75  C <- (n - 1) * (n - 2) * (n - 3) * S0^2
76
77  EI2 <- (A - n*B) / C # ***** error in the online formula! B has to be multiplied by n *****
78
79  VI <- EI2 - EI^2
80
81  Zi <- (I - EI) / sqrt(VI)
82
83
84  output_list <- list("moransi" = I, "zscore" = Zi, "expectedI" = EI, "vi" = VI)
85
86  return(output_list)
87 }

```

call_moran_threshold_multipleimages_new2.R

```
1  #install.packages("spdep")
2
3  # load EBImage package
4  library(EBImage)
5  library(stringr)
6  library(spdep)
7
8  # Import all sample image-stacks in greyscale
9  chosen <- tcltk::tk_choose.files()
10
11 threshold <- 255 #threshold applied in other evaluation
12 thr <- (1/255)*threshold
13
14 img <- list() #List of images
15 MoransI <- list() #List of Moran's I results
16 Zi <- list() #List of Z-values I results
17 Ei <- list() # list of expected morans value
18 Vi <- list()
19 SDi <- list()
20 p <- list()
21 cells <- list() #List of cell names corresponding to Moran result
22
23 for(i in 1:length(chosen)){
24   img[[i]] <- readImage(chosen[i])
25   img[[i]][img[[i]]<thr]=0
26
27   output_list <- moran3D_final(img[[i]], w = 1)
28   MoransI[[i]] <- output_list$moransi
29   Zi[[i]] <- output_list$zscore
30   Ei[[i]] <- output_list$expectedi
31   Vi[[i]] <- output_list$vi
32
33   z_i <- (-1)*abs(output_list$zscore) # lower Zi
34   expected_i <- output_list$expectedi
35   sd_i <- sqrt(output_list$vi)
36   SDi[[i]] <- sd_i
37
38   pvalue <- 2*pnorm(q=z_i, mean = 0, sd = 1, lower.tail = TRUE)
39
40   p[[i]] <- pvalue
41
42   cells[i] = str_sub(chosen[i],start=-30,end=-5)
43 }
44
45 output = cbind(cells, MoransI, Zi, Ei, Vi, SDi, p)
46 write.csv2(output, file = as.character(tcltk::tkgetSaveFile()))
```

B.3 Programme for determining clusters

B.3.1 Programme for determining clusters by OPTICS

optics_version.py

```
1  # -*- coding: utf-8 -*-
2  """
3  Created on Wed Mar 18 15:10:04 2020
4
5  @author: tinas
6  """
7  import pandas as pd
8  import numpy as np
9
10 from sklearn.cluster import OPTICS
11
12 import matplotlib.pyplot as plt
13 from mpl_toolkits.mplot3d import Axes3D
14
15
16 def nr_of_clusters(excel_doc, subdir):
17     cluster_dict = {}
18     label_names = list(excel_doc['Source.Name'].unique())
19     rows, columns = excel_doc.shape
20
21     for i, element in enumerate(label_names):
22         celle = element
23         column_names = list(excel_doc.columns)
24         dataset = pd.DataFrame(columns = column_names)
25
26         for index in range(rows):
27             if excel_doc.iat[index, 0] == celle:
28                 position_row = excel_doc.iloc[index]
29                 dataset = dataset.append(position_row, ignore_index = True)
30             else:
31                 continue
32
33         X = dataset.drop(columns=['Source.Name'])
34         antall_markers, columns = dataset.shape
35
36         if antall_markers == 1:
37             cluster_dict[celle] = '1' #only one marks
38
39         elif antall_markers == 2:
40             cluster_dict[celle] = '1' #only two marks
41
42         elif antall_markers >= 3:
43
44             if antall_markers == 3:
```



