HELLENIC REPUBLIC
National and Kapodistrian
University of Athens

INTER-INSTITUTIONAL, INTER-DEPARTMENTAL POSTGRADUATE STUDIES PROGRAM "MEDICAL PHYSICS- RADIOPHYSICS"
MEDICAL SCHOOL, NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

Characterization of kinases involved in DNA damage response checkpoints after low and high doses of low-LET irradiation and detection of structural DNA variations after induction of complex DNA damage using Nanopore sequencing.

MASTER THESIS

of

CHRISTINA VASILEIOU

Supervisor: Anastasios Siountas.
Associate Professor, School of Medicine, Aristotle University of Thessaloniki

In collaboration with: Alexandros Georgakilas, Associate Professor, Physics Department, School of Applied Physical Sciences, National Technical University of Athens.

and

Professor George Iliakis, Institute of Medical Radiation Biology, University of Duisburg-Essen, Medical School

Athens, July 2019

HELLENIC REPUBLIC
National and Kapodistrian University of Athens

INTER-INSTITUTIONAL, INTER-DEPARTMENTAL POSTGRADUATE STUDIES PROGRAM "MEDICAL PHYSICS- RADIOPHYSICS"

MEDICAL SCHOOL, NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

Characterization of kinases involved in DNA damage response checkpoints after low and high doses of low-LET irradiation and detection of structural DNA variations after induction of complex DNA damage using Nanopore sequencing.

MASTER THESIS

OF

CHRISTINA VASILEIOU

Supervisor: Anastasios Siountas
Associate Professor, School of Medicine, Aristotle University of Thessaloniki

Approved by the three-member examination committee:

Siountas Anastasios	**Georgakilas Alexandros**	**Terzoudi Georgia**
Associate Professor	Associate Professor	Research Scientist A
AUTH	NTUA	NCSR
(Signature)	*(Signature)*	*(Signature)*
.....	.....	.....

Athens, July 2019

Acknowledgements

First of all, I would like to thank all those who contributed to the successful completion of this master thesis. In particular, I would like to thank my Professor, Dr. Anastasios Siountas, Associate Professor in AUTH for the assignment and supervision of this work, for his valuable help and for the trust that he showed me during his elaboration. I would like to thank, my Professor Dr Alexandros Georgakilas, Associate Professor in NTUA for his constant guidance, support and assistance throughout this entire project. I would like to acknowledge Professor George Iliakis for the unique opportunity that he provided me to work in his laboratory in Essen, for the cooperation we had and the help he offered me, providing me with the necessary cognitive tools and the right guidance. I would also like to thank all the members of Prof. Iliakis' research team at Medical radiation Biology department of Uniklinikum Essen and in particular Dr Simon Magin and Dr Emil Mladenov for their contribution, guidance and patience during our experiments. Thanks again to Associate Professor. Dr. A.Georgakilas and Research Scientist Dr. G.Terzoudi who accepted to participate in my three-member examination committee.

Abstract

Cancer treatments such as radiotherapy and chemotherapy mainly act by inducing severe (complex) damage on DNA of cancer cells. In presence of DNA damage, cells implement their multifaceted DNA Damage Response (DDR) system to process the damage and repair it. Checkpoint control plays a major role during this response, therefore targeted inhibition of the aforesaid can lead to increased cell lethality and potentially to more efficient treatments (Bartek and Lukas, 2001; Iliakis et al., 2003; Jackson and Zhou, 2004).

Here, using data generated by cells exposed to low and high doses of low LET ionizing radiation that introduce a low number of double strand breaks (DSBs) in the genome, we cross validate the regulation of the G2 checkpoint mainly by ATR and ATM (results first published by (Mladenov et al., 2019)). We also provide evidence that these kinases interphase with cell cycle machinery mainly through Chk1 kinase. At high doses, we present preliminary results of the dominant role of Chk1 in checkpoint control.

High number of DSBs is known to cause even more severe cell damage (Khanna and Jackson, 2001; van Gent et al., 2001). One of the exogenous agents inducing multiple DSBs is the high Linear Energy Transfer (LET) ionizing radiation. One of the characteristics of high LET irradiation is the introduction of highly complex lesions. This complexity is traditionally thought to be the combination of lower level lesions like one stranded lesions, base changes or losses, SSBs, DSBs etc. or a combination of the previous mentioned types of damage. The combination of low level lesions with DSBs poses major threats to genome integrity according to literature (Campa et al., 2005; Gupta et al., 2009; Seluanov et al., 2010). Here we have used a model based on the hypothesis that increased complexity of DNA damage is manifested through the colocalization or combination of an increased number of DSBs in a relative small stretch of sequence (Iliakis et al., 2018; Mladenova et al., 2016).

The cellular clones used for this part of the analysis are CHO and human clones that were created with the aim to model DSB-clusters. The I-*Sce*I recognition sequence was inserted into cell lines using molecular biology techniques (Jasin, 1996) and clones as needed were selected and characterized. I-*Sce*I recognition sequences introduced in a genome can then be cut, to generate from simple DSBs up to more complex DSB clusters. Constructs carrying different combinations of I-SceI sites were engineered at specific distances and orientations, subsequent transient expression of *I-Sce*I induces the relevant number of DSBs in predetermined loci. Therefore these clones provide the opportunity to measure the effect of several DSBs within a relatively small part of a sequence, excluding postulations about other variables and types of DNA damage that arise when high LET IR is used. A characterization of these clones was performed using long-read sequencing with MinION platform by ONT. Localization and identification of the I-*Sce*I sites is presented on both human and CHO clones and a novel model is proposed to study induced structural variations. Results reveal significantly more integrations

sites in the clones tested in comparison to other techniques. In addition, compromised DSB repair as well as increased translocation formation were evident.

Keys-words: ionizing radiation, complex DNA damage, DNA damage response, low doses, Nanopore sequencing, MinION, ONT, bioinformatics, high LET, low LET, ATM, ATR, Chk1, Chk2, G2 checkpoint, cell cycle**.**

Περίληψη

Αντικαρκινικές θεραπείες, όπως η ακτινοθεραπεία και η χημειοθεραπεία, δρουν κυρίως προκαλώντας βλάβες στο DNA των καρκινικών κυττάρων. Σε περίπτωση βλάβης του DNA, τα κύτταρα κινητοποιούν ένα πολύπλοκο σύστημα απόκρισης στις βλαβες DNA (DDR) για να αξιολογήσουν το είδος και το επίπεδο των βλαβών και να τις επιδιορθώσουν. Τα διάφορα σημεία ελέγχου στο κυτταρικό κύκλο διαδραματίζουν σημαντικό ρόλο κατά τη διάρκεια αυτής της απόκρισης, επομένως η στοχευμένη αναστολή των προαναφερθέντων μπορεί να οδηγήσει σε αυξημένη θνησιμότητα των καρκινικών κυττάρων και ενδεχομένως σε πιο αποτελεσματικές θεραπείες (Bartek and Lukas, 2001; Iliakis et al., 2003; Jackson and Zhou, 2004).

Στη παρούσα εργασία, χρησιμοποιώντας δεδόμενα που παράγονται από κύτταρα που εκτίθενται σε χαμηλές και υψηλές δόσεις ιονίζουσας ακτινοβολίας χαμηλού LET, που εισάγουν χαμηλό αριθμό δίκλωνων θραύσεων στο γονιδίωμα, επιβεβαιώνουμε τη ρύθμιση του σημείου ελέγχου G2/M κυρίως μέσω των κινασών ATR και ATM (αποτελέσματα που δημοσιεύθηκαν αρχικά από τους (Mladenov et al., 2019)). Παρέχουμε επίσης αποδείξεις ότι αυτές οι κινάσες αλληλεπιδρούν με το κυτταρικό κύκλο κυρίως μέσω της κινάσης Chk1. Στις υψηλές δόσεις παρουσιάζουμε προκαταρκτικά αποτελέσματα του κυρίαρχου ρόλου του Chk1 στην εποπτεία των σημείων ελέγχου.

Μεγαλύτερος αριθμός δίκλωνων θραύσεων είναι γνωστό ότι προκαλεί ακόμα πιο σοβαρή κυτταρική βλάβη (Khanna and Jackson, 2001; van Gent et al., 2001). Ένας από τους παράγοντες επαγωγής πολλαπλών δίκλωνων θραύσεων (DSBs) θεωρείται ότι είναι ιονίζουσα ακτινοβολία υψηλού LET. Ένα από τα χαρακτηριστικά της ακτινοβολίας υψηλού LET είναι η εισαγωγή στο γονιδίωμα σύνθετων αλλοιώσεων. Αυτή η πολυπλοκότητα παραδοσιακά θεωρείται ότι είναι ο συνδυασμός αλλοιώσεων χαμηλότερου επιπέδου όπως βλάβες της μίας αλυσίδας, μεταβολές ή απώλειες βάσης, θραύσεις μονής έλικας, δίκλωνες θραύσεις κλπ ή συνδυασμός των προαναφερθέντων τύπων βλαβών. Ο συνδυασμός των αλλοιώσεων χαμηλού επιπέδου με δίκλωνες θραύσεις αποτελεί τη κύρια απειλή για την ακεραιότητα του γονιδιώματος σύμφωνα με τη βιβλιογραφία (Campa et al., 2005, Gupta et al., 2009, Seluanov et al., 2010). Εδώ χρησιμοποιούμε ένα μοντέλο που βασίζεται στην υπόθεση ότι η αυξανόμενη πολυπλοκότητα της βλάβης του DNA προκύπτει απο τη συνεγκατάσταση ή το συνδυασμό αυξημένου αριθμού DSBs σε ένα σχετικά μικρό εύρος αλληλουχίας (Iliakis et al., 2018; Mladenova et al., 2016).

Οι κυτταρικοί κλώνοι που χρησιμοποιήθηκαν για αυτό το τμήμα της ανάλυσης προέρχονται από άνθρωπο και χάμστερ και δημιουργήθηκαν με σκοπό να μοντελοποιήσουν τις σύνθετες βλάβες DNA επαγόμενες μονο απο τις δίκλωνες θραύσεις DNA. Περιέχουν πολλαπλές γενετικά τροποποιημένες αλληλουχίες που δημιουργήθηκαν με τη βοήθεια πλασμιδιακού DNA κατασκευασμένες μέσω κατάλληλων θέσεων αναγνώρισης της μεγανουκλεάσης I-*Sce*I , επακόλουθη έκφραση του ενζύμου προκαλεί διαφορετικο αριθμό DSBs σε προκαθορισμένες γεννετικές θέσεις. Έτσι, παρέχεται η δυνατότητα ποσοτικοποίησης της επίδρασης πολλαπλών DSBs σε ένα σχετικά μικρό τμήμα μιας αλληλουχίας και ταυτόχρονα εξαιρούνται άλλες

μεταβλητές και άλλου τύπου βλάβες που εισέρχονται όταν χρησιμοποιείται ιοντίζουσα ακτινοβολία υψηλού LET. Ένας χαρακτηρισμός αυτών των κλώνων πραγματοποιήθηκε υλοποιώντας αλληλούχιση του DNA στην πλατφόρμα MinION από την ONT. Ο εντοπισμός και η ταυτοποίηση των θέσεων της I-*Sce*I παρουσιάζεται ταυτόχρονα και σε κλώνους χάμστερ και σε ανθρώπινους. Επιπροσθέτως προτείνουμε ένα καινοτόμο μοντέλο για τη μελέτη επαγόμενων δομικών αλλοιώσεων. Τα αποτελέσματα αποκαλύπτουν περισσότερες τοποθεσίες ενσωμάτωσης της γενετικα τροποποιημένης αλληλουχίας σε σύγκριση με άλλες τεχνικές. Επιπλέον, στα αποτελέσματα της ανάλυσής μας, ήταν εμφανής η υποβάθμιση της λειτουργίας επιδιόρθωσης και αποκατάστασης των δίκλωνων θραύσεων καθώς και ο αυξημένες μετατοπίσεις σε τμήματα του γονιδιώματος.

Λέξεις-κλειδιά: ιοντίζουσα ακτινοβολία, σύνθετες βλάβες DNA , σύστημα απόκρισης στις βλάβες DNA , χαμηλές δόσεις, υψηλές δόσεις, αλληλούχιση, Nanopore, MinION, ONT, βιοπληροφορική, υψηλό LET, χαμηλό LET, ATM, ATR, Chk1, Chk2, G2 σημείο ελέγχου κυτταρικός κύκλος.

Play is the highest form of research.

(Albert Einstein 1879-1955)

To my parents and my dog Perla.

Contents

Introduction

Role of LET in Radiotherapy

Linear Energy Transfer or LET (KeV/μm) is a term used in medical physics in order to quantify the average amount of energy that is deposited from an ionizing particle to a material, when it passes through the latter, per unit distance. In other words LET can be described as the density of the ionization events that occur per unit of distance when an ionizing particle is travelling through matter. High density ionization events are considered to be high-LET radiation (e.g. α-particles, high energy protons or neutrons) while more sparse ionization events are considered to be low-LET radiation (e.g. X-Rays).

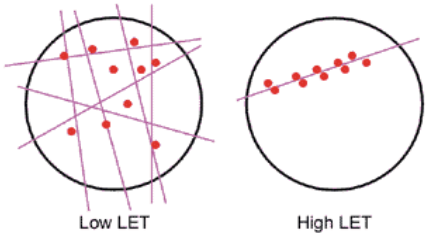


Figure 1: Distribution of ionizing events in high and low LET radiation. ource:https://www.rerf.or.jp/en/programs/roadmap_e/health_effects-en/basickno-en/radcell-en/

High LET radiation is considered to be more dangerous for the cells because it results in localized ionizations in a smaller volume thus making it more difficult for the cell to repair it. Consequently it is possible to use irradiation with relevant LET in radiotherapy aiming at more efficient cell killing while minimizing the dose to the surrounding healthy tissue (Niemantsverdriet et al., 2012). It has been shown that the physical differences between high LET radiation and low LET radiation have a distinct impact on the induction of DNA damage, causing unambiguous differences in the properties of the damaged DNA (Goodhead, 2006; Jakob et al., 2009).

Biological effects of ionizing radiation, DSBs

Radiation-induced ionizations may act directly on the cellular component molecules or indirectly on water molecules, causing water-derived free radicals. Radicals have the ability to react with nearby molecules, resulting in their oxidation. The major effect in cells is DNA breaks. Breaks of either a single strand or both strands can occur depending on the type of radiation. However, DNA double strand breaks (DSBs) are considered the most cytotoxic type of DNA lesion. Accumulated DSBs will be referred as clustered DSBs. Clustering of DNA lesions highly depends on LET. Most single-strand breaks can be repaired normally thanks to the double-stranded nature of the DNA molecule (the two strands complement each other, so that an intact strand can serve as a template for repair of its damaged, opposite strand). In the case of double-strand breaks, however, repair is more difficult and inaccurate rejoining of broken ends may occur. Unrepaired or misrepaired DSBs or DSB clusters may lead to the loss of genomic stability by triggering the

formation of translocations, deletions, mutations and by destabilizing chromatin. These repair failures can constitute key driving forces for the development of various human syndromes, neurological disorders, immunodeficiency and cancer (Bohgaki et al., 2010; Jackson and Bartek, 2009).

DNA Damage Response (DDR) system to DSBs low LET

Repair pathways

DDR system's pathway choice in repair in higher eukaryotes is highly dependent on which phase of the cell cycle the cell is undergoing the damage (Saha et al., 2017). There are four repair pathways available during DSB processing that are divided into two main categories. On the one hand we have complete restoration of the DNA sequence in an error-free way that is implemented by Homologous Recombination. Alternatively, cells deploy end-joining repair pathways. These pathways include classical non-homologous end joining (c-NHEJ), alternative end joining (alt-EJ) and Single Strand Annealing (SSA) (Mladenov et al., 2016). C-NHEJ is a homology independent pathway which is found to be active in all phases of the cell cycle and it is the main pathway in response to DSBs while HRR –homology dependent pathway- directs repair in late S and G2 phases when a DNA template via a sister chromatid is available (Figure 2).

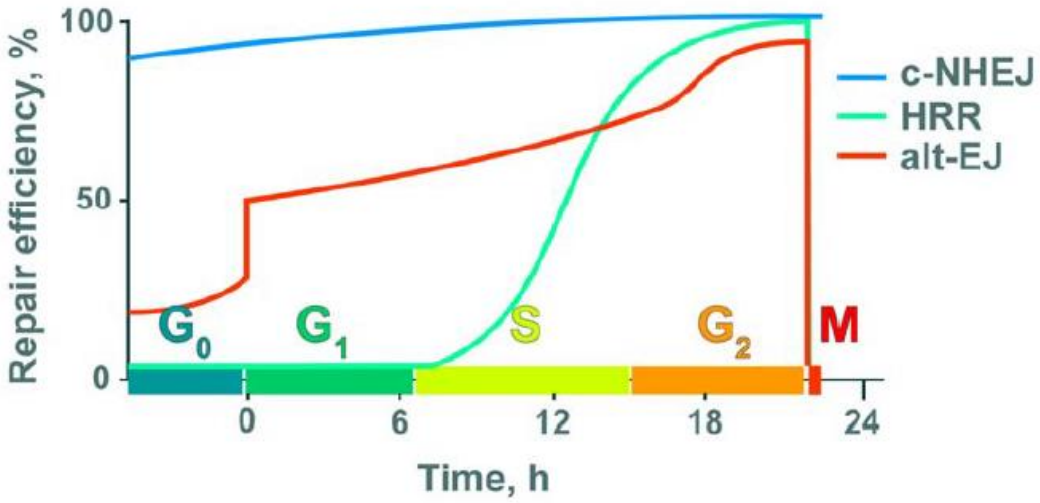


Figure 2: Activity of HR, c-NHEJ and alt-EJ in the different phases of cell cycle Source: (Mladenov et al., 2016)

Main role after activation of DDR is the regulation, among other processes of cell cycle progression. Cell cycle control operates like a timer that triggers the events of the cell cycle in

time sequence through cell cycle checkpoints. Cell cycle checkpoints are based on a series of biochemical switches which initiate specific events. Cell cycle checkpoints are highly conserved and constitute critical mechanisms that cells utilize to ensure that the progression from one phase of the cell cycle to the other is performed properly otherwise the cycle can be halted until conditions are favorable (Stark and Taylor, 2004).

Thus to ensure that healthy cells are produced after each round of cell division, the cell cycle consists of three checkpoints with distinct functions: G1/S, G2/M, and Spindle (M) checkpoints (Figure 3). The first checkpoint is in late G1 where the cell commits to cell cycle entry and chromosome duplication and among other processes the integrity of the DNA is assessed at the so called G1/S checkpoint (Bartek and Lukas, 2001). The second one is the G2/M checkpoint where proper chromosome duplication is assessed and early mitotic events start to occur. Finally the last checkpoint takes place during metaphase to anaphase transition in M phase and in this checkpoint the control system starts stimulating sister-chromatid separation to the end of mitosis and cytokinesis. (Barnum and O'Connell, 2014).

It is known that exposure of cells to ionizing radiation (IR) and other external agents delays the normal progression through the cell cycle (Iliakis, 1997). These delays take place mainly in G2 phase but also in in G1 and S in a lesser extent. DSBs in particular activate DDR signaling, which results in the activation of cell cycle checkpoint the main role of which is to allow and promote DNA-repair mechanisms in the different phases of the cell cycle. The transitions from one phase of the cell cycle to the next one are controlled mainly by the activation of cyclin-dependent kinases (CDKs). Other regulatory mechanisms such as alterations in chromatin structure are also present during these transitions and all in combination form a complex, multiparametric control system available in cell's palette for managing DNA damage (Duursma and Cimprich, 2010; Solomon and Kaldis, 1998). Cells arrested in a checkpoint resume cell-cycle progression once damage has been repaired, whereas cells with un-repairable DNA lesions undergo checkpoint adaptation, permanent cell-cycle arrest, or apoptosis. In the present thesis we study the effects of low LET radiation on G2 checkpoint.

G2 checkpoint

The activation of the G2 checkpoint delays entry into the proliferative Mitotic (M) phase when DNA damage is present. The G2 checkpoint is the last resort for cells to evaluate their genomic integrity and to address and correct any potential genomic damage prior to cell division thus making the importance of this checkpoint obvious (Figure 3).

Central components of cell cycle control are members of a family of protein kinases known as cyclin-dependent kinases (Cdks). During the entrance of the cell in mitosis M-cyclins activate Cdks that stimulate this transition. The G2-checkpoint is implemented by suppressing Cdc25C that normally functions to activate Cdk1 at the end of the G2 phase to enable progression of cells into mitosis (Nilsson and Hoffmann, 2000). Cdk1 is associated with cycling B1 and forms the so called

mitosis promotion factor (MPF). M-Cdk activation begins with accumulation of M-cyclin. The increase of M-cyclin leads to accumulation of M-Cdk, the complex of Cdk1 and M-cyclin. Those accumulated complexes at the end of G2 are kept inactive by phosphates. This complex gets activated during mitosis by the phosphatase cdc25 which removes the inhibitory phosphates from the M-Cdks.

Normally cells during transition from G2 to M phase are in a state of extensive chromatin condensation which is accompanied by the phosphorylation of Histone H3ps10 which can then be used as a marker in our case by targeting it with a specific antibody.

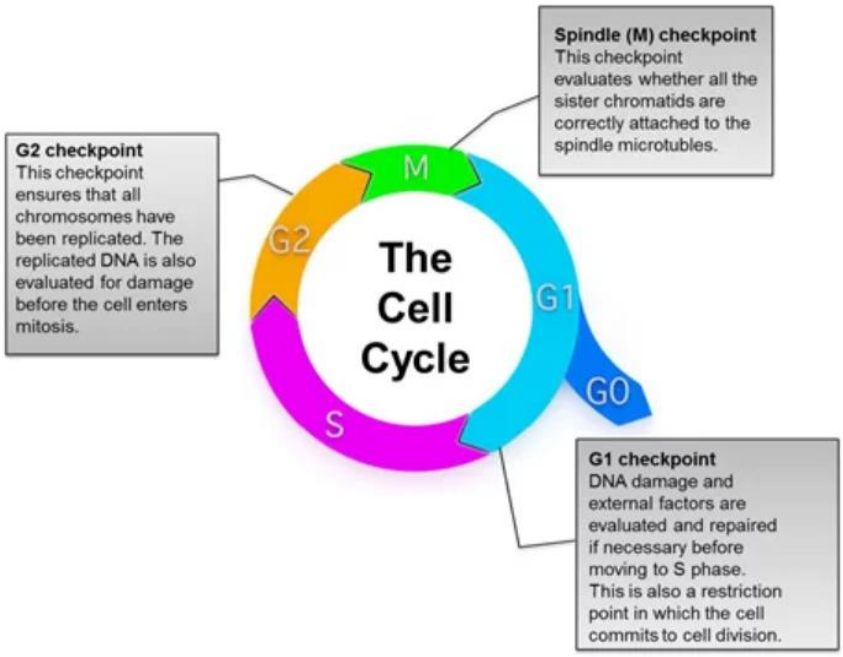


Figure 3: Overview of the eukaryotic cell cycle checkpoints. Source: https://www.bio-rad-antibodies.com/

The proteins participating in the activation of the G2 checkpoint are highly conserved in eukaryotes. DSBs are very effective to activate checkpoints. The DNA double-strand breaks (DSBs) are the principle cytotoxic lesions induced by ionizing radiation (IR). They DDR signaling collectively involves sensing of DNA damage, signal its presence and mediate the cell cycle responses.