```

46     m_s = 3
47     elif antall_markers == 4:
48         m_s = 4
49     elif antall_markers == 5:
50         m_s = 5
51     elif antall_markers == 6:
52         m_s = 6
53     elif antall_markers == 7:
54         m_s = 7
55     else:
56         m_s = 8
57
58     op = OPTICS(min_samples=m_s)
59     op_preds = op.fit_predict(X)
60     op_n_clusters_used = len(np.unique(op_preds))
61
62     cluster_dict[celle] = "{0}".format(op_n_clusters_used)
63
64     threed = plt.figure()
65     threed = threed.gca(projection='3d')
66     threed.scatter(X['X'], X['Y'], X['Z'], c=op_preds)
67     threed.set_xlabel('X')
68     threed.set_ylabel('Y')
69     threed.set_zlabel('Z')
70     threed.set_title("OPTICS clustering on cell {0}".format(celle.split('for ')[1][:4]))
71     plt.savefig(subdir+"\{0}_clusters.png".format(celle.split('for ')[1][:4]))
72
73     return(cluster_dict)
74
75
76 import os
77 rootdir = 'C:\\Users\\tinan\\Documents\\Suspension_data\\Uke11'
78
79 for subdir, dirs, files in os.walk(rootdir):
80     for file in files:
81         if file == 'coordinates3.xlsx':
82             df = pd.read_excel(os.path.join(subdir, file))
83             resultat = nr_of_clusters(df, subdir)
84             resultater = pd.DataFrame.from_dict(resultat, orient='index')
85             resultater = resultater.reset_index()
86             resultater.columns = ['Source.Name', 'Result']
87             resultater.to_excel(subdir+"\OPTICS3TEST.xlsx", index=False)
88

```

B.3.2 Programme for determining cluster by k-means

KMEAN_gap.py

```
1  # -*- coding: utf-8 -*-
2  """
3  Created on Thu May 28 10:52:37 2020
4
5  @author: tinas
6  """
7  # -*- coding: utf-8 -*-
8  """
9  Created on Wed Mar 18 15:10:04 2020
10 """
11
12 ### Importing packages ###
13
14 import pandas as pd
15 from sklearn.cluster import KMeans
16 from sklearn.metrics import silhouette_score
17 import numpy as np
18 import math
19 import copy
20
21
22 ### import functions ###
23 from gap import gap
24
25
26 ### Function to find distance ###
27 def distance(x1, y1, z1, x2, y2, z2):
28
29     d = math.sqrt(math.pow(x2 - x1, 2) +
30                   math.pow(y2 - y1, 2) +
31                   math.pow(z2 - z1, 2) * 1.0)
32     return(d)
33
34
35 ### Create clustering function ###
36
37 def nr_of_clusters(excel_doc):
38     cluster_dict = {}
39     label_names = list(excel_doc['Source.Name'].unique())
40     rows, columns = excel_doc.shape
41
42     for i, element in enumerate(label_names):
43         celle = element
44         column_names = list(excel_doc.columns)
```

```

45 dataset = pd.DataFrame(columns = column_names)
46
47 for index in range(rows):
48     if excel_doc.iat[index, 0] == celle:
49         position_row = excel_doc.iloc[index]
50         dataset = dataset.append(position_row, ignore_index = True)
51     else:
52         continue
53
54 X = dataset.drop(columns=['Source.Name'])
55 antall_markers, columns = dataset.shape
56 if antall_markers == 0:
57     cluster_dict[celle] = '0.0'
58 elif antall_markers == 1:
59     cluster_dict[celle] = '0.1'
60 elif antall_markers == 2:
61     cluster_dict[celle] = '0.2'
62 elif antall_markers >= 3:
63     X2 = copy.copy(X)
64     X2 = X2.to_numpy(dtype='float')
65     rader, kolonner = X2.shape
66
67     if rader == kolonner:
68         """
69         Checking if dataset is unclustered based on a simple threshold
70         by distance between points
71         """
72         x_1, y_1, z_1 = X2[0,0], X2[0,1], X2[0,2]
73         x_2, y_2, z_2 = X2[1,0], X2[1,1], X2[1,2]
74         x_3, y_3, z_3 = X2[2,0], X2[2,1], X2[2,2]
75         d12 = distance(x1=x_1, y1=y_1, z1=z_1, x2=x_2, y2=y_2, z2=z_2)
76         d13 = distance(x1=x_1, y1=y_1, z1=z_1, x2=x_3, y2=y_3, z2=z_3)
77         if 1/3 <= d12/d13 <= 3:
78             cluster_dict[celle] = "1"
79         else:
80             cluster_dict[celle] = "2"
81     else:
82         liste=[]
83         for i in range(100):
84             resultat = gap(X2, refs=None, nrefs=20, ks=range(1,3))
85             liste.append(resultat.argmax())
86         n_unclustered = liste.count(0)
87         n_clustered = liste.count(1)
88
89         if n_unclustered >= n_clustered:
90             cluster_dict[celle] = "1".format(resultat[0])
91

```