DNA damage sensing, the PI3k family of kinases.

Molecular components involved in these mechanisms vary from one pathway to the other but looking in a more general view, in response to DSBs mainly three members of the phosphoinositide-3 kinase (PI3K)-related family of protein kinases have been shown to be activated, ATM, ATR and DNA-PKcs (Marechal and Zou, 2013; Nilsson and Hoffmann, 2000; Yang et al., 2003) with the latter being more involved in DNA repair(Figure 4).

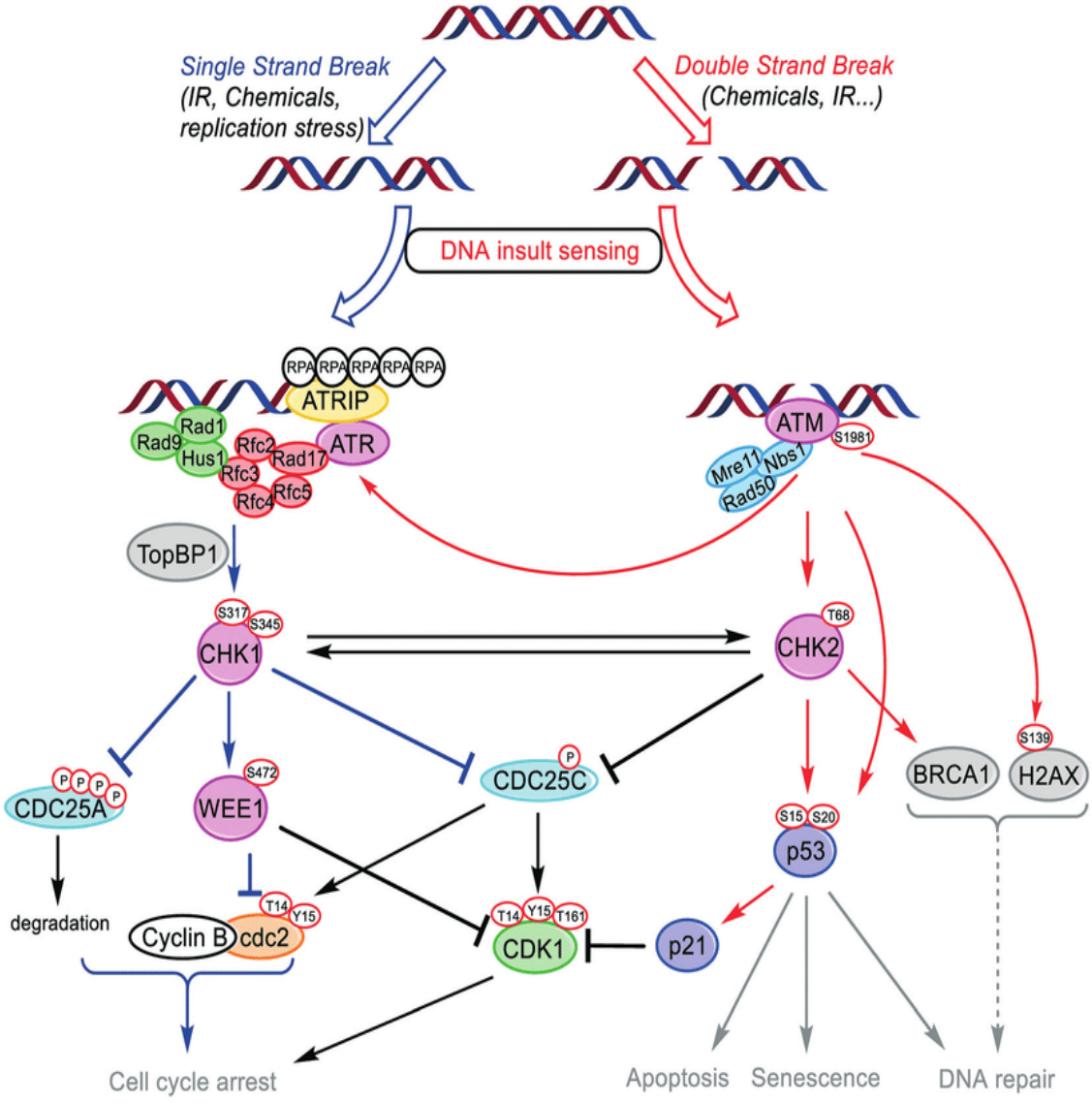


Figure 4: DNA damage sensing by the PI3k family of kinases.

As it is shown in Figure 4 in the presence of DNA damage the main sensors are the: phosphoinositide 3-kinase like kinases (PI3KKs) a ATM and ATR. ATM on the one hand phosphorylates various substrates including Chk2, p53, BRCA1 and itself whereas ATR

phosphorylates and activates CHk1 which in G2 inactivates Cdc25c phosphatase which removes inhibitory mechanisms of Cdks who initiate cell cycle transitions (Bartek et al., 2001; Lukas et al., 2001).

Recent results have shown that at low IR doses, ATM and ATR regulate the G2-checkpoint as an integrated module in which the two kinases are functionally paired in a fully equivalent and epistatic manner with ATR located downstream of ATM (Mladenov et al., 2019). Under physiological conditions the G2-checkpoint is implemented by suppressing Cdc25C that normally functions to activate Cdk1 at the end of the G2 phase to enable progression of cells into mitosis (Nilsson and Hoffmann, 2000).

When DNA damage is present at this phase of the cell cycle a signaling pathway is initiated mainly through the activation of the ATM and ATR kinases. Their key role is to associate with the site of the damage and phosphorylate various target proteins including the kinases Chk1 and Chk2. Together all these kinases phosphorylate other proteins that lead to cell cycle arrest. A major target protein is the regulatory protein *p53.*

As previously mentioned, ATM and ATR can convey inhibitory signals to Cdc25C through the checkpoint kinases Chk2 and Chk1, respectively. Therefore here we implemented a series of experiments in different cell lines treated with inhibitors of kinases involved in DNA damage response and repair of DSBs. The experiments were based on implementation of bivariate Flow Cytometry of DNA content and mitotic index to detect the G2 checkpoint activation and recovery after exposure to IR.

The hypothesis that the G2–checkpoint is implemented by exclusively connecting ATR to the cell cycle engine using Chk1 and other as of yet unidentified M-kinases and similar experiments were repeated as in (Mladenov et al., 2019) and cross-validated. The project investigated therefore functional interactions between Chk1 and Chk2 and searched for additional kinases implementing the G2 checkpoint. We focused on low doses of IR. The experiments studied carefully both the inception, as well as the recovery from the checkpoint.

DNA Damage Response system to DSBs high LET

Source of DSBs, the I-*Sce*I endonuclease

High LET IR-induced DSBs are of critical importance but their random distribution in the genome as well as their generation at different genomic sites in each individual cell, obscure precise characterization and analysis. Targeted genome modification can be used to mimic the effects of high LET IR. This effect can be achieved using specific endonucleases. Homing endonucleases (HEs) are practically enzymes that are mainly encoded in introns or in inteins (Belfort and Roberts,

1997) Homing endonucleases like I-*Sce*I have a specific role under physiological conditions to create a site specific double strand break by recognizing 12-40bp DNA sequences. The I-*Sce*I HE endonuclease is originally found in yeast and specifically in the mitochondria of Saccharomyces cerevisiae with an 18 bp non-palindromic recognition sequence and leaves a 4 base pair 3' hydroxyl overhang (Figure 5)(Iliakis et al., 2018; Mladenova et al., 2016).

Figure 5: 18 bp cognition sequence of I-SceI endonuclease.

I-*Sce*I model, DSB clusters

The sequencing data used for this purpose were either produced directly from sequencing the DNA or after transfecting the corresponding cell lines with vectors expressing the restriction enzyme I-*Sce*I, based on a model that mimics the effects of radiation and specifically that generates double strand breaks (DSBs) clusters (Iliakis et al., 2018; Mladenova et al., 2016), one of the presumed most common effects of high LET radiation (Figure 6). The I-*Sce*I-Mediated Double-Strand Break model implements an approach to induce DSBs using nucleases with the expression of I-*Sce*I homing endonuclease, whose 18-bp long recognition sequence is not present in the mammalian genome; transient expression of the enzyme was used during the construction of those cell clones(Mladenova et al., 2016). The I-SceI recognition sequence was inserted into cell lines using molecular biology techniques (Jasin, 1996) and clones as needed were selected and characterized. I-*Sce*I recognition sequences introduced in a genome can then be cut, to generate from simple DSBs up to more complex DSB clusters (Figure 6.). Constructs carrying different combinations of I-SceI sites were engineered at specific distances and orientations. Figure 6 illustrates constructs that allow the generation of single-DSBs, of DSB-pairs at different distances (100 and 200 bp), as well as of DSB quadruplets. In constructs harboring two of more I-SceI sites, the orientation between the sites (Direct: D or Reversed: R) is also indicated, as it results in the generation of compatible or incompatible apical DNA ends after loss of the intervening sequence. The grey arrows represent the locations of forward and reverse primers utilized to amplify the corresponding DNA segment for junction analysis by sequencing (Iliakis et al., 2018; Mladenova et al., 2016). Further developments of the assay will utilize inducible forms of the endonuclease to avoid the limitations of transfection (Bindra et al., 2013). The biological consequences of these DSBs can then be analyzed using molecular biology approaches.

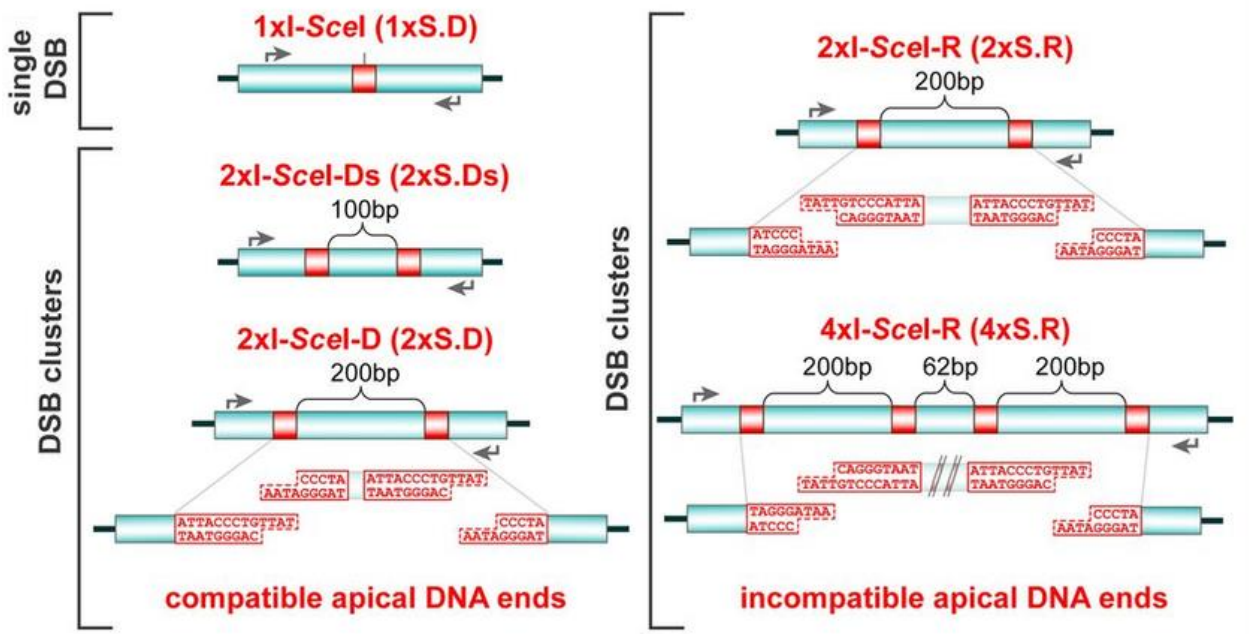


Figure 6: Genomic integrated constructs containing various I-SceI recognition sites result in different forms of DSB clusters.

The main difference in this model in comparison with previous similar methods is the fact that with this approach we can introduce as many I-*Sce*I sites as possible into the mammalian cell genomein contrast with the simple transfection methods that yield only one or two integrations (Iliakis et al., 2018; Mladenova et al., 2016). Thus a Transposon-based system approach is applied which allows the integration of the sequence of interest at various locations in the genome randomly.

Transposable elements (TEs) are DNA sequences that can change their position within a genome, by moving, replicating, creating or reversing mutations and therefore result in alteration the cell's genetic identity (Aronovich et al., 2011; Dupuy et al., 2006).

TEs are classified into two main categories Class I TEs or retrotransposons and Class II TEs or DNA transposons. In the model investigated in this project we will deal with data created with class II TEs and specifically the Sleeping Beauty transposon system which is classified as Tc1/mariner-type system (Hermanson et al., 2004; Ivics et al., 1997). The Sleeping Beauty transposon system is composed of a Sleeping Beauty (SB) transposase and a transposon that was designed to insert specific sequences of DNA into genomes (Figure 8). DNA transposons translocate from one DNA site to another in a simple, 'cut-and-paste' manner. The mechanism of the transposition is illustrated in Figure 7 (Ivics et al., 1997).

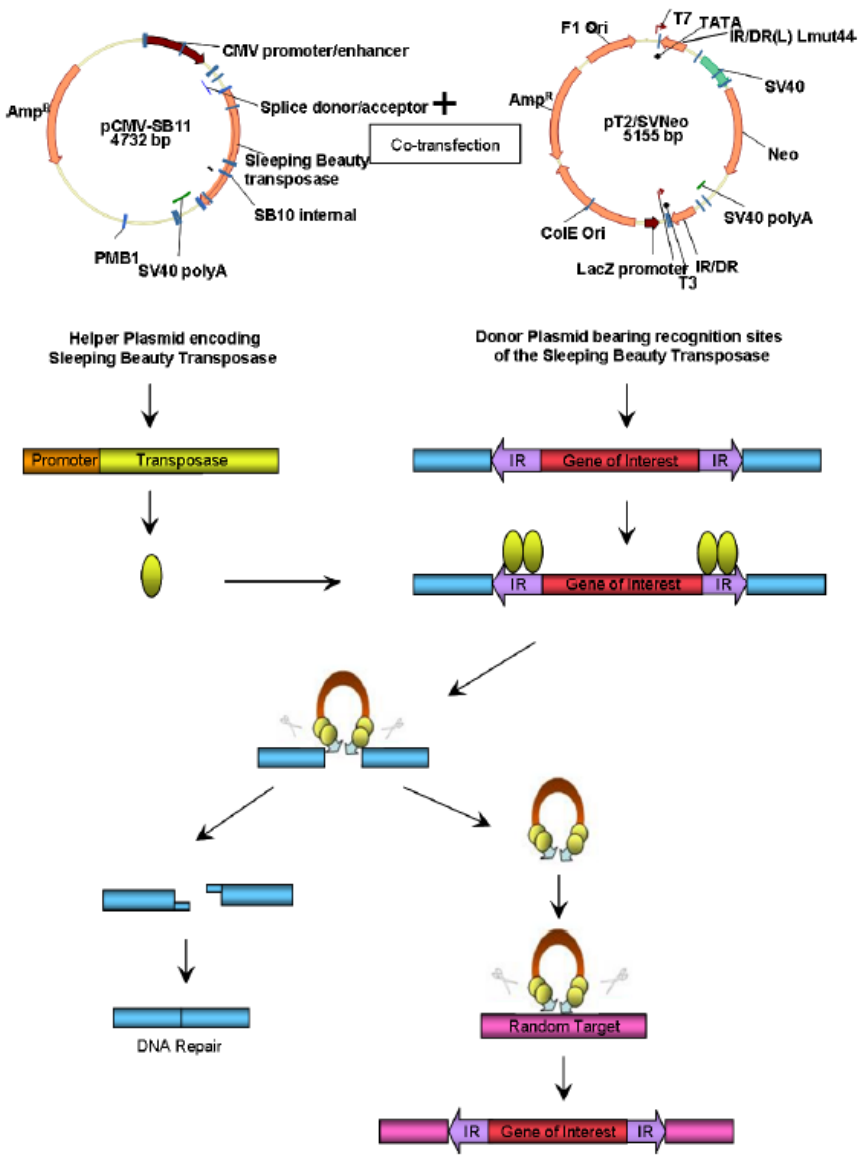


Figure 7: Sleeping Beauty Transposition mechanism.

Common consequences of these DSBs include structural variations (SVs). SVs, such as large deletions, duplications, inversions and most importantly translocations are abundant in large genomes and get enhanced on a repair pathway choice dependent manner (Iliakis et al., 2015). Their detection requires long reads for precise characterization in order to confidently detect alterations that can span kb of DNA (Norris et al., 2016).

The majority of SVs are poorly assayed using currently dominant short read sequencing technologies, but can be detected using long read sequencing technologies from Pacific Biosciences and Oxford Nanopore Technologies (ONT) ((De Coster et al., 2019; Gong et al., 2018; Norris et al., 2016)). Long read sequencing technologies have a lower single nucleotide accuracy

(~90%) but have the advantage of a better mapping in repetitive and larger in terms of kbases regions, further extending the part of the genome in which variation can be called reliably(De Coster et al., 2019).

In this direction we created two models to first detect the I-*Sce*I integration sites in the genome and analyze translocation formation from sequencing data produced on the Nanopore platform (MinION), which allows to generate lengths in our reads of the appropriate size for the ongoing analysis. Our work resulted in different handling of the CHO and human clones mainly due to differences in the accuracy of the respective reference genomes. The first method developed was designed for the human clones and is implemented through identification and localization of integration sites using a bioinformatics analysis method using topological screening of existing translocations. The second method developed emerged by the inevitably different handling of CHO clones. It is based on an initial topological approach, as with human clones, but this time in combination with an innovative experimental investigation based on product-tailored PCR reactions.

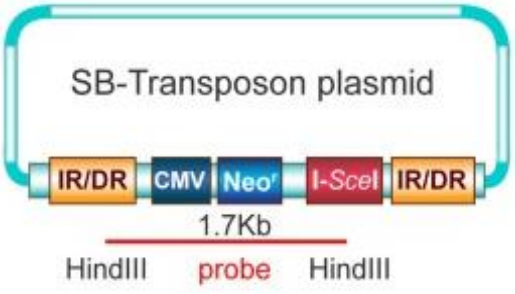


Figure 8: Map of the SB-transposon plasmid carrying the I-SceI construct. The transposase binding sites comprising the IR/DR regions are shown. CMV – cytomegalovir, Neor – neomycin resistance. (Iliakis et al., 2018; Mladenova et al., 2016)

Nanopore sequencing

MinION is a third-generation sequencing (TGS) technology and it is a portable real-time device for DNA and RNA sequencing provided by Oxford Nanopore Technologies (ONT). Each flow cell of minion can generate up to 30 Gb of sequencing data in long reads (up to 2Mb/read depending on user's fragment length choice). Its portable size, very low cost and very fast and high throughput are some of the long list of advantages that this technology has to offer. The long reads generated by this technology can help to overcome limitations of short read technologies, although they present higher error rates and they pose other challenges that require tailored algorithms and specific workflows. Thus, with the appropriate bioinformatics handling they allow for the first time to perform applications such as *de novo* assembly of larger genomes or structural variations (SVs) detection studies that were thought to be cumbersome if not impossible using alternative sequencing technologies(Gong et al., 2018; Lu et al., 2016). Furthermore MinION, considering the benefits of real time analysis, can provide rapid insights

into infectious disease, cancer research and many other areas of biomedical research and therefore attracts great interest in clinical applications (Patel et al., 2018).

The physical principle behind Nanopore is electrophoresis. Biological pores are attached in a lipid membrane (Jou and Muthukumar, 2017). A voltage is applied across this electrically resistant polymer membrane which creates an ionic current that guides DNA through the pores. The Nanopore is a protein-complex (CsgG) spanning a membrane in an electrolyte solution. The biological pores consist of proteins and one can tailor their surface charge by inducing mutations in their genes. This way their structure, size and position are essentially engineered at angstrom level (Maglia et al., 2008). One key advantage of biological nanopores is also the high reproducibility from one pore to the other. One disadvantage that should be mentioned is that the biological Nanopore as their biological analogues require a lipid membrane which in some cases is difficult in handling as one of the characteristics of this membrane is that it can easily collapse when voltage is applied on it.

SsDNA is then passed though the only available passage in the membrane -the nanopores- creating a characteristic disruption in current (Figure 9b). Measurements of that 'current disruption' result in the identification of the combination of bases (A, T, G, and C) of the DNA under interrogation.

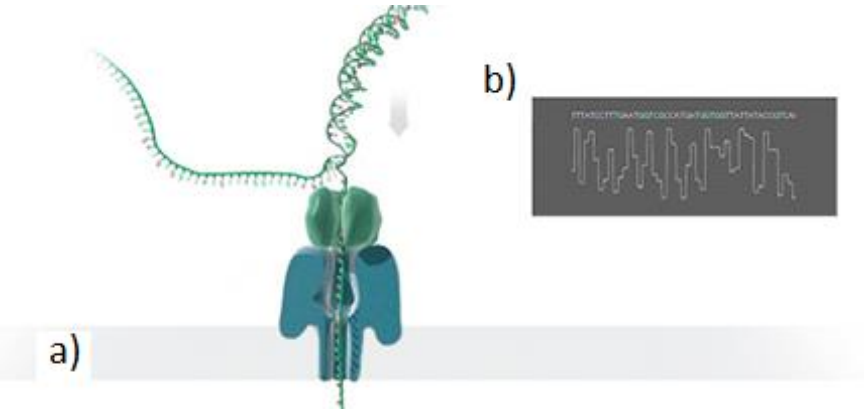


Figure 9: Source: https://nanoporetech.com/ a) SsDNA is passed through the protein pore. b)Disruptions in current from each base are translated in electrical signal.

A parameter that influences our measurements is the so called length of constriction site which describes the number of bases that influence the signal in other words how many bases are read simultaneously. The length of this constriction site is desirable to be kept low to be able to discriminate between the signal of different combination of bases and at the same time it is desirable to be kept high so that overlap between current and later combination of bases is allowed so that each base is read multiple times and we result in increased basecalling accuracy.

Biological nanopores function in very high speeds. ($(2x10^6$-10) bases/sec), these velocities though are hard to be detected by the sensor. For this reason a regulation of speed is done with a motor protein. This motor protein is nothing else but a DNA polymerase or alternatively a helicase. In a

initial stage the double strand of DNA has at the beginning a poly-T tail or a leader adapter sequence and on these places the motor protein gets attached. In MinION the motor protein is modified helicase which regulates the speed of base flow at 450 bases per second.