```

92         else:
93             highest_score = 0
94             best_cluster = 0
95             for n_clusters in range(antall_markers-2):
96                 clusterer = KMeans(n_clusters=n_clusters+2)
97                 preds = clusterer.fit_predict(X)
98                 score = silhouette_score(X, preds)
99                 if score > highest_score:
100                     highest_score = score
101                     best_cluster = n_clusters+2
102                 cluster_dict[celle] = "{}".format(best_cluster)
103
104     return(cluster_dict)
105
106
107 import os
108 rootdir = 'C:\\Users\\tinast\\Documents\\ferdigbehandlet-masterdata'
109 for subdir, dirs, files in os.walk(rootdir):
110     for file in files:
111         if file == 'coordinates.xlsx':
112             df = pd.read_excel(os.path.join(subdir, file))
113             resultat = nr_of_clusters(df)
114             resultater = pd.DataFrame.from_dict(resultat, orient='index')
115             resultater = resultater.reset_index()
116             resultater.columns = ['Source.Name', 'Result']
117             resultater.to_excel(subdir+"\\KMEANS_gap.xlsx", index=False)
118
119

```

B.3.2.1 Gap Statistics function

```
# gap.py
# (c) 2013 Mikael Vejdemo-Johansson
# BSD License
#
# SciPy function to compute the gap statistic for evaluating k-means clustering.
# Gap statistic defined in
# Tibshirani, Walther, Hastie:
# Estimating the number of clusters in a data set via the gap statistic
# J. R. Statist. Soc. B (2001) 63, Part 2, pp 411-423

import scipy
import scipy.cluster.vq
import scipy.spatial.distance
dst = scipy.spatial.distance.euclidean

def gap(data, refs=None, nrefs=20, ks=range(1,11)):
    """
    Compute the Gap statistic for an n*m dataset in data.

    Either give a precomputed set of reference distributions in refs as an (n,m,k) scipy array,
    or state the number k of reference distributions in nrefs for automatic generation with a
    uniformed distribution within the bounding box of data.

    Give the list of k-values for which you want to compute the statistic in ks.
    """
    shape = data.shape
    if refs==None:
        tops = data.max(axis=0)
        bots = data.min(axis=0)
        dists = scipy.matrix(scipy.diag(tops-bots))

        rands = scipy.random.random_sample(size=(shape[0],shape[1],nrefs))
        for i in range(nrefs):
            rands[:,i] = rands[:,i]*dists+bots
    else:
        rands = refs

    gaps = scipy.zeros((len(ks),))
    for (i,k) in enumerate(ks):
        (kmc,kml) = scipy.cluster.vq.kmeans2(data, k)
        disp = sum([dst(data[m,:],kmc[kml[m],:]) for m in range(shape[0])])

        refdisps = scipy.zeros((rands.shape[2],))
        for j in range(rands.shape[2]):
            (kmc,kml) = scipy.cluster.vq.kmeans2(rands[:,j], k)
            refdisps[j] = sum([dst(rands[m,:j],kmc[kml[m],:]) for m in range(shape[0])])
        gaps[i] = scipy.mean(scipy.log(refdisps))-scipy.log(disp)
    return(gaps)
```

C Additional results

C.1 PI incubation test

Images of samples irradiated with 2 Gy and incubated for 48 hours, but incubated with PI for 30 min and 70 min to check effect of oxygen exposure on the PI intensity. The PI signal intensity is consistently low.

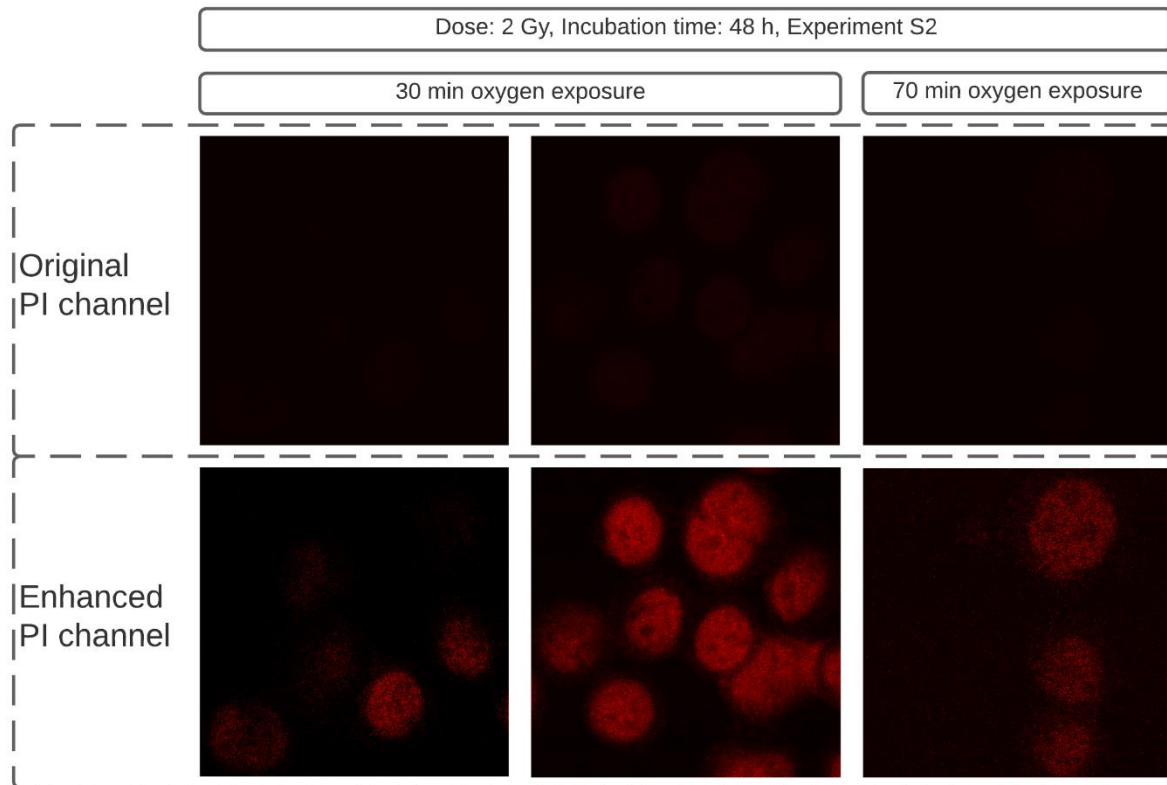


Figure 40 Images of PI channel for sample irradiated with 2 Gy and incubated for 48 hours.

C.2 Antibody tests

The conjugated antibody did not produce considerable γ -H2AX foci. Figure 40 below, shows samples irradiated with 2 Gy and 10 Gy and incubated for 0.5 hours. The experiment used the highest antibody concentration recommended, 4 $\mu\text{g/ml}$. The conjugated foci does not produce DSB foci sufficiently.