Each flow cell in MinION is made of 512 channels allowing this way up to 512 DNA molecules to pass simultaneously. Each channel is connected with 4 wells (4 wells constitute 1 MUX), which are all tested at the beginning of each run of MinION and the most active well is used each time to obtain the sequencing data. Higher throughput can be achieved with synchronous sequencing by multiple pores at the same time. A typical image from MinION's default software, MinKNOW during regular checking of the channels is shown in Figure 10. MinKNOW performs at the beginning of each run a first quality control where unusable wells are marked. Unusable well may occur due to false positioning on the pore, defects in the membrane or blockades. Then a second quality control is done where the wells are examined, three out of four wells on MUX are ranked in terms of signal quality and sequencing starts from the best one. MinKNOW among other offers the following: data acquisition, real time analysis, basecalling, data streaming, device control etc. In Figure 10 report is shown about a sequencing run, we can see details concerning the size of throughput till the observation time, the temperature that is developed in the flowcell as well as the voltage applied. In general the voltage applied is -180mV and it only increases buffer changes from one well to the other. In the MinION report it is also shown how many pores are active at the moment many are sequencing and how many are either recovering or completely inactive. These numbers depend strongly on the conditions on which the flow cell is kept. On Figure 11 another part of regular MinION report is shown. There we can see a real-time analysis of the distribution of the length in kilobases of the reads that are generated. This of course is defined in a step of library preparation during DNA fragmentation and can be adjusted in the desired length.

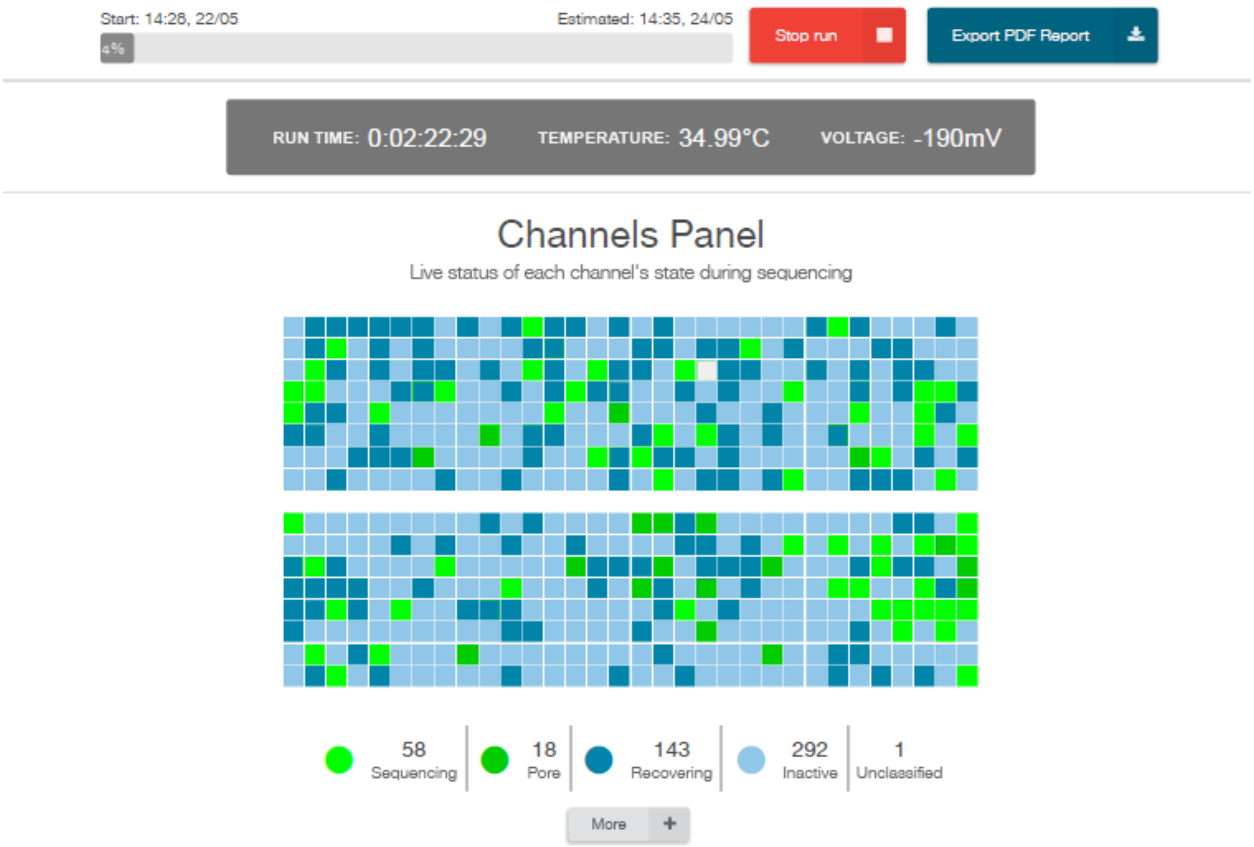


Figure 10: Regular check in channel activity during sequencing on MinION.

The first step while processing this electrical signal is a procedure called *basecalling.* Basecalling is the process of assigning to raw data (values of signal intensity), through M-HAP algorithms, the corresponding base during or after sequencing. M-HAP algorithms is a probabilistic sequence overlapping algorithm, designed to efficiently detect all overlaps between noisy long-read sequence data. It efficiently estimates Jaccard similarity by compressing sequences to their representative fingerprints composed on min-mers (minimum k-mer). The Jaccard similarity coefficient is a statistic used for gauging the similarity and diversity of sample sets. Equation 1

Equation 1: Jaccard similarity coefficient formula

$$\mathrm{J}(A,B) = \frac{|A \cap B|}{|A \cap B|}$$

MinKnow acquisition software can perform in-run basecalling but that of course poses some computational issues when we have high throughput. The data/reads that are received from MinION are in .fastq format and are ready for further bioinformatical analysis In-run analysis also provides various information like for example details concerning read-length distribution (Figure 11).

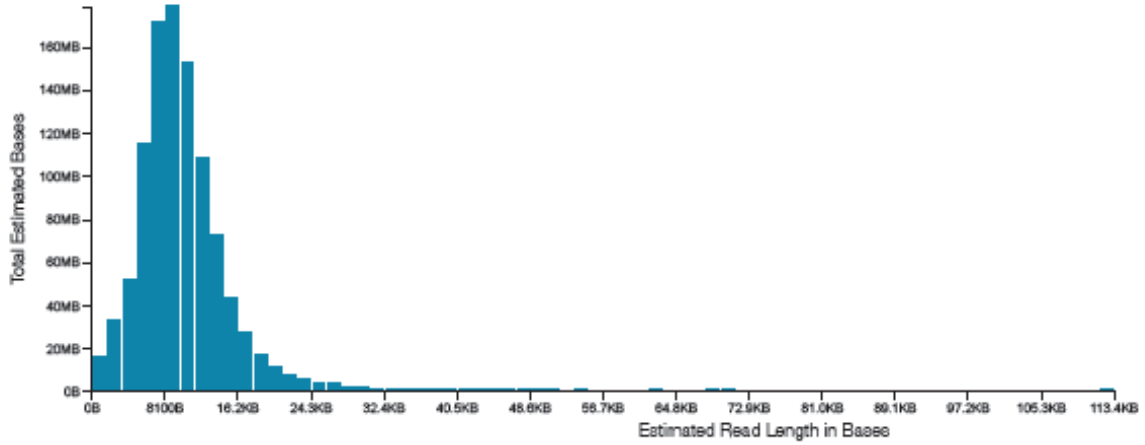


Figure 11: Read length distribution on MinION report.

Materials and methods:

Cell culture

The human cell lines that were used during the experiments were the A549 cell line and the 82.6 cell line. A549 cells are adenocarcinomic human alveolar basal epithelial cells and they express wild type p53 and also grow in a nice plateau phase with the majority of cells accumulated in G1 phase.(Giard et al., 1973) Cells are grown in McCoy's 5A medium (M4892, Sigma-Aldrich) supplemented with 10% Newborn Calf Serum (NCS-15) and 0.5 μg/ml Fe(II)SO_4. We next tested 82-6 fibroblasts that were immortalized with hTERT, the catalytic subunit of human telomerase. Cells were grown in minimum essential medium (MEM) supplemented with 10%fetal bovine serum(FBS 24, (S0615, BIOCHROM)) and 1% non-essential amino acids, Non-Essential Amino Acids are used as a supplement for cell culture medium, to increase cell growth and viability. All cells were mediated at 37°C in a humidified atmosphere of 5% CO2 in air in SANYO MCO-18 O2/CO2 incubators. Cells were grown in 100 ml tissue culture dishes (664160, CELLSTAR) with 15ml growth media. Exponentially growing cells were passaged every 2 days while avoiding confluency levels above 80%. For passage, media were removed and cells were rinsed with cold PBS. Cells were rinsed with 2 ml trypsin-EDTA solution and incubated for 2 min at 37°C. Detached cells were resuspended in 10 ml cold media supplemented with 10% FBS. Single cell suspensions were obtained by passing cells through the Pasteur pipette. Cells were counted and appropriate numbers of cells were further incubated for experiments or for subculture. When frozen cells were taken to subculture, they were passed at least 5 times every 2 days before being used in experiments. Cells were discarded after about 50 passages, since their genomic stability could not be guaranteed.

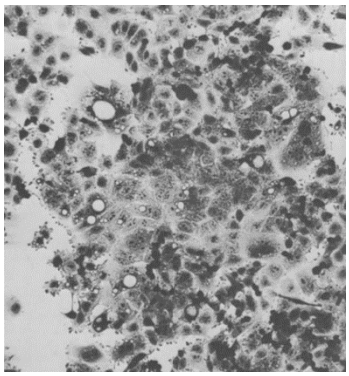


Figure 12: Cells basically appear epithelial-like but many have bizarre shapes and contain numerous vacuoles (Giard et al., 1973)

X-Ray Irradiation

For experiments designed to determine the mitotic index, cells were exposed to IR at 37°C using a 320 kV X-ray machine with a 1.65 Al filter (GEHealthcare). The dose rate at 500 mm distance from the source was 2.7 Gy/min. A uniform radiation distribution was ensured by rotating the radiation table.

Treatment with kinase inhibitors

The Chk1 inhibitor ,7-hydroxystaurosporine (UCN-01, Chk1i, Calbiochem) was dissolved in DMSO at 100 µM and was used at 100 nM final concentration. The **Chk2 Inhibitor** (Chk2 Inhibitor-II/BML-277, Chk2i, Calbiochem) was dissolved in DMSO at 1 mM and used at a final concentration of 400 nM. 8-(4-Dibenzothienyl)-2-(4-morpholinyl)-4H-1-benzopyran-4-one (NU7441, **DNA-PKi**, Tocris Bioscience) was dissolved in DMSO at 10 mM and was used at a 10 µM final concentration. 3-Amino-6-[4-(methylsulfonyl)phenyl]-N-phenyl-2-pyrazinecarboxamide (**ATRi**, VE-821, Haoyuan Chemexpress) was dissolved in DMSO at 10 mM concentration and was used at a 5 µM final concentration. All inhibitors were added to the cells 1 h prior to irradiation and were maintained until collection for analysis.

H3-pS10 staining

Mitotic cells quantification was performed through measurements of the phosphorylated Histone H3 at Serine 10 (H3-pS10).

The following solutions were prepared:

1) Permeabilization buffer

This buffer is used to make the cell membrane penetrable for the antibody. In order to prepare it we collected Triton-X detergent and PBS and created a 0.25% Triton solution.

2) PBG blocking buffer

The blocking buffer is necessary to provide an abundance of proteins (so called ballast proteins like BSA or cold fish gelatin for our experiment) so that in case excess quantity of our antibody is present in our solution these proteins will cover all the 'unspecific sites' of our 'sticky cells' removing unwanted binding places for the antibody. They are in fact blocking the antibody from binding in the wrong places so that way we can avoid any background noise in our Flow Cytometry signal. The preparation method of this buffer consists of 0.2 % Gelatine which we boil at 50°C to dissolve and then we let it cool down, 0,5% BSA fraction V, PBS. All of the above mentioned were diluted in distilled water.
For 500ml I added 1g Gelatin, 2.5g BSA and 50ml 10xPBS and filled up the rest with water. The buffer was maintained for 10 minutes in a warm bath (~40-45°C)
Blocking buffer (PBS +0.05% Tween-20 + 1% BSA (8076.2, ROTH))

3) PI stain

100X Propidium iodine was prepared with final concentration 4 mg/ml
400 mg of PI was added and the volume was adjusted to 100 ml with ddH2O.
PI stain was stored at -20°C in dark.

The following steps were conducted:

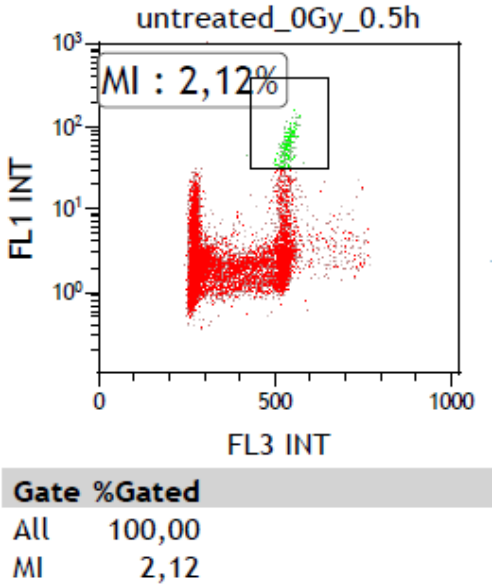
Cells were irradiated two days after incubation in 75cm^2 flasks (~10^6 cells) with a dose of 2 Gy (2.7 Gy/min) for around 40-50 seconds. Cells were washed with cold PBS and collected by trypsinization. Collection timepoints to monitor the MI were set to 30 min, 1 hour, 2 hours, 4 hours, 6 hours and 8 hours post irradiation. Cells were centrifuged at 1500 rpm for 5 min at 4°C and fixed with 70% Et-OH on ice and the pellets were stored at -4°C overnight. Supernatant was removed the next day by centrifugation at 1500 RPM for 5 min. Cells were permeabilized for 15 min in 500µl of ice-cold PBS supplemented with 0.25% Triton X-100. Cells were resuspended to a single cell suspension and then collected by centrifugation at 1500 rpm for 5 min at 4°C.

Blocking of the unspecific antibody binding sites was performed by incubation of the cells in 500 µl of PBG (0.05% Tween-20, 1% BSA in PBS solution) for 1 h at room temperature (RT) with gentle agitation. Supernatant was removed and cell were collected by centrifugation 1500 rpm for 5 min at 4°C.

Primary antibody anti-H3-pS10 (Abcam) was diluted 1:2000 in blocking buffer. 150 µl was used per sample quantity sufficient for up to 3x10^6 cells. The pellets were suspended in 100 µl diluted primary antibody. Cells were extensively resuspended in a single cell suspension and incubated for 2 h on a rotation shaker with simultaneous agitation every 30 minutes. Cells were washed with 500 µl of PBS per sample and centrifuged at 2000 rpm for 5 minutes at 4°C. Supernatant was removed and wash with PBS was repeated with the above settings two more times.

Secondary antibody AlexaFluor 488-conjugated goat-anti rabbit-IgG (Thermo Fisher Scientific) was diluted 1:300 in blocking buffer. The pellets were suspended in 150 µl diluted secondary antibody per sample. Cells were extensively resuspended in a single cell suspension and incubated for 1.5 hours on a rotation shaker with simultaneous agitation every 20 minutes. Cells were washed with 500 µl of PBS per sample and centrifuged at 2000 rpm for 5 minutes at 4°C. Supernatant was removed and wash with PBS was repeated with the above settings one more time.

300-500 µl of PI staining solution was added and cells were incubated for 15 min at 37°C and then we proceeded immediately to Flow Cytometry acquisition. Alternatively, PI stained cells could be left at 4°C overnight and could be measured on FACS next day.



Flow Cytometry analysis of mitotic index

Bivariate flow Cytometry was employed to simultaneously measure DNA content by propidium Iodide (PI) staining and mitotic cells by quantification of the phosphorylated Histone H3 at Serine 10 (H3-pS10).
Analysis was carried out in a flow cytometer (Gallios, Beckman Coulter) by measuring $2x10^4$ cells per sample; proper gating was applied to select H3-pS10 positive events that represent mitotic cells (**Error! Reference source not found.**). The mitotic index (MI) was determined as the fraction of cells in mitosis and is shown normalized to the MI of non-irradiated controls. The actual MIs of the controls used for the normalization are given in the legends of the corresponding figures.

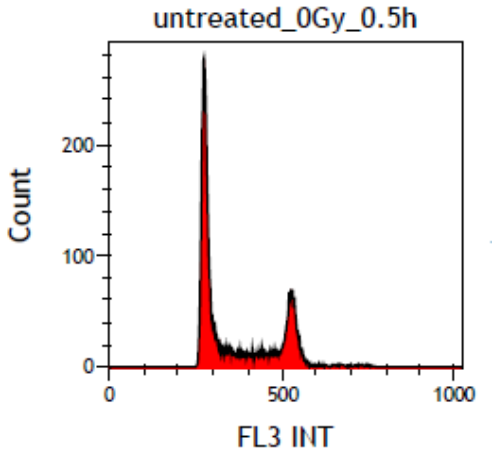


Figure 13:: Mitotic Index measurement with bivariate Flow Cytometry. Gating (green) shows the percentage of Mitotic cells in non-irradiated sample of 20.000 cells (human fibroblasts)

Library preparation for MinION

gDNA isolation

A) gDNA isolation with silica columns.

Cells are freezed directly after the desirable experimental conditions at -150°C. To start with, a thorough genomic DNA isolation (RNAfree) preparation takes place with commercially available Nucleospin kit used for forensic analysis of gDNA. Cells were collected in PBS, spinned down and redissolved in PBS. Cells were pre-lysed by incubation in ProteinaseK solution for 2 hours at 56°C. Proteinase K is used for the destruction of proteins in cell lysates (tissue, cell culture cells) and for the release of nucleic acids, since it very effectively inactivates DNases and RNases. Cells were then incubated in RNase for one hour in 20°C. Lysis was performed with 30 minutes incubation at 70°C

All lysis buffers used contain a high concentration of chaotropic salts. Chaotropes when used in nucleic acid extraction destabilize hydrogen bonds, van der Waals forces and hydrophobic interactions, leading to destabilization of proteins, including nucleases and also disrupt the association of nucleic acids with water, thereby providing optimal conditions for their transfer to silica. Chaotropic salts in our buffer include guanidine HCL. In addition to chaotropes, a detergent is often present in the lysis buffer to aid protein solubilization and cell lysis. As discussed above, chaotropic salts are critical for lysis and binding to the column. So in presence of chaotropic salts DNA is separated from the water molecules and it becomes less soluble (more hydrophobic).The addition of alcohol (or sometimes isopropanol) will further enhance and influence the binding of nucleic acids to the silica. Spin columns contain a silica resin that selectively binds DNA (or RNA), depending on salt conditions and other factors influenced by the extraction method.

After centrifuging our lysate through the silica membrane the desired nucleic acids should be bound to the column and impurities such as protein and polysaccharides should be in the flow-through. However, the membrane will contain protein and salt residues, therefore 4 washing steps with in between centrifugation steps are performed one with buffer BW which contains low concentration of chaotropic salts to remove residual proteins and pigments and three extra washes with buffer B5 to remove any residuals salts. Last a dry wash is performed to remove any residual solution for a cleaner eluent.

The final step in the DNA extraction protocol is the release of pure DNA or RNA from the silica.

For DNA extraction, 10 mM Tris at pH 8-9 is typically used. DNA is more stable at a slightly basic pH and will dissolve faster in a buffer than water. Water tends to have a lower pH of 4-5, and high molecular weight DNA may not completely rehydrate in the short time used for elution.

For maximal DNA elution, we allow the buffer to stand in the membrane for a few minutes before centrifugation. Results with our intact high molecular weight DNA, are shown in Figure 14.

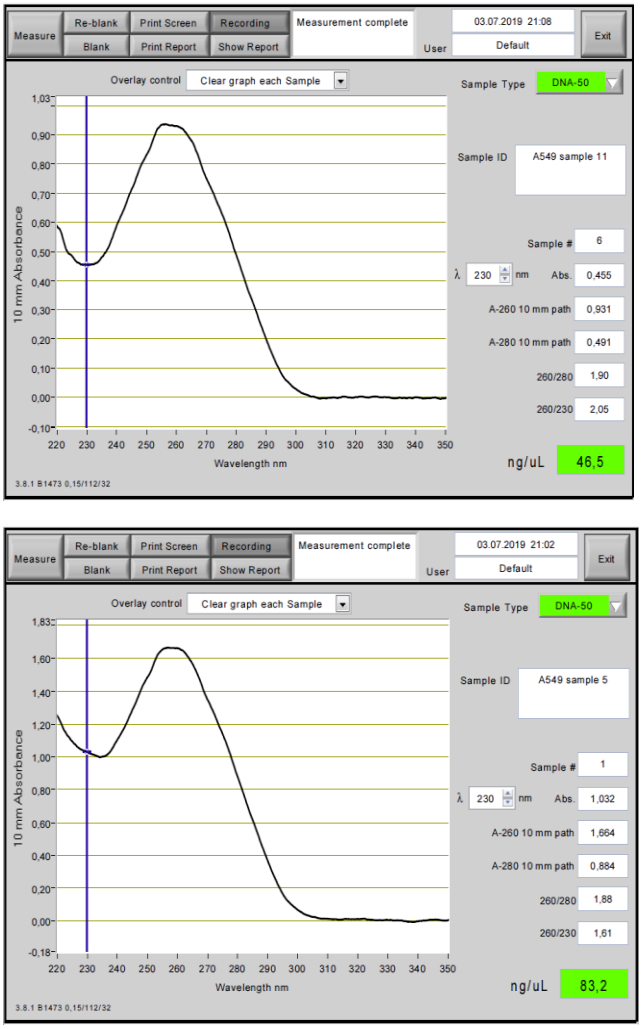


Figure 14: gDNA isolation with silica columns (on top) and increase in DNA concentration of the same sample using magnetic beads binding (bottom picture)

B) Increase of DNA concentration with magnetic beads.

With long read sequencing technologies like Nanopore a DNA concentration of minimum 40ng/µl is desired. High quantities and high purity and integrity of the DNA are essential for the device and also directly related to the quality of downstream bioinformatical analysis. In this case when the previous gDNA isolation technique fails to deliver these results we implement further increase in DNA concentration using magnetic beads. For samples with low yield in DNA concentration the following steps where performed:

AMPure XP beads were add to the solution with the isolated DNA and incubated in a rotator for 20 minutes in 20°C, this way the DNA fragments can bind on the magnetic beads. Next we pellet the beads on a magnet and remove any residual supernatant, followed up by two washes with 70% ethanol. The pellets were resuspended in prewarmed elution buffer and incubated in

20 °C for 20 minutes on a rotator. The eluted DNA was collected and quantified. Related results are shown on Figure 14 (on the right) where 78% recovery was achieved.

Library preparation is completed with end preparation and adapter ligation. Flow cell priming mix was prepared and then the samples were loaded.

Clones used for sequencing

The clones that were used to produce the sequencing data originated both from human and CHO cell lines and they are described in Table 1. Concerning the names of the clones created for example in CHO-2xS.R14 the first letters indicate the name of the cell line. Then the description of the I-SceI constructs follows, whereas the number 2x stands for the number of I-SceI sites in the integrated construct. After the description of the construct, the integration site number and the reverse (R) or direct (D) orientation of the I-SceI sites is indicated. The clone number completes the name.

Table 1: Clones used to produce sequencing data for high LET effects analysis. Number of integrations presented here are from Southern Blot results for comparison.

Clone	**Cell origin**	**DSBs per cluster**	**Integrations**	**Distance between DSBs**	**Orientation**
CHO-2xS.R14	Chinese Hamster Ovary	2	14	200 bp	Reverse
CHO-4xS.R12	Chinese Hamster Ovary	4	12	200,62,200 bp	Reverse
RPE1-2xS100.R11	Human Retinal Epithelial	2	11	100 bp	Reverse
RPE1-2xS200.R12	Human Retinal Epithelial	2	12	200 bp	Reverse

Localization through expressed vectors

Following the sequencing experiments the data produced from MinION were filtered in order to obtain the relevant reads that contain the I-*Sce*I integrations sites. This procedure was achieved by using the flanking reference sequences that surround the integration sites as it can be seen in Figure 8. During filtering, different vectors were used for the human and CHO clones respectively. The corresponding vectors for each clone can be found in Appendix 1. Filtering was implemented through k-mer search. Each scan was performed using as query 6 “18-mers” originating from the vectors (Figure 34,Figure 35,Figure 36).

One related example can be found in the commands that follow

>java -cp "C:\Program Files\bbmap\current" jgi.BBDuk2 in="F:\MinION\Data\Reads\reads\2xSR14_merged_files_all_runs\2xSR14_allrunsmerged_withoutskipped.fastq" outm="L:\Exp.xxx_MinION_2xS.R14\Hits_Exp.xxx_2xSR14_allrunsmerged_withoutskipped_search1.fastq" k=18 mm=f

literal=tagggataacagggtaat,GTCTGCCGTAGATATCTC,atcgccatgggtcacgac,catttaatcctagtaaaa,gctgcctcgt cctgcagt,cacaaagctctgacctca overwrite=true

k=18
kfiltering using 6 literals.
Initial:
Memory: max=910m, free=46m, used=16m
Added 6 kmers; time: 0.033 seconds.
Memory: max=910m, free=47m, used=15m
Input is being processed as unpaired
Started output streams: 0.058 seconds.
Processing time: 232.575 seconds.

Input: 513224 reads 3521109518 bases.
Contaminants: 13 reads (0.00%) 333018 bases (0.01%)
Total Removed: ***13 reads*** *(0.00%) 333018 bases (0.01%)*
Result: 513211 reads (100.00%) 3520776500 bases (99.99%)

Time: 232.686 seconds.
Reads Processed: 513k 2.21k reads/sec
Bases Processed: 3521m 15.13m bases/sec

Reads containing at least one 18-mer pass as positive results. The results from each search resulted in our first pool of data that were processed.

Those data were compared to similar results using the online Blast tool with input all the sequencing data and query for the alignment the corresponding vector (see Appendix 1). Results were filtered using E-value and BitScore values adjusted appropriately for each alignment.

Results were extracted in fasta format out of the pool of the total sequencing data for each experiment under a Linux-terminal the following way

>awk 'BEGIN [n_seq=1;] /^>/ [if(n_seq%1==0)[file=sprintf("Exp-Hit-%d.fa",n_seq);] print >> file; n_seq++; next;] [print >> file;]' < Exp-BLAST-Hits.fasta

The combined data (‘reads’) were then examined (Figure 15) individually using the open source software ‘A plasmid Editor’ obtained online from Department of Biology, University of Utah, Salt Lake City, UT 84112, USA in order to decide whether they contain part or all of the vector.

False positive results were excluded and the final results are shown in the last column of Table 2.

Table 2: General information about the sequencing data

Clone	#Sequencing experiments	Sequencing data (Gbases)	#Reads representing Integrations	#Reads under interrogation
CHO-2xS.R14	5	20	31	36
CHO-4xS.R12	6	32	35	186
RPE1-2xS100.R11	6	29,8	23	
RPE1-2xS200.R12	1	14,6	14	

The reads that were examined and were identified as reads that represent the integration (partially or in total) were then examined against all the other sequencing experimental data so that duplicates can be excluded. In general for reference the biggest (in terms of kbases) read each time was chosen as representative of the corresponding integration.

Furthermore the reads were examined with the respective reference genomes in order to localize them in the relevant chromosomes. Reference genomes were downloaded from National Center for Biotechnology Information (NCBI) and the most recent version was selected. For analysis concerning Chinese Hamster the following reference genome was chosen:

CriGri-PICR

Organism name: Cricetulus griseus (Chinese hamster)

Infraspecific name: Strain: 17A/GY

Sex: female

BioSample: SAMN07140313

BioProject: PRJNA389969

Submitter: CHO Genome Community

Date: 2018/10/21

Assembly level: Scaffold

Genome representation: full

GenBank assembly accession: GCA_003668045.1 (latest)

RefSeq assembly accession: GCF_003668045.1 (latest)

RefSeq assembly and GenBank assembly identical: no

WGS Project: RAZU01

Assembly method: AllPaths v. 2017-05-01; HGAP v. 2017-05-01; MetaAssembler v. 2017-08-01

Genome coverage: 45.0x

Sequencing technology: Illumina HiSeq; PacBio

IDs: 2027221 [UID] 7620648 [GenBank] 8069188 [RefSeq]

For analysis concerning the human (RPE) clones the following reference genome was chosen:

GRCh38.p13

Description: Genome Reference Consortium Human Build 38 patch release 13 (GRCh38.p13)

Organism name: Homo sapiens (human)

BioProject: PRJNA31257

Submitter: Genome Reference Consortium

Date: 2019/02/28

Assembly type: haploid-with-alt-loci

Release type: patch

Assembly level: Chromosome

Genome representation: full

RefSeq category: reference genome

GenBank assembly accession: GCA_000001405.28 (latest)

RefSeq assembly accession: GCF_000001405.39 (latest)

RefSeq assembly and GenBank assembly identical: no

IDs: 2334371 [UID] 8687898 [GenBank] 8765528 [RefSeq]

The previous process was performed under a typical command line interpreter application using the local blast tool (Coordinators, 2013; Wheeler et al., 2000). First a database was created with the following command

>makeblastdb -in CHO-4xSR14_All_Int-Reads_joined.fasta -dbtype nucl -out CHO4xSR

Building a new DB, current time: 01/21/2019 11:23:03

New DB name: C:\Users\ChristinaVasileiou\Desktop\Blast\DB\CHO2X

New DB title: CHO-2xSR14_All_Int-Reads_joined.fasta

Sequence type: Nucleotide

Keep MBits: T

Maximum file size: 1000000000B

Adding sequences from FASTA; added 35 sequences in 0.0267036 seconds.

Then the reads were tested as followed and the output file was used for further investigations.

> blastn -db DB\NewCHOrefgenome\RefGenome -query Query\E654_E653_joinedFPclean.fasta -out Results\LocalisedE654E653.out -evalue 1e-100

Following the annotation of the reads in chromosomes, the next step in the analysis constituted of examination of multiple reads that represent the same integration (eg integration AA) in public available genome browsers, which resulted in the localization of the integration in their exact 'coordinates' in the genome. Each integration-read was then aligned and compared against the remaining to follow through with the analysis.

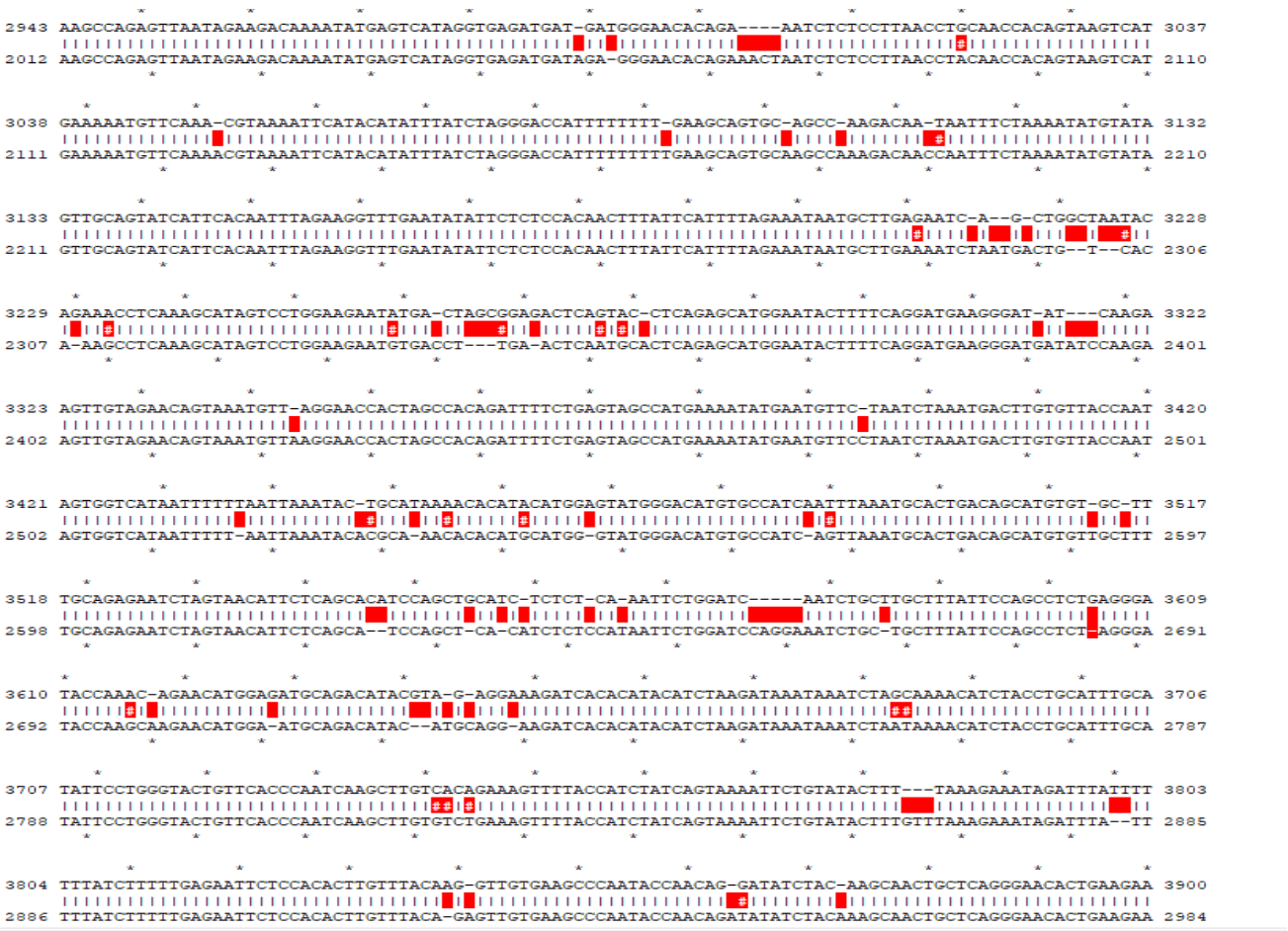


Figure 15: Example of identification of the integration site using alignment with the corresponding vector (A plasmid Editor)

Primers Design

Primers to test and cross validate our experimental hypotheses through PCR reactions were designed through 'Blast primers' (Ye et al., 2012) search optimizing for various parameters in order to avoid DNA amplification, such as melting and annealing temperature , self-complementarity (which describes the probability that the forward primer will bind to the reverse one of his pair), 3' self-complementarity (which describes the probability that each primer will bind to himself), product length, GC content etc. Melting temperatures were kept around (58-60)°C and the maximum difference between the primers of each pair were kept between 1-2°C, primers were also chosen that were separated at least by one intron assuring this way that each primer will bind to a different exon. Intron and product lengths were kept to minimum. GC content was selected to be 50%.Repeated nucleotides were also avoided.

PCR

The Polymerase chain reaction was used for the amplification of single DNA sequences to obtain DNA fragments to cross-validate various hypotheses. The PCR reaction that were designed fit in the following categories:

A) PCR verification of Integrations
   In this category PCR reactions were designed with the respective primers in order to validate the integrations that were found with previous described bioinformatical analysis. Taq DNA polymerase was used at first (Table 3, Figure 28) but due to the relative big length products that are expected in our project Phusion Polymerase was also tested (Table 4, Figure 29). Scientific® Phusion High-Fidelity DNA Polymerase offers higher fidelity and robust performance. In particular it is speculated that it is 50 times higher fidelity is expected than Taq and it also offers robust reactions with minimal optimization for bigger products. In the same direction, to avoid unspecific products PCR reactions were designed where one 'internal primer' was used in the pair tested (Table 5,). By the term internal primer, it is indicated that the primer was designed to bind to sequences on the construct (probe as shown in Figure 8).
B) PCR investigation of potential translocations.
   According to the model that is proposed in following sections PCR reactions are anticipated to verify the existence of potential translocations. These reactions were designed with the previous protocol (Table 4, Figure 29) but further optimization was adopted. First a blocking oligonucleotide was used to avoid unspecific products. Blocking oligonucleotides bind to library adapter sequences to reduce off-target capture during library enrichment and this way increased specificity is achieved. The proper concentrations in this reactions were tested (Table 6, Figure 32).

PCRs was set up in a total Volume of 20-25µL with concentrations of template DNA ranging from 1ng-10pg. Furthermore the reaction contained 100mM dNTP Mix, 1 x PCR-Buffer, 2.5U NEB_Taq- or PhusionR Polymerase and distilled H_20. For primer sequences, refer Appendix, Table 2.

The amplification reactions were performed in a Thermo-Cycler with the following programs:

Table 3: PCR scheme and thermocycler settings for NEB_Taq Polymerase

Standard PCR-Scheme

Component	25 µl reaction	5x MasterMix
10X Reaction Buffer	2,5	12,5
dNTPs (10mM)	0,5	2,5
Forward Primer (10µM)	0,5	
Reverse Primer (10µM)	0,5	
Template DNA	1	5
NEBTaq DNA Polymerase*	0,25	1,25
Nuclease-free water	19,75	98,75
Total Volume	**25**	**120**

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	94°C	30 seconds	1
Denaturation	94°C	30s	
Annealing	58°C	30s	35
Extension	68°C	300s	
Final Extension	68°C	5 minutes	1
Hold	4°C	∞	

Primers

CHO4x-AA-FW	CCCTCAAAGGCTAGCCAACA
CHO4x-AA-REV	GCAGACAGGCAGGAAGTGAT
CHO4x-AB-FW1	ATGTCAGGCCCTGGAAATGG
CHO4x-AB-REV	CCAGTCCACCTTTGCGGTAT
CHO4x-AE-FW	GGGCTCAGTGGACAAAATGC
CHO4x-AE-REV	AGGCTATCATGCTTGGGGTG
CHO4x-AF-FW	TGGAGCTGGAGGTACAGACA
CHO4x-AF-REV	CCAGTCCACCTTTGCGGTAT

Table 4: PCR Scheme with Phusion Polymerase.

Phusion PCR-Scheme

Component	20 µl reaction	5x MasterMix
5X Phusion HF	4	20
dNTPs (10mM)	0,4	2
Forward Primer (10µM)	1	
Reverse Primer (10µM)	1	
Template DNA	1	5
Phusion DNA Polymerase*	0,2	1
Nuclease-free water	12,4	62
Total Volume	**20**	**90**

CYCLE STEP	TEMP	TIME	CYCLES		
Initial Denaturation	98°C	30s	1		
Denaturation	98°C	10s	35		
Annealing	64°C	20s			
Extension	72°C	160s			
Final Extension	72°C	5 min	1		
Hold	4°C	∞			
1.25U for both NEB_Taq with Thermo Pol or DreamTaq					
Extension:68° for NEB-Taq, 72°C for DreamTaq;72°C for Phusion					

Primers			with Int.	w/o Int
R1	CHO4x-AA-FW	CCCTCAAAGGCTAGCCAACA	3715	1015
R1	CHO4x-AA-REV	GCAGACAGGCAGGAAGTGAT		
R2	CHO4x-AB-FW1	ATGTCAGGCCCTGGAAATGG	3142	442
R2	CHO4x-AB-REV	CCAGTCCACCTTTGCGGTAT		
R3	CHO4x-AE-FW	GGGCTCAGTGGACAAAATGC	3595	895
R3	CHO4x-AE-REV	AGGCTATCATGCTTGGGGTG		
R4	CHO4x-AF-FW	TGGAGCTGGAGGTACAGACA	3720	1020
R4	CHO4x-AF-REV	CCAGTCCACCTTTGCGGTAT		

Table 5: Experimental design with internal primer

Primers			Product length
Sceinsert460ASRev	ACTTTCCACACCCTAACT		1547
CHO4x-AA-Rev	GCAGACAGGCAGGAAGTGAT		
Sceinsert460ASRev	ACTTTCCACACCCTAACT		1334
CHO4x-AB-REV	CCAGTCCACCTTTGCGGTAT		
Sceinsert460ASRev	ACTTTCCACACCCTAACT		1489
CHO4x-AE-Rev	GGGCTCAGTGGACAAAATGC		
CHO4x-AH-FW	GCAATCTCTATCTCACTGGCCT		1284
Sceinsert460ASRev	ACTTTCCACACCCTAACT		
ID	**Name**	**Seq**	
277	Sceinsert460ASFwd	GACTGTGCCTTTAAACAGC	ISceI-construct-FW
278	Sceinsert460ASRev	ACTTTCCACACCCTAACT	ISceI-construct-Rev

CYCLE STEP	**TEMP**	**TIME**	**CYCLES**
Initial Denaturation	94°C	30 seconds	1
Denaturation	94°C	30s	35
Annealing	58°C	30s	
Extension	68°C	300s	
Final Extension	68°C	5 minutes	1
Hold	4°C	∞	

Table 6:PCR oligo nucleotide optimization the concentrations tested were 0M, 25nM, 62.5nM,125nM,187.5nM,250nM,1.25µM,2.5µM.

Standard PCR-Scheme		
Component	**25 µl reaction**	**10x MasterMix**
10X Reaction Buffer	2,5	25
dNTPs (10mM)	0,5	5
Forward Primer (10µM)	0,5	5
Reverse Primer (10µM)	0,5	5
Template DNA	1	10
Blocking-Oligo	1	
NEBTaq DNA Polymerase*	0,25	2,5
Nuclease-free water	18,75	187,5
Total Volume	**25**	**240**

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	94°C	30 seconds	1
Denaturation	94°C	30s	
Annealing	58°C	30s	35
Extension	68°C	300s	
Final Extension	68°C	1 minutes	1
Hold	4°C	∞	

1.25U for both NEB_Taq with Thermo Pol or DreamTaq
Extension:68° for NEB-Taq, 72°C for
DreamTaq

Primers

CHO4x-AB-FW1
CHO4x-AB-REV

Blocking-Oligo

AB-BlockShort-S1

PCR results for individual cases were further analyzed with the free software Image Studio Lite Ver 5.2 which allows analysis based on optical density measurements.

Investigation of Translocation formation model

In this project we present a model tested for CHO-4xS.R12 clones to test through product-tailored PCR reactions, structural variations and specifically translocation formation. As we can see in Figure 16 the product of a PCR reaction of a construct that contains the integrative part will be equal to the relevant length shown for the selected primers plus 2,7 kilobases (length of integrative part of CHO-4xS.R12, see Appendix 1). In case there is some deletion or an unrepaired DSB we expect no product as it also shown in Figure 16. In the next step, through testing and through indicative results from mFISH for common translocations between specific chromosomes (eg a common translocation formed in CHO-4xS.R12 clones is found between chromosome 1 and chromosome 2), we designed pair of relevant primers (eg forward primer of chromosome 1 paired with reverse primer of chromosome 2) and the PCR results in case they were found to be around the length of the integrating part plus the distance that the primers cover for the specific region then the results will confirm the formation of a translocation in the DNA under interrogation.

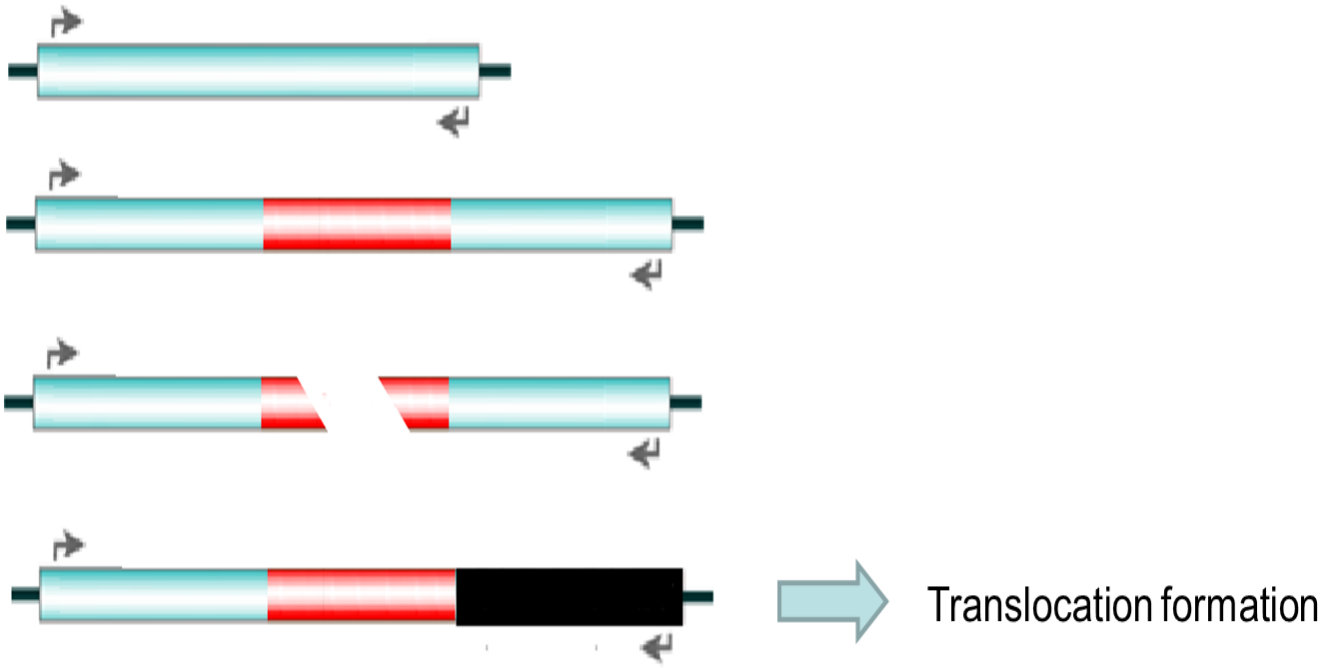


Figure 16: Proposed model to experimentally identify translocation formation illustrated

A number of combinations of the previous discussed model were tested but not all are presented here. We will focus in the experimental PCR scheme presented in Table 7.

Table 7: Example of experimental design of proposed model to investigate SVs

Component	20 µl reaction	5x MasterMix
5X Phusion Buffer	4	20
dNTPs (10mM)	0,4	2
Forward Primer (10µM)	1	
Reverse Primer (10µM)	1	
Template DNA	1	5
Phusion DNA Polymerase*	0,2	1
Nuclease-free water	12,4	62
Total Volume	**20**	**90**

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	98°C	30s	1
Denaturation	98°C	10s	35
Annealing	64°C	20s	
Extension	72°C	300s	
Final Extension	72°C	5 min	1
Hold	4°C	∞	

Reaction	**Primer-1**	**Primer-2**
AS>J-FW	CHO4x-AS-FW	CHO4x-J-REV
J>AS-FW	CHO4x-J-FW	CHO4x-AS-REV
J>AS-FF	CHO4x-J-FW	CHO4x-AS-FW
AS>J-RR	CHO4x-J-REV	CHO4x-AS-REV
AB>AE-FW	CHO4x-AB-FW	CHO4x-AE-REV
AE>AB-FW	CHO4x-AE-FW	CHO4x-AB-REV
AE>AB-FF	CHO4x-AB-FW	CHO4x-AE-FW
AB>AE-RR	CHO4x-AB-REV	CHO4x-AE-REV

Human clones analysis Vector identification with alignment tools

Concerning the human cell lines from which data were analyzed identification and localization of the integration sites was performed in an identical way as it is previously described for CHO clones but this time a different approach was adopted to study the existence of SVs. Human genome is pretty nice documented which allows us through a bioinformatics approach to investigate the potential translocation formation. To this end a topological approach was implemented and results that were generated from alignment tools and genome browsers (discussed previously) were further analyzed. In particular statistically significant results for the reads under interrogation that refer to different chromosomes were examined and potential translocations were revealed.

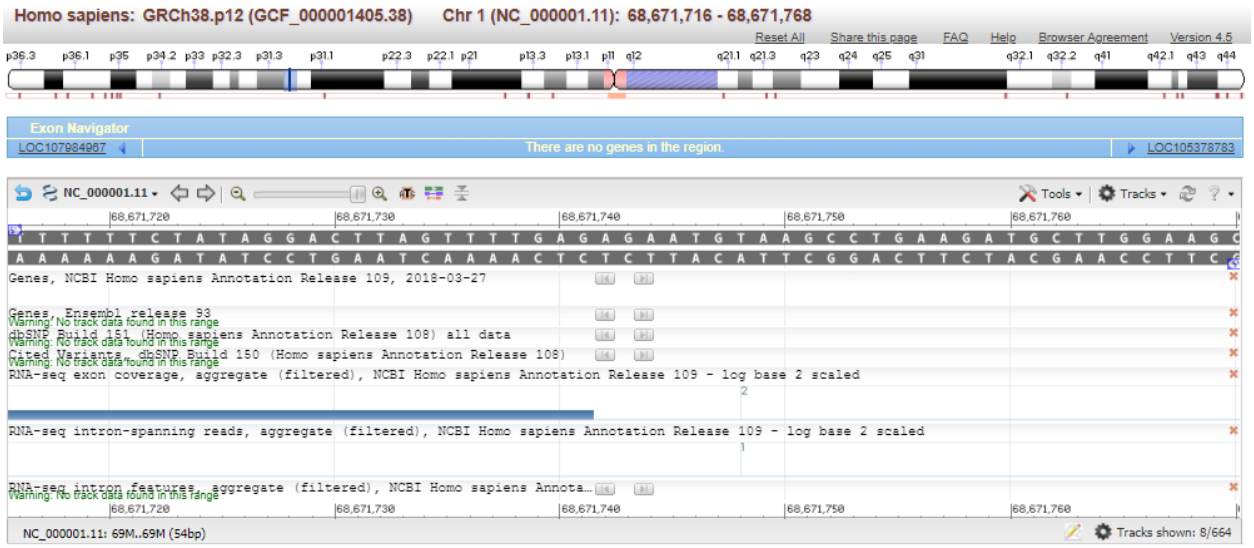


Figure 17:Genome browser analysis in Blast.

Results

Flow Cytometry analysis of the Mitotic Index

ATM and ATR inhibition in 82.6 cell line at low doses (2Gy).

After exposure to IR, cells experience a delay in their progression from the G2 into the M-phase of the cell cycle. As discussed above, the ATM-Chk2 and the ATR-Chk1 signaling pathways are thought to be integral to the signal transduction underlying the development of the G2 checkpoint. The contributions of both ATM and ATR was previously tested (Mladenov et al., 2019). Here we set out a series of experiments to repeat and validate the previously discussed contribution of these two kinases in the inception of G2/M checkpoint. 82.6 cells were used as control in these experiments. The results obtained with this cell line are summarized in Figure 18.

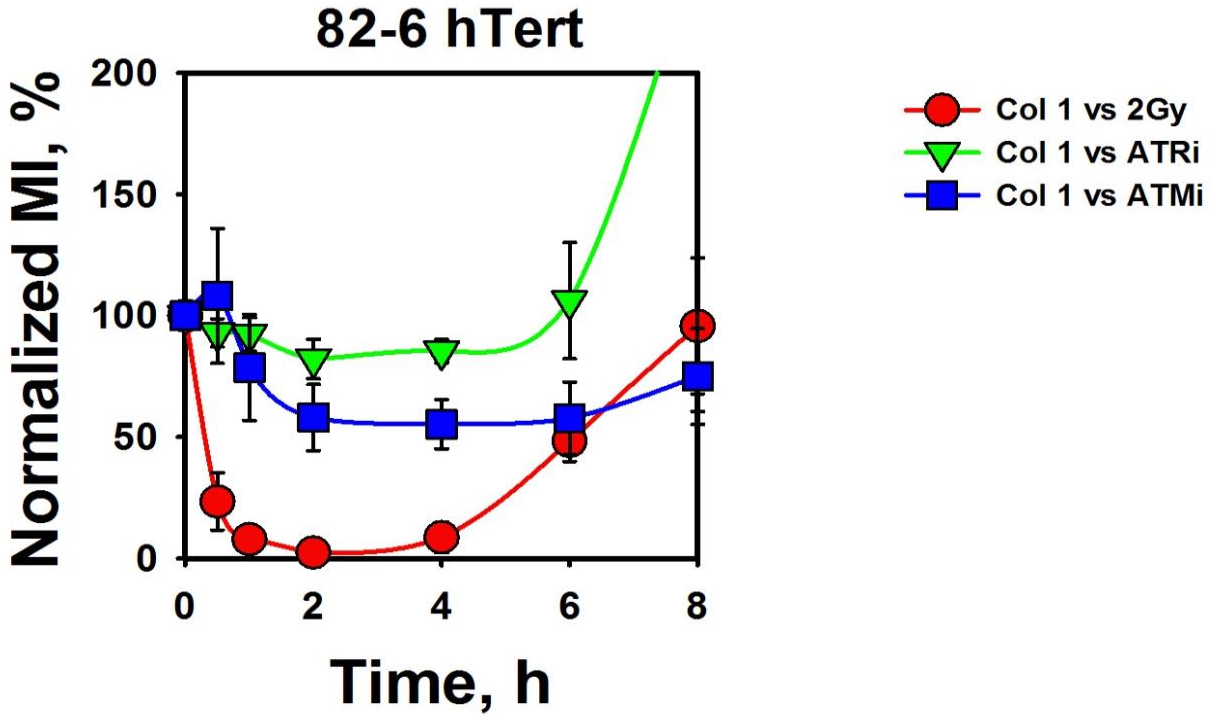


Figure 18: Exponentially growing 82.6 cells were pre-treated with ATM and ATR inhibitors 1 h before 2Gy IR, and the drugs were maintained in the duration of the experiment. Cell cycle distributions are calculated. G2 fractions obtained are plotted against time. (A), 10 µM KU-55933; (B)5 µM VE-821.The results are from 3 experiments. Error bars stands for standard deviation.

Exposure of control samples of 82.6 cells at 2Gy IR reveals a significant drop in MI 1 hour post irradiation which indicates the prompt activation of the G2/M checkpoint. The checkpoint is

maintained for up to 4 h, but cells start re-entering mitosis at later times indicating recovery from the checkpoint that is nearly complete at 8 h.

ATM inhibition partially abrogates the G2/M checkpoint. As it can be seen in Figure 18, 2 hours post irradiation ATM inhibition drops the MI around 50% revealing underlying mechanisms that function in order to maintain the G2 checkpoint. A slight increase but not a complete recovery of the MI is shown 8 hours post irradiation. Key results are shown concerning ATR inhibition which completely abrogates the G2/M checkpoint and irradiated cells enter mitosis without delay.

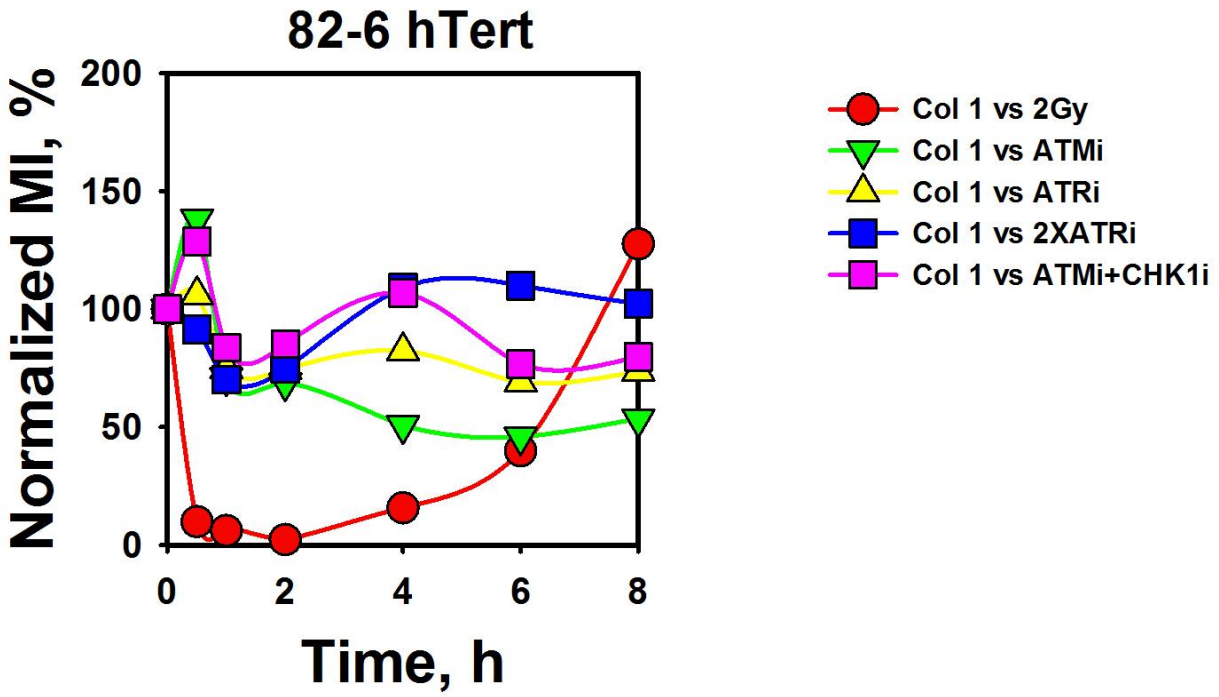


Figure 19: Exponentially growing 82.6 cells were pre-treated with various concentrations of ATM, ATR, CHK1 inhibitors 1 h before 2Gy IR, and the drugs were maintained in the duration of the experiment. Cell cycle distributions are calculated. G2 fractions obtained are plotted against time.

Chk1 and Chk2 inhibition in 82.6 cell line at low doses (2Gy).

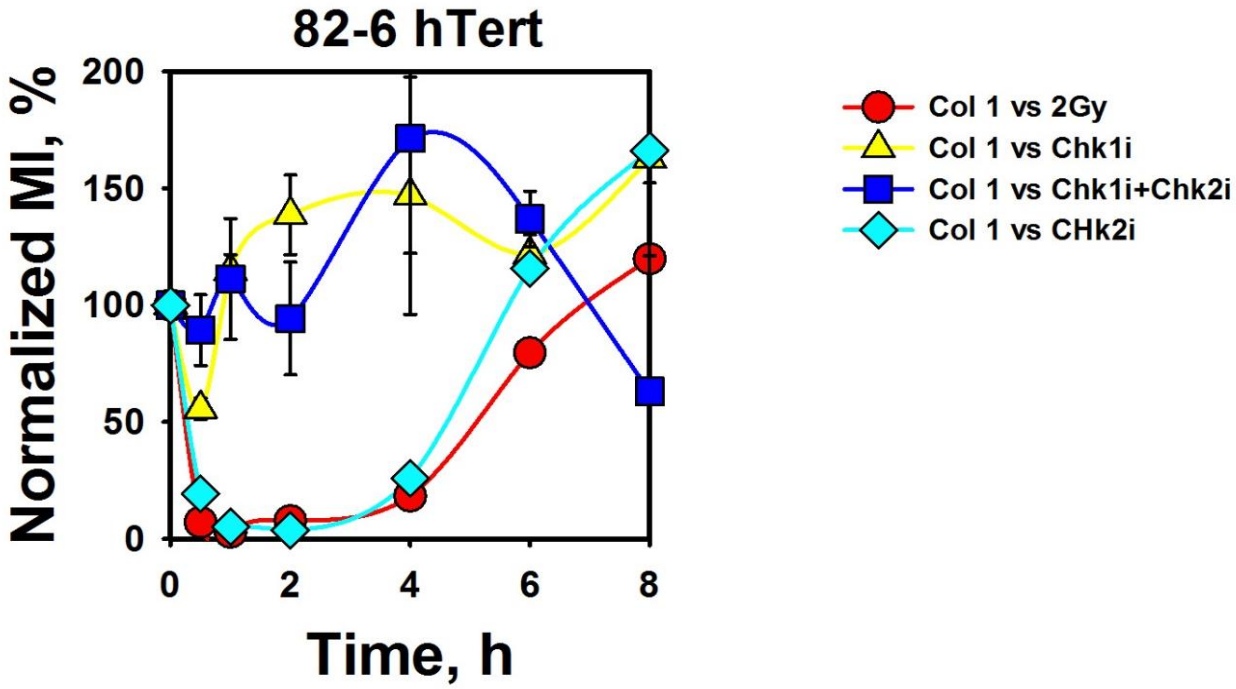


Figure 20: Normalized MI as a function of time after exposure to 2 Gy of 82-6 hTert cells, treated with Chk1 or Chk2 inhibitors added 1 h before 2Gy IR, and the drugs were maintained in the duration of the experiment. Cell cycle distributions are calculated..

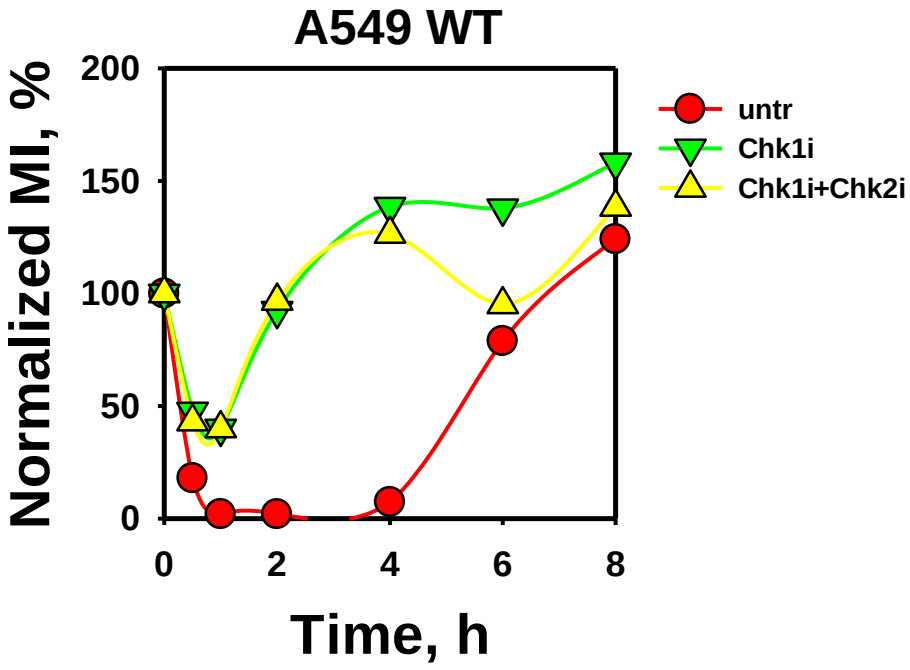


Figure 21: Normalized MI as a function of time after exposure to 2 Gy of A549 cells, treated with Chk1 or Chk2 inhibitors added 1 h before 2Gy IR, and the drugs were maintained in the duration of the experiment. Cell cycle distributions are calculated.

The G2-checkpoint is implemented by ATM and ATR which suppress Cdc25C through Chk1 and Chk2 at the end of the G2 phase who are responsible to enable progression of cells into mitosis (Blackford and Jackson, 2017; Hoekstra, 1997; Iliakis, 1997; Mladenov et al., 2019; Nakanishi, 2001). We examined therefore the effect of both those kinases, Chk1 and Chk2 on the G2-checkpoint under the low IR dose conditions by repeating the experiments (Mladenov et al., 2019) as previously described. Treatment of 82-6 hTert cells with the Chk2 inhibitor (Chk2 inhibitor II, BML-277) fails to generate any effects on the induction of the checkpoint and has only a small effect on its recovery 6 and 8 hours post irradiation. Similar results are shown also in Figure 22 where Chk2 inhibition was combined on top of ATM and ATR inhibition. On the other hand, UCN-01, inhibitor of Chk1 causes a significant suppression of the G2-checkpoint in 82-6 hTert but not a complete abrogation of it. These results combined with the previous results discussed about ATR function in G2 checkpoint (Figure 18), suggest as it is well reported in (Mladenov et al., 2019) that ATR controls the G2 checkpoint through Chk1 and other yet unidentified kinases. Combined inhibition of both kinases seem to slightly increase the suppression of the G2 checkpoint as described for Chk1 inhibition alone. Similar results are repeated and validated for A549 cells too (Figure 21).

Chk2 in combination with other kinases inhibition in 82.6 cell line at low doses (2Gy).

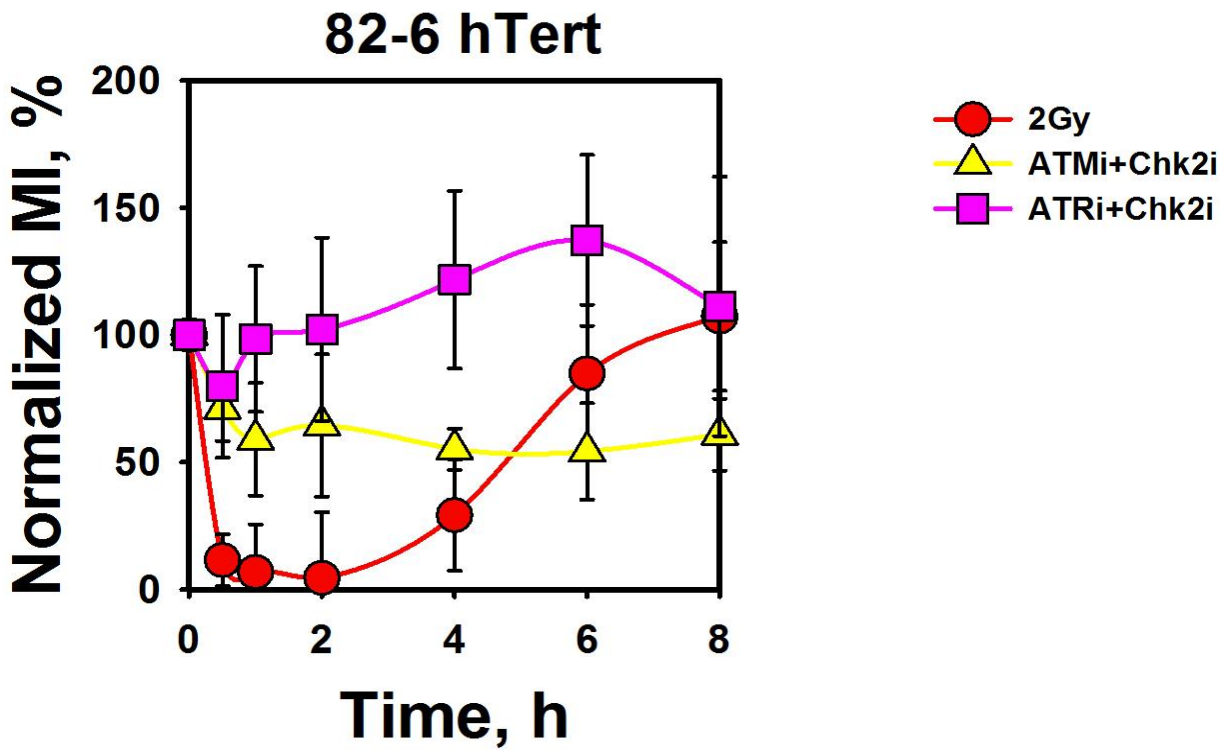


Figure 22:Exponentially growing 82.6 cells were pre-treated with ATM, ATR and CHK2 inhibitors 1 h before 2Gy IR, and the drugs were maintained in the duration of the experiment. Cell cycle distributions are calculated.

Chk2 inhibition on top of ATM and ATR doesn´t seem to generate additional effects on G2/M checkpoint results that are in harmony with results shown in Figure 20 and Figure 21.

Chk1 and Chk2 inhibition in 82.6 cell line at high doses (10Gy).

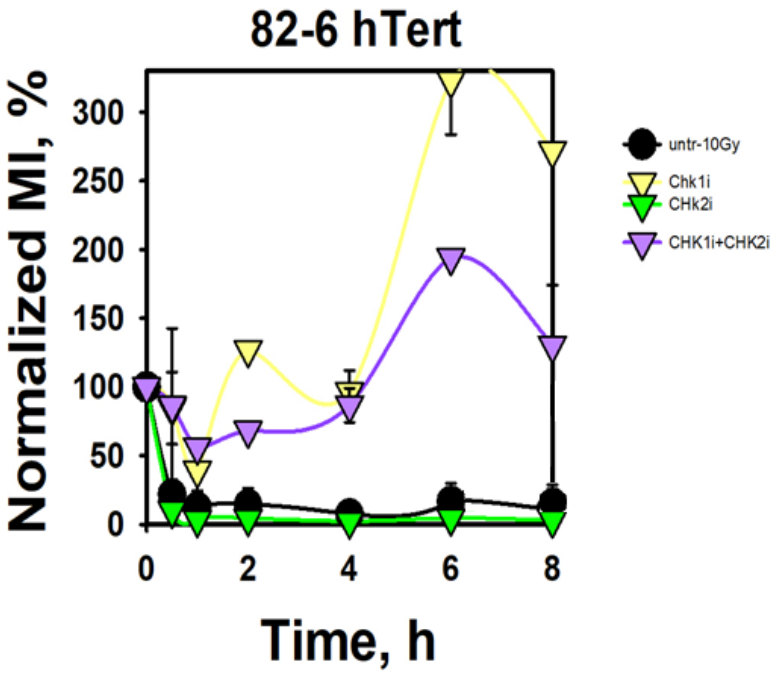


Figure 23: Normalized MI as a function of time after exposure to 10 Gy of 82-6 hTert cells, treated with Chk1 or Chk2 inhibitors added 1 h before 10Gy IR, and the drugs were maintained in the duration of the experiment. Cell cycle distributions are calculated

High IR dose (10Gy) initiates a strong checkpoint that prevents any cells from entering mitosis during the same time of observation (up to 8 hours).This shows a completely different response of the cells indicating a coherent dose dependent response mechanism from the cells. Chk1 inhibition causes a significant suppression of the G2-checkpoint in 82-6 hTert cells although it quickly recovers after 2 hours and cells enter mitosis. Treatment of 82-6 hTert cells with the Chk2 inhibitor fails to generate any effects on the induction of the checkpoint results that are similar to the response in low doses of IR. Combination of above mentioned inhibitors of kinases shows a much stronger suppression of the G2 checkpoint and complete abrogation of it after 4 hours indicating a potential synergistic action between the two kinases in the inception and recovery of the checkpoint.

Identification and localization of I-*Sce*I sites in CHO clones

Concerning the CHO clones analysis through sequencing data from the platform MinION, two clones were analyzed CHO-2xS.R14 and CHO 4xS.R12 (see Figure 6 for further details on these clones). Chinese hamsters have 22 chromosomes (10 pairs of autosomes and 1 pair sex-chromosomes). Although the karyotype of the CHO cell lines under investigation might slightly differ from the one used as reference. The CHO genome has a size of approximately 2.6 Gbases.

CHO-2xS.R14 clones analysis

In CHO-2xS.R14 5 sequencing experiments from MinION were analyzed. 36 reads were analyzed in total resulting in 31 integration sites.

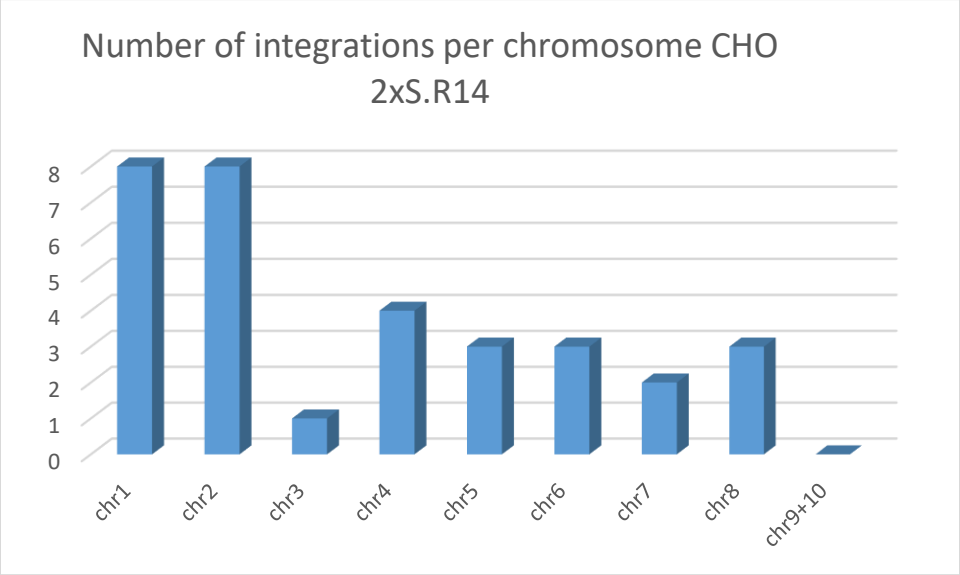


Figure 24: Distribution of Integration sites per chromosome in CHO 2xS.R14 clone

This clone was the last to be analyzed and further sequencing experimental data are required A different approach was adopted during library preparation of some samples (eg no ligation) resulting in lower yield of throughput. In general since CHO genome is 2.6 Gbases and integration sites might belong either to the paternal or maternal chromosome at least 2Xcoverage is required to estimate the real number of integration sites, which in not the case on most sequencing data that are attributed to this analysis group. A general pattern is shown in Figure 24 where it is evident that the number of integrations follows the chromosome size. We analyzed 3 integrations that were mapped by different reads either from the same or different sequencing experiments. Further gene analysis was not performed at this stage.

CHO 4xS.R12clones analysis

In CHO-4xS.R12, 6 sequencing experiments from MinION were analyzed. 186 reads were analyzed in total resulting in 35 integration sites. In particular 36 Gbases of sequencing data was examined which results in 14X haploid coverage or 7X diploid coverage giving us a high level of confidence in our results. 80% of these reads were assigned with high confidence to stable recurring integrations. 10% could not be assigned due to either insufficient amounts of flanking chromosomal DNA or because they are embedded in yet unknown sequences. The 35 Integrations that were confirmed with multiple reads in multiple independent sequencing experiments are shown in Figure 25 with the respective number of reads for each integration. There is an evident bias towards some integrations/DNA regions that can possibly be assigned to the difficulty to detect certain integrations and their probability to be cut.

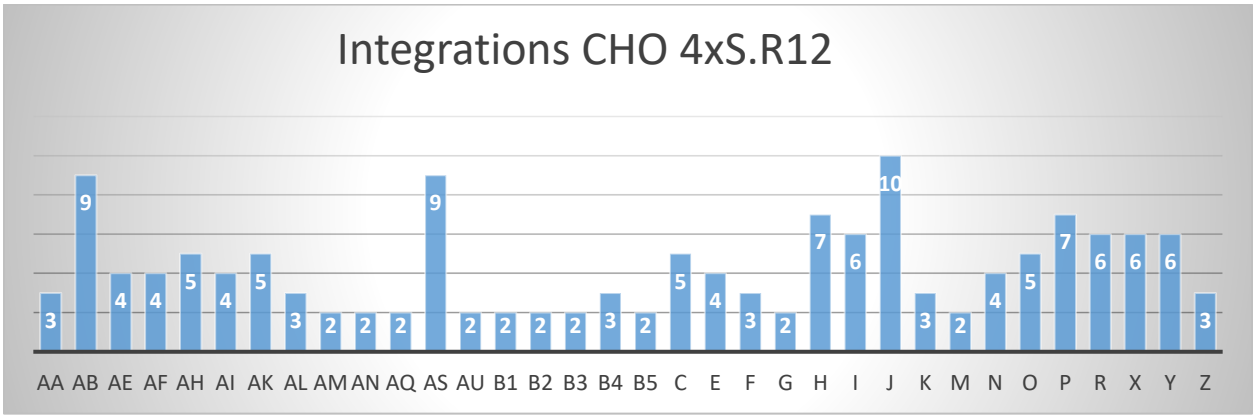


Figure 25: Number of reads that represent the same Integration site in CHO 4xS.R12 clone

There are another 22 Integration sites that only appear once in a single experiment. These integration sites can be clearly assigned to a genomic location and unambiguously contain the Integrating-construct. A key aspect of the analysis that shows the accuracy of this sequencing technique is that in general in all analyzed clones the alignment allowed us to place the integrations with an accuracy of 1-2 bases. An example of this accurate allocation of the integration and a general overview of one clone analysis can be found on Appendix 3. Again worth mentioning is the fact that the number of integrations (except chr6) follows the size of the chromosome. So DNA content could be used as a prognostic tool to the number of integrations expected (Figure 26, Figure 27).

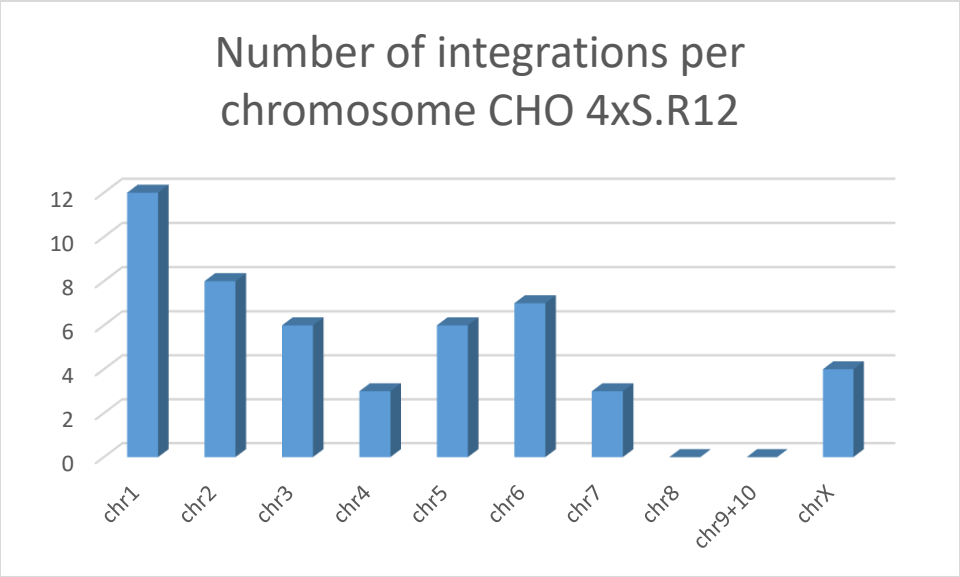


Figure 26: Integrations per chromosome

Figure 27: Chinese Hamster Genome

PCR verification of I-*Sce*I sites

For CHO 4xS.R12 and for Integrations that were confirmed at least from 3 or more (to increase the level of confidence) primers were designed (Table 9). From 36 Integration sites we concluded to 24 thus 24 pairs of primers were used in to verify our results. In Figure 28 and Figure 29 some indicative results for the Integrations AA, AB, AE, and AF. Results with Phusion polymerase reveal less unspecific products validating the hypothesis that this polymerase is more suitable for bigger (in terms of length) products. Regarding the verification of the Integration, referring to Figure 30 , the expected products without the integration construct for

all 4 reactions are easily observed on the lower bands. On the other hand the expected bands that contain the integration construct are only observed in AF reaction.

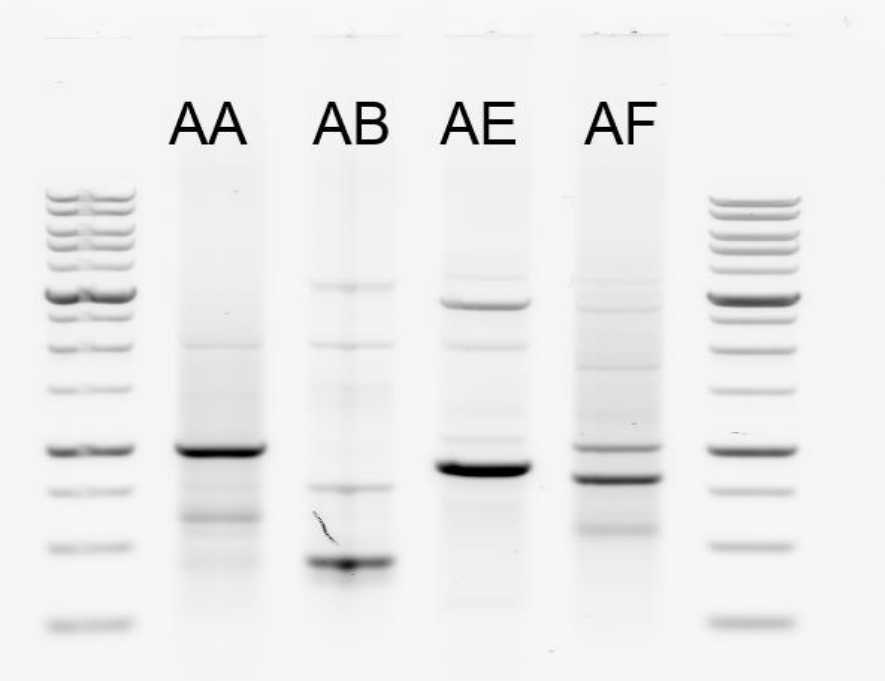


Figure 28: PCR results for NEB_Taq Polymerase (see Table 3)

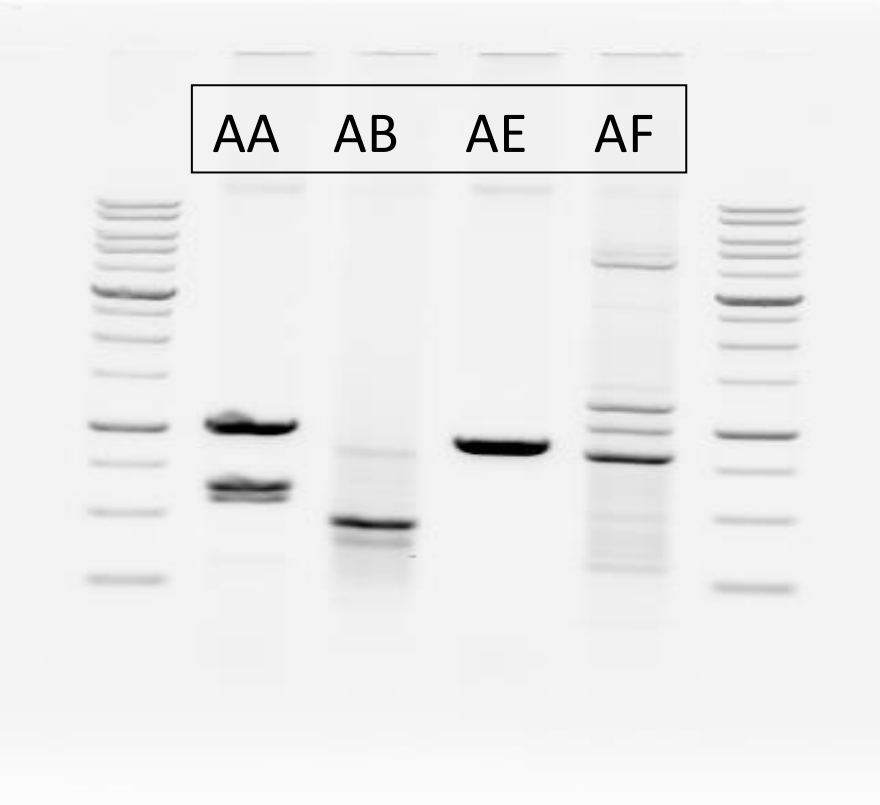


Figure 29: PCR Results with Phusion Polymerase (see Table 4)

		Expectations with Int.	wo Int
CHO4x-AA-FW	CCCTCAAAGGCTAGCCAACA	3715	1015
CHO4x-AA-REV	GCAGACAGGCAGGAAGTGAT		
CHO4x-AB-FW1	ATGTCAGGCCCTGGAAATGG	3142	442
CHO4x-AB-REV	CCAGTCCACCTTTGCGGTAT		
CHO4x-AE-FW	GGGCTCAGTGGACAAAATGC	3595	895
CHO4x-AE-REV	AGGCTATCATGCTTGGGGTG		
CHO4x-AF-FW	TGGAGCTGGAGGTACAGACA	3720	1020
CHO4x-AF-REV	CCAGTCCACCTTTGCGGTAT		

Figure 30: Expected product lengths for PCR verification of Integration sites for CHO 4xS.R12 and Integrations AA, AB, AE and AF.

An internal primer on the construct and a blocking oligo-nucleotide was used to address these issues where it is evident that during the PCR reaction the two products for each integration (with and without the construct) are antagonizing during the strand synthesis and amplification. The proper concentrations of the blocking oligo-nucleotide was thoroughly tested. In Figure 32 we see the action of this blocking oligo-nucleotide with a PCR on AB integration. The concentrations used in this reaction are 0M, 25nM, 62.5nM, 125nM, 187.5nM, 250nM, 1.25µM, 2.5µM from left to right and it is obvious that with increasing concentration the lower band

(442 bases) significantly decreases while at the same time the upper band that is expected to represent the integration (3595 bases) starts to appear, validating our hypotheses.

Further optimization was performed (Figure 31). Usage of one primer in the construct causes some specificity issues.

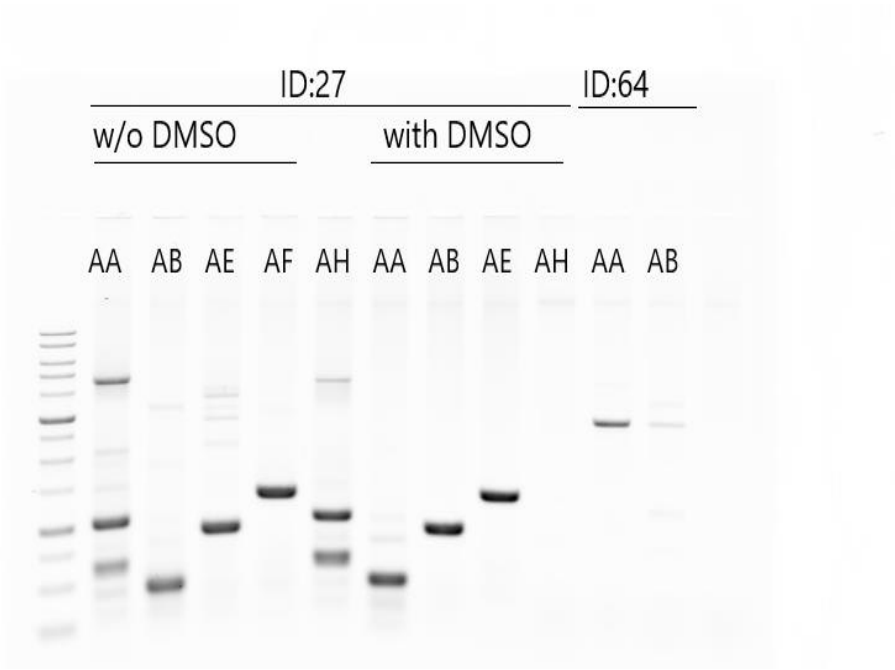


Figure 31: Optimization using DMSO and further testing on different DNA samples.

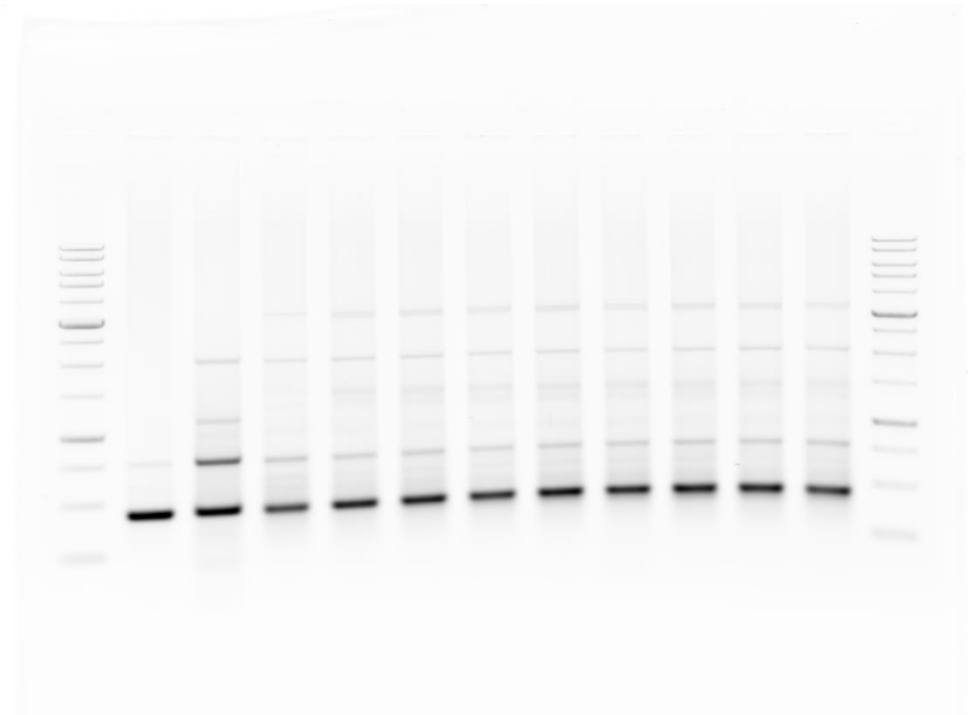


Figure 32: PCR Results for various concentrations of blocking oligo-nucleotide and AB integration (see Table 6)

Model to study translocations

Following the verification of the integrations sites, structural variations were examined. DNA induced with I-*Sce*I double strand breaks was tested this time in order to have a closer look to possible genomic rearrangements. Various combinations of primers were tested randomly or guided by mFISH results. One example of these combination is shown in Figure 33. AS and AB integrations are allocated on chromosome 1 while AB and AE integrations are found on chromosome 6. All the possible combinations of the respective primers were examined. A result indicating a potential translocation was found between the following pairs of primers:

- Chromosome 6 forward primer with chromosome 1 forward primer

- Chromosome 1 forward primer with chromosome 6 reverse primer

The PCR results were repeated to exclude random finding and this sample of potential translocation will examined further with other techniques too (eg Sanger sequencing).

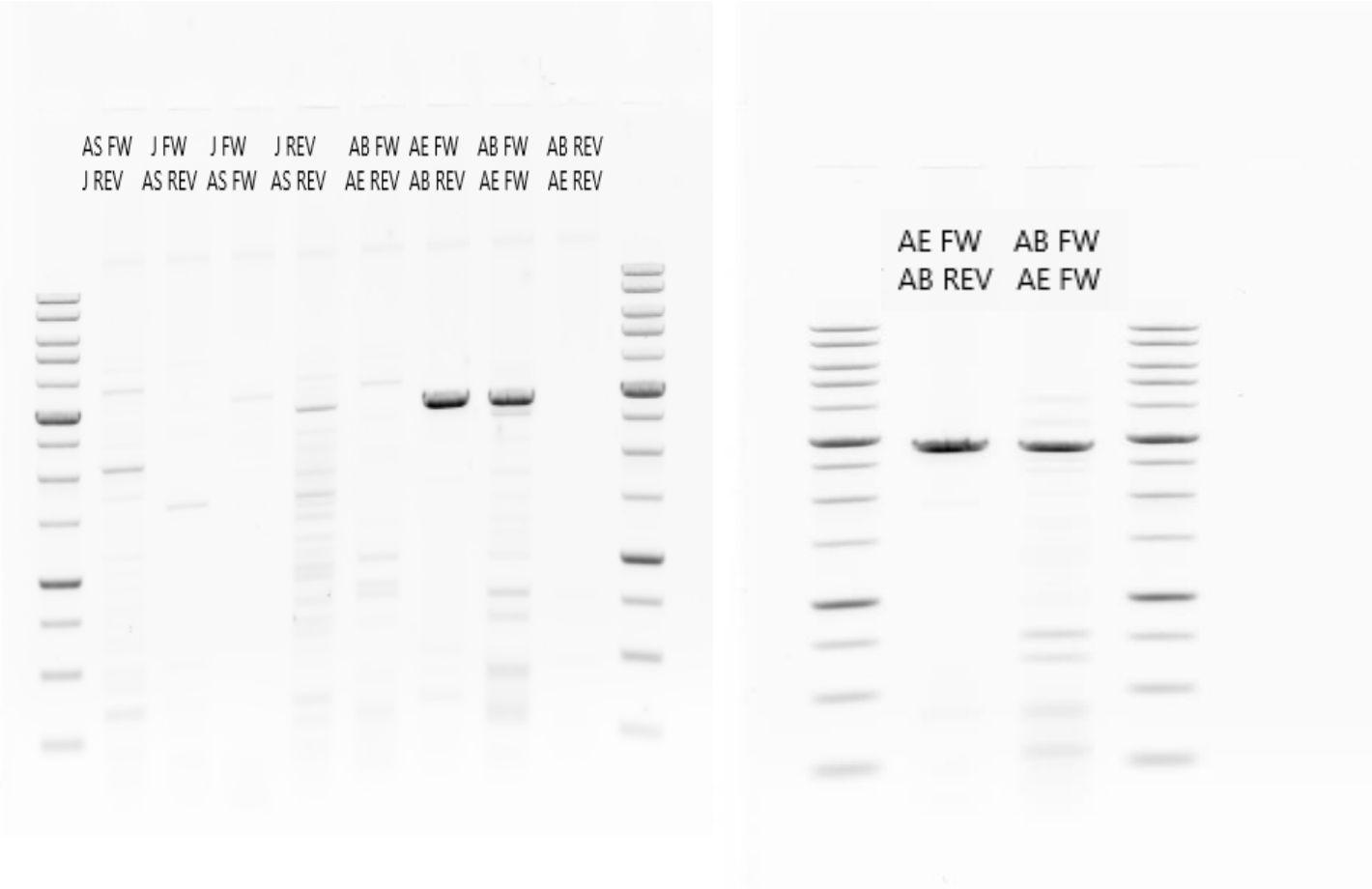


Figure 33:Results based on the proposed model with DNA induced with I-SceI DSBs that reveal a translocation between chromosome 1 and chromosome 6.

Identification and localization of I-*Sce*I sites in human clones

Concerning the human clones analysis through sequencing data from the platform MinION, two clones were analyzed RPE1-2xS100.R11 and RPE1-2xS200.R12 (see Figure 6 for further details on this clones). Humans have 46 chromosomes (22 pairs of autosomes and 1 pair sex-chromosomes).The human genome is very well mapped and documented and this offers a lot of possibilities during the analysis. The human genome has a size of approximately 3.23 Gbases.

RPE1-2xS100.R11 clones analysis

In RPE1-2xS100.R11 6 sequencing experiments from MinION were analyzed. 85 reads were analyzed in total resulting in 22 integration sites while with Southern blot only 11 were detected. Analysis of the last experiment is presented below

6 integration hits were found from BBduk from which 2 were found to be false positives and 4 from Blast tool where one was false positive. 1 common read was identified between the two searches.so in total we have 6 new reads containing the Integrations, two of them are really small reads with even smaller Small flanking sequences of two reads that matched to various chromosomes lead to exclusion of these reads from the analysis.

The local blast alignment gave the following results:
>blastn -db Blast/DB/HS/HSRefgenome -query Blast/Query/E715_joined.fasta -out HumanGenomeVSE715.out -evalue 1e-100 -culling_limit 3

>blastn -db Blast/DB/HS/HSRefgenome -query *Blast/Query/E715_Hit8_right_flanking_sequence_only.fasta -out HumanGenomeVSE715_Hit8_right_flanking_sequence_only.out -evalue 1e-100*

The analysis on the Human genome browser revealed 3 sites in intergenic region, one in an exon and one intronic. Next step was the inter-experimental analysis and comparison with the previous sequencing experiments.

>C:\Users\Christine Vasileiou\Desktop\Blast>makeblastdb -in DB\ allhitscontainingtheintegrationjoinedE671_E694_E668_E666_E664_E715.fasta -parse_seqids -dbtype nucl -max_file_sz 2500000000B

Buildlding a new DB, current time: 03/19/2019 15:50:45

New DB name: C:\Users\Christine Vasileiou\Desktop\Blast\DB\allhitscontainingtheintegrationjoinedE671_E694_E668_E666_E664_E715.fasta

New DB title: DB\allhitscontainingtheintegrationjoinedE671_E694_E668_E666_E664_E715.fasta

Sequence type: Nucleotide

Keep MBits: T

Maximum file size: 1000000000B

Adding sequences from FASTA; added 64 sequences in 0.0306132 seconds.

Followed by the intra-allignment of the reads

>blastn –db Blast/DB/allhitscontainingtheintegrationjoined_RPE_2xS100_E671_E694_E668_E666_E664_E715.fasta -query Blast/Query/allhitscontainingtheintegrationjoined_RPE_2xS100_E671_E694_E668_E666_E664_E715.fasta -out IntrarunuptoE715_allhitscontainingtheintegrationjoined_RPE_2xS100_E671_E694_E668_E666_E664_E715.fasta.out -evalue 1e-100

Results are shown in the following table:

Int.No.	Chr.	Location	gene	type	Description
1	13	q12.1	MICU2	Protein coding intronic	MICU2=mitochondrial calcium uptake
2	12	p11.2	FAR2	Protein coding intronic	FAR2=fatty acyl-CoA reductase 2
3	8	q22.1	RAD54B	Protein coding intronic	RAD54 homolog B
4	21	q22.1	no	n.a	
5	4	p15.3	no	lncRNA	
6	19	q11	ENST00000592806.1	miscRNA/Pseudogene	Centromeric Region
7	9	p22/p21	SLC24A2	Protein coding intronic	sodium/potassium/calcium exchanger 2 isoform 1
8	3	q22	CPNE4	Protein coding intronic	Copine 4; 16 Exons; zwischen Exons 14 und 15
9	9	q31	no	n.a	
10	1	p21	no	n.a	
11	5	q31.1	C5orf66	putative Protein intronic	putative uncharacterized protein; 6 Exons
12	19	q12	LOC100420587	Pseudogene intronic	
13	10	q22.1	no	n.a	
14	6	p22.2	no	n.a	
15	1	p.36.1	LOC105376819/3'UTR?		
16	1	p.33/p32.2	no	n.a	
17	2	q.34	VWC2L		Von Willebrand Factor C Domain Containing Protein 2 Like
18	5	q.31.3/q.32	no	n.a	
19	7	p.12	TNS3	Protein coding intronic	Tensin 3
20	12	q.22?	no	n.a	
21	13	q.13	no	n.a	
22	21	q.13.1	no	n.a	

RPE1-2xS200.R12 clones analysis

In RPE1-2xS200.R12 2 sequencing experiments from MinION were analyzed. 34 reads were analyzed in total resulting in 16 integration sites while with Southern blot only 12 were detected. Results of this analysis of all the experiments are presented below:

Integration.No.	Chr.	Location	Protein coding gene	type	Description
1	2	q37.3	no	intergenic	
2	20	q11.2	CBFA2T2	intronic	CBFA2/RUNX1 translocation partner 2; Plays a Role in AML [t(8;21)(q22;q22)]
3	6	q13	no	intergenic	
4	6	q13	no	intergenic	
5	X	q22.1	no	intergenic	gene-duplication of t-complex protein 11 (TCP11)
6	7	q34	KDM7A	intronic	lysine-specific demethylase 7A
7	10	q24.3	GBF1	intronic	Golgi-specific brefeldin A-resistance guanine nucleotide EF
8	1	q32.1	CACNA1S	intronic	voltage-dependent L-type calcium channel subunit alpha-1S
9	9	q22.2	no	intergenic	
10	7	q36	no	intergenic	
11	14	q12	no	intergenic	
12	7	p21	ISPD	intronic	isoprenoid synthase domain containing
13	7	q36	no	intergenic	
14	X	p22.3	PRKX	intronic	cAMP-dependent protein kinase catalytic subunit
15	7	p.14	NT5C3A	intronic /5'UTR?	NT5C3A (5'-Nucleotidase, Cytosolic IIIA)
16	19	q.13.4	no	intergenic	

Discussion

1. *Regulation of checkpoint activation at low doses by ATM and ATR.*

ATM and ATR are known to be the main kinases activated in response to double strand breaks and generally in response to DNA damage (Bensimon et al., 2011; Bhatti et al., 2011). Their function is mainly through cell cycle regulation depending mainly on the phase of the cell cycle that the cells undergo the damage (Beishline and Clifford, 2014). It is mainly reported that cell cycle arrest and strong G2/M checkpoint activation are predominantly due to ATM activation (Barnum and O'Connell, 2014; Falck et al., 2001; Khanna et al., 2001). Prompt activation of the G2/M checkpoint is evident following exposure of control samples of 82.6 cells at 2Gy and we observe a significant drop in the Mitotic Index which takes up to 4 hours to recover. We provide evidence that the functional interactions of ATM and ATR go way further than the traditional concept of their activation (Blackford and Jackson, 2017; Lavin et al., 1995). Our results strongly suggest that ATM is responsible for strong checkpoint responses as it is already known but this kinase is not able to lead to complete abrogation of the checkpoint, showing us the way to examine more deeply other 'cooperators' in the inception of the checkpoint. We reveal a synergistic action with high degree of overlap between ATM and ATR in the implementation of the G2/M checkpoint. In particular the influential role of ATR in response to DSBs through checkpoint regulation is shown in contrast to the common conception that ATR is activated only when replication stress and ssDNA is present (Nam and Cortez, 2011; Smits et al., 2010). Results actually reveal novel findings like complete checkpoint abrogation in absence of ATR. To conclude we demonstrate a functional coupling between the two kinases in the regulation of the G2-checkpoint activation in cells sustaining small loads of DSBs through exposure to low IR doses.

2. *Connections to the cell cycle machinery through ATR to Chk1*

Downstream targets of the damage response kinases ATM and ATR have proven to be important in activation of cell cycle checkpoints, most notably, the checkpoint kinases, Chk1 and Chk2 (Busch et al., 2017; Demidova et al., 2009; Nam and Cortez, 2011). The main mechanism known so far is that ATM interphases with cell cycle through Chk2 and ATR interphases with cell cycle through Chk1, but the mechanisms require further elucidation (Chaturvedi et al., 1999; Tominaga et al., 1999). Here we present our results that clearly show that in low doses of ionizing radiation and thus small DSB loads the regulation of cell cycle through strong checkpoint responses is due to ATR –Chk1 signaling. Experimental data also show no Chk2 entanglement in checkpoint activation indicating the need to examine further kinases that may act as downstream signaling targets of ATM considering the role of ATM in checkpoint regulation. Combined inhibition of both kinases seem to slightly increase the suppression of the G2 checkpoint as described for Chk1 inhibition alone. Similar results are

repeated and validated for A549 cells too. Our results point to the direction that some yet unidentified kinases that complement the ATR-Chk1 signaling should be examined in future projects to completely unveil the ATR/ATM signaling to cell cycle regulation.

3. IR-dose-dependent functional adaptation of Chk1.

At high doses cells are observed to react in a completely different way compared to that of low doses. In particular a strong checkpoint activation is observed and no recovery is shown even 8 hours post irradiation (Hoeijmakers, 2001; Hoekstra, 1997). Similar alterations in response are expected to be found for the ATM and ATR kinases activation too. Indeed, at high IR-doses that induce high DSB-numbers in the genome, the tight ATM/ATR coupling relaxes and independent outputs to G2-checkpoint. Cells have limited checkpoint response in the combined absence of ATM and ATR activities even after exposure to high doses of Ionizing Radiation (Mladenov et al., 2019). Here we take these results one step further by examining once again the downstream kinase Chk1 and Chk2. Similarly to the response to low doses Chk2 fails to generate any detectable effects on checkpoint regulation. On the other hand we observe a completely different behavior of the downstream kinase Chk1 comparing to low doses of ionizing radiation indicating a dose-dependent adaptation on cell's behalf. In particular complete checkpoint abrogation of the G2 checkpoint is observed in the absence of Chk1 kinase suggesting that mainly the cell cycle regulation is controlled by this kinase downstream ATM and ATR activation. These dose dependent adaptations make the necessity, to study more dose ranges more epistatically, obvious and at the same time pave the way to take into consideration in similar studies the dose-alterated DDR responses.

4. Characterization of clones on model system for simple and clustered DSBs

We tried to elucidate the type of lesion that high-LET IR induces using an established I-*Sce*I based model systems in CHO and human cells (Mladenova et al., 2016). One of the major advantages of these systems is that DSBs are exclusively induced in contrast to the plethora of lesion types induced by IR. In addition to the control over the type of lesion the spacing between consecutive DSBs can also be defined and studied. We therefore examined those clones by sequencing on MinION. Bioinformatics analysis revealed the integration of the I-SceI recognition sequence in those clones as expected. In the majority of the filtered reads the vector from the SB-Transposon plasmid was found completely intact and integrated in various locations of both CHO and human genome. The location of these integrations and their spatial proximity was also determined revealing interesting results. There is an obvious bias towards some DNA regions indicating either inclined selection inherent for the Sleeping Beauty transposase system or on the other hand it could indicate structures that are formed and 'protect' other potential region candidates, e.g. chromatin condensation ,during the expression

of the transposase. There might also be that there is a difficulty in the detection of some sites using sequencing techniques as those integrations sites are mainly located on repetitive regions which rises some questions on the accuracy of sequencing on such regions.

Results also reveal in some cases loss of the intervening sequence between the two I-S*ce*I sites and in some cases the loss was accompanied by loss of both I-SceI sites at the same time and other deletions both upstream and downstream of these sites. Such losses are also reported in other studies for two homing endonuclease sites when excised simultaneously (Boubakour-Azzouz and Ricchetti, 2008).

The results from this analysis open the way to a lot of opportunities to study damage and repair of DSBs on well documented clones and in a system that excludes the contribution of other lesions combined with the numerous advantages of Nanopore sequencing. These novel technologies that are available promise to revolutionize the ways we address fundamental questions of radiation damage and the associated radiation responses. Other potential applications can also be analyzed like for example the features of the CHO-4x I-*Sce*I cells also make them a potential candidate as a model system for chromothripsis that arises during cancer development. Chromothripsis is the accumulation of locally clustered DNA rearrangements with known losses of sequence fragments. (Forment, Kaidi et al. 2012; Korbel and Campbell 2013). Long read sequencing like MinION and the I-*Sce*I model described here make it an ideal candidate for such kind analysis.

5. *The number of Integration sites defined by sequencing is clearly elevated compared to previous techniques*

We present here a table that summarizes the number of Integration found by Southern Blot on the related clones compared to the number of Integrations that we identified and localized for the same clones.

Table 8: Comparative illustration of Integrations found with different techniques Southern Blot (column 4) and Nanopore sequencing (column 5)

Clone	Cell origin	DSBs per cluster	Integrations found by Southern Blot	Integrations found by Nanopore sequencing	Ratio
CHO-2xS.R14	Chinese Hamster Ovary	2	14	31	2.2
CHO-4xS.R12	Chinese Hamster Ovary	4	12	35	2.9
RPE1-2xS100.R11	Human Retinal Epithelial	2	11	22	2
RPE1-2xS200.R12	Human Retinal Epithelial	2	12	34	2.83

There is an evident elevated number of identified integration sites using Nanopore sequencing on all clones. This trend could be explained by the known higher resolution that is achieved in these analysis results using sequencing. Higher accuracy when studying in sequence level is expected. It is worth to mention that in some cases different integration sites were found to be in close proximity. This addresses the fact with this high resolution technique we were able to discriminate between integration sites that were otherwise found to be the same due to intrinsic restraints of other techniques. Other possible explanations of this phenomenon, that we are currently investigating, are the possibility of some kind of mix-up in clones during their preparation. Furthermore complicated rearrangements due to duplications or "jumping" transposons could also clarify these results. It is worth mentioning that some integrations were only found in one experiment represented by one read but with great confidence. It therefore becomes clear that some integrations are more easily detected than others. This mechanism in addition to the fact that on some chromosomes no integration sites were located need further elucidation.

6. The probability for misrepair of clustered DSBs is clearly elevated

As discussed previously we examined on induced clones the effects of DSBs after their repair. A quick comparison concerning the biological consequences between clones with two single DSBs in close proximity and clones with DSB-clusters reveals that processing of the latter during repair leads in increase number of deletions. This becomes obvious both by the PCR results using primers that target the specific construct but also it is evident during the manual handling of these sequences. These effects are observed both on CHO and human clones suggesting that DSB- clusters are more difficult to process during DNA repair.

7. *Clustered DSBs lead to translocation formation in both CHO and human clones.*

Our model to examine structural variants on DNA gave some promising results. It provides the possibility to assess various kinds of structural variations on DNA which are currently poorly assayed using dominant short read sequencing technologies. Indeed translocation formation was found in clones that were carrying constructs connected to clustered DSBs, linking this type of damage to DSBs. This indicates the severity of such lesions and the difficulties that arise in their repair. Other genomic rearrangements should now be examined using a combination of these two techniques and further in depth analysis should be carried out in the other clones as well. These results should then be compared to high-LET irradiation data and this way we can possibly proceed to identifying and isolating the effects of DSBs separated from other confounding factors. Furthermore existing data should now be expanded and analyses from different sequencing experiments should be integrated. The optimization of the proposed models should be taken to the next step and results from different sequencing experiments analyzed and compared to each other. It is important to increase the sensitivity and efficiency of detection of these random transgenic integrations as this will be important in the analysis of the genomic alterations generated during the processing of the DSBs of different complexity and the analysis of translocations needs to be refined and key translocations confirmed with different techniques including PCR and alternative sequencing methods. The results need to be carefully analyzed and consolidated to reach conclusions regarding the genomic stability of the affected cells.

Table of figures

Abbreviations:

TEs: Transposable elements

ONT: Oxford Nanopore Technology

TGS: Third Generation Sequencing

WGS: Whole Genome Sequencing

SV: Structural Variation

ssDNA : single stranded DNA

LET: Linear Energy Transfer

DSB: Double Strand Break

DDR: DNA damage response

ATM: Ataxia telangiectasia mutated

ATR: Ataxia telangiectasia RAD3-related

HR: Homologous Recombination.

CHK1: Checkpoint kinase 1

CHK2: Checkpoint kinase 2

BSA: Bovine serum albumin

DMEM: Dulbecco's Modified Eagle Medium

DMSO: Dimethyl sulfoxide

DNA-PKcs: DNA-dependent protein kinase catalytic subunit

FACS: Fluorescence Activated Cell Sorting

FBS: Fetal Bovine Serum

Gy: gray

IR: Ionizing radiation

MQ: Milli-Q

H3-pS10: Phosphorylation of Histone H3 at Serine 10

KU-55933 KUDOS55933 (2-morpholin-4-yl-6-thianthren -1-yl-pyran-4-one)

MI Mitotic index

MPF Mitosis promotion factor

NHEJ: Nonhomologous end-joining

NU-7026:2-(Morpholin-4-yl)-benzo[h]chomen-4-one

PBG: PBS, BSA, Gelatine

EtBr :Ethidium Bromide

eV: Electron Volt

kDa; Kilodalton

keV: Kilo electron volt

KV: Kilo Volt

CHO: Chinese hamster ovary

PBS: Phosphate Buffered Saline

PBST: PBS with Tween 20

PcAb: Polyclonal antibody

PCC Prematurely condensed chromosomes

PI: Propidium iodine

PI-3Ks Phosphoinositide-3 lipid family of kinases

RPM: Rounds per minute

RT: Room temperature

UCN-01: 7-hydroxystaurosporine

SB: Sleeping Beauty

Cdk: cyclin-dependent protein kinase

(M)bp(s): (Mega) base pair (s)

cDNA : Complementary DNA

dNTP: Deoxynucleoside triphosphate

gDNA: Genomic DNA

HGP: Human Genome Project

NGS: Next-generation sequencing

PacBio: Pacific Biosciences

PCR: Polymerase chain reaction

RNA: Ribonucleic acid

SD: Standard deviation

Appendix 1

Flanking sequences used for I-SceI site localization for all relevant clones

CAGTTGAAGTCGGAAGTTTACATACACCTTAGCCAAATACATTTAAACTCACTTTTTCACAATTCCTGACATTT
AATCCTAGTAAAAATTCCCTGTCTTAGGTCAGTTAGGATCACCACTTTATTTTAAGAATGTGAAATATCAGAA
TAATAGTAGAGAGAATGATTCATTTCAGCTTTTATTTCTTTCATCACATTCCCAGTGGGTCAGAAGTTTACATA
CACTCAATTAGTATTTGGTAGCATTGCCTTTAAATTGTTTAACTTGGGTCAAACATTTCGAGTAGCCTTCCACA
AGCTTCTGATAGACTAATTGACATCATTTGAGTCAATTGGAGGTGTACCTGTGGATGTATTTCAAGGGATCCA
GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTG
AAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTGGGGTGGGCGAAGAACT
CCAGCATGAGATCCCCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCA
TAGAAGGCGGCGGTGGAATCGAAATCTCGTGATGGCAGGTTGGGCGTCGCTTGGTCGGTCATTTCGAACCCCAG
AGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTA
AAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCT
GATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTC
GGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTCGCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAG
TTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACG
TGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCA
TTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCG
CCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCGTGGCC
AGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAAC
CGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAGC
CGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCT
CATCCTGTCTCTTGATCAGATCCGAAAATGGATATCCAAGCTCCCGGGAGCTTTTTGCAAAAGCCTAGGCCTCC
AAAAAAGCCTCCTCACTACTTCTGGAATAGCTCAGAGGCCGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAAA
TTAGTCAGCCATGGGGCGGAGAATGGGCGGAACTGGGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGGCGG
GACTATGGTTGCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCAC
ACCTGGTTGCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACAC
CCTAACTGACACACATTCCACAGTAGGGATAACAGGGTAATAGCTCTACGCTTGGCAATCTACCCGCGAGGCTC
TCTTGAATTCTAGGTCAGAGACGGCACATTAGGCTCCCACCGGTGCTGTTAGCTTCCATACCAGCTGATTACCC
TGTTATCCCTACCATCACAAAGCTCTGACCTCAATCCTATAGAAAGGAGGAATGAGCCAAAATTCACCCAACTT
ATTGTGGGAAGCTTCTAAAGCCATGACATCATTTTCTGGAATTTTCCAAGCTGTTTAAAGGCACAGTCAACTTA
GTGTATGTAAACTTCTGACCCACTGGAATTGTGATACAGTGAATTATAAGTGAAATAATCTGTCTGTAAACAA
TTGTTGGAAAAATGACTTGTGTCATGCACAAAGTAGATGTCCTAACTGACTTGCCAAAACTATTGTTTGTTAAC
AAGAAATTTGTGGAGTAGTTGAAAAACGAGTTTTAATGACTCCAACTTAAGTGTATGTAAACTTCCGACTTCA
ACTGTATA-

Figure 34: Integrating part of vector RPE 2xS100-inversed

acagttgaagtcggaagtttacatacacttaagttggagtcattaaaactcgtttttcaactactccacaaatttcttgttaacaaacaatagt
tttggcaagtcagttaggacatctactttgtgcatgacacaagtcatttttccaacaattgtttacagacagattatttcacttataattcactgt
atcacaattccagtgggtcagaagtttacatacactaagttgactgtgcctttaaacagcttggaaaattccagaaaatgatgtcatggcttt
agaagcttcccacaataagttgggtgaattttggctcattcctcctttctataggattgaggtcagagctttgtgatggaattctacccgcgga
gatctcccaaccctgaatagggataacagggtaatatcattttcggtagcaagtttaagcatcaattctaggatgaaggagggaatttgtaa
atcttcagcccattgcagatgacgcggccgcgatactagtgacaagcttgatggatcctgagatatctacggcagaccgcaattaaagacct
tgctgtgtcttcctgacttgtaaggagtccacatacgatggtgcttgtctaattaccctgttatccctaaaccctgaacccctcgagtctagaccc
gaattctgtggaatgtgtgtcagttagggtgtggaaagtccccaggctccccaggcaggcagaagtatgcaaagcatgcatctcaattagt
cagcaaccaggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccatagtcccgc
ccctaactccgcccatcccgcccctaactccgcccagttccgcccattctccgccccatggctgactaatttttttatttatgcagaggccgagg
ccgcctcggcctctgagctattccagaagtagtgaggaggctttttggaggcctaggctttgcaaaaagctcccgggagcttggatatcca

ttttcggatctgatcaagagacaggatgaggatcgtttcgcatgattgaacaagatggattgcacgcaggttctccggccgcttgggtggag
aggctattcggctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttcttttt
gtcaagaccgacctgtccggtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcag
ctgtgctcgacgttgtcactgaagcgggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctg
ccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccaccaagcgaaacatcgc
atcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagagcatcaggggctcgcgccagccgaactg
ttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacccatggcgatgcctgcttgccgaatatcatggtggaaaatg
gccgcttttctggattcatcgactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctacccgtgatattgctgaagagc
ttggcggcgaatgggctgaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcatcgccttctatcgccttcttgacgagttct
tctgagcgggactctggggttcgaaatgaccgaccaagcgacgcccaacctgccatcacgagatttcgattccaccgccgccttctatgaaa
ggttgggcttcggaatcgttttccgggacgccggctggatgatcctccagcgcggggatctcatgctggagttcttcgcccaccccaacttgtt
tattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtcca
aactcatcaatgtatcttatcatgtctggatcccttgaaatacatccacaggtacacctccaattgactcaaatgatgtcaattagtctatcaga
agcttgtggaaggctactcgaaatgtttgacccaagttaaacaatttaaaggcaatgctaccaaatactaattgagtgtatgtaaacttctg
acccactgggaatgtgatgaaagaaataaaagctgaaatgaatcattctctctactattattctgatatttcacattcttaaaataaagtggt
gatcctaactgacctaagacagggaatttttactaggattaaatgtcaggaattgtgaaaaagtgagtttaaatgtatttggctaaggtgta
tgtaaacttcc

Figure 35: Integrating part of vector RPE 2xS200-inversed

tatacagttgaagtcggaagtttacatacaccttagccaaatacatttaaactcacttttcacaattcctgacatttaatcctagtaaaaattc
cctgtcttaggtcagttaggatcaccactttattttaagaatgtgaaatatcagaataatagtagagagaatgattcatttcagcttttatttctt
tcatcacattcccagtgggtcagaagtttacatacactcaattagtatttggtagcattgcctttaaattgtttaacttgggtcaaacatttcga
gtagccttccacaagcttctgatagactaattgacatcatttgagtcaattggaggtgtacctgtggatgtatttcaagggatccagacatgat
aagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgta
accattataagctgcaataaacaagttggggtgggcgaagaactccagcatgagatccccgcgctggaggatcatccagccggcgtcccg
gaaaacgattccgaagcccaacctttcatagaaggcggcggtggaatcgaaatctcgtgatggcaggttgggcgtcgcttggtcggtcattt
cgaaccccagagtcccgctcagaagaactcgtcaagaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgataccgtaaagc
acgaggaagcggtcagcccattcgccgccaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcggtccgccacaccca
gccggccacagtcgatgaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctc
gccgtcgggcatgcgcgccttgagcctggcgaacagttcggctggcgcgagcccctgatgctcttcgtccagatcatcctgatcgacaagac
cggcttccatccgagtacgtgctcgctcgatgcgatgtttcgcttggtggtcgaatgggcaggtagccggatcaagcgtatgcagccgccgc
attgcatcagccatgatggatactttctcggcaggagcaaggtgagatgacaggagatcctgccccggcacttcgcccaatagcagccagt
cccttcccgcttcagtgacaacgtcgagcacagctgcgcaaggaacgcccgtcgtggccagccacgatagccgcgctgcctcgtcctgcagt
tcattcagggcaccggacaggtcggtcttgacaaaaagaaccgggcgcccctgcgctgacagccggaacacggcggcatcagagcagcc
gattgtctgttgtgcccagtcatagccgaatagcctctccacccaagcggccggagaacctgcgtgcaatccatcttgttcaatcatgcgaaa
cgatcctcatcctgtctcttgatcagatccgaaaatggatatccaagctcccgggagcttttgcaaaagcctaggcctccaaaaaagcctcct
cactacttctggaatagctcagaggccgaggcggcctcggcctctgcataaataaaaaaaattagtcagccatggggcggagaatgggcg
gaactgggcggagttaggggcgggatgggcggagttaggggcgggactatggttgctGACTAATTGAgatgcatgctttgcatacTG
CCTGCTGGGGagcctggggactttccacacctggttgctgactaattgagatgcatgctttgcatacttctgcctgcctggggagcctggg
gactttccacaccctaactgacacacattccacagaattCGGGTCTAGACTCGAGGGGTTCAGGGTTTAGGGATAACAGGGT
AATTAGGCAAGCACCATCGTATGTGGACTCCTTACAAGTCAGGAAGACACAGCAAGGTCTTTAatTGCGGTCTG
CCGTAGATATCTCAGGATCCATCAAGCTTGTCACTAGTATCGCGGCCGCGTCATCTGCAATGGGCTGAAGATTT
ACAAATTCCCTCCTTCATCCTAGAATTGATGctTAAACTTGCTACCGAAAatgATATTACCCTGTTATCCCTATT
CAGGGTTGGGAAGATCTCCGCGGGTGAGAGCTCGGGTCTACAGACTCGAGGGTTCAGGAGTTTAGGGATAACA
GGGTAATGAACAAGCACCATCGTATGTGGACTCCTtacaAGTCAGGAAGACACAGCAAGGTCTTTAaTTGCGGTC
TGCCGTAgATATCTCAGGATCCaTCAAGCTTGTCACTAGTATCGCGGCCGCGTCATCtgcAaTGGGCtGAAGATT
TACAAATTCCCTCCTTCATCCTAGAATTGATGTAAACTTGCTaCCGAAAATGATATTACCCTGTTATCCCTAccc
aGGGtTGGGAGAATTCcacaaagctctgacctcaatcctatagaaaggaggaatgagccaaaattcacccaacttattgtgggaagct
tctaaagccatgacatcattttctggaattttccaagctgtttaaaggcacagtcaacttagtgtatgtaaacttctgacccactggaattgtga

tacagtgaattataagtgaaataatctgtctgtaaacaattgttggaaaaatgacttgtgtcatgcacaaagtagatgtcctaactgacttgc
caaaactattgtttgttaacaagaaatttgtggagtagttgaaaaacgagttttaatgactccaacttaagtgtatgtaaacttccgacttca
actgtata

Figure 36: Vector for CHO-4xS.R12

Appendix 2

Name	Sequence	Product length	Vector orientation
CHO4x-AA-FW	CCCTCAAAGGCTAGCCAACA	1015	north
CHO4x-AA-REV	GCAGACAGGCAGGAAGTGAT		south
CHO4x-AB-FW1	ATGTCAGGCCCTGGAAATGG	442	north
CHO4x-AB-REV	CCAGTCCACCTTTGCGGTAT		south
CHO4x-AE-FW	GGGCTCAGTGGACAAAATGC	895	north
CHO4x-AE-REV	AGGCTATCATGCTTGGGGTG		south
CHO4x-AF-FW	TGGAGCTGGAGGTACAGACA	3718	-
CHO4x-AF-REV	CCAGTCCACCTTTGCGGTAT		-
CHO4x-AH-FW	GCAATCTCTATCTCACTGGCCT	1308	south
CHO4x-AH-REV	AAGGCGAAGCCCTGATTGTT		north
CHO4x-AI-FW	ACCTCTTGTCATCTGCAGCC	2424	north
CHO4x-AI-REV	GGCTCAGTTGCAGCCTTCTA		south
CHO4x-AK-FW	TTTCAGGAGTGCCCTGGTTC	1503	north
CHO4x-AK-REV	TTTCTTGGGTGACCGTGGTT		south
CHO4x-AL-FW	CACCAGCTGGGCAAGACATA	1111	south
CHO4x-AL-REV	TGTTTCAAATGCCCAAGC		north
CHO4x-AS-FW	GGTTGCCTCCCACCACTAAA	1272	south
CHO4x-AS-REV	GCCCACTGGTCTACTTTCCC		north
CHO4x-B-FW	TTCAACACACCCAAGGCAGA	1245	south
CHO4x-B-REV	ACTGCAGTTCCCAGGATTGG		north
CHO4x-C-FW1	TCTCAGGATGGGGTTCAGAGA	1227	north
CHO4x-C-REV1	CCCTCCCCCAACACTCTTTC		south
CHO4x-E-FW1	TTACCCTGCAGGATGGGTTG	956	north
CHO4x-E-REV1	CTTGTTGCCTGTCCACAGAGT		south
CHO4x-F-FW	CAGCCAGGCTAGAACACCAA	1724	south
CHO4x-F-REV	GGGAACCTGAACCAGAAGCA		north
CHO4x-H-FW	ACACCAGAGCCAAGTTAGCC	1222	south

CHO4x-H-REV1	CCTCCTCCTCCCTGTTTTCAG		north
CHO4x-I-FW1	TGATCCCCCACAACAAAGCA	1183	north
CHO4x-I-REV1	TGGCTGGGGTACATCAACAG		south
CHO4x-J-FW	CCCCTTTTCACGTAAGTGCC	1371	north
CHO4x-J-REV	GGCACATGTCCCATACCCAT		south
CHO4x-K-FW	GGGAGGGACTGCTAGAAGGA	987	north
CHO4x-K-REV	CCCTTCCCTGGTCTGACAAC		south
CHO4x-N-FW	GGCTGTCCAGACTAGCCAAA	1147	north
CHO4x-N-REV1	AAGGTGGTAGGCAAGCACTC		south
CHO4x-O-FW	GCCAGCATGGCAGTTTTAGG	1173	north
CHO4x-O-REV	CTTCGGACGGCTGTTGAAAA		south
CHO4x-P-FW	TGTCCCTGTGGATTTCAGCC	1162	south
CHO4x-P-REV	CCACCAGTGAGAAGCAAGGA		north
CHO4x-R-FW	TGTGTTGGTGGAGGTTGGAG	1115	north
CHO4x-R-REV	GGGCATCAGACACCCTGAAA		south
CHO4x-X-FW1	AGGTACAGGTGGTTGTTGCC	1396	north
CHO4x-X-REV	CAGTGCCTGTTCTCAAATGGC		south
CHO4x-Y-FW	TTGAGTACCAGAAGCTCGGC	1696	south
CHO4x-Y-REV	TCCTCAGCTCTGGCATTTCA		north
CHO4x-Z-FW	GGCAAATCTCCAGGACACCA	1269	south
CHO4x-Z-REV	TGTGTCTTGTGAGTTGCTGC		north

Table 9: List of Primers

Bibliography

Aronovich, E. L., McIvor, R. S., and Hackett, P. B. (2011). The Sleeping Beauty transposon system: a non-viral vector for gene therapy. Human molecular genetics *20*, R14-20.

Barnum, K. J., and O'Connell, M. J. (2014). Cell cycle regulation by checkpoints. Methods in molecular biology *1170*, 29-40.

Bartek, J., Falck, J., and Lukas, J. (2001). CHK2 kinase--a busy messenger. Nature reviews Molecular cell biology *2*, 877-886.

Bartek, J., and Lukas, J. (2001). Mammalian G1- and S-phase checkpoints in response to DNA damage. Current opinion in cell biology *13*, 738-747.

Beishline, K., and Clifford, J. (2014). Interplay Between the Cell Cycle and Double-Strand Break Response in Mammalian Cells, Vol 1170).

Belfort, M., and Roberts, R. J. (1997). Homing endonucleases: keeping the house in order. Nucleic acids research *25*, 3379-3388.

Bensimon, A., Aebersold, R., and Shiloh, Y. (2011). Beyond ATM: the protein kinase landscape of the DNA damage response. FEBS letters *585*, 1625-1639.

Bhatti, S., Kozlov, S., Farooqi, A. A., Naqi, A., Lavin, M., and Khanna, K. K. (2011). ATM protein kinase: the linchpin of cellular defenses to stress. Cellular and molecular life sciences : CMLS *68*, 2977-3006.

Bindra, R. S., Goglia, A. G., Jasin, M., and Powell, S. N. (2013). Development of an assay to measure mutagenic non-homologous end-joining repair activity in mammalian cells. Nucleic acids research *41*, e115.

Blackford, A. N., and Jackson, S. P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. Molecular cell *66*, 801-817.

Bohgaki, T., Bohgaki, M., and Hakem, R. (2010). DNA double-strand break signaling and human disorders. Genome integrity *1*, 15.

Boubakour-Azzouz, I., and Ricchetti, M. (2008). Low joining efficiency and non-conservative repair of two distant double-strand breaks in mouse embryonic stem cells. DNA repair *7*, 149-161.

Busch, C. J., Kroger, M. S., Jensen, J., Kriegs, M., Gatzemeier, F., Petersen, C., Munscher, A., Rothkamm, K., and Rieckmann, T. (2017). G2-checkpoint targeting and radiosensitization of HPV/p16-positive HNSCC cells through the inhibition of Chk1 and Wee1. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology *122*, 260-266.

Campa, A., Ballarini, F., Belli, M., Cherubini, R., Dini, V., Esposito, G., Friedland, W., Gerardi, S., Molinelli, S., Ottolenghi, A.*, et al.* (2005). DNA DSB induced in human cells by charged particles and gamma rays: experimental results and theoretical approaches. International journal of radiation biology *81*, 841-854.

Chaturvedi, P., Eng, W. K., Zhu, Y., Mattern, M. R., Mishra, R., Hurle, M. R., Zhang, X., Annan, R. S., Lu, Q., Faucette, L. F.*, et al.* (1999). Mammalian Chk2 is a downstream effector of the ATM-dependent DNA damage checkpoint pathway. Oncogene *18*, 4047-4054.

Coordinators, N. R. (2013). Database resources of the National Center for Biotechnology Information. Nucleic acids research *41*, D8-D20.

De Coster, W., De Rijk, P., De Roeck, A., De Pooter, T., D'Hert, S., Strazisar, M., Sleegers, K., and Van Broeckhoven, C. (2019). Structural variants identified by Oxford Nanopore PromethION sequencing of the human genome. Genome research.

Demidova, A. R., Aau, M. Y., Zhuang, L., and Yu, Q. (2009). Dual regulation of Cdc25A by Chk1 and p53-ATF3 in DNA replication checkpoint control. The Journal of biological chemistry *284*, 4132-4139.

Dupuy, A. J., Jenkins, N. A., and Copeland, N. G. (2006). Sleeping beauty: a novel cancer gene discovery tool. Human molecular genetics *15 Spec No 1*, R75-79.

Duursma, A. M., and Cimprich, K. A. (2010). Checkpoint recovery after DNA damage: a rolling stop for CDKs. EMBO reports *11*, 411-412.
Falck, J., Mailand, N., Syljuasen, R. G., Bartek, J., and Lukas, J. (2001). The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. Nature *410*, 842-847.
Giard, D. J., Aaronson, S. A., Todaro, G. J., Arnstein, P., Kersey, J. H., Dosik, H., and Parks, W. P. (1973). In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. Journal of the National Cancer Institute *51*, 1417-1423.
Gong, L., Wong, C. H., Cheng, W. C., Tjong, H., Menghi, F., Ngan, C. Y., Liu, E. T., and Wei, C. L. (2018). Picky comprehensively detects high-resolution structural variants in nanopore long reads. Nature methods *15*, 455-460.
Goodhead, D. T. (2006). Energy deposition stochastics and track structure: what about the target? Radiation protection dosimetry *122*, 3-15.
Gupta, A., Yang, Q., Pandita, R. K., Hunt, C. R., Xiang, T., Misri, S., Zeng, S., Pagan, J., Jeffery, J., Puc, J.*, et al.* (2009). Cell cycle checkpoint defects contribute to genomic instability in PTEN deficient cells independent of DNA DSB repair. Cell cycle *8*, 2198-2210.
Hermanson, S., Davidson, A. E., Sivasubbu, S., Balciunas, D., and Ekker, S. C. (2004). Sleeping Beauty transposon for efficient gene delivery. Methods in cell biology *77*, 349-362.
Hoeijmakers, J. H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366-374.
Hoekstra, M. F. (1997). Responses to DNA damage and regulation of cell cycle checkpoints by the ATM protein kinase family. Current opinion in genetics & development *7*, 170-175.
Iliakis, G. (1997). Cell cycle regulation in irradiated and nonirradiated cells. Seminars in oncology *24*, 602-615.
Iliakis, G., Mladenova, V., Sharif, M., Chaudhary, S., Mavragani, I. V., Soni, A., Saha, J., Schipler, A., and Mladenov, E. (2018). Defined Biological Models of High-Let Radiation Lesions. Radiation protection dosimetry.
Iliakis, G., Murmann, T., and Soni, A. (2015). Alternative end-joining repair pathways are the ultimate backup for abrogated classical non-homologous end-joining and homologous recombination repair: Implications for the formation of chromosome translocations. Mutation research Genetic toxicology and environmental mutagenesis *793*, 166-175.
Iliakis, G., Wang, Y., Guan, J., and Wang, H. (2003). DNA damage checkpoint control in cells exposed to ionizing radiation. Oncogene *22*, 5834-5847.
Ivics, Z., Hackett, P. B., Plasterk, R. H., and Izsvak, Z. (1997). Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell *91*, 501-510.
Jackson, J. R., and Zhou, B. B. (2004). G2 checkpoint control: checking in to the last resort for DNA damage. Cancer biology & therapy *3*, 314-316.
Jackson, S. P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071-1078.
Jakob, B., Splinter, J., and Taucher-Scholz, G. (2009). Positional stability of damaged chromatin domains along radiation tracks in mammalian cells. Radiation research *171*, 405-418.
Jasin, M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. Trends in genetics : TIG *12*, 224-228.
Jou, I., and Muthukumar, M. (2017). Effects of Nanopore Charge Decorations on the Translocation Dynamics of DNA. Biophysical journal *113*, 1664-1672.
Khanna, K. K., and Jackson, S. P. (2001). DNA double-strand breaks: signaling, repair and the cancer connection. Nature genetics *27*, 247-254.
Khanna, K. K., Lavin, M. F., Jackson, S. P., and Mulhern, T. D. (2001). ATM, a central controller of cellular responses to DNA damage. Cell death and differentiation *8*, 1052-1065.

Lavin, M. F., Khanna, K. K., Beamish, H., Spring, K., Watters, D., and Shiloh, Y. (1995). Relationship of the ataxia-telangiectasia protein ATM to phosphoinositide 3-kinase. Trends in biochemical sciences *20*, 382-383.
Lu, H., Giordano, F., and Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. Genomics, proteomics & bioinformatics *14*, 265-279.
Lukas, C., Bartkova, J., Latella, L., Falck, J., Mailand, N., Schroeder, T., Sehested, M., Lukas, J., and Bartek, J. (2001). DNA damage-activated kinase Chk2 is independent of proliferation or differentiation yet correlates with tissue biology. Cancer research *61*, 4990-4993.
Maglia, G., Restrepo, M. R., Mikhailova, E., and Bayley, H. (2008). Enhanced translocation of single DNA molecules through alpha-hemolysin nanopores by manipulation of internal charge. Proceedings of the National Academy of Sciences of the United States of America *105*, 19720-19725.
Marechal, A., and Zou, L. (2013). DNA damage sensing by the ATM and ATR kinases. Cold Spring Harbor perspectives in biology *5*.
Mladenov, E., Fan, X., Dueva, R., Soni, A., and Iliakis, G. (2019). Radiation-dose-dependent functional synergisms between ATM, ATR and DNA-PKcs in checkpoint control and resection in G2-phase. Scientific reports *9*, 8255.
Mladenov, E., Magin, S., Soni, A., and Iliakis, G. (2016). DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. Seminars in cancer biology *37-38*, 51-64.
Mladenova, V., Mladenov, E., and Iliakis, G. (2016). Novel Biological Approaches for Testing the Contributions of Single DSBs and DSB Clusters to the Biological effects of High LET Radiation. Front Oncol *6*.
Nakanishi, M. (2001). [Regulation of cell cycle checkpoints in mammalian cells]. Seikagaku The Journal of Japanese Biochemical Society *73*, 343-350.
Nam, E. A., and Cortez, D. (2011). ATR signalling: more than meeting at the fork. The Biochemical journal *436*, 527-536.
Niemantsverdriet, M., van Goethem, M. J., Bron, R., Hogewerf, W., Brandenburg, S., Langendijk, J. A., van Luijk, P., and Coppes, R. P. (2012). High and low LET radiation differentially induce normal tissue damage signals. International journal of radiation oncology, biology, physics *83*, 1291-1297.
Nilsson, I., and Hoffmann, I. (2000). Cell cycle regulation by the Cdc25 phosphatase family. Progress in cell cycle research *4*, 107-114.
Norris, A. L., Workman, R. E., Fan, Y., Eshleman, J. R., and Timp, W. (2016). Nanopore sequencing detects structural variants in cancer. Cancer biology & therapy *17*, 246-253.
Patel, A., Belykh, E., Miller, E. J., George, L. L., Martirosyan, N. L., Byvaltsev, V. A., and Preul, M. C. (2018). MinION rapid sequencing: Review of potential applications in neurosurgery. Surgical neurology international *9*, 157.
Saha, J., Wang, S. Y., and Davis, A. J. (2017). Examining DNA Double-Strand Break Repair in a Cell Cycle-Dependent Manner. Methods in enzymology *591*, 97-118.
Seluanov, A., Mao, Z., and Gorbunova, V. (2010). Analysis of DNA double-strand break (DSB) repair in mammalian cells. Journal of visualized experiments : JoVE.
Smits, V. A., Warmerdam, D. O., Martin, Y., and Freire, R. (2010). Mechanisms of ATR-mediated checkpoint signalling. Frontiers in bioscience *15*, 840-853.
Solomon, M. J., and Kaldis, P. (1998). Regulation of CDKs by phosphorylation. Results and problems in cell differentiation *22*, 79-109.
Stark, G. R., and Taylor, W. R. (2004). Analyzing the G2/M checkpoint. Methods in molecular biology *280*, 51-82.

Tominaga, K., Morisaki, H., Kaneko, Y., Fujimoto, A., Tanaka, T., Ohtsubo, M., Hirai, M., Okayama, H., Ikeda, K., and Nakanishi, M. (1999). Role of human Cds1 (Chk2) kinase in DNA damage checkpoint and its regulation by p53. The Journal of biological chemistry *274*, 31463-31467.
van Gent, D. C., Hoeijmakers, J. H., and Kanaar, R. (2001). Chromosomal stability and the DNA double-stranded break connection. Nature reviews Genetics *2*, 196-206.
Wheeler, D. L., Chappey, C., Lash, A. E., Leipe, D. D., Madden, T. L., Schuler, G. D., Tatusova, T. A., and Rapp, B. A. (2000). Database resources of the National Center for Biotechnology Information. Nucleic acids research *28*, 10-14.
Yang, J., Yu, Y., Hamrick, H. E., and Duerksen-Hughes, P. J. (2003). ATM, ATR and DNA-PK: initiators of the cellular genotoxic stress responses. Carcinogenesis *24*, 1571-1580.
Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics *13*, 134.