Conjugated antibody: Anti-phospho-Histone H2A.X (Ser139) Antibody, clone JBW301, FITC

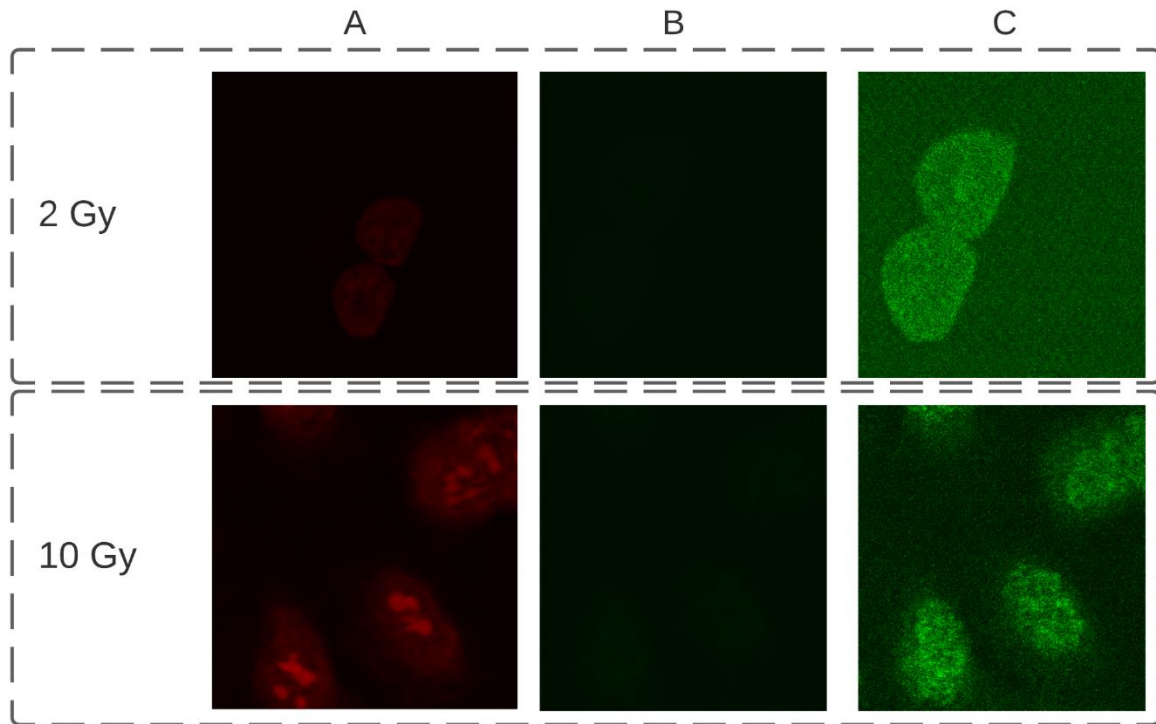


Figure 42 Results of conjugated antibody test. The cells was irradiated with 2 Gy and 10 Gy

Figure 41 shows comparison of the two secondary antibodies tested (table 6) on cells with same treatment. Polyclonal Rabbit Anti-Mouse Ig/FITC performed best in terms of affinity and specificity.

Primary antibody: Anti-phospho-Histone H2A.X (ser139) clone JBW301

Ab6785 Goat Anti-Mouse IgG
(FITC)

Polyclonal Rabbit Anti-Mouse
Immunoglobulins/FITC

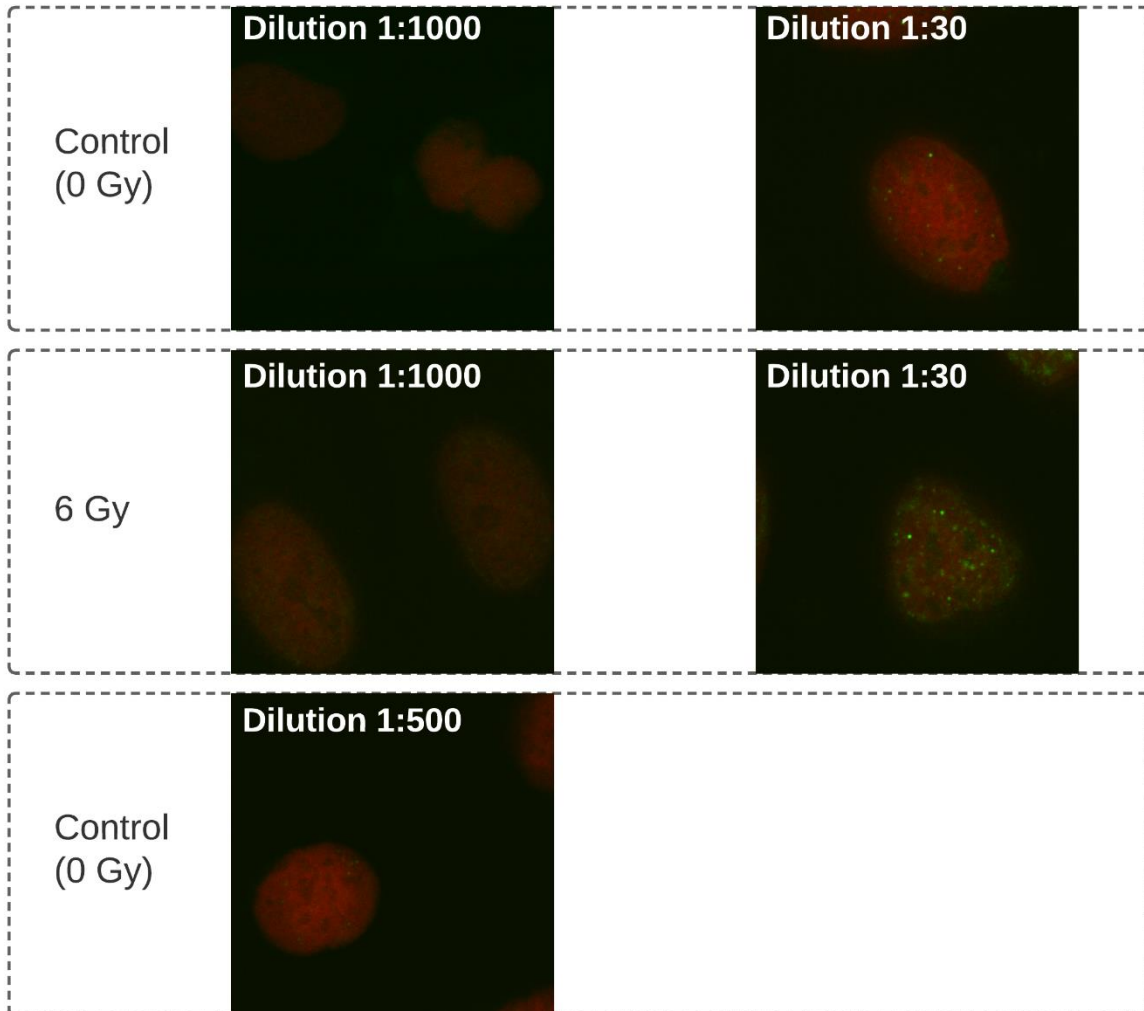


Figure 41 Test conducted with same primary antibody and two different secondary antibodies. Polyclonal Rabbit Anti-Mouse Ig/FITC performed best in terms of affinity and specificity.

C.3 Fixation and permeabilization test

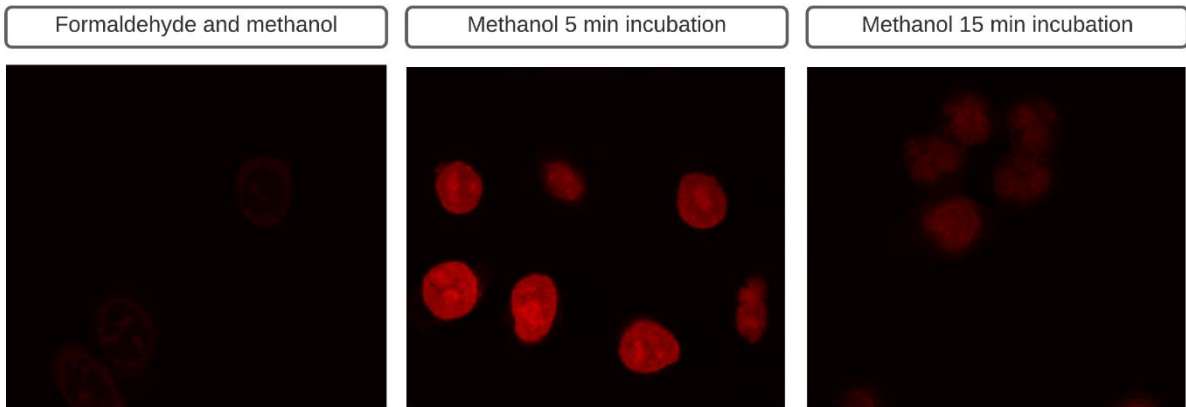


Figure 42

C.3 Cell seeding confluence

Table 7 displays confluence level for the number of cells seeded on slide-flask according to irradiation dose and incubation time.

Tabell 7 Confluence level specified for the number of cells seeded on slide-flasks according to different doses given and incubation time.

Dose [Gy]	Incubation time [hours]	Number of cells seeded	Confluency [%]
0	72	150 000	32.8
0	72	175 000	35.6
0	72	200 000	53.1
2	72	150 000	32.9
2	72	175 000	38.0
2	72	200 000	40.7
5	72	150 000	28.9
5	72	175 000	26.3
5	72	200 000	38.3
8	72	150 000	28.8
8	72	175 000	32.4
8	72	200 000	35.8
10	72	150 000	30.4
10	72	175 000	35.4
10	72	200 000	37.2

C.4 X-ray dose rate at SSD 40

Dose rate was estimated to be 1.21 ± 0.04 Gy/ min.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitenskapelige universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway