Development of a novel translocation reporter system

Dissertation

zur Erlangung des Grades

"Doktor der Naturwissenschaften"

am Fachbereich Biologie

der Johannes Gutenberg-Universität Mainz

Marta Cipińska
geboren am 20.01.1989 in Pabianice

JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Mainz, 2020

Abstract

Genome integrity is vital for the propagation of genetic information. However, genomes are under the constant exposure to DNA damaging factors. One of the most deleterious types of DNA damage are DSBs. Illegitimate repair of broken chromosomes may result in the formation of chromosome rearrangements, including chromosome translocations. Their occurrence may lead to the deregulation of gene expression or activation of oncogenes that can be causal in a variety of cancers. The formation of chromosome translocations is a multistep process that involves the chromosome breakage, chromosome-end synapsis and misrepair to form fusions. However, still little is known about the factors that contribute to those individual steps. Due to their rare occurrence, probing and quantifying chromosome translocations is difficult. Current tools, such as FISH and PCR are laborious and not sensitive enough for such low-frequency events. The major focus of the field has been to understand the repair mechanisms and studying the characteristics of individual factors. The purpose of this study was to develop the system that would enable the identification of novel factors involved in the formation of chromosome translocations through performing a genome-wide screening. Created translocation reporter assay is based on the CRISPR/Cas9 technology for inducing DSBs, which is a new golden standard for creating translocation relevant cancer models. In addition, the custom-made selection cassette was integrated into the genome of the haploid cell line. It contains the antibiotic resistance gene, which expression is activated only upon translocation. Therefore, the main advantage of the system is that it enables the selection and enrichment of cells positive for translocations. Our results show that the system is suitable for the generation of different types of chromosome rearrangements, such as deletions, inversions and translocations. It is also compatible with PCR based methods, which allows quantification of frequencies of occurring events. We anticipate our assay to be a starting point for more sophisticated studies on the processes behind the translocation biogenesis.

Zusammenfassung

Die Integrität des Genoms ist für die Vermehrung der genetischen Information von entscheidender Bedeutung. Allerdings sind Genome ständig schädlichen Faktoren ausgesetzt. Eine der gefährlichsten Arten von DNA-Schäden sind DSBs. Illegitime Reparaturen von beschädigten Chromosomen können zur Entstehung von Chromosomen-Rearrangements, einschließlich Chromosomentranslokationen, führen. Ihr Vorkommen kann zur Deregulierung der Genexpression oder zur Aktivierung von Onkogenen führen. Tatsächlich sind Translokationen das Merkmal vieler Krebsarten, psychiatrischer Erkrankungen und Unfruchtbarkeit. Der Prozess der Entstehung von Translokationen ist bekannt und umfasst: den Bruch von Chromosomen, die Migration in den Kernraum, Synapsis und falsche Reparaturen. Über die Faktoren, die zu diesen einzelnen Schritten beitragen, ist jedoch noch wenig bekannt. Aufgrund ihres seltenen Vorkommens ist die Untersuchung und Quantifizierung von Chromosomentranslokationen schwierig. Die derzeitigen Methoden wie FISH und PCR sind aufwendig und nicht empfänglich genug für solche niederfrequenten Vorgänge. Der Schwerpunkt in diesem Gebiet liegt auf dem Verständnis der Reparaturmechanismen und der Untersuchung der Eigenschaften einzelner Einflussfaktoren. Das Ziel dieser Studie war es, ein System zu entwickeln, dass die Identifizierung neuer Faktoren, die an der Bildung von Chromosomentranslokationen beteiligt sind, durch die Durchführung des genomweiten Screenings ermöglicht. Der hergestellte Translokationsreporter-Assay basiert auf der CRISPR/Cas9-Technologie zur Induktion von DSBs, welches einen neuen Goldstandard für die Herstellung von translokationsrelevanter Krebsmodellen ist. Darüber hinaus wurde die maßgeschneiderte Auswahlkassette in das Genom der haploiden Zelllinie integriert. Sie enthält das Antibiotikaresistenzgen, dessen Expression erst bei der Translokation aktiviert wird. Der Hauptvorteil des Systems besteht somit darin, dass es die Auswahl und Anreicherung von Zellen ermöglicht, die für Translokationen positiv sind. Unsere Ergebnisse zeigen, dass das System für die Generierung verschiedener Arten von Chromosomen-Rearrangements, wie Deletionen, Inversionen und Translokationen, geeignet ist. Es ist auch mit PCR-basierten Methoden kompatibel, wodurch eine Quantifizierung der Häufigkeit der auftretenden Ereignisse möglich ist. Wir gehen davon aus, dass unser Assay ein Beginn für anspruchsvollere Studien zu den Prozessen im Hintergrund der Translokationsbiogenese sein wird.

Table of contents

1. Introduction

1.1. Triggers of genome instability

The right balance between genome integrity and instability ensures cellular survival, the evolution of adaptive immune systems and prevents the development of genomic disorders. However, genomes are under the constant exposure to endo- and exogenous damaging factors, which can influence their organisation and structure. Structural damage is the most deleterious because it leads to loss of genomic information by creating chromosomal rearrangements such as insertions, deletions, inversions, duplications and translocations (Vitelli et al. 2017).

Endogenous damage can be triggered by metabolic products of cellular processes. Most frequently occurring damage with the frequency of up to 3x10^4 damaging events per cell per day (Cannan et al. 2017) is generated by reactive oxygen species (ROS) which are associated with mitochondrial respiration or inflammatory response (De Bont and van Larebeke 2004). Another instance are stalled replication forks which can collapse as a result of the accumulation of single-stranded DNA (ssDNA) and formation of double-strand breaks (DSBs) (Casper et al. 2002). Besides, head-on collisions of replication and transcription machineries are a potential threat to DNA (Helmrich, Ballarino, and Tora 2011). Furthermore, the formation of co-transcriptional R-loops and G-quadruplexes or resolution of topological constraints by topoisomerase activity are associated with increased risk of damage (Jiang, Lucas, and Beau 2015; Hangas and Pohjoismäki 2019). External agents causing damage to DNA such as UV light, ionising radiation, cigarette smoke and chemicals found in food are commonly present in our environment. Moreover, chemotherapeutics, such as topoisomerase II inhibitors, although primarily applied to eradicate cancer cells might as well cause therapy-related secondary cancers (Jiang, Lucas, and Beau 2015; Strick et al. 2000).

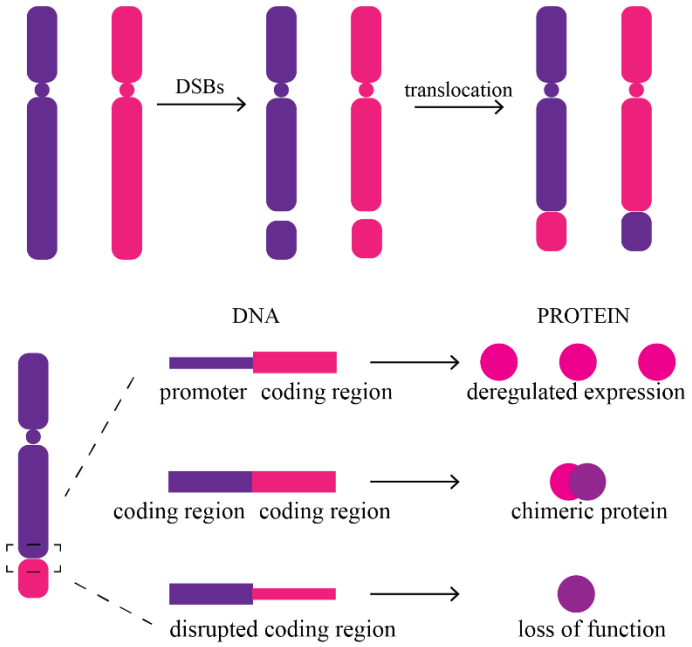
Breaks on both strands of DNA phosphate backbone (DSBs) are the most cytotoxic type of DNA damage. They occur with a frequency of up to 50 breaks per cell per day (White and Vijg 2017). Due to misuse of repair processes or lack of complementary template necessary for error-free repair they may result in chromosome translocations (Vilenchik and Knudson 2003; H. Wang and Xu 2017).

1.2. Chromosome translocations

1.2.1. Definition and consequences

Translocations are genome rearrangements in which ends of broken chromosomes are illegitimately joined. Balanced reciprocal translocations involve an exchange of telomeric and centromeric fragments between both chromosomes and are often a cause of primary abnormalities leading to cancer. When only one chromosomal segment is transferred to another chromosome, such as in nonreciprocal translocation may result in, for example, a formation of dicentric chromosomes. This type of aberrations is often acquired during the development of tumours (Bunting and Nussenzweig 2013).

Formation of translocation may deregulate the expression of the affected gene. It can be constitutively activated as in case of transcription factor and a potent proto-oncogene MYC, which translocates with immunoglobulin loci *IgH* t(8,14)(q24,q32) in mature B-cells resulting in Burkitt lymphoma. Gene expression can be also repressed like in the case of Serine/threonine-protein kinase TEL1 (TEL1) in *TEL1-AML*(t(12,21)) fusion in acute lymphocytic leukaemia (ALL). Fusion proteins may also gain novel functions and display an oncogenic potential. As an example the translocation of breakpoint cluster region protein (BCR) and tyrosine-protein kinase ABL1 (ABL) genes resulting in (t(9,22)(q34,q11)), well-known as Philadelphia chromosome, leads to constant activation of tyrosine kinase and is a common characteristic of chronic myelogenous leukaemia (Gunji et al. 2004; Bemark and Neuberger 2000; Clark, Crist, and Witte 1989).



Modified from (Roukos and Misteli 2014)

Figure 1 Formation of chromosome translocations and its consequences.

Translocations are a characteristic feature of many cancers, mostly haematological disorders, but also solid tumours and account for up to 20% of cancer morbidity (Mitelman, Johansson, and Mertens 2007). Translocations have been also associated with psychiatric illness and infertility. Disruption of DISC1 (Disrupted-In-Schizophrenia 1) by a t(1;11)(q42.1;q14.3) translocation results in its reduced expression and production of chimeric transcripts and is associated with depression, schizophrenia, bipolar disorders and autism (Eykelenboom et al. 2012). Balanced translocations have been shown to reduce testicular volume, testosterone levels and to impair spermatogenesis. They are also linked to the risk of offsprings with aneuploidy (Y. Dong et al. 2012).

Formation of fusion proteins and changes in transcriptional levels of affected genes can lead to increased genetic instability. Therefore, recurrent translocations can be drivers of tumorigenesis. Introduced mutations may deregulate proliferation-controlling pathways, which lead to increased proliferative capacity of cells and decreased death rates. As a consequence, the replication rates and generation of reactive oxygen species (ROS) are elevated, which can trigger additional DNA damage (Bohlander & Kakadia, 2015). The example of fusion genes that result in increased ROS production is BCR-ABL (Sallmyr, Fan, & Rassool, 2008). Furthermore, products of chromosome rearrangements can influence cell cycle checkpoint

genes and repair proteins, which may cause alterations in repair pathway choice. The presence of mentioned BCR-ABL downregulates c-NHEJ factors such as Artemis and Lig4 and increases activity of a-EJ pathway, including Lig3 and WRN (Sallmyr, Tomkinson, & Rassool, 2016). Genome instability may also arise from changes in epigenetic landscape. One example is hypomethylation of H3K79 observed in the presence of CALM-AF10 fusion. Cells harbouring this translocation have increased radiosensitivity, whereas patients display more secondary chromosomal rearrangements (Y. Lin et al., 2019). Acquisition of genetic lesions happens during clonal selection of cells with growth advantage, all of which accompanies tumour development.

1.2.2. Cellular and molecular mechanisms contributing to the formation of translocations

Although the involvement of chromosome translocations in cancer development has been widely studied, the actual mechanisms of their formation within the nucleus are still poorly described. The biogenesis of chromosome translocations can be divided into separate steps, which involve the initial breakage of two distinct chromosomes, their migration in the nuclear space, followed by temporal joining with possible translocation partner (synapsis) and final inaccurate ligation. Therefore the frequency of translocations depends on factors that contribute to the generation of DSBs, their persistence, and joining (Roukos and Misteli 2014).

1.2.3. Factors influencing translocation formation

1.2.3.1 The role of genomic and chromatin features

The occurrence of DSBs at two chromosomes at the same time is the first requirement for the formation of translocation. It can be triggered by endo- and exogenous factors as described before. It is has been shown that some sequence features and chromatin modifications make specific loci more susceptible to breakage and affect translocation frequency. Among sequences most prone to damage are CpG islands, repetitive elements (Alu and LINE repeats) and common fragile sites (CFS). Also DNA regions that adopt non-B secondary structure , e.g. Z-DNA, H-DNA are more affected by genomic rearrangements (Tsai et al. 2008; Elliott, Richardson, and Jasin 2005; Glover et al. 2005; G. Wang et al. 2008).

One of the requirements of translocation formation is spatial proximity of broken ends, which increases their probability to synapse. The chromosome proximity is influenced by higher-order organization in the nuclear space since chromosomes are not randomly distributed,

but occupy a predefined space called chromosome territories (Cremer and Cremer 2001). Interestingly, it positively correlates with translocation frequencies (Zhang et al. 2012). For example, frequent translocation partners of *MYC* – *IgH* and *IgL* are found in closer distance than its less frequent partner – *IgK* (Roix et al. 2003).

In many cancers, recurrent translocation partners are found in a close distance. It is in agreement with the established notion that proximity is cell-type and tissue-type specific. It was demonstrated that RET and H4 genes, which rearrangements are found in thyroid tumours, are distributed much closer than expected by random and this observation was specific to the thyroid, but not epithelial cells (Nikiforova et al. 2000). Also, fusion protein which is the product of *PML-RARA* translocation is found only in the hematopoietic precursors and results in promyelocytic leukemia (APL). Another example is BCR-ABL fusion which leads to chronic myeloid leukemia (CML) and is specific for haematopoietic progenitor cells. Distances between mentioned translocation partners were also examined during cell cycle progression, which showed that they are variable and smallest during the transition between S and G2 phases (Neves et al. 1999).

It has been shown that proximity can be induced by various signalling pathways. There is experimental evidence that transcription is spatially distributed into so-called transcription factories. Their occurrence may lead to gene-clustering or formation of chromatin loops, which both result is smaller distances between transcribed loci (Osborne 2014; Rieder, Trajanoski, and McNally 2012). Another example is an androgen-induced transcription, which in the presence of DHT mediates the interaction of TMPRSS2-ERG genes (Mani et al. 2009).

Studies on locus mobility showed that in mammalian cells DSBs have a stable position and display a limited movement of approximately 1 μm^2/h, which is comparable with the one of an intact locus. However, there is also an example of the long-range movement of DSBs induced by α-particles (Kruhlak et al. 2006; Aten et al. 2004). Another study showed that translocating breaks move faster than non-translocating and the majority of breaks translocate within the distance of 2,5 μm with sporadic cases of more than 5 μm (Roukos et al. 2013). It has been also shown that the mobility might be cell-cycle depended since chromosome territories in HeLa cells displayed increased positional shifts in early G1 phase (Walter et al. 2003). Moreover, in S-phase, the movement of ionizing radiation-induced foci (IRIFs) within euchromatin was increased over heterochromatin (Krawczyk et al. 2012). Interestingly, the synapsis rates of DSBs measured across different cell cycle phases did not display any changes,

which suggests that chromosome mobility is not a crucial determinant of translocations formation (Roukos et al. 2013).

Chromatin environment can impact the susceptibility to breakage and the motion of DSBs, therefore changing the propensity to create translocations. It has been shown that chromatin density is the most significant predictor of translocation breakpoints. The analysis of thousands of karyotypes from human blood and solid cancers revealed that chromosome breaks leading to translocations occur mostly within open chromatin regions, in acrocentric chromosomes and in the regions rich in CTCF/cohesin binding sites (Lin et al. 2018). A growing body of evidence suggests that chromatin remodelling associated with active transcription contributes to genome instability. Application of high-throughput genomic techniques helped to identify promoters of transcriptionally active genes that replicate early as particularly fragile. This might be associated with the occurrence of co-transcriptional R-loops and G4s or with the activity of TOP2B which includes generation of transient DSBs to release paused RNA Pol II (Marnef, Cohen, and Legube 2017). Systematic study in CD34+ cells revealed that translocation-prone regions are enriched with histone modifications associated with transcriptionally active and open chromatin, such as H3K4 mono and trimethylation (Burman et al. 2015). It has been also shown that binding of transcription factors can trigger alterations in chromatin structure bringing some genomic regions in proximity and increasing the probability of their translocations (Mathas et al. 2009). Positive association of transcription and formation of chromosome rearrangements was also observed in high-throughput sequencing-based studies (Klein et al. 2011; Chiarle et al. 2011). Since it was exemplified that chromatin containing DSBs is more mobile (Krawczyk et al. 2012) and translocating broken ends move faster than non-translocating ones (Roukos et al. 2013), the motion of DSBs is an important phenomenon contributing to translocations. Changes in motion can be induced by chromatin remodelers, e.g. VP16 and INO80 (shown in yeast) (Neumann et al. 2012) or by inhibitors such as TSA (histone deacetylase inhibitor) (shown in human cells) (Krawczyk et al. 2012).

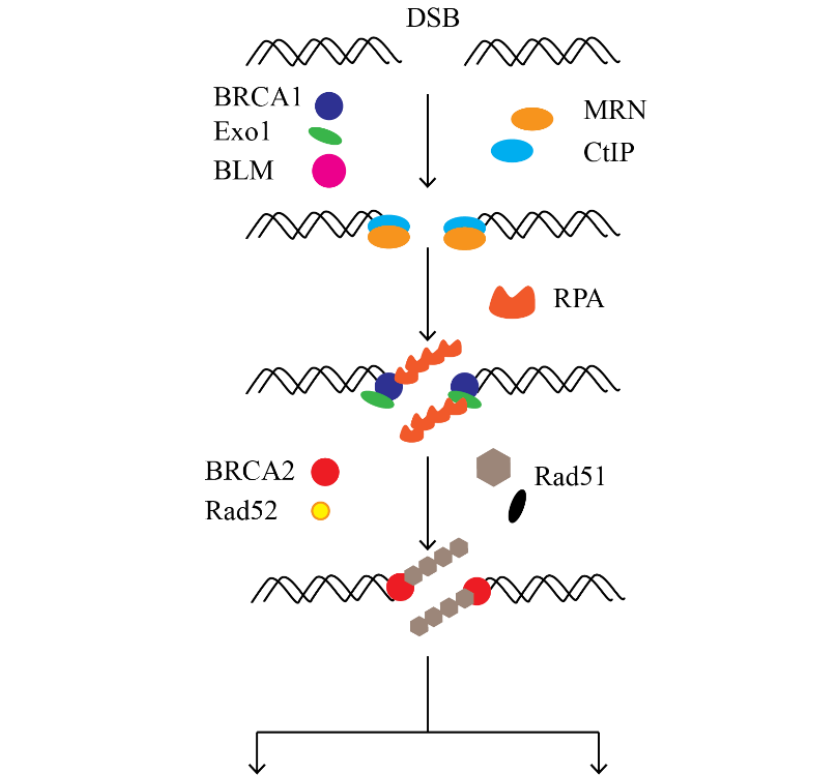
1.2.3.2. The role of DSB repair pathways

Cellular systems developed a plethora of DNA repair pathways designated for specific types of lesions. DSBs are repaired by two major ones distinguished by their requirement of homology for repair: homologous recombination (HR) and non-homologous end-joining (NHEJ). NHEJ is a rapid and most prominent pathway in mammalian cells. It is active throughout the cell cycle and is template-independent since it has an ability to religate even

unprocessed ends. On the other hand, it is error-prone and often leads to the formation of short insertions or deletions. Instead, HR is error-free but due to the requirement of homology template, it is active only during S/G2 when the sister chromatid is present (Lieber 2008; San Filippo, Sung, and Klein 2008). DSB repair pathways process broken ends to maintain genomic stability, thereby preventing the formation of translocations. Failures in repair pathways may occur on a global level by mutations in repair factors or on a local level through problems with assembling the repair machinery, limited availability of factors, decreased accessibility of damage site due to chromatin compaction, and finally complexity of the break (Giglia-Mari, Zotter, and Vermeulen 2011; Dueva and Iliakis 2013). In HR pathway DSBs are recognised and bound by MRN complex (Petrini and Stracker 2003). It is followed by activation of ATM by TIP60 and phosphorylation of γH2AX (Sun et al. 2005; Rogakou et al. 1998). First, to enable the homology search, the broken ends are resected by CtIP nuclease (Sartori et al. 2007). This leads to the formation of ssDNA bound by RPA, which is further replaced by nucleoprotein Rad51, in a process mediated by BRCA2 (C. J. Ma et al. 2017). Afterwards, the strand invasion takes place and D-loop is generated. The missing sequence is extended using the homologous template, after which DNA double strands are ligated and junctions are resolved (Brandsma and Gent 2012).

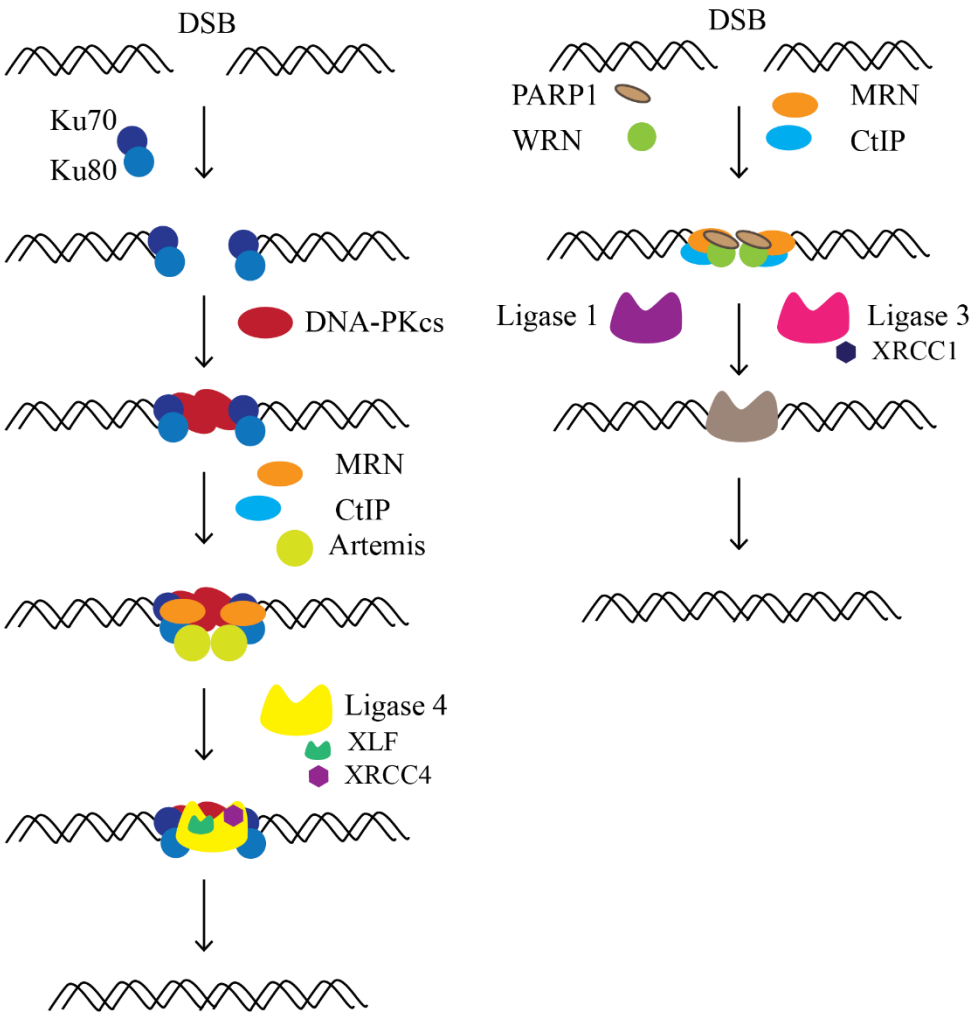
HR is thought to not play a major role in simple chromosomal translocations. However, as shown by (Piazza, Wright, and Heyer 2017), some forms of HR intermediates such as multi-invasions may give rise to chromosomal rearrangements. The abundance of repetitive sequences in the human genome gives the possibility that non-allelic repeats can be used for recombination in a non-allelic homologous recombination (NAHR) pathway. Analysis of translocation junctions of interchromosomal HERV-HERV recombination showed that the sequence was consistent with NAHR mechanism (Hermetz et al. 2012). There is also evidence of involvement of Rad51-independent sub-pathway of HR – single-stranded annealing SSA in the generation of translocations between Alu elements (Elliott, Richardson, and Jasin 2005).

Analysis of breakpoint junctions from many tumours revealed that they are mostly created by NHEJ and therefore this pathway and its components will be described in more detail. NHEJ is a predominant pathway involved in V(D)J and class switch recombinations (CSR). First, DSBs are bound by Ku70/80 heterodimer. Then the DNA-PK is recruited and through phosphorylation activates downstream factors. Any required end-processing involves Artemis and MRN. Final ligation is done by Lig4-XRCC4 complex (Ghezraoui, et al., 2014).



Modified from (Dueva and Iliakis 2013)

Figure 2. Repair of double-strand breaks by homologous recombination.



Modified from (Dueva and Iliakis 2013)

Figure 3. Repair of double-strand breaks by non-homologous end joining pathways.

Ku heterodimer is one of DSB sensors, which consists of Ku70 and Ku80 polypeptides of 70 and 83 kDa respectively. Its rapid binding, independent of the sequence, functions as a bridge and stabilizator of the broken ends, which facilitates ligation (Brown, Jackson, and Jackson 2015). It also acts as a docking platform for other components on NHEJ (G. Lu et al. 2016). Presence of Ku protects DNA ends from resection by preventing the activity of exonucleases. Ku is highly abundant in human cells and plays a critical role in joining RAG-generated DSBs and in the prevention of dramatic telomere loss. It was shown to be an essential protein but only in humans (Indiviglio and Bertuch 2009; Liang and Jasin 1996). Studies by Difilippantonio et al. implicate that Ku suppresses chromosomal aberrations since in Ku80 $^{-/-}$ mice increased occurrence of breakage, translocations and aneuploidy were observed (M. J. Difilippantonio et al. 2000). Depletion of Ku80 in mouse cells also leads to the increased levels of translocations arising from I-SceI endonuclease-induced breaks (Roukos et al. 2013),

similarly to knock-out of Ku70 in mouse embryonic stem cells (Weinstock, Brunet, and Jasin 2007). There is also evidence that Ku80 constrains the local mobility of broken chromosome ends (Soutoglou et al. 2007).

Due to high affinity to Ku, DNA-PKcs is recruited to the site of damage and forms the DNA-PK complex. DNA-PKcs is a member of the phosphatidylinositol 3-kinase-related kinase (PIKK) family. Two molecules of DNA-PKcs bind at opposite sides of a DSB and bring them together (synapsis), which also activates the kinase. On one hand, it phosphorylates downstream factors such as Artemis, which is involved in DNA-end processing. On the other hand its autophosphorylation plays a role in the regulation of end processing, enzyme inactivation and dissociation of the whole DNA-PK complex from DNA (Dobbs, Tainer, and Lees-Miller 2010; Uematsu et al. 2007; Neal and Meek 2011; Chu et al. 2002). It has been shown that chemical or siRNA mediated inhibition of this kinase results in increased levels of translocations and low levels of end processing in mouse experimental system (Roukos et al. 2013) and in human cells (Shibata et al. 2017; Wray et al. 2013).

Artemis is a structure-specific endonuclease. It is activated by autophosphorylated DNA-PK and acts with it in a complex by processing broken ends prior to ligation. The complex is involved in hairpin nicking in V(D)J recombination and cleaves 5'-3' overhangs in NHEJ (Y. Ma et al. 2002). Experiments in Artemis-deficient human fibroblasts showed that it is involved in promoting IR-induced translocations through slow cNHEJ process occurring in G1 (Shibata et al. 2017).

The final ligation step and resolution are executed by the Lig4/XRCC4 complex promoted by XLF. Deficiency of Lig4/XRCC4 in mice results in abolishment of V(D)J recombination and erroneous events in CSR (G. Lu et al. 2016). Lig4/XRCC4 does not have any known function outside NHEJ. Its binding to the DNA ends is promoted by Ku. Lig4 is responsible for recruitment and assembly of other ligation factors. It has been shown that it influences the cellular localization and protein levels of XRCC4. The Lig4/XRCC4 complex also contributes to DNA-PKcs autophosphorylation and DNA-PKcs mediated DNA end synapsis (Simsek and Jasin 2010; Fonslow et al. 2015).

Studies in mouse cells by (Simsek and Jasin 2010; Simsek, Brunet, et al. 2011) showed that cells deficient in Lig4 and XRCC4 experience increased levels of translocations and characteristics of translocation breakpoint junctions including microhomology usage were not altered. Also, a significant increase of translocations in Lig4-deficient CH12 cells but with

longer microhomologies was observed by (G. Lu et al. 2016). In the study by Ghezraoui et al. different types of endonucleases were used (ZFN/TALENs/CRISPR/Cas9) to induce DSBs in absence of Lig4 and XRCC4 in human HCT116 cells. Obtained translocation frequencies were lower than in WT, whereas translocation junctions had frequent microhomologies and longer deletions (Ghezraoui, et al., 2014). Similarly, knockdown of Lig4 and XRCC4 decreased androgen-induced translocations of TMPRSS2-ERG in human cancer cell lines. In contrast, Soni et al. showed an increase of translocations after IR in Lig4 mutated MEFs and HCT116 cells (Soni et al. 2014).

PAXX and XLF are the newest discovered factors of NHEJ. They share structural homology to XRCC4 and are functionally redundant. PAXX was shown to promote the accumulation of Ku at DSBs, whereas XLF works in complex with Lig4 and XRCC4 to promote ligation. Although knock-out of XLF in mouse cells resulted in increased levels of chromosomal aberrations including translocations, PAXX-deficient cells lack c-NHEJ phenotypes, such as shift towards microhomology usage, increased sensitivity to radiomimetic drugs or unproficiency of V(D)J recombination (Zha et al. 2007; Tadi et al. 2016).

The observation that loss of cNHEJ factors didn't completely abrogate the end joining and that translocations were detected in their absence led to the discovery of alternative end joining (a-EJ) pathway. For example, mouse cells lacking Lig4 or XRCC4 have translocations of IgH in CSR (Yan et al. 2007). A-EJ is activated only when major repair pathways - HR and cNHEJ are compromised, therefore it is considered to function as a backup pathway. However, it still shares either factors or mechanisms of other processes. As in HR, resection machinery is activated to process the broken ends. On contrary, extensive sequence homology is not required to ligate them, which resembles ability of cNHEJ to directly join broken ends. a-EJ is therefore defined as any pathway active in absence of cNHEJ pathway.

A-EJ process is initiated by PARP1 recruiting MRN and complex of Lig3/XRCC1. The broken ends are processed by MRN and CtIP due to their endonucleolytic function. This resection creates up to 5 bp long stretches of complementarity between broken ends called microhomologies. Deletions and extensive usage of microhomologies are frequent characteristics of the junctions processes by this error-prone pathway. Lig3 (or Lig1) and XRCC1 complete the end-joining process by performing ligation (Frit et al. 2014).

CtIP was first identified as an interactor of transcriptional co-repressor CtBP. Interestingly, it was later shown to also control cell cycle checkpoint and that it is recruited to

DNA damage sites, where it interacts with MRN. Now, it is known as an essential component of a-EJ responsible for end resection (Zhang and Jasin 2011; Sartori et al. 2007). Study in mouse cells showed that depletion of CtIP, resulted in a decrease in translocations frequency accompanied by lower microhomology usage (Zhang and Jasin 2011). However, another work in human cells showed that CtIP acts only in absence of factors form c-NHEJ, since its depletion in WT cells had no effect on translocation frequency and junction characteristics (Ghezraoui et al. 2014).

The primary role of Lig1 is in replication and ligation of Okazaki-fragments. It is also involved in NER. Lig3 is present both in mitochondria, where it maintains the mtDNA and in the nucleus. The latter form of Lig3 interacts with XRCC1 and as a complex is a component of BER and SSBR (Y. Gao et al. 2011; Simsek, Furda, et al. 2011). Both ligases were shown to mediate the ligation of broken ends in absence of Lig4 (G. Lu et al. 2016). Absence of Lig3 impacts translocations frequency and decreases microhomology usage in mouse cells. Additional depletion of Lig1 further inhibits translocations formation. Loss of Lig1 alone has no impact on translocations frequency (Simsek, Brunet, et al. 2011). On contrary, Lig3 deficiency in human HCT116 cells didn't alter translocations occurrence, length of deletion or microhomology bias (Ghezraoui et al. 2014). XRCC1 increases the competitive advantage of Lig3 against Lig1 (Soni et al. 2014).

The PARP family of 17 proteins has 3 members (PARP1, PARP2 and PARP3) that have been implicated in DNA repair (Day et al. 2017). PARP-1 competes with Ku for binding free DNA ends and initiates a-EJ (M. Wang et al. 2006). In human cells, PARP-1 inhibition by olaparib, rucaparib or siRNA resulted in decreased translocation levels after enzymatic induction of breaks with I-SceI or ZFN and with more clinically relevant causes such as IR and VP16 (Wray et al. 2013). Extended studies on mechanisms of PARP1 activity in DSBs repair led to the identification of highly potent inhibitors, which exert synthetic lethality in BRCA1/BRCA2 tumours. For example, olaparib has been approved by FDA for therapy of ovarian cancer and by EMEA also for fallopian tube and the peritoneum cancers. It is a major step in the development of personalised medicine therapies for cancer.

PARP3 was shown to be involved in the recruitment of cNHEJ factors, e.g. it accelerates the DNA ligation process by promoting the loading of Lig4/XRCC4 complex on chromatin (Rulten et al. 2011). It was identified in the shRNA based screening as a positive modulator of translocations. Its function is to facilitate the repair process by suppressing the formation of G-

quadruplex DNA, which could impair the deposition of repair factors. It was also shown that PARP3 promotes binding of CtIP and RPA at DSBs (Day et al. 2017).

53BP1 is known to promote c-NHEJ and inhibit HR by blocking 5' end resection. It also mediates the synapsis of broken ends and promotes their mobility (Bunting et al. 2010; S. Difilippantonio et al. 2008). It was shown that 53BP1 suppresses chromosome rearrangements to intergenic regions, since its loss led to an increased number of translocations and more extensive resection (Jankovic et al. 2013). 53BP1 also suppresses the formation of new ultrafine DNA bridge-breakage process, which leads to translocations (Tiwari, Addis Jones, and Chan 2018).

PolQ is a component of a-EJ pathway mediating the joining of resected 3' ends. The two main signatures of PolQ-dependent pathway are the usage of microhomology-primed DNA synthesis and generation of insertions at the repair junctions (Wood and Doublié 2016). PolQ has been shown to be able to both suppress and promote chromosome translocations. Mice defective in PolQ had an increase in *Myc-IgH* translocation (Yousefzadeh et al. 2014), but a decrease of CRISPR/Cas9 induced translocations (Mateos-Gomez et al. 2015). Interestingly, no significant changes in the frequencies were detected by (Wyatt et al. 2016). The same study shows that PolQ is able to protect from chromosome rearrangements, but only in the absence of c-NHEJ component-Ku70. Mateos-Gomez et al. described the breakpoint junctions characteristics in PolQ-depleted cells, showing the absence of insertions and decrease in microhomologies (Mateos-Gomez et al. 2015).

In summary, extensive studies on understanding how DNA repair pathways work, also helped to shed light on the mechanisms of translocation formation. Currently, there is a discrepancy in the field in terms of which pathway suppresses and which one promotes translocations. Majority of results suggest that translocations in mouse cells are created through a-EJ, whereas in human cells by c-NHEJ. However, due to variability in applied systems of DSBs induction and translocations detection, more systematic approaches are required to address this question. This also raises the need for better translocation reporter systems. The majority of factors identified so far are directly linked to the DNA repair processes. However, there are already some indications of other possible players outside those pathways. For example, senataxin is an R-loop helicase, and apart from its role in transcription termination, it is also involved in the genome maintenance. It has been shown that it is recruited to the DSBs induced in active loci and limits translocations formation by promoting the loading of Rad51 and inhibiting joining of distal ends (Cohen et al. 2018).

The key function of the multiprotein cohesin complex is tethering of sister chromatids until anaphase, to prevent the premature division of chromosomes. Besides, it acts in the organization of 3D chromatin structure and regulation of transcription. Cancer-associated SA2 mutation of cohesin results in failure to repress transcription at DSBs, which leads to the formation of genomic rearrangements. Interestingly, cohesin also has been shown to influence DSBs repair process and protect from the occurrence of large chromosomal rearrangements, such as translocations. It is proposed that it limits the mobility of broken ends, suppressing the joining of distant, but not adjacent ones. Additionally, the positive correlation of translocation breakpoints and CTCF/cohesion sites was observed in many cancer karyotypes (C.-Y. Lin et al. 2018; Gelot et al. 2016; Meisenberg et al. 2019).

In conclusion, presented evidence indicates complexity and diversity of mechanisms and pathways participating in the formation of chromosome fusions.

1.3. Methods of induction, detection and quantification of translocations

Studies of translocation-associated cancers aim to prevent their occurrence or decrease the negative functional consequences. It is done by investigating the mechanisms of their formation and creating cancer models based on ectopic or endogenous expression of fusion genes. Fundamental research is commonly applying endonucleases to induce DSBs, which misrepair may lead to translocations, generating tumour models suitable to study those mechanisms. Other ways to induce DSBs, more clinically relevant are IR and chemotherapeutics, such as etoposide.

1.3.1. Induction of translocations by designer nucleases

I-SceI is a yeast homing rare-cutting endonuclease, targeting 18 bp recognition sequence. Since this sequence is not present in humans, it became a popular tool for induction of breaks in mammalian genomes. Its biggest drawback is the necessity to introduce the restriction sites into cell model, which limits performed experiments to just one targeted genomic loci and makes it a laborious technique (Erika Brunet and Jasin 2018). I-SceI is commonly used in translocation reporter assays, to create breaks at nonhomologous chromosomes, which translocation reconstitutes a functional gene, e.g. fluorescent reporter – GFP or antibiotic resistance gene – Neomycin (Pierce et al. 1999; Wray et al. 2013). This methodology was also successfully combined with modern genome-wide approaches to study translocations mechanisms (Chiarle et al. 2011; Hakim et al. 2012).

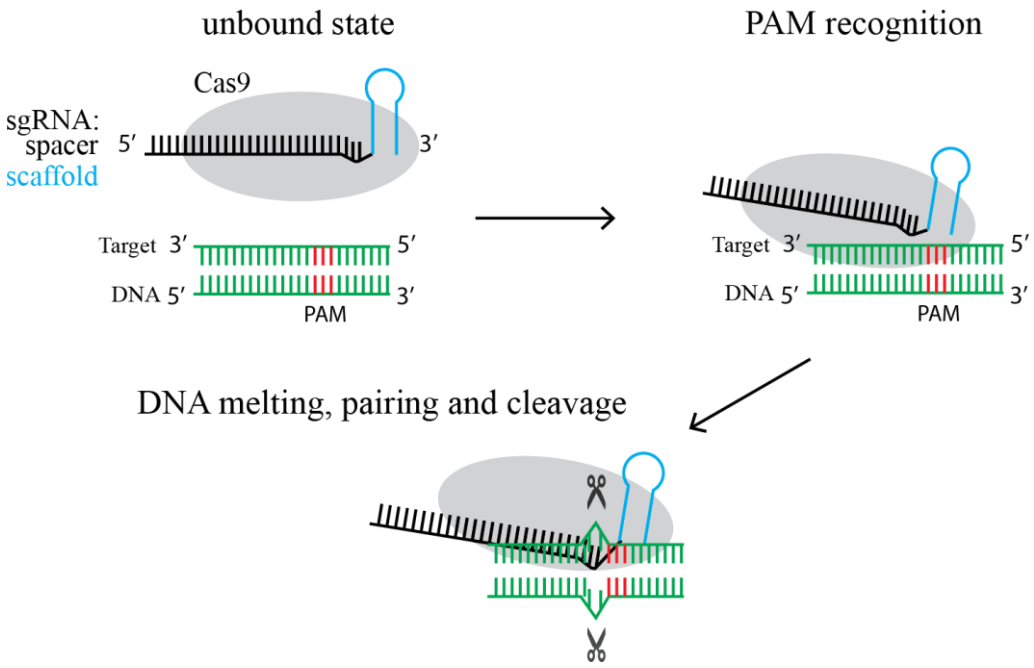
Programmable nucleases, such as zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALENs) give much more flexibility in choosing the locus for induction of DSBs. ZFNs were developed in the lab of Chandrasegaran, who studied restriction enzymes as potential tools for genome editing and they were the first programmable nucleases used for targeting custom sites (Y. G. Kim, Cha, and Chandrasegaran 1996; Chandrasegaran and Carroll 2016). The first discovery of TALEs comes from the studies of genus Xanthomonas - pathogenic bacteria damaging crop plants. They secrete into the host plant cells DNA binding proteins which activate the expression of targeted genes (Boch and Bonas 2010). Both ZFNs and TALENs consists of DNA-binding modules, which recognise the target site, and FokI domain, which plays an endonuclease function. Zinc finger array consists of amino-acid modules which recognise 3 bp of a sequence. Therefore combining various types of modules in tandems enables targeting of specific loci (Gupta et al. 2012). Single modules of TALENs are built of 34 amino acids and each one recognises just one nucleotide. Again assembly of arrays gives specificity to the sequence that is to be modified (Joung and Sander 2013).

The induced cleavage provokes the DSBs repair and often results in the occurrence of insertions and/ deletions which modify the locus. Using those nucleases enables precise cut at translocation breakpoints, required to induce cancer-relevant translocations. They can be also used for a target integration of a transgene into the genome (Piganeau et al. 2013). However, the experimental design is very complex and laborious and the specificity and affinity are limited. Cell toxicity associated with usage of those nucleases was another drawback.

The solution to some of the genome engineering problems came with the discovery and development of CRISPR/Cas9 methodology. It was initially shown to play a role in bacterial adaptive immune system, protecting from invading virus and plasmid DNA (Wiedenheft, Sternberg, and Doudna 2012). In a very short time, because of it is simplicity and efficiency it was applied in genome editing field and revolutionised it. In contrary to ZFNs and TALENs, which action is based on DNA-protein binding, CRISPR is using complementarity of base pair between RNA and DNA. The first component of a system is sgRNA (single-guided RNA) responsible for site recognition. It consists of crRNA sequence complementary to target locus and tracrRNA part interacting with the second element - Cas9. This *Streptococcus pyogenes* endonuclease recognises the 5'-NGG-3' motif called PAM sequence located directly downstream the target site. The formed CRISPR/Cas9 complex induces DSBs (or DNA nicks in case of nuclease dead dCas9 form) within a sequence of interest (Martin Jinek et al. 2012;

Sander and Joung 2014). So far this technique was successfully used for inducing rearrangements found in different types of cancer, e.g. NPM1-ALK and RANX1-ETO associated with AML, EWSR1-FLI1 present in Ewing Sarcoma or CD74-ROS1 in lung cancer, and is a new golden standard for creating translocation relevant cancer models (Ghezraoui et al. 2014; Torres et al. 2014; Choi and Meyerson 2014).

1.3.2. Induction of translocations by agents used in cancer therapy



Modified from (Xuebing Wu et al., 2014)

Figure 4. Principle of CRISPR-based induction of DSBs.

Cancer treatment with radio and chemotherapeutics often relies on the induction of DNA damage. However, apart from eradicating tumour cells, it also can affect the healthy ones. This unwanted effect of therapy can result in chromosome aberrations, including translocations and lead to the development of secondary cancer (Allan and Travis 2005). Therefore, it is of high clinical relevance to understand how those agents act on genome stability. For this reason low doses of IR and anti-cancer drugs, such as etoposide, doxorubicin, mitoxantrone (a topoisomerase II inhibitors), or olaparib (PARP1 inhibitor) are also used as tools to study DNA repair and translocation mechanisms (Soni et al. 2014; Libura et al. 2005; Wray et al. 2013).

1.3.3. Tools for detection of translocations formation and quantification of frequencies

Detection of chromosome rearrangements is used as a diagnostic and prognostic tool in the treatment of cancer patients. First translocations were detected at the beginning of the 20th century with cytogenetic techniques. They were based on banding patterns, which enabled identification of distinct chromosomes. Their biggest limitation is low resolution and lack of information about the breakpoint. Further advancement came with the development of fluorescent *in situ* hybridization (FISH). Using locus-specific probes made it possible to label breakpoints on metaphase spread chromosomes. It greatly improved the resolution, however, is still limited by the number of fluorochromes that can be simultaneously applied (Mertens et al. 2015).

The biggest breakthrough was the application of high-throughput technologies. Initially, as array-based platforms, they enabled copy number profiling and analysis of expression of fusion genes. Joined with deep sequencing technologies made it possible to identify gene fusions on both DNA and RNA level. Thousands of samples from patients harbouring cancers of different origin were tested and until now more than 20,000 gene fusions have been identified. This number includes primary driver mutations as well as alterations acquired during tumorigenesis. The study of Gao et al. indicates that fusions drive the cancer development in 16.5% of cases and they are unique drivers in >1% of them. The biggest advantage of such techniques is its high resolution to nucleotide-level and unbiased genome-wide approach. However, it also has some limitations, which include short read length, possible errors in sequencing libraries or problems with a reference genome assembly, which are constantly being improved. Recently advancements in whole-genome sequencing (WGS) technique has been developed and include low-pass and mate-pair WGS. Improvement of bioinformatic tools and software to analyse the huge amount of sequencing data is an important factor, which influences the performance of those assays (Tran et al. 2018; Z. Dong et al. 2018; Q. Gao et al. 2018). Predicted fusions are compiled in catalogues, such as Mitelman database, Chimer 4.0., FusionGDB and more (Lee et al. 2017; P. Kim and Zhou 2019; Jang et al. 2020).

Translocations which result in transcription of a fusion gene can be detected at RNA level with standard RT-PCR (Salto-Tellez et al. 2003). Rearrangements which are not characterised by the presence of chimeric protein can be still detected on the DNA level by a broad range of PCR variants, which include: nested, multiplex, inverse, ligation-mediated, long-range, ΔPCR just to mention a few. PCR can be further modified to obtain quantifiable data on translocation frequency. For that purpose qPCR or ddPCR based on SybrGreen,

EvaGreen or fluorochrome-labelled probes are applied (Germini et al. 2017; Williams, Rainville, and Nicklas 2002; P. Wang et al. 1997; M. T. Lin et al. 2011; E. Brunet et al. 2009; Shuga et al. 2013). Although very high specificity and sensitivity, those techniques have some limitations. They require the knowledge about exact breakpoint in at least one of translocation partners. Successful amplification depends also on deletion/resection size, GC content and presence of repeats within the template. Also, multiplexing is dictated by the number of fluorochrome detection channels in a thermocycler (Abildinova et al. 2016).

Another interesting approach to detect both balanced and unbalanced rearrangements was to apply in-nucleus Hi-C. It has base pair resolution and lower cost than standard sequencing. It allowed to detect previously not known rearrangements and can be as well applied for analysis of copy number variations (Harewood et al. 2017).

Detection of translocations induced by both I-SceI and etoposide is limited by the lack of information about the translocation partner. Fortunately, there are alternative methods to overcome this difficulty, but they are much more laborious than standard PCR. One possibility is to apply inverse PCR (Varga and Aplan 2005) or more sensitive and specific LAM-PCR, which are used to identify unknown fragments flanking the known sequence (Gabriel et al. 2014). LAM-PCR is also compatible with genome-wide high- throughput sequencing, which allows studying the repair processes on the DNA level with single-nucleotide resolution. LAM-HTGTS is a powerful method for identification of translocation partners via bait-prey DSBs approach and has been already applied for detection of intra- and interchromosomal rearrangements. It is also a tool for recognition of off-target cleavage induced by designer nucleases (Hu et al. 2016).

In pure tumour samples, all cells harbour the same somatic driver mutation responsible for cancer development, e.g. translocation. However, only occasionally DNA samples from solid tumours are completely pure, which can negatively influence the analysis. The difficulties increase with the heterogeneity of cell pool which is due to tumour evolution and presence of non-cancerous cells (Tong et al., 2018). Still, the frequencies of translocation events measured in clinical samples are much higher than the ones induced in research settings ($\sim 10^{-3}$-10^{-4}) (Vanoli and Jasin 2017). Therefore, routine detection of structural aberrations in patients' samples relies on low resolution, cytogenetic and FISH-based assays complemented by high-resolution molecular methods. The latter becomes more accessible for diagnostic labs with decreasing sequencing costs (Abel, Pfeifer, and Duncavage 2015). Low-frequency events (<1%) induced e.g. by nucleases require more precise and sensitive tools. The gold standard

PCR-based methods require the knowledge about breakpoint sequence and are not suitable for high-throughput approaches. For instance, quantification of a large number of cells can be performed with imaging method - High-throughput break-apart FISH (HiBA-FISH), which has been used to detect clinically relevant NPM1-ALK translocation (Burman, Misteli, and Pegoraro 2015).

1.4. Open questions in the field

While a lot of effort has been put forth to understand how certain chromosome translocations drive oncogenesis, less is known about the molecular and cellular pathways that contribute to their formation. What determines where chromosomes break and why specific sites undergo translocations whereas others not, how broken ends move and find each other in nuclear space and what is the contribution of the chromatin context in this process are still open questions.

Information about gene fusions is used for a correct diagnosis, selection of treatment and prognosis. Studies of gene fusions and the mechanisms behind their formation are necessary for the development of novel therapeutic strategies to improve patients' lifespan and quality of life. Successful treatment may lead to a decrease or eradication of chimeric protein causing the disease. There are already known success stories of targeting the product of gene fusion with drugs, e.g. imatinib targeting the product of BCR-ABL fusions in patients with CML or tretinoin used in PML-RARA–positive patients (Druker 2008; Warrell et al. 1991). A huge advancement in oncological practice would be to reduce the risk of metastases in cancer patients. Application of clinical inhibitors, such as PARP1 inhibitors, has a potential to reduce levels of translocation formation and therefore occurrence of secondary cancers (Wray et al. 2013; Nickoloff 2016).

Due to their rare occurrence, probing and quantifying the formation of chromosome translocations is difficult. Available gold standard techniques, such as FISH and PCR are laborious and not suitable for low-frequency events. Additionally, these methodologies were valuable for the identification of individual factors that influence translocation frequency, but are not applicable for unbiased genome-scale screens. Therefore, to identify novel factors involved in the formation of translocations, this project aims to develop a new system that is compatible with a genome-wide screen for translocation suppressors.

1.5. Goals of the project

The goal of our work is to shed light on molecular pathways that contribute to biogenesis of translocations by identifying novel factors that play role in this process. More specifically, I aimed to:

1) Identify suppressors of chromosome translocations by developing a novel translocation reporter system compatible with genome-wide screen utilizing CRISPR/Cas9 pooled libraries and applying it.
2) Access in details how selected candidates affect individual steps of translocations formation.
3) Evaluate the findings in models of recurrent cancer initiating translocations, which are frequently found in patients.

1.5.1. Generation of novel translocation reporter system.

Chromosome translocations form at very low frequencies, which is the major limitation for genome-wide screenings. To overcome it, there is a strong need for robust translocation reporter system, which would enable to identify novel factors in unbiased and high throughput way. Its generation and optimization were the main subjects of this thesis.

The performance of informative screening requires a compatible translocation assay. First of all, it has to display a selectable phenotype, which in this case is the presence of translocations. The second important aspect is the selection of a cell line. The libraries available for CRISPR based screenings are usually in the form of plasmid pooled libraries and require lentiviral infections. Since the MOI applied in such systems depends on virus titer and transduction efficiency, therefore the cell line selected for assay need to be easily transduced. Depending on the library size, the number of cells needed to perform a screening varies and may involve many resources, including consumables and storage for cell culture, as well as the time necessary for harvesting and seeding. Another critical requirement is the high level of Cas9 expression in the cell line. Last but not least, the selectable markers present in the library cannot overlap with the ones used in the assay (Miles, Garippa, and Poirier 2016).

1.5.2. Advantages of CRISPR/Cas9 based screens and compatibility of translocation assay design

The rapid evolution of gene editing methods including CRISPR/Cas9 technology opened new possibilities to study many cellular processes in an unbiased manner. Previous state of the art shRNA based screenings had some limitations, such as the incompleteness of protein

depletion and decreased sensitivity due to off-target effects. The superiority of CRISPR/Cas9 pooled screenings is associated with high sensitivity and specificity. The lack of CRISPR/Cas9 in eukaryotes means that it does not compete with endogenous pathways. It is characterised by lower noise, minimal off-target effects and high experimental reproducibility. Low costs and ease of oligo design for preparation of libraries make them an affordable alternative. The bottlenecks of this technique include the fact that loss of function screenings are limited to only non-essential factors. Additionally, while creating a gene knock-outs by WT Cas9 in same cases in-frame mutations can be created, which generates a mixed population of cells and noisy results. CRISPR/Cas9 screen can be a good basis for studying biological processes. Generated hits can be further verified with the help of computational tools to create smaller focused libraries for secondary validation screens. The versatility of the system enables to perform viability-based positive and negative screens, marker-based selection screens or even combinatorial ones with mixing WT and dCas9 for simultaneous activation and deletion of two genes (Doudna and Charpentier 2014; Sharma 2018).

1.5.3. Functional characterisation of the selected candidates

1.5.3.1. Cellular system to study translocations formation in living cells

After performing the screening the second aim of the project would be to validate the hits and access the impact of selected factors on spatiotemporal aspects of translocation biogenesis: DSBs motion, synapsis and frequency of illegitimate repair. To perform those experiments I could apply the cell-based system described in (Roukos, Burgess, and Misteli 2014), which enables visualisation and tracing of broken ends. It is possible due to the integration of I-SceI endonuclease sites flanked by Lac/TetO operator arrays, which are bound by fluorescently labelled respective repressors. This system can be used for observations in living cells using time-lapse microscopy and for quantification of translocations frequency by PCR. Unfortunately, it is limited to array integration loci. Another alternative or complementary technique could be CRISPR-based genomic imaging, which uses dCas9 fused with fluorescent protein and arrays of multiple sgRNAs. However, this method is still in its infancy (Wu et al. 2019).

1.5.3.2. Models of recurrent cancer initiating translocations

The final part of the project will focus on applying obtained results in models of recurrent cancer-initiating translocations, to elucidate mechanisms of their formation. The selected models include translocations identified in patients with prostate cancer (PCA)

and anaplastic large cell lymphoma (ALCL). Prostate cancer is a second, after lung cancer, a major cause of male cancer-related death. Gene rearrangements involved in PCA include androgen-regulated gene TMPRSS2 and ETS transcription factor family members: ERG, ETV1, ETV4, with TMPRSS2-ERG fusion being the most predominant molecular subtype of prostate cancer (identified in 48,5% clinically localized PCAs). It is also the most common rearrangement identified in humans so far, being exclusive just for PCA (Tomlins et al. 2008). ALCL is a rare, aggressive T-cell lymphoma. It accounts for up to 15% lymphoma cases in children and up to 8% in adults. About 50% of patients with ALCL carry the translocation between chromosomes 2 and 5 t(2;5)(p23;q35), which results in expression of NPM-ALK (nucleophosmin-anaplastic lymphoma kinase) fusion protein. ALK^+ patients (50-85%) have better prognosis after chemotherapy treatment. Understanding the mechanisms of ALK-translocations will help to develop new treatment strategies (Roukos and Mathas 2015).

The more detailed evaluation of candidates selected in screening, including the estimation of synapsis rates and translocation frequencies, will be based on multi-coloured FISH and PCR methodologies. The sequence flanking the breakpoint junctions specific to each translocation partner can be labelled with fluorescent probes of different colours, which enables their detection. DSBs within the sequences of partner genes as well as candidate gene knock-outs would be induced by CRISPR/Cas9.

2. Results

2.1. Genome-wide screening for novel factors involved in the formation of chromosome translocations – general experimental approach

The experimental strategy is based on the idea to enrich the pool of translocation positive cells through antibiotic selection. In order to rich this goal, the antibiotic selection cassette is integrated into the genome of human cell line. Then, by using the CRISPR/Cas9 system, DSBs at specific locations are induced. The first one is upstream the promoterless antibiotic resistance gene (Hygromycin) (Chr.A) and the second one is downstream the active promoter (Chr.B). Cells in which those broken chromosomes translocate, acquire antibiotic resistance, which gives them a selective advantage (Figure 5).

Those DSBs are induced in the absence of individual candidate factors, which are knocked-out with barcoded sgRNA pool library. The library consists of thousands of members and include nontargeting controls. Every sgRNA is labelled with the distinct barcode, allowing its quantification at key steps of the screening by High-throughput Illumina sequencing (HiSeq 2500). Comparing the abundance of sgRNAs present in the initial pool of cells and after antibiotic selection allows to identify suppressors of translocations. Negative controls are non-targeting sgRNAs and cells without Hygromycin selection.

The experimental setup depends on the usage of a haploid human cell line. This characteristic serves perfect conditions for generating cell lines with complete loss of function phenotypes since gene inactivation can be facilitated by the lack of a second gene copy.

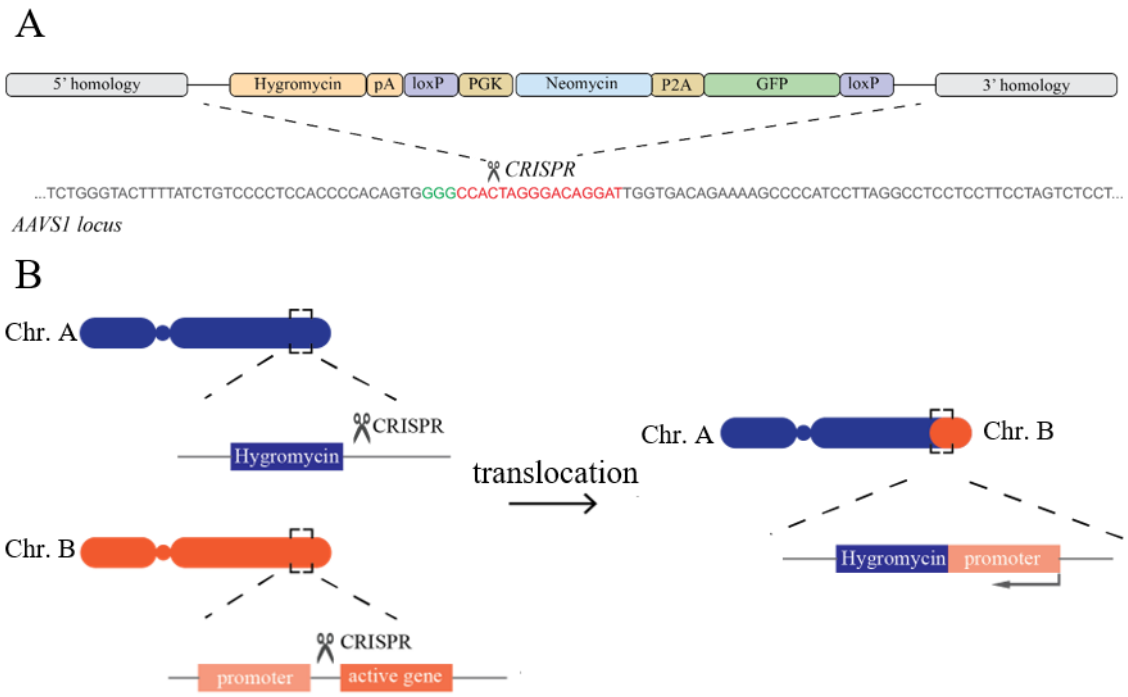


Figure 5. Principles of the translocation reporter system

A. Selection cassette integrated into the genome of HAP1 cells by the CRISPR/Cas9-mediated homology-directed repair. The cassette consists of: homology arms specific to the integration site, which is AAVS1 locus (in grey), promoterless Hygromycin resistance gene with polyadenylation sequence (pA), Neomycin resistance gene and GFP expressed as a multigene under PGK promoter and cleaved by P2A (porcine teschovirus-1 2A) peptide, 2 loxP sites positioned in the same orientation to be recognised by Cre recombinase. The sgRNA target sequence for AAVS1 locus is marked with red and its PAM sequence in green. Scissors icon shows the position of a CRISPR cut. **B.** General strategy of selecting translocation positive cells. DSBs are induced by CRISPR/Cas9 targeting sequences upstream the promoterless Hygromycin resistance gene (Chr. A) and downstream the active promoter (Chr. B). Translocation between those two broken chromosome ends results in activation of Hygromycin resistance gene expression under the target promoter, giving a selective advantage to cells positive for rearrangement.

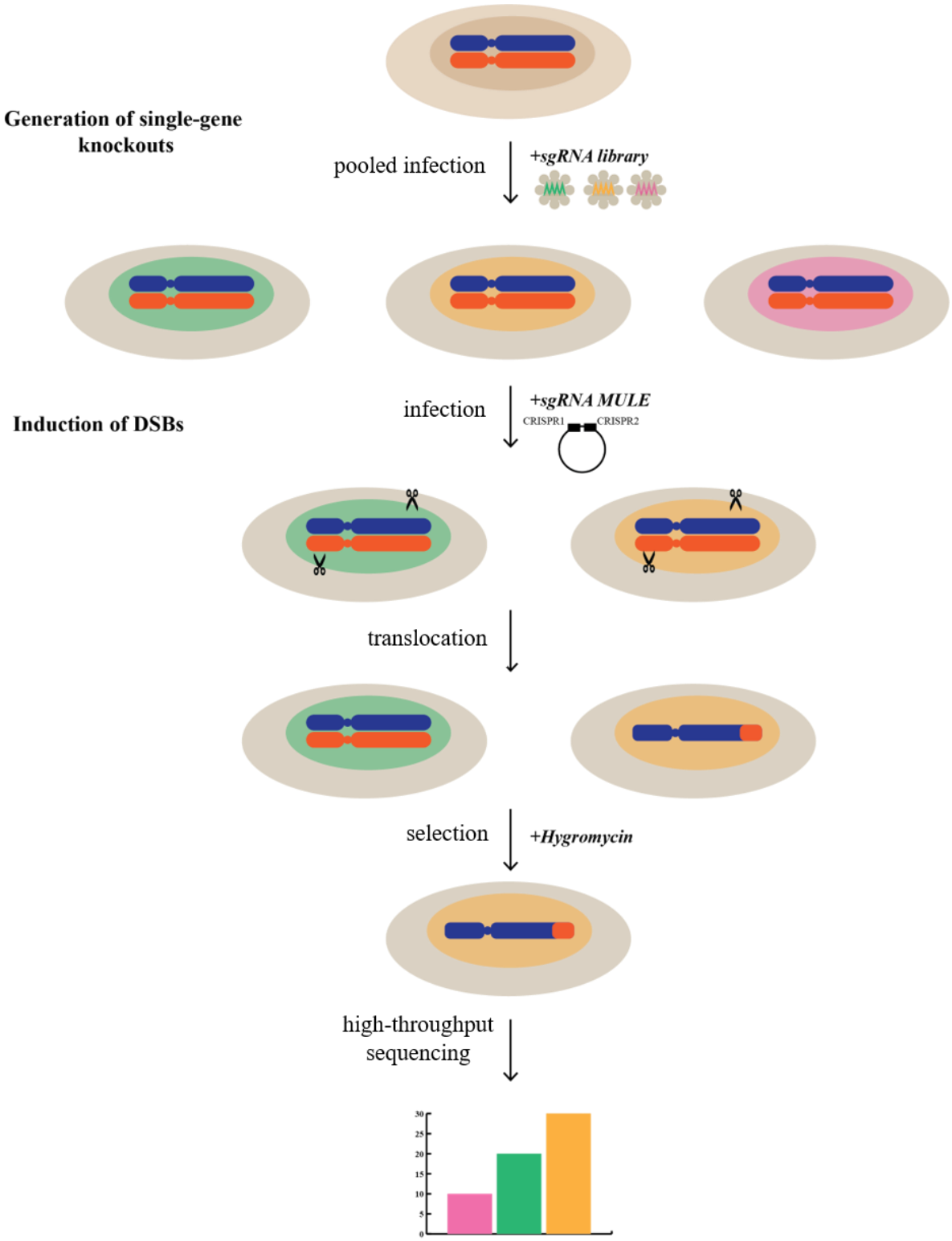


Figure 6. Design of the genome-wide loss of function screening for novel factors involved in the formation of chromosome translocations.

Genome-wide screening workflow. In cells with an integrated selection cassette, single-gene knockouts are generated by barcoded sgRNA pooled library through lentiviral infection. In knockout cells DSBs are induced as described in B. Cells positive for translocations are selected with Hygromycin. Abundance of each sgRNA is monitored in the initial pool, after infection and after antibiotic selection by high-throughput sequencing.

2.1.1. Generation of HAP1 cell line stably expressing the Cas9 protein

The cell line selected for performing a screen is the human near-haploid (HAP1). It was yielded in reprogramming experiments aiming to induce pluripotency in KBM-7 cell line isolated from a chronic myeloid leukaemia patient carrying the BCR-ABL fusion. HAP1 cells are adherent with fibroblast-like morphology. They have one copy of each chromosome, except heterozygous 30 Mbp fragment of Chromosome 15, which is fused to Chromosome 19 (Essletzbichler et al. 2014). HAP1 cells have been successfully used in multiple genome-wide screens, e.g. for the identification of host factors in virus infection, regulators of glutathione abundance and ferroptosis sensitivity, genes required for glycosylphosphatidylinositol biosynthesis (Fessler and Jae 2018; Cao et al. 2019; Rong et al. 2015).

CRISPR/Cas9 technology is a novel powerful genome editing tool that could be applied for multiple purposes throughout the project. In combination with sgRNA pool library, it can be used for generation of single-gene knockouts for a loss of function screening. Furthermore, it is used to induce the breakage within AAVS1 locus – a selection cassette integration site, to enhance the probability of the integration. Another application is to induce DSBs within translocation partner genes, which incorrect repair can lead to chromosome rearrangements. Finally, it is a technique of choice for creating frameshift mutations to obtain knock-out cell lines depleted in factors previously described to be involved in the formation of translocations.

To be able to utilise the CRISPR/Cas9 system, HAP1 cell lines expressing Cas9 protein, stably or in an inducible manner, were generated through lentiviral infections. Positive cells were identified by antibiotic selection and IF analysis of protein expression levels. All clones were tested for localisation of a protein exclusively in the nucleus. The most efficient control of inducible expression of Cas9 was achieved by 24h induction with 1 µg/ml of doxycycline and inhibited by washing it away from culture media, which resulted in complete loss of signal 72h later. T7 mismatch cleavage assay confirmed positive induction of breaks at targeted (AAVS1) locus, which is visualised as digested PCR product on a gel. Therefore, both cell lines can be successfully used for CRISPR/Cas9 based genetic engineering, though stable Cas9 protein achieved higher cutting efficiency and will be the version of choice in all the following experiments (Figure 7).

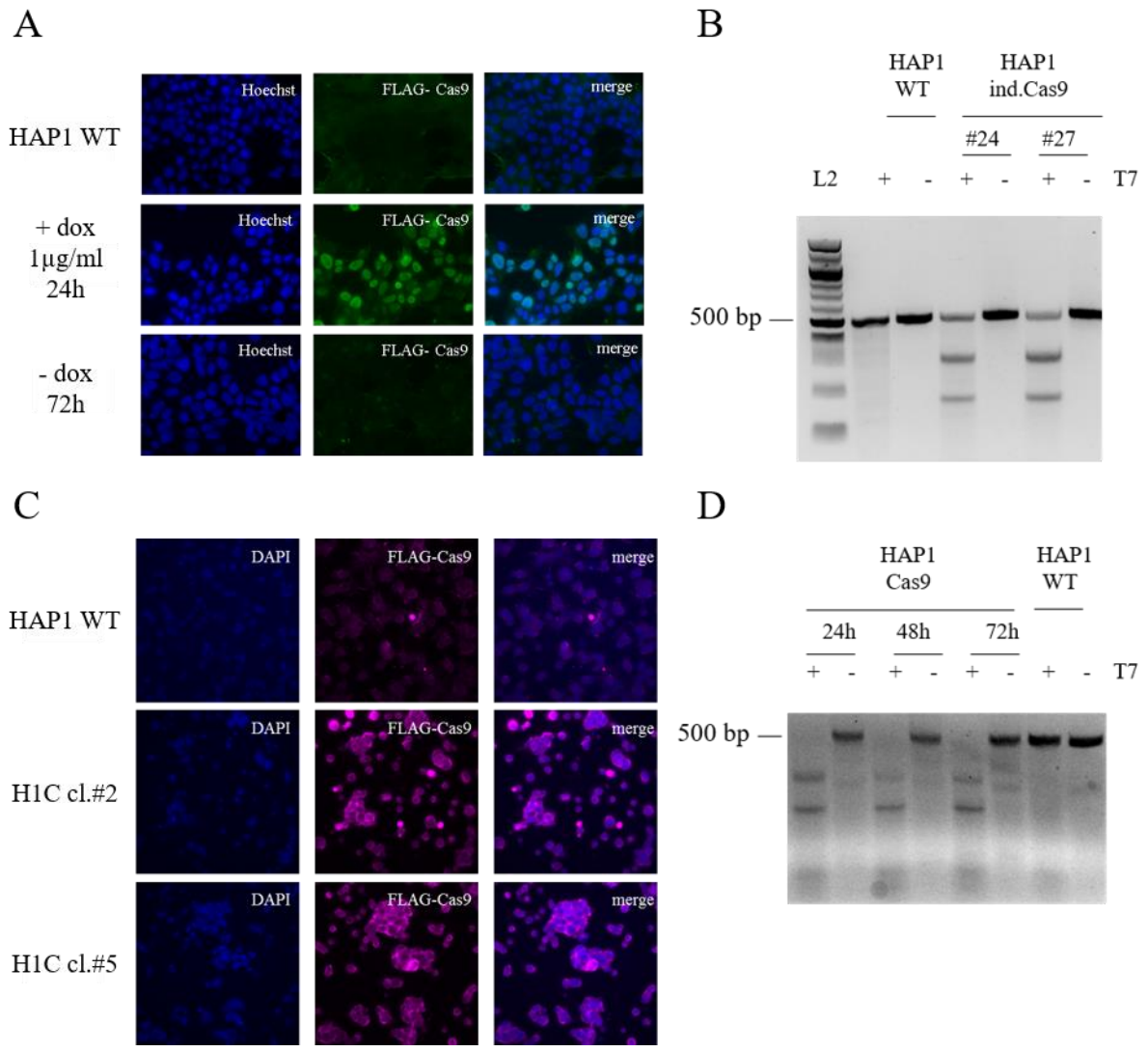


Figure 7. HAP1 cells stably express Cas9 protein.

A. and **C**. - the fluorescent images of IF staining using anti-FLAG M2 antibody. **A.** Induction and inhibition of a doxycycline-inducible expression of Cas9 protein in HAP1 cells. HAP1 cell line was infected with pCW-Cas9 encoding TetON-Flag-Cas9-Puro. The Cas9 expression was tested upon incubation with doxycycline (1 µg/ml /5 µg/ml) for 24h and 48h. The bottom panel shows the loss of expression 72h after washing doxycycline away. **C.** HAP1 cell line was infected with lentiCRISPR v2 encoding for EFS-NS-Cas9-FLAG-Puro. Positive clones were identified through antibiotic selection followed by IF. Images show Cas9 expression in WT (negative ctrl) and two example clones - #2 and #5. Enzyme activity for inducible **(B.)** and stably **(D.)** expressed Cas9 was validated by T7 assay. DSBs were induced by CRISPR with sgRNA targeting AAVS1 locus. Negative controls: HAP1 WT cells and reaction in absence of T7 enzyme.

2.2. Generation of the translocation reporter system

2.2.1. The selection cassette is integrated into AAVS1 locus of HAP1 genome

To make the selection of translocation positive cells possible, first the selection cassette containing promoterless Hygromycin-resistance gene had to be integrated into the genome of HAP1 cells.

The selection cassette is integrated into adeno-associated virus integration site 1 (AAVS1) locus by CRISPR/Cas9-mediated homology-directed repair. AAVS1 is present on chromosome 19q13. It has been chosen due to its open chromatin structure, transcriptionally competent environment and presence of native insulators (J. R. Smith et al. 2008). The cassette consists of homology arms specific to AAVS1 locus, promoterless Hygromycin-resistance gene, reporter genes - Neomycin resistance and GFP, and 2 loxP sites recognized by Cre recombinase for excision of selected cassette fragment (Figure 5A).

To integrate the selection cassette into the genome of previously generated HAP1-Cas9 cell line cells were cotransfected with a plasmid containing a specific sgRNA to induce a breakage within AAVS1 locus (pLX-sgRNA) and a second plasmid containing the cassette (AAVS1hygroNEOP2aGFP). Clones positive for integration (now named H1C for HAP1 with Cassette) were selected with Neomycin, assessed for GFP expression and further tested by PCR amplification. Primers were specific for sequence within the selection cassette or for sequences upstream and downstream of the integration site. Additionally, the sequences of those PCR products were analysed confirming the correct insertion (Figure 9). It is noteworthy to mention that the possibility of the off-target cassette integration in another genomic location was not tested.

2.2.2. Excision of neomycin resistance gene and GFP from the selection cassette by Cre recombinase

Due to multiple infection steps included within a screening setup, various antibiotics for mammalian resistance are required: Hygromycin is used for selection of translocation positive cells, Blasticidin is used in sgRNA pool libraries, Puromycin was applied in creating stable Cas9 cell line. In order to be able to utilise Neomycin resistance and GFP at later stages of a project, both genes were excised from the genome by Cre recombinase, which recognises loxP sites present in the selection cassette. H1C cells were transfected with pLM-cmv-R-Cre plasmid. Positive clones (now named H1Cx for HAP1 with Cassette excised) were identified by fluorescence analysis as a loss of GFP signal and gain of mCherry expression. Additionally, a fragment flanked by loxP sites was amplified by PCR and successful modification was proven by a reduced size of PCR product in comparison with control cells (H1C), which were positive for cassette integration but not expressing Cre (Figure 10).

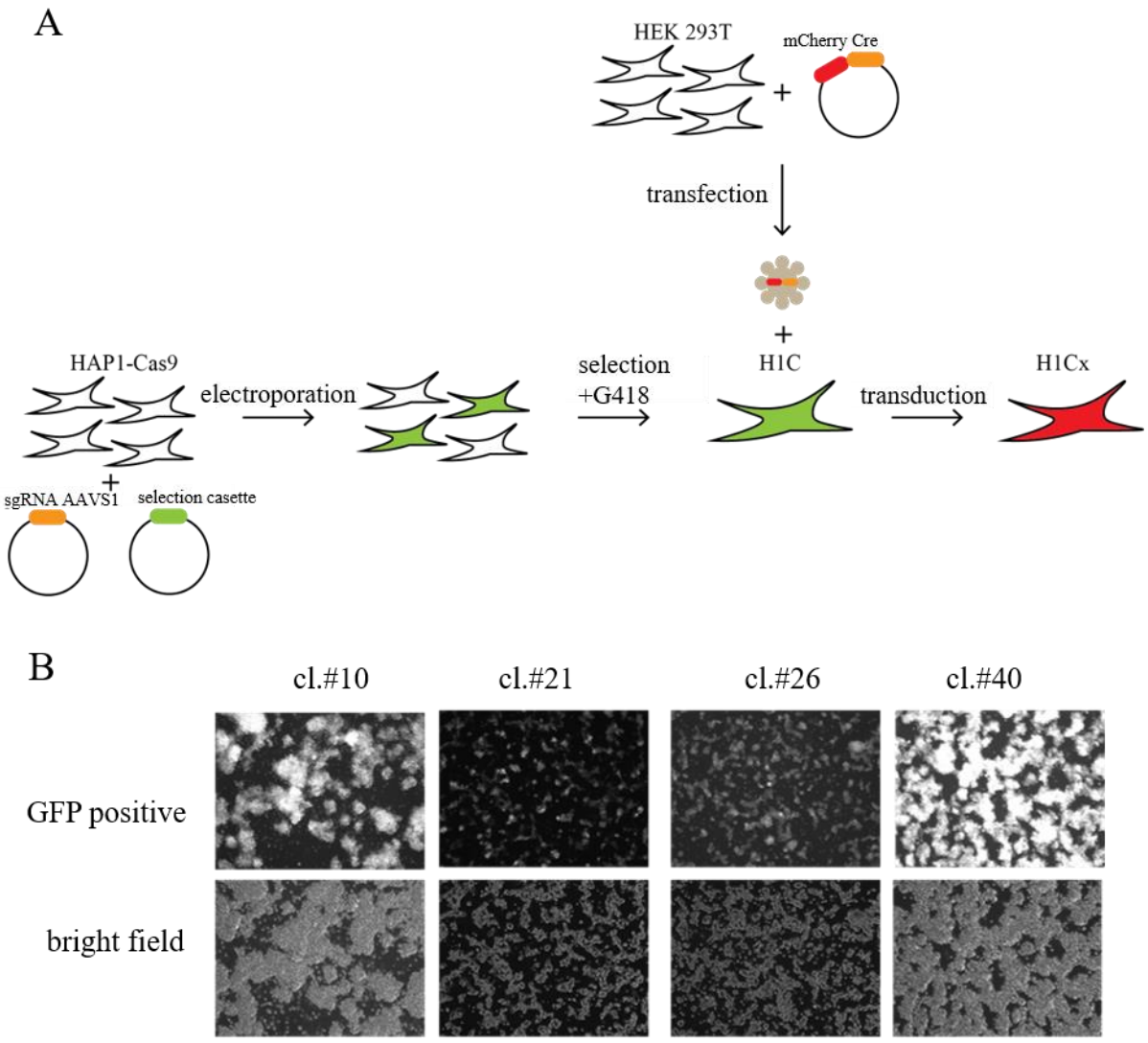


Figure 8. The selection cassette was integrated into HAP1 genome.

A. Workflow of cassette integration. HAP1-Cas9 cells were cotransfected with plasmid carrying sgRNA targeting AAVS1 locus (pLX-sgRNA) and second one carrying the selection cassette (AAVS1hygroNEOP2aGFP) via electroporation with Neon® Transfection System. Positive clones were identified through selection with G418 (800 µg/ml) and presence of GFP signal. Integrated cassette in H1C cells was modified by transduction with virus containing mCherry-Cre (pCMV_R_Cre). Positive cells were selected based on loss of GFP expression and gain of mCherry signal. **B.** H1C cell line - fluorescent and bright field images of example GFP positive clones after selection with G418.

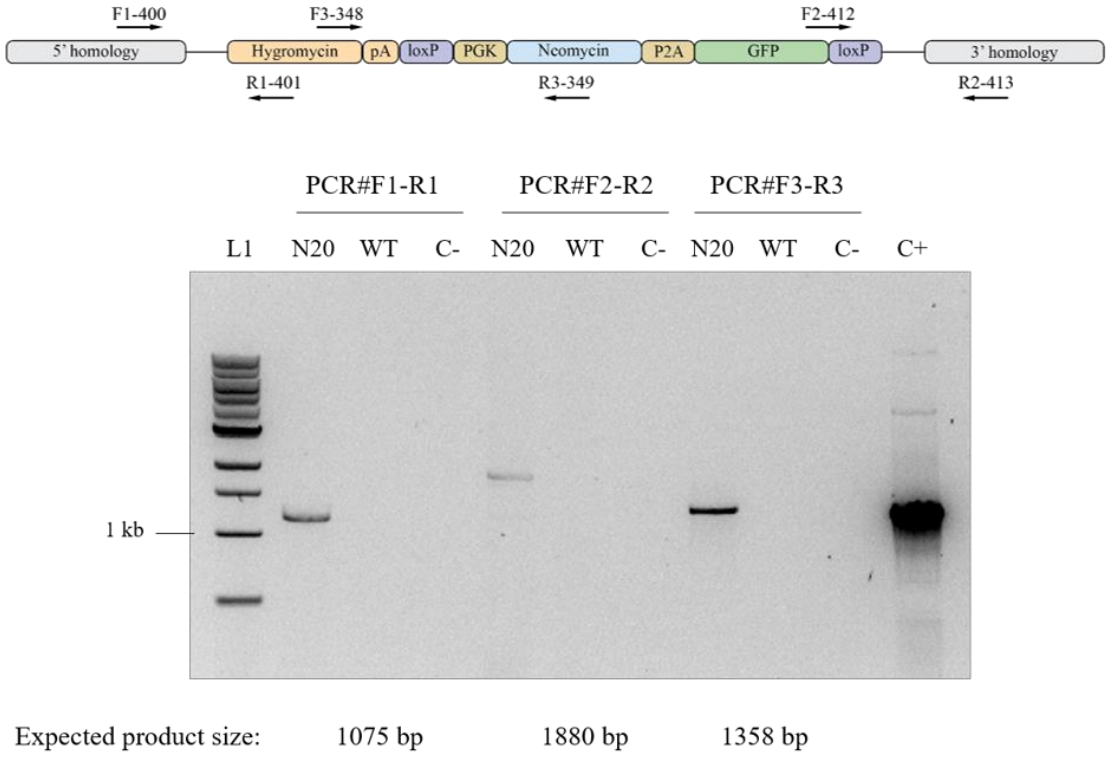


Figure 9. Genomic analysis of H1C clone positive for integration of selection cassette.

gDNA from a Neomycin resistant and GFP positive H1C clone was analysed by PCR to check the correct integration into AAVS1 locus. Primers for PCRs are located within the cassette sequence (PCR#3) and at the flanking regions at the 5' (PCR#1) and 3' (PCR#2) end of a cassette. Positive control (C+) is an amplification of a plasmid sequence containing the selection cassette (AAVS1hygroNEOP2aGFP).

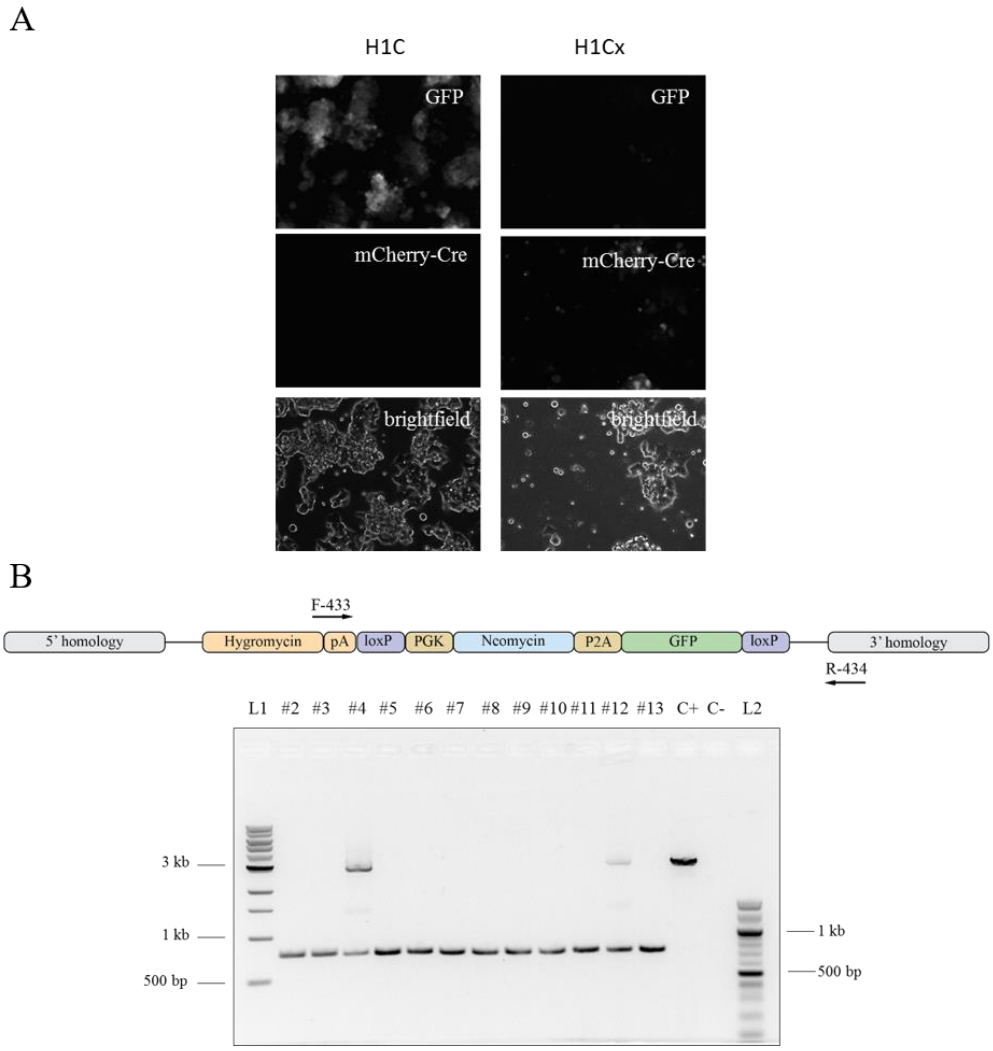


Figure 10. The selection cassette integrated into HAP1 genome was modified with Cre recombinase.

A. H1Cx cell line – fluorescent and bright field images of example clone with loss of GFP and gain of mCherry signal. **B.** Gene excision was validated by PCR analysis on genomic DNA from H1C and H1Cx clones (L1-1 kb DNA ladder, L2 – 100 bp DNA ladder, #2-#13 – H1Cx tested clones, C+ (positive control-H1C) – amplification of the target locus in the clone with full length cassette integration, C- (negative control) – no template PCR reaction).

2.3. Generation of cell lines depleted in DSBs repair factors

In order to prove that the established system can be applied to identify novel factors involved in translocations biogenesis, positive control cell lines were created. They were impaired in well-studied DSBs repair factors, such as Lig4, nucLig3 and Parp1, which were previously shown to decrease the frequency of translocations (Lig4 and Parp1 in human and Lig3 in mouse cells). Lig3 is translated from two start sites resulting in existence of mitochondrial and nuclear isoforms. Mitochondria isoform of Lig3 is essential for cell viability, therefore could not be completely depleted (Simsek and Jasin 2010; Simsek, Brunet, et al. 2011; Wray et al. 2013).

The knock-out cell lines were obtained by applying CRISPR/Cas9 system to create frame-shift mutations in targeted genes sequence. The correct clones were selected with Blasticidin (5 μg/ml) and further confirmed by immunofluorescence assay with antibodies against specific factor, showing a significant reduction of their endogenous expression levels. Additionally, the regions targeted by sgRNAs were amplified by PCR on genomic DNA, and the sequencing confirmed the introduction of frameshift mutations (Figure 11).

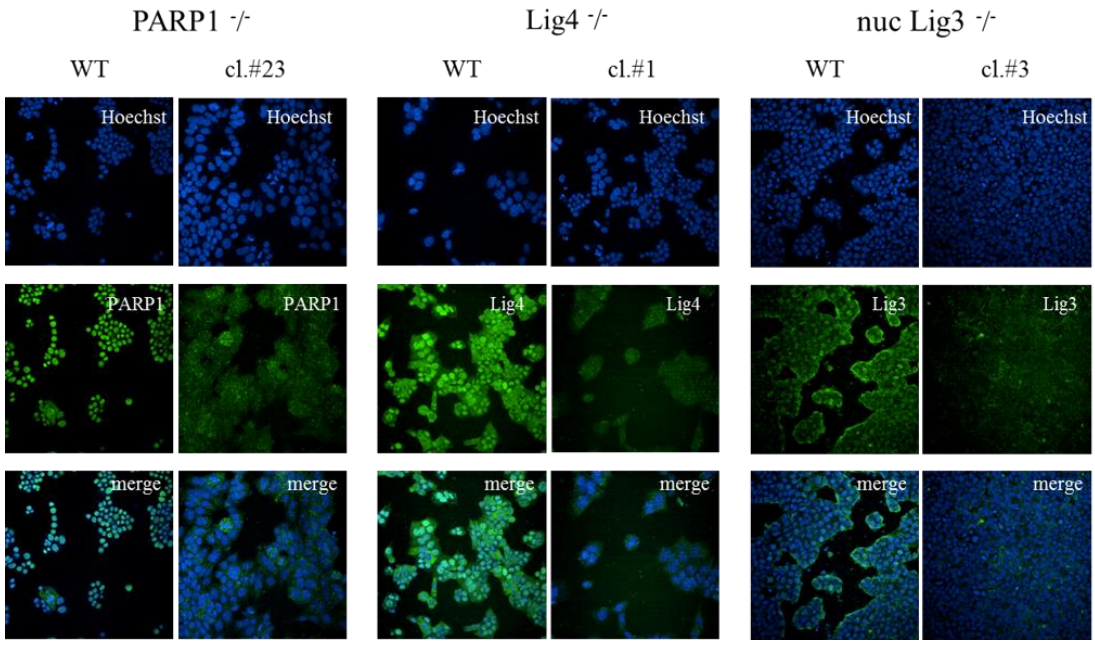


Figure 11. H1Cx knockout cells are deficient in Parp1, Lig4 and nuclear Lig3.

Frame shift mutations in the coding regions for targeted genes were induced by CRISPR/Cas9. Cells were infected with pLX-sgRNA encoding for U6-sgRNA-Blast by lentivirus. Positive clones were identified by Blasticidin selection and positive depletion of protein expression, measured by IF. The figure shows immunofluorescence staining of Parp1, Lig4 and nuclear Lig3 using anti-PARP, anti-Lig4 and anti-Lig3 antibodies respectively in WT and tested mutants.

2.4. Optimisation of the screening conditions

The successful performance of the screening depends on various factors, which include correct system design and optimal experimental conditions. Distinct steps of the procedure had to be planned in detail and tested in order to ensure the best possible parameters.

First of all, I selected the specific active promoters, which serve as translocation partners. Secondly, since the rates of translocations depend on DSBs occurrence, I designed various sgRNAs and tested their cutting efficiency. Additionally, the lentiviral transduction of cells with plasmids carrying those sgRNAs has to happen with high efficacy, so I established the parameters of this method. The important part of the screening process is the selection with Hygromycin, therefore I set up its optimal concentration and duration of the treatment. The resistant colonies need to be healthy and exhibit a growth potential to be able to reliably estimate translocations frequencies and collect material for high-throughput sequencing. To ensure proper growth conditions for HAP1 cells, I studied various types of medium and culturing parameters. Finally, I determined the frequency of translocations that can be achieved with the system.

2.4.1. Selection of promoter partners

In order to induce genomic rearrangements, DSBs at a specific locus are created with CRISPR/Cas9 system. First one is induced upstream of the promoter-less Hygromycin-resistance gene and the second one downstream of the promoter of an active gene.

To choose promoters which serve as translocation partners the following selection criteria were applied:

- the promoter should drive the expression of an active and not essential gene
- it should be located on a chromosome in close proximity to chromosome 19, where the selection cassette is integrated (due to the fact that translocations occur with a higher frequency between broken ends which are within shorter distance (Meaburn, Misteli, and Soutoglou 2007; Zhang et al. 2012)
- it should be in correct orientation on a chromosome arm

To identify active genes, available RNA-seq data for HAP1 cells was analyzed. 10% of the most highly expressed genes in this cell line were identified (Figure 12). The essentiality of genes was determined in a genome-wide trap insertion study, in which disruptive mutations are

integrated into the intronic regions and prevent exon-exon splicing (Blomen et al. 2015). Using this information, I prepared the list of most highly expressed and nonessential genes in HAP1 genome. To measure the spatial proximity of chromosomes, FISH using individual probes for the genes located on different chromosomes was applied (chr4-AF4, chr9-AF9, chr12-C1S, chr6-DSP, chr19-ENL, chr11-MLL): one probe always labelling AAVS1 locus and another one specific for a gene on a different chromosome (Figure 12A). Figure 12B shows the analysis of distances between probes for different chromosome combinations. Among tested one's, chromosome 11 and 12 are in the closest proximity to chromosome 19. Therefore, I selected the genes located on those chromosomes as possible translocation partners.

Various types of chromosome rearrangements, such as insertions, deletions or translocations can be created by targeting promoters in different locations in the genome. In order to create inversion or deletion, DSBs are induced at two sites on the same chromosome. On the other hand, inducing a translocation requires breaks at two distinct chromosomes. Initially, I selected genes on chromosome 12 as translocations partners – FKBP4 and PTMS. However, due to poor performance of a system associated with low translocation frequencies I decided to change the target promoter. The final promoter selected as a translocation partner drives expression of LDHA gene located on Chr11, which is one of the most highly expressed nonessential genes in HAP1 genome. Inversion partner comes from the PPP5C gene on Chr19, downstream the selection cassette and on the opposite strand. The ~300 kb long deletion is obtained by targeting the promoter of ISOC2 gene, which lies upstream the selection cassette and on the same strand on Chr19. The expression levels of LDHA, PPP5C and ISOC2 in HAP1 cells are 507, 77 and 51 respectively.

sgRNAs targeting each promoter were designed using CRISPR Design Tool and cloned into MULE plasmids with either G418 or GFP resistance. Lentiviral infection of cells with plasmids carrying both sgRNAs enables simultaneous induction of DSBs upstream the Hygromycin-resistance gene and downstream the promoter partner.

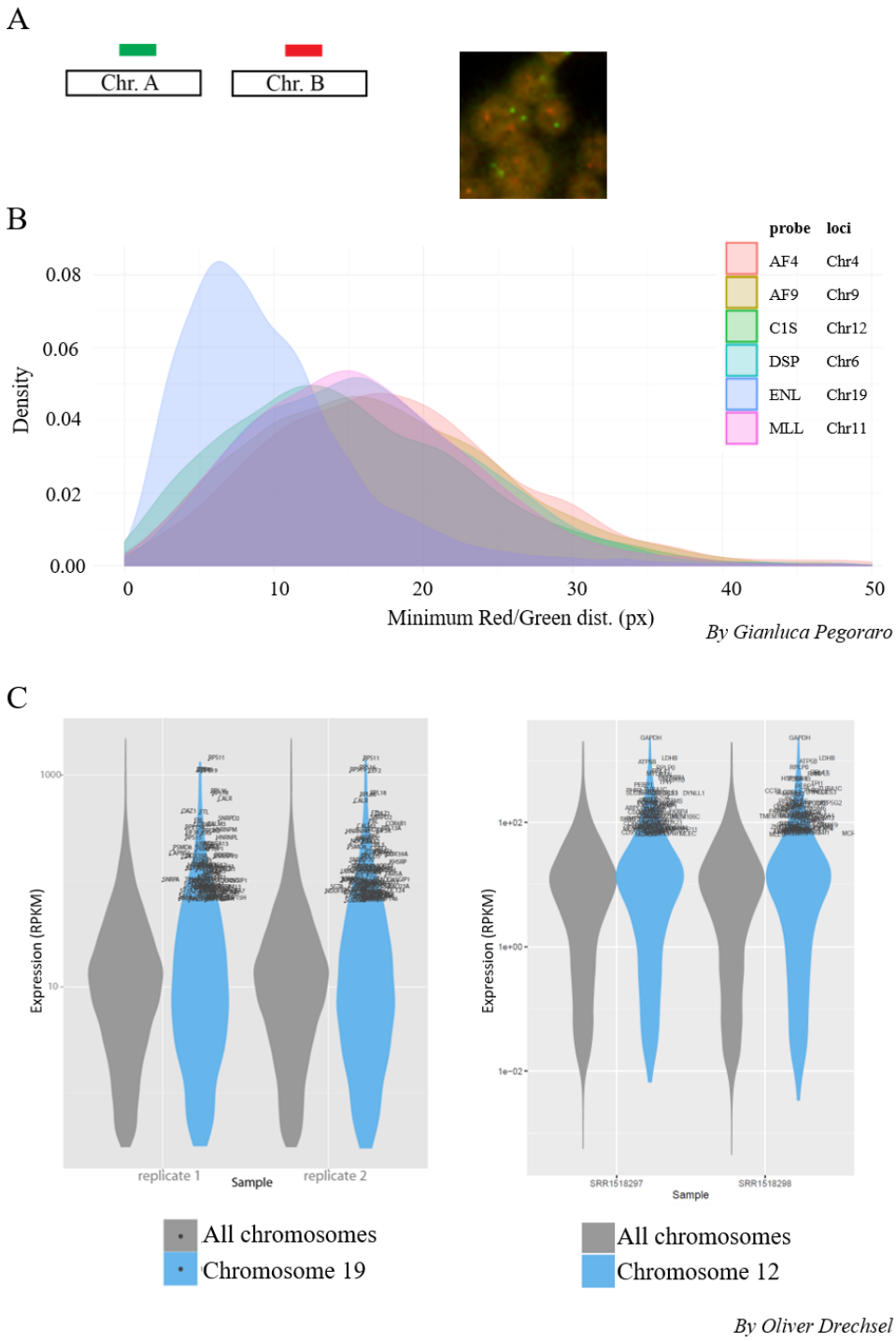


Figure 12. The selection of translocation promoter partners.

A. and **B.** The proximity events between AAVS1 locus and other chromosome loci on different chromosomes were analysed by FISH. The minimum distance between green (AAVS1) and red (other loci) probes was measured. **C.** Bioinformatics analysis of RNAseq data of HAP1 cells performed to provide a list of highly expressed genes. Violin plots present the expression of genes (with RPKM>0.3) in HAP1 cells (grey area: genes in all chromosomes, blue area: genes in chromosome 19/12, names of the 10% highest expressing genes of Chr19/12 are shown.

2.4.2. Improving efficiency of transduction with lentiviral particles carrying sgRNAs targeting sequence of interest

Translocations are very rare events, so increasing their frequency in a system is an important optimisation step to improve the performance of a screening. The main prerequisite is an efficient induction of DSBs. This can be ensured by obtaining high rates of transduction efficiency with plasmids carrying sgRNAs targeting the sequence of interest. H1Cx cell line was infected with plasmids carrying sgRNAs targeting 2 selected loci, which also encode for GFP. 5 days after infection successfully transduced cells expressed GFP and the protein expression levels were analysed by fluorescent microscopy and automatic image analysis. Efficiencies of transduction with different plasmids in two different cell lines (WT and Lig4 $^{-/-}$) were compared. I observed only small differences in percentage of GFP positive cells between different plasmids (<30%), but the results between two cell lines were almost identical (differences <10%). This shows that regardless of the plasmid used and the cell line transduced, I was able to obtain high and reproducible transduction rates reflected as 70%-99% of GFP positive cells (Figure 13).

2.4.3. Confirmation of efficient DSBs induction was confirmed by T7 assay

The successful transduction does not directly prove the efficient DSBs induction. It is important to confirm that applied sgRNAs can effectively create a breakage at the targeted locus. For that purpose, I applied the T7 mismatch cleavage assay. In this test, a genomic DNA extracted from a pool of cells, 5 days after infection, was used for amplification of the modified sequence. PCR products were denaturated and reannealed, so that homo- or heteroduplexes were formed. T7 endonuclease recognises and cleaves only mismatched strands. Digested products can be visualized on an agarose gel. I compared positive samples to the ones not treated with T7 endonuclease and to not transfected controls in which no digestion of PCR product should be observed. The results confirm that regardless of the plasmid used and the cell line transduced, efficient cut of targeted loci was achieved in all positive samples. Because of high transduction efficiencies and good sgRNA cuts, I expected to obtain sufficient frequency of genomic rearrangements in assay condition (Figure 13 and Figure 14).

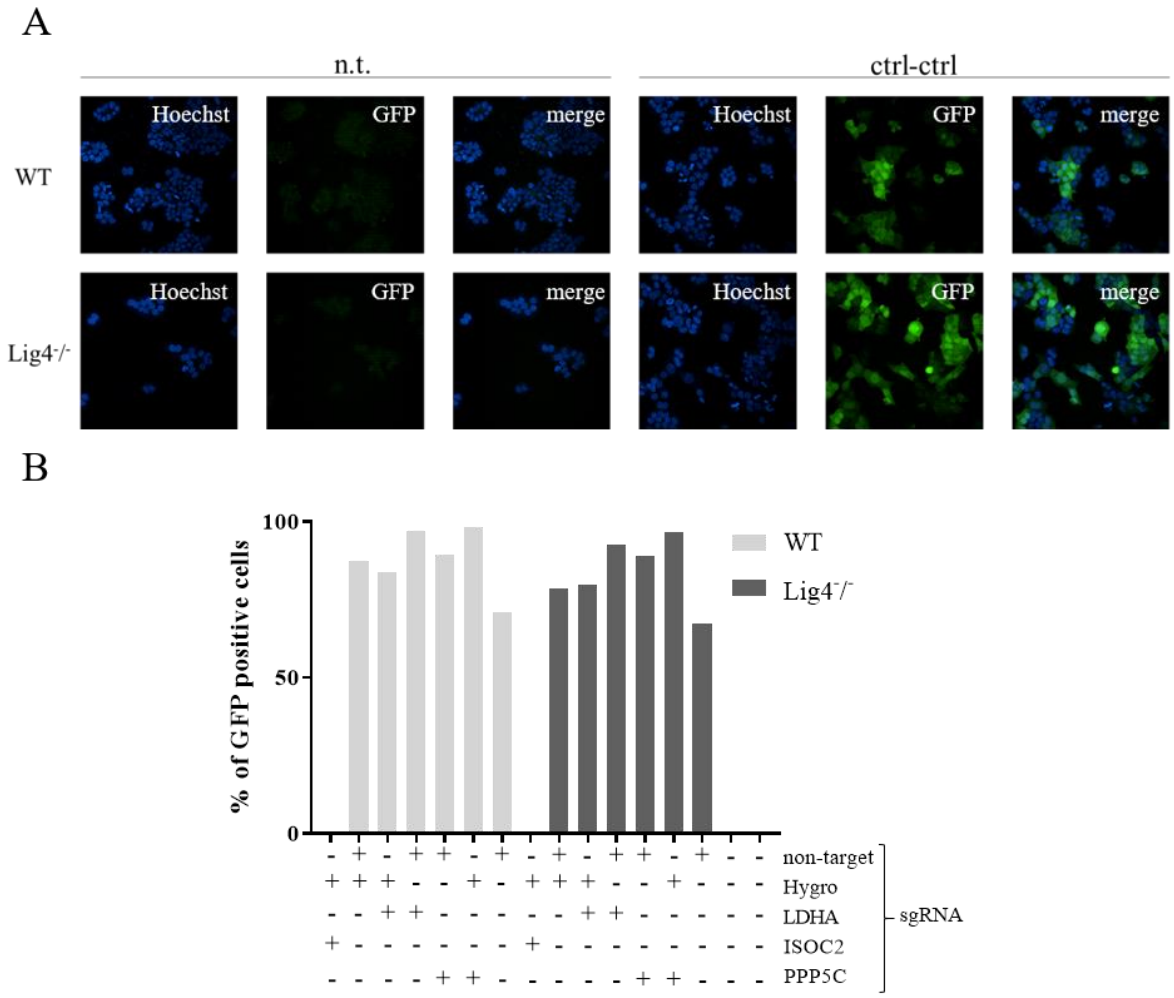


Figure 13. Efficient transduction with lentiviral particles carrying sgRNAs targeting sequence of interest.

A. and **B.** Efficiency of transduction was measured by percentage of GFP positive cells in WT and Lig4 $^{-/-}$ samples transduced with distinct GFP MULE expression plasmids encoding for 2 sgRNAs: first one targeting region upstream Hygromycin resistant gene (Hygro) and second one targeting the promoter of translocation partner (ISOC2/LDHA/PPP5C). Figure **A.** shows example fluorescent images of GFP negative (not transduced) and GFP positive (2 ctrl sgRNAs) cells. **B.** Comparison of transduction efficiency in WT and Lig4 $^{-/-}$ cells.

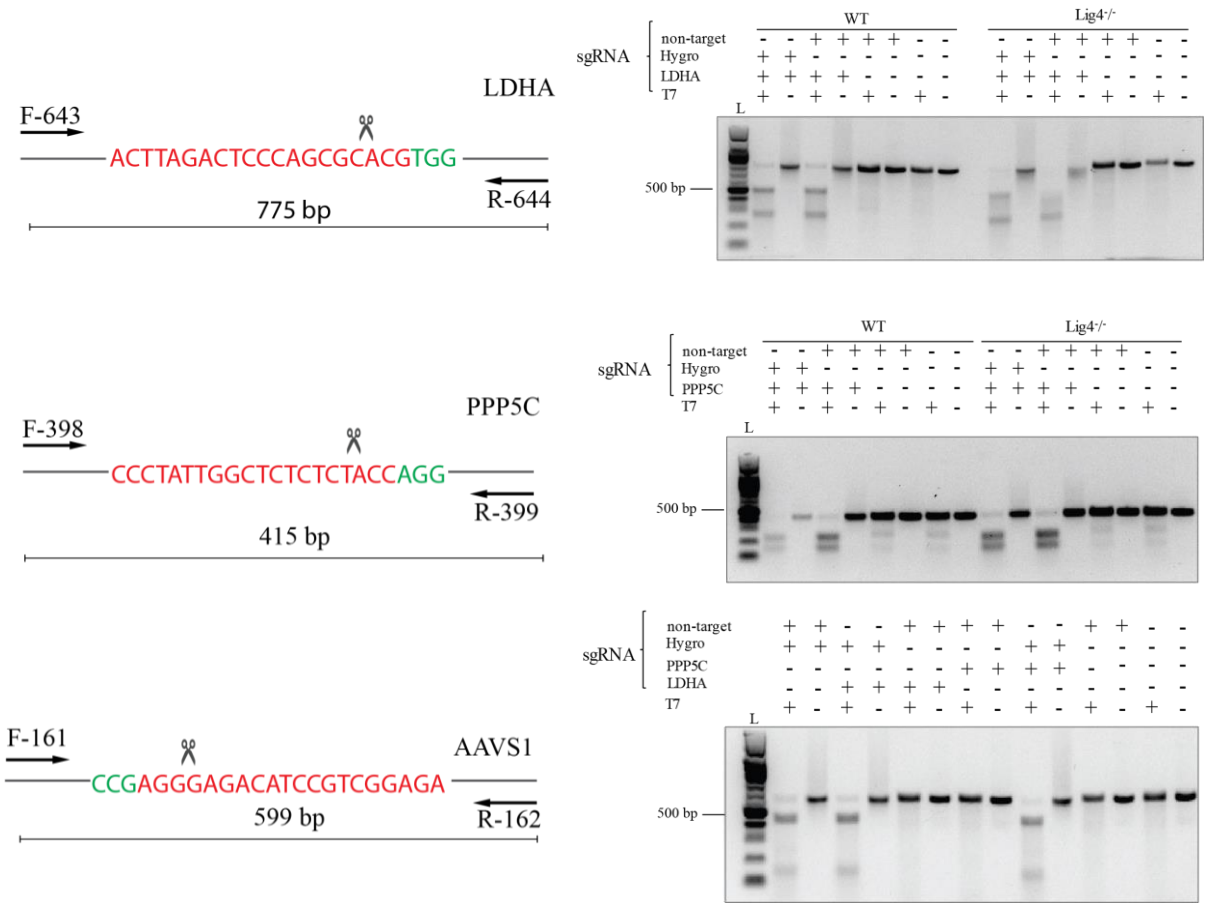


Figure 14. Efficient induction of DSBs with CRISPR/Cas9 was confirmed by T7 assay.

T7 mismatch cleavage assay on gDNA from samples described in Fig. 13. sgRNA target sequence is marked in red and PAM sequence in green. Scissors icon shows the Cas9 cutting site. Length of T7 PCR product in bp is displayed below. Effective induction of a breakage in targeted locus is detected on an agarose gel as digested PCR product in the presence of T7 endonuclease. Negative controls are: samples without T7 treatment, infected with non-target sgRNAs or not infected at all.

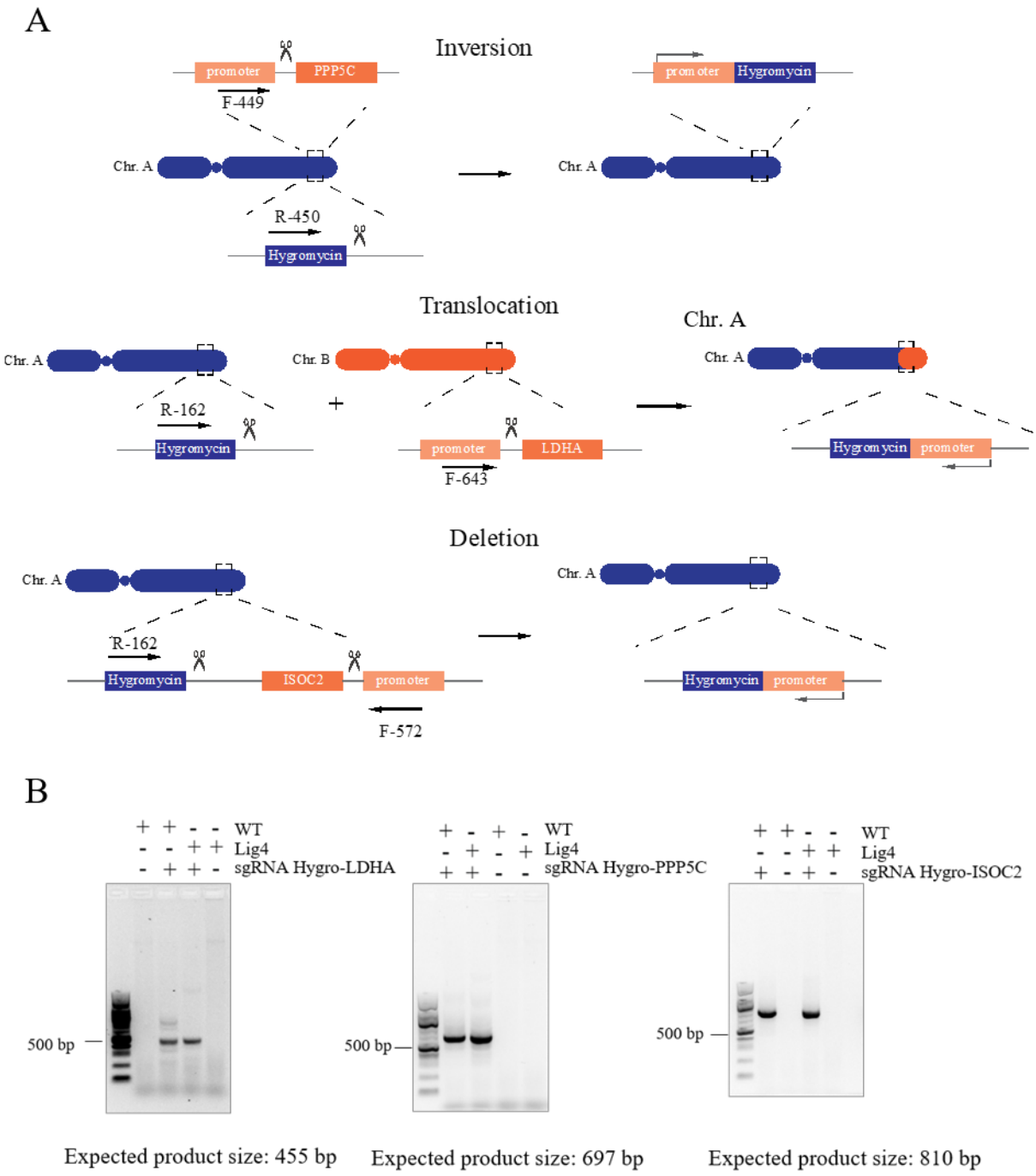


Figure 15. Induction of chromosome rearrangements by CRISPR/Cas9.

A. and **B.** Translocation, inversion and deletion were successfully induced which was detected by PCR amplification on genomic DNA, collected from the pool of cells 4 days after viral transduction with sgRNAs. Forward primers were specific to promoter partner and reverse primers were specific for selection cassette. Samples from not infected WT and Lig4 $^{-/-}$ cells state as negative controls.

2.5. Determination of optimal growth condition of Hygromycin-resistant colonies

The experimental workflow of the system optimisation started with the delivery of MULE sgRNA plasmid into H1Cx cells. Both transfection rates and sgRNA cutting efficiencies were shown to be optimal. Five days later selection with Hygromycin was initiated. The duration of antibiotic treatment, which was 14 days, was established based on the provided product data sheet and confirmed experimentally as a minimum time required for the death of control cells (translocation negative). At the end of the selection process, the expression of the reconstituted Hygromycin-resistance gene would confer resistance only to cells with the translocation between active promoter and selection cassette.

In the course of the experiment I encountered multiple challenges at this step. First of all, the number of obtained resistant colonies was much lower than required for performing the screening, suggesting a low frequency of induced translocations. Secondly, the number of colonies was not reproducible between technical replicates of the experiment. Another problem was that the majority of surviving colonies was unhealthy and was easily detaching from the plate, making their count unreliable. To overcome those challenges and improve growth conditions of resistant colonies, I further optimised the Hygromycin concentration and treatment duration. For that purpose, a cell line stably expressing Hygromycin-resistance gene was generated as a positive control (Hygro+). Antibiotic concentrations from 200 to 500 μg/ml were tested, and 350 μg/ml was chosen as an optimal one, since it effectively eradicated WT cells and didn't negatively influence the growth of Hygro+ cells. The experimental workflow includes multiple passaging rounds. To improve the general health of cells, I examined different ways of detaching them from the plate surface, which included detachment factors such as: Trypsin-EDTA, Accutase and mechanical forces. Application of either Trypsin-EDTA or Accutase resulted in smaller loses of cell number. To improve the cell adhesion I compared coating substances like poly-L-lysine (PLL) and 1% gelatine. In both conditions, cells were more attached to the plate, but PLL coating made it difficult for detachment. For this reason, 1% gelatine was selected for coating. All the tests were performed on uninfected cells and the positive impact of any implemented changes wasn't observed in cells after delivery of MULE sgRNAs. The next attempt was to exchange the cell culture medium from regular liquid to soft-agar, in order to eliminate the problem of losing cells by detachment from the surface. However, I have not found the conditions leading to the concomitant survival of Hygro+ cells and death of WT cells. This observations led me to the conclusion that current workflow conditions: cell

number, cell culturing, virus titer or Hygromycin concentration are not optimal and require a revisit in a more systematic way.

2.6. Detection and quantification of chromosome rearrangements by PCR

Because growing and expanding resistant colonies was difficult, I decided to confirm that chromosome rearrangements formed on genomic level. I tested this by PCR amplification of genomic DNA from the pool of cells harvested 5 days after transduction. For each type of rearrangement (translocations, insertions, deletions) a different set of primers was designed: the forward primer was always specific to a sequence within a cassette and the reverse primer specific to a promoter partner. All types of rearrangements were detected in both WT and Lig4$^{-/-}$ cells. This result proves that the unsuccessful selection of resistant colonies was not due to lack of rearrangements induced and it required further optimisation of assay conditions (Figure 15).

Problems with maintaining healthy cells made it also impossible to estimate the frequencies of rearrangements by counting resistant colonies. To overcome this difficulty, knowing that rearrangements can be detected by regular PCR, I quantified their frequencies by digital droplet PCR. The results show that translocations created with the developed system occur in 1:200 cells. This value was reproducible between 3 biological replicates and comparable between WT and Lig4$^{-/-}$ cells (Figure 17). This points out again that low numbers of resistant colonies obtained were strictly due to the experimental parameters and not caused by small rate of creating rearrangements.

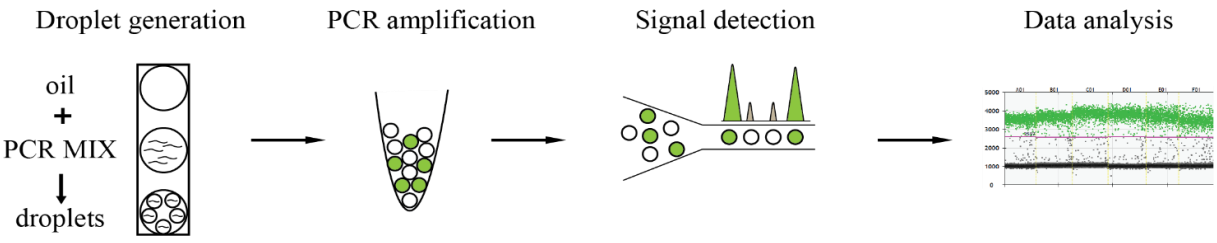


Figure 16. Workflow of DNA amplification with ddPCR.

PCR reaction mix containing template DNA and fluorescent probes is partitioned into water-oil droplets with a Droplet Generator. The amplification takes place in each individual droplet and is performed in regular thermocycler. The presence of a product is detected by fluorescent signal read by Droplet Reader. The absolute copy numbers are quantified with QuantaSoft software.

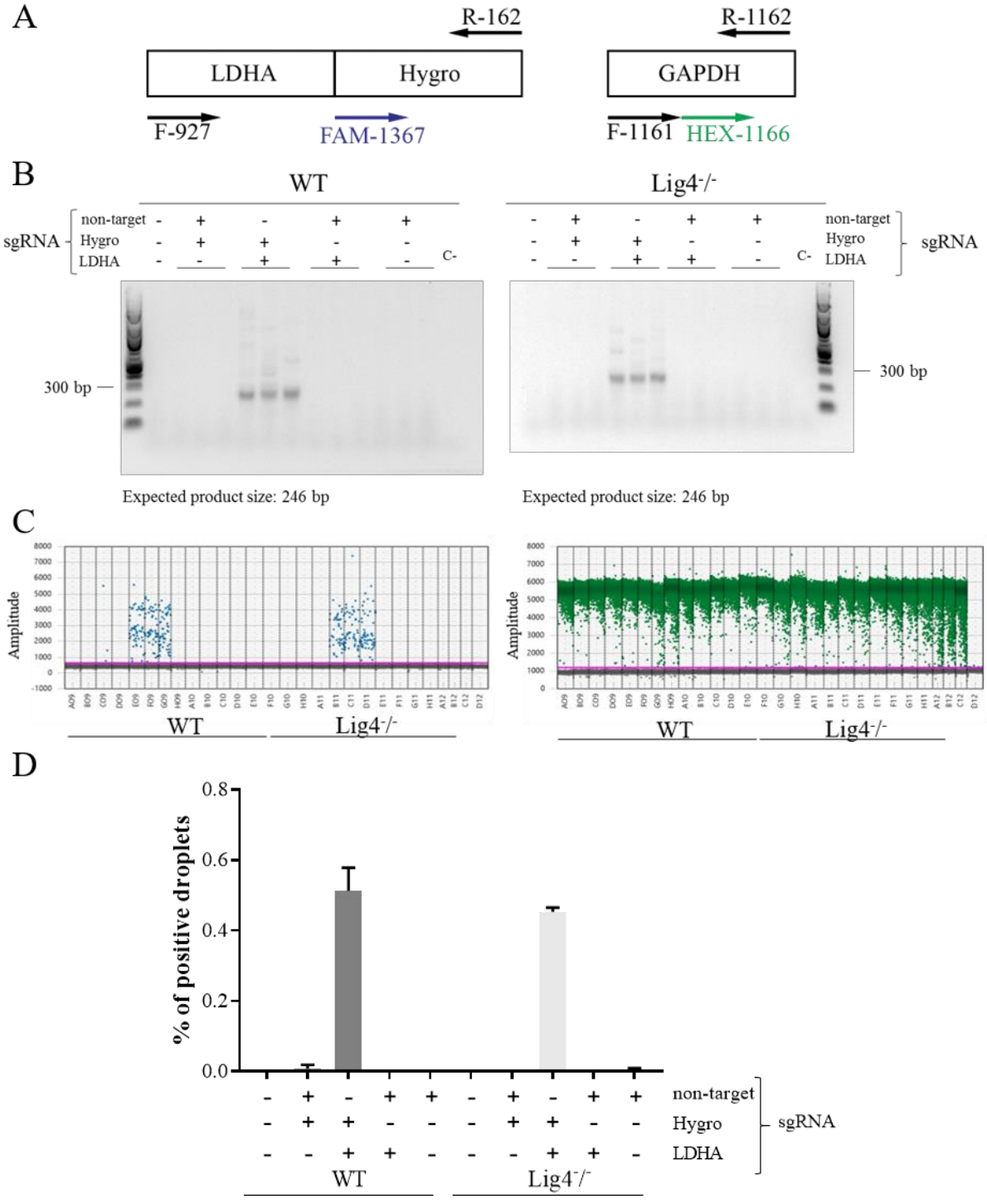


Figure 17. Quantification of the frequency of chromosome translocations by ddPCR.

A. Strategy of PCR primers design. FAM-labeled probe is specific to translocation product (Hygro-LDHA) and HEX-labeled probe to GAPDH used for normalisation. **B.** Conventional PCR amplification of genomic DNA from pool of cells transduced with plasmids carrying two sgRNAs. **C.** Visual representation of target detection using TaqMan-based ddPCR. The positive (blue for FAM-Hygro, green for HEX-GAPDH) and negative droplets (grey), classified by the QuantaSoft analysis of individual wells with manually adjusted threshold (pink), are shown. **D.** Absolute quantification of target positive droplets presented as a fraction after normalization to GAPDH signal. Error bars are the standard deviation of 3 biological replicates.

2.7. Analysis of breakpoint junctions

The ultimate goal of this project was to shed light on the mechanisms behind the formation of chromosome translocations. Looking at the repair products gives additional information about the possible pathways involved in their processing. For this reason, I analysed the sequences of T7 and translocation PCRs products. In the case of intrachromosomal repair, all types of repair outcomes such as perfect repair and introduction of insertions and/or deletions were observed. Majority of them (64-80%) being deletions of up to 400 bp in size. Insertions and accompanying presence of insertion and deletion (indels) account for ~20% of all types of outcomes. Only a small minority of up to 7% was the perfect repair of the locus. The only difference observed between WT and Lig4 $^{-/-}$ condition was the slightly bigger length of deletion in mutant cells. Interestingly, for the interchromosomal junction, represented here as a translocation between the Hygromycin-resistance gene and promoter of the LDHA gene, no events of perfect repair or insertions were recorded. Same as for intrachromosomal repair, the majority of sequences included deletions of up to 200 bp of length, slightly longer in Lig4 $^{-/-}$ mutant samples than in WT (Figure 18). Those results are in agreement with current state of knowledge that chromosomal translocations are not primarily formed by HR, but are mostly products of NHEJ, since majority of obtained outcomes included insertions and deletions. The fact that deletion length in Lig4 mutant cells was larger than in WT proves again that in absence of Lig4, extensive resection occurs and a-EJ is a dominant repair pathway (Ghezraoui et al. 2014). The differences observed between intra- and interchromosomal junctions have been seen before in mouse cells (Simsek, Brunet, et al. 2011) but not in human cells (Ghezraoui et al. 2014) and it was suggested that their repair is mediated by distinct mechanisms.

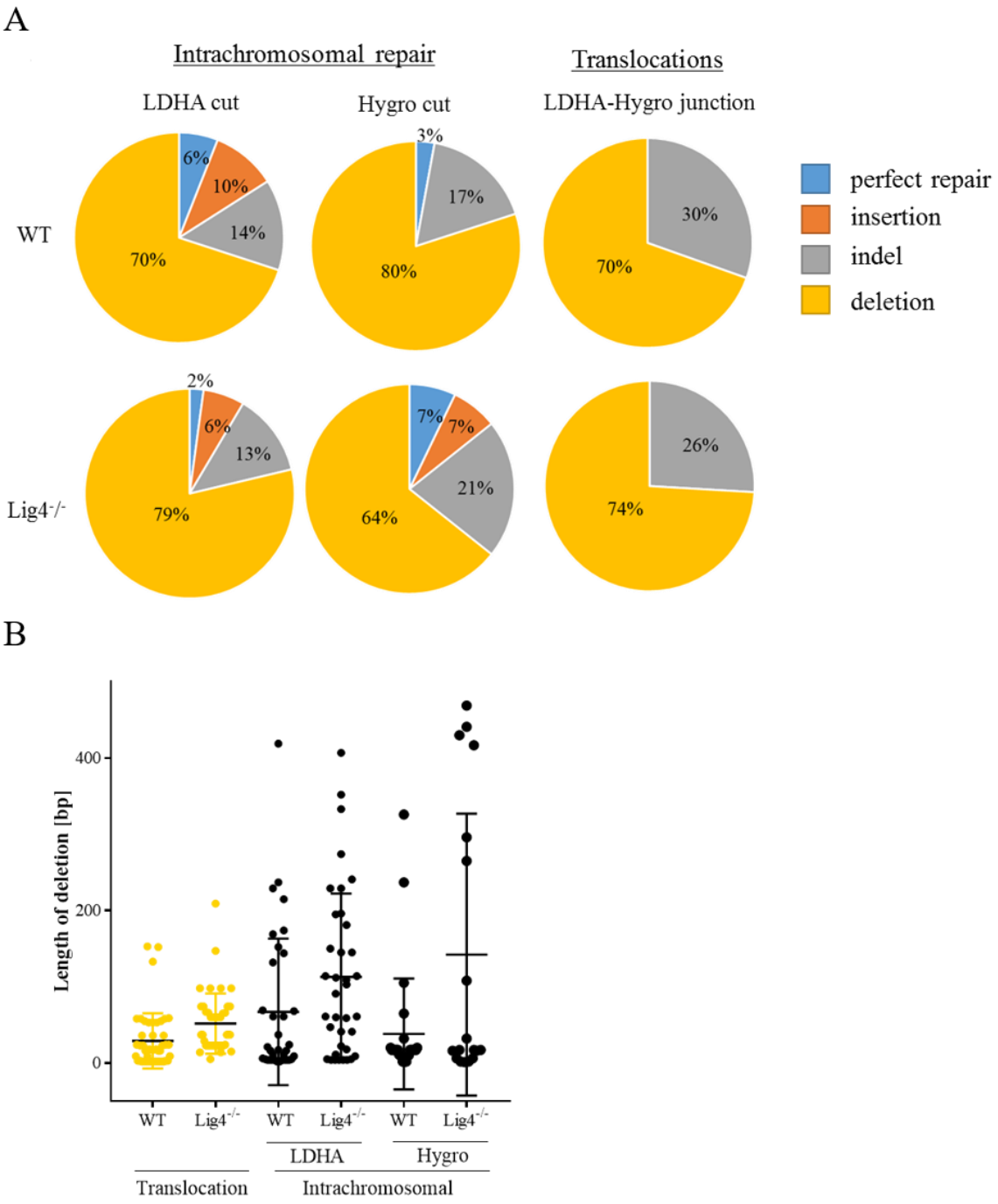


Figure 18. Analysis of breakpoint junctions' sequences.

A. Comparison of different junction processing outcomes such as insertions, deletions and indels between WT and Lig4 $^{-/-}$ samples presented as a fraction. Between 18 and 48 sequences were analysed. **B.** Quantification of the resection length in different products of intra- (black) and interchromosomal (yellow) rearrangements, compared between WT and Lig 4$^{-/-}$ samples.

2.8. Alterations in initial project set-up

At this stage of the project I was able to apply the system to induce and detect genomic rearrangements with CRISPR/Cas9 and determine their frequency by ddPCR. However, I did not manage to optimise the conditions for growth of Hygromycin-resistant colonies, which made it impossible to perform the screening. For that reason, the initial experimental set-up was modified to avoid high mortality of cells, which couldn't be resolved at that point. The new approach takes advantage of the presence of I-SceI recognition sites in the selection cassette upstream promoterless Hygromycin-resistance gene. Therefore, I-SceI endonuclease can be used to induce DSBs at chromosome 19. Active promoters are targeted genome-wide by etoposide. It was shown that DSBs induced in the presence of etoposide - a Topoisomerase II (Top2)-poisoning agent, are mainly located at promoters of active genes, which corresponds to preferential Top2 binding (Yang, Kemp, and Henikoff 2015). As previously described translocation events result in expression of Hygromycin-resistance gene and cells with this genomic rearrangement can be selected. This modified set up does not require the step of lentiviral transduction with MULE plasmids, which should reduce the mortality of cells during screening procedure.

2.9. Generation of the HAP1 cell line with selection cassette stably expressing I-SceI endonuclease

I-SceI is the *Saccharomyces cerevisiae* mitochondrial endonuclease with 18 bp restriction site: TAGGGATAACAGGGTAAT, which leaves 3' hydroxyl overhang of 4 base pairs. This sequence does not occur in the human genome, so can be integrated into the specific locus. Its characteristics make it a common tool to induce DSBs and study mechanisms of translocations.

A stable cell line expressing I-SceI was created in H1Cx cells by transduction using a lentiviral system with the plasmid containing the construct of I-SceI fused with GR receptor and DD degron. Expression of the construct is under a doxycycline-inducible promoter. Rapid nuclear localization of the HA- DD-I-SceI-GR fusion protein is obtained in the presence of dexamethasone - an agonist of the glucocorticoid receptor (GR). Shield-1 binds to the DD and protects the protein from degradation (Figure 20 B-C).

The I-SceI expression level and efficient nuclear transfer were monitored after 6h induction with doxycycline (4 μg/ml), followed by addition of dexamethasone (5 μM) for

different periods of time -30min/1h/2h/3h/4h in presence of Shield-1 (1μM). Immunofluorescence assay was performed on multiple clones using anti-HA antibody. The results show that already 2h incubation with dexamethasone is sufficient for complete nuclear transfer of expressed fusion protein and it does not increase global DNA damage, as checked by the formation of γH2AX foci. The clones with the highest expression levels - #14 and #26 were selected for further experiments (Figure 21 and Figure 22).

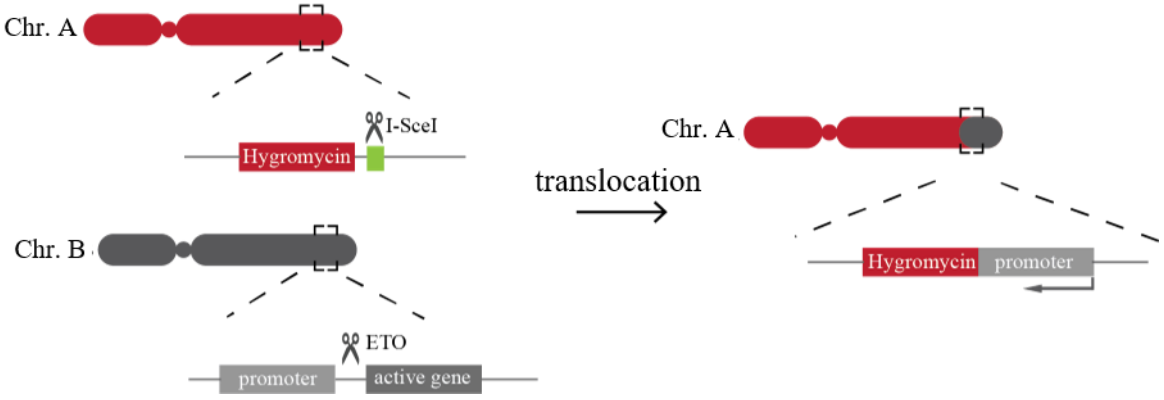


Figure 19. Alternative strategy for utilizing the translocation reporter system with I-SceI endonuclease and etoposide -general principle of selecting translocation positive cells.

The selection cassette integrated into AAVS1 locus on chromosome 19 contains restriction sites recognized by I-SceI endonuclease to induce DSBs upstream Hygromycin resistance gene. DSBs downstream active promoters are generated genome-wide by etoposide. As a result of a translocation between those chromosomes expression of Hygromycin resistance gene gets activated and cells with this genomic rearrangement can be selected.

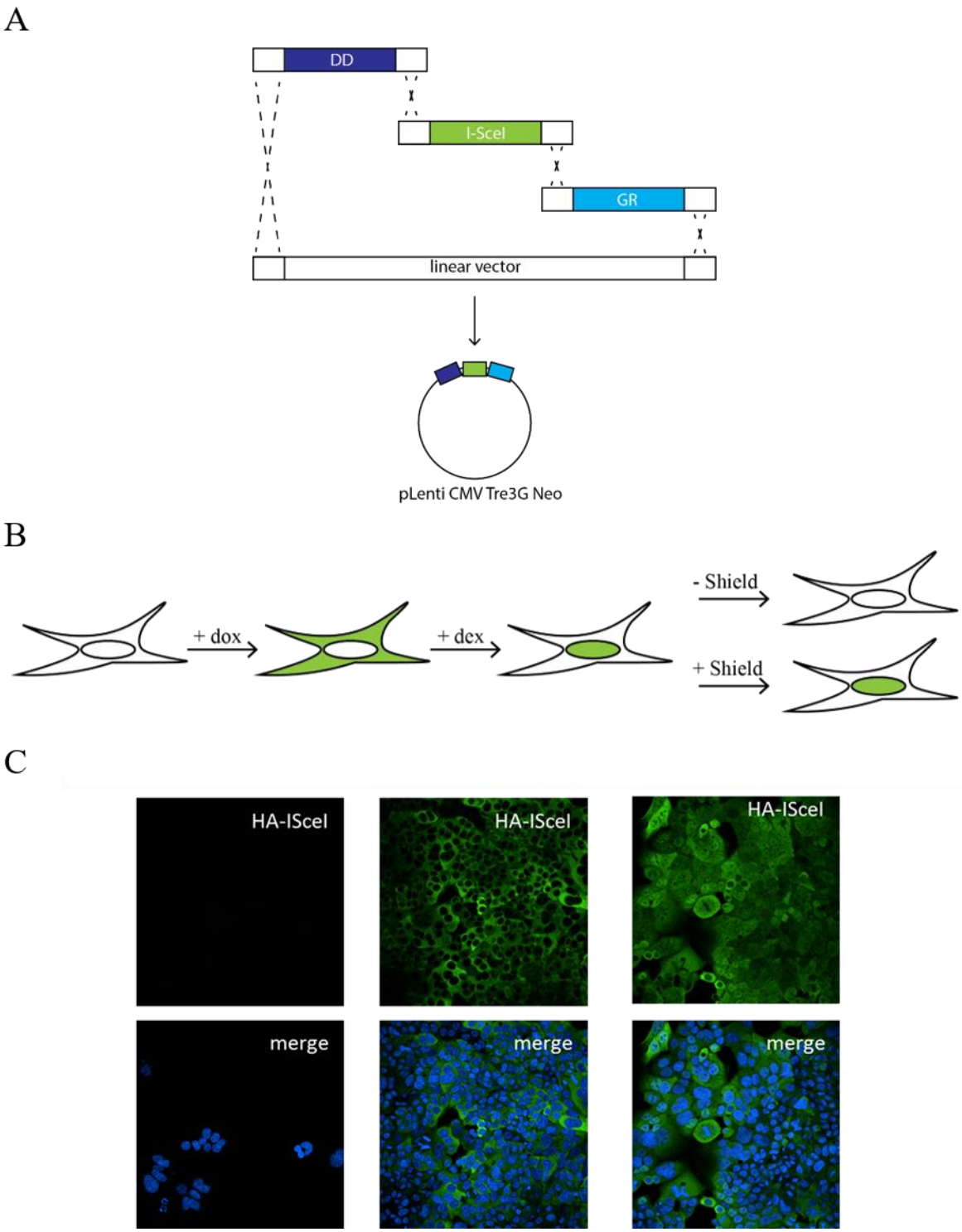


Figure 20. I-SceI endonuclease is expressed in H1Cx cell line.

A. The I-SceI endonuclease was cloned as a fusion protein HA-DD-I-SceI-GR under an inducible Tet-ON promoter into pLenti CMV Tre3G Neo using the GeneArt® Seamless Cloning. **B.** and **C.** I-SceI is stably expressed in H1Cx cell line. Addition of doxycycline induces the expression of the fusion protein, which in presence of dexamethasone is relocated to the nucleus. The stability of the protein is maintained by constant administration of Shield-1.

2.10. Induction of DSBs with I-SceI and etoposide

I validated the functionality of expressed I-SceI by T7 assay. Cells were treated with doxycycline only or with a combination of doxycycline, dexamethasone and Shield-1 for 24h/48h/72h. Subsequently, the region targeted by I-SceI endonuclease was amplified by PCR on genomic DNA and the products were digested with T7 endonuclease. DNA from H1Cx (without I-SceI) cells was used as a negative control. Results show that the obtained cut efficiency for both tested clones was around 70% and was comparable with the one previously obtained by CRISPR/Cas9 (Figure 22B). Efficient cut followed by repair was detected the earliest at the 48h time point.

We applied the knowledge that etoposide targets active promoters to induce DSBs at promoter partners. In order to find the perfect conditions of etoposide treatment, I tested different concentrations (0/1/2,5/5/10 μM) and durations (0h-4h). The optimal concentration should be high enough to create DSBs but low enough to keep cells viable and able to repair induced damage. Induction of DSBs was measured by IF as an increase of mean intensity of γH2AX signal in the cell nucleus (Figure 23). Additionally, I measured the viability of cells with the Cell Titer Blue after various release times (0h/24h/48h). At lower concentrations up to 2,5 μM, individual repair foci were detectable. At higher concentrations, the γH2AX signal was spread in the whole nucleus, which suggests increased cellular toxicity. Based on the obtained results, I chose the 5h treatment (slightly longer than the tested one) with 1μM etoposide and release of 48h for further experiments.

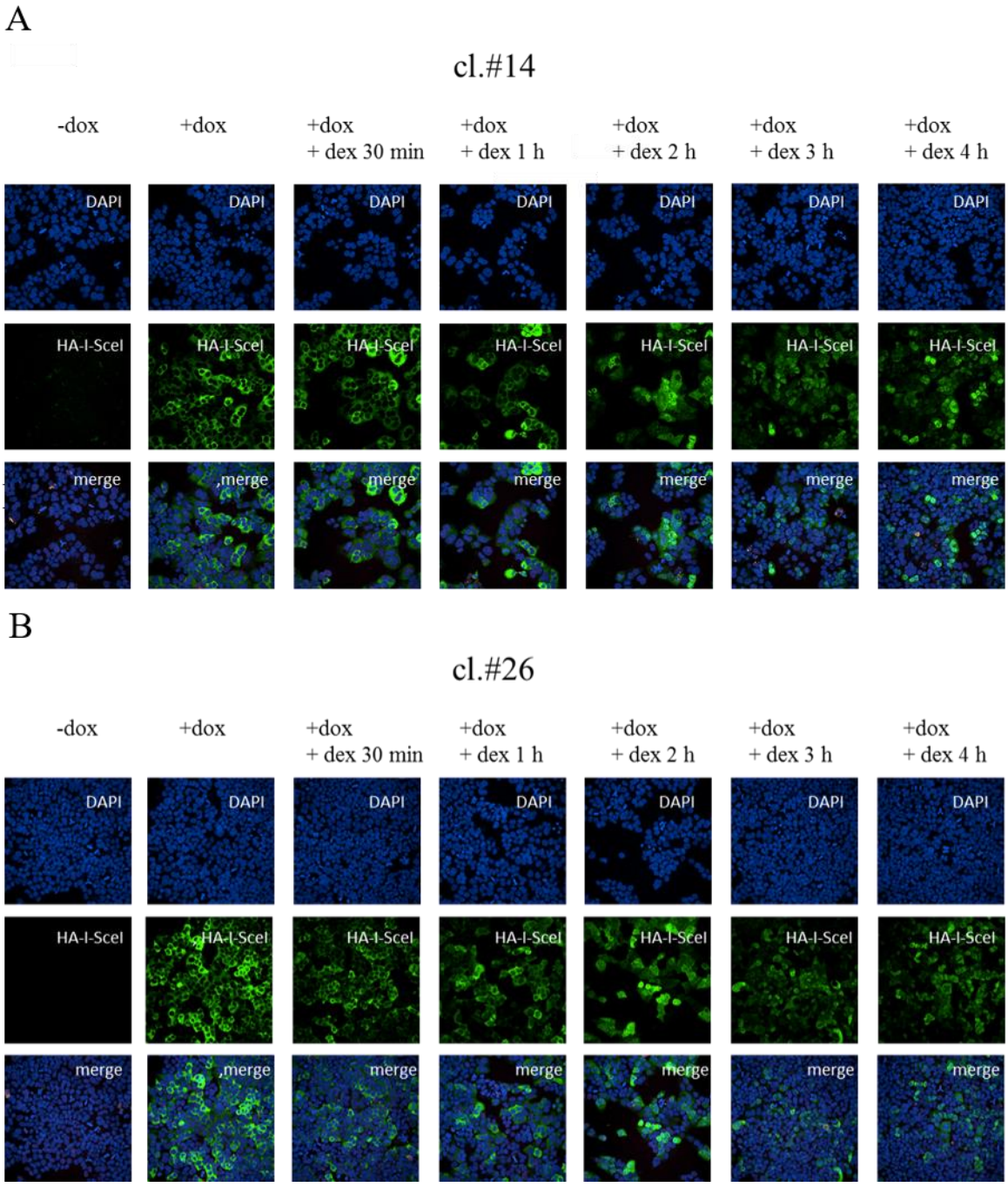


Figure 21. Fusion protein HA-I-SceI-GR-DD is expressed in the nucleus.

Expression of I-SceI and its nuclear transfer were induced by addition of doxycycline for 6h and dexamethasone for 0,5/1/2/3/4h. The figure shows fluorescent images of IF staining with anti-HA antibody in example clones #14 (**A**) and #26 **(B).**

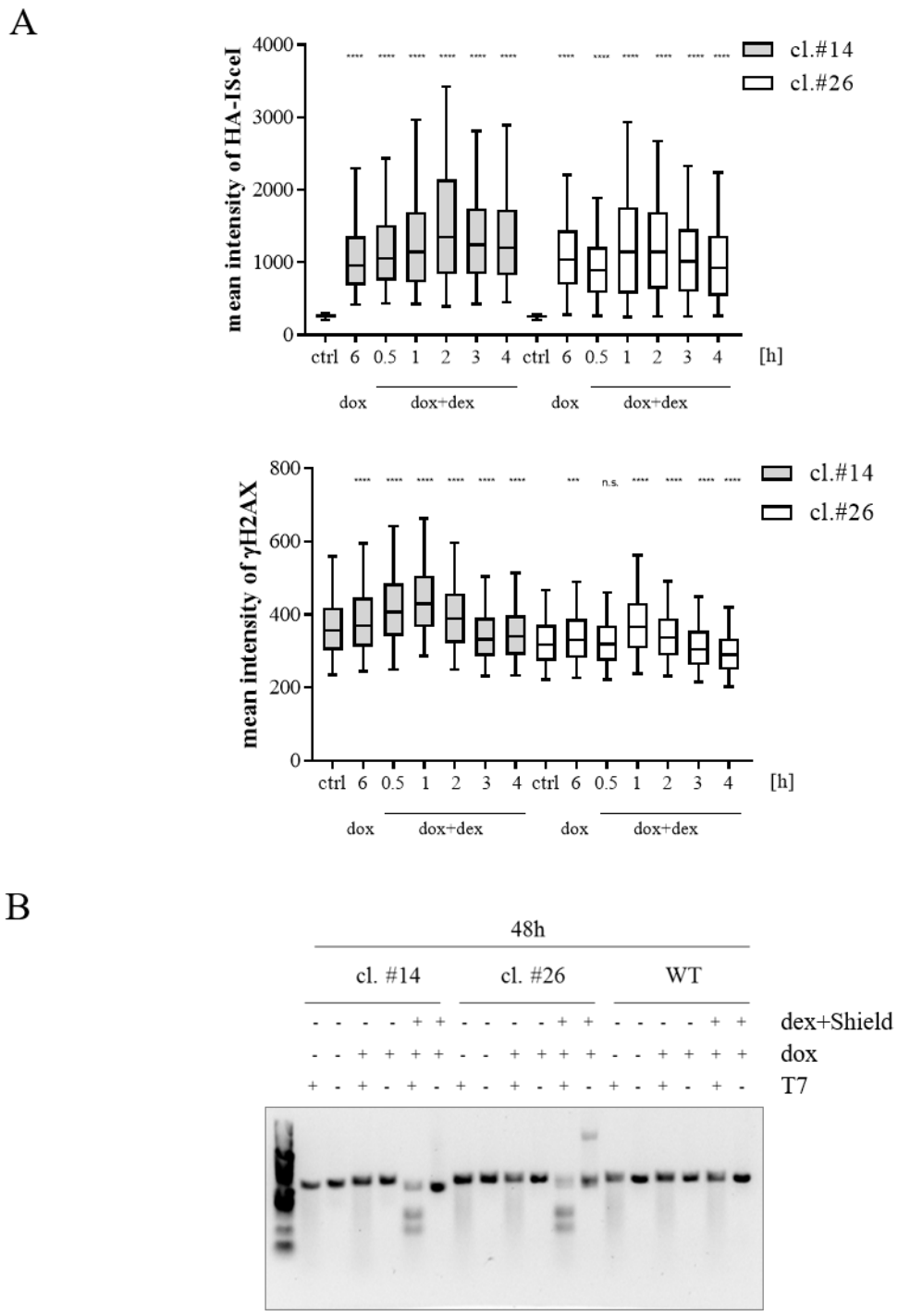


Figure 22. Fusion protein HA-I-SceI-GR-DD efficiently induces DSBs.

A. Quantification of the mean intensity of HA-I-SceI and γH2AX signals within the nucleus after induction with doxycycline only or doxycycline with dexamethasone. **B.** I-SceI activity in clones #14, #26 and negative control-WT was checked by T7 assay on gDNA collected 48h after the start of induction with doxycycline.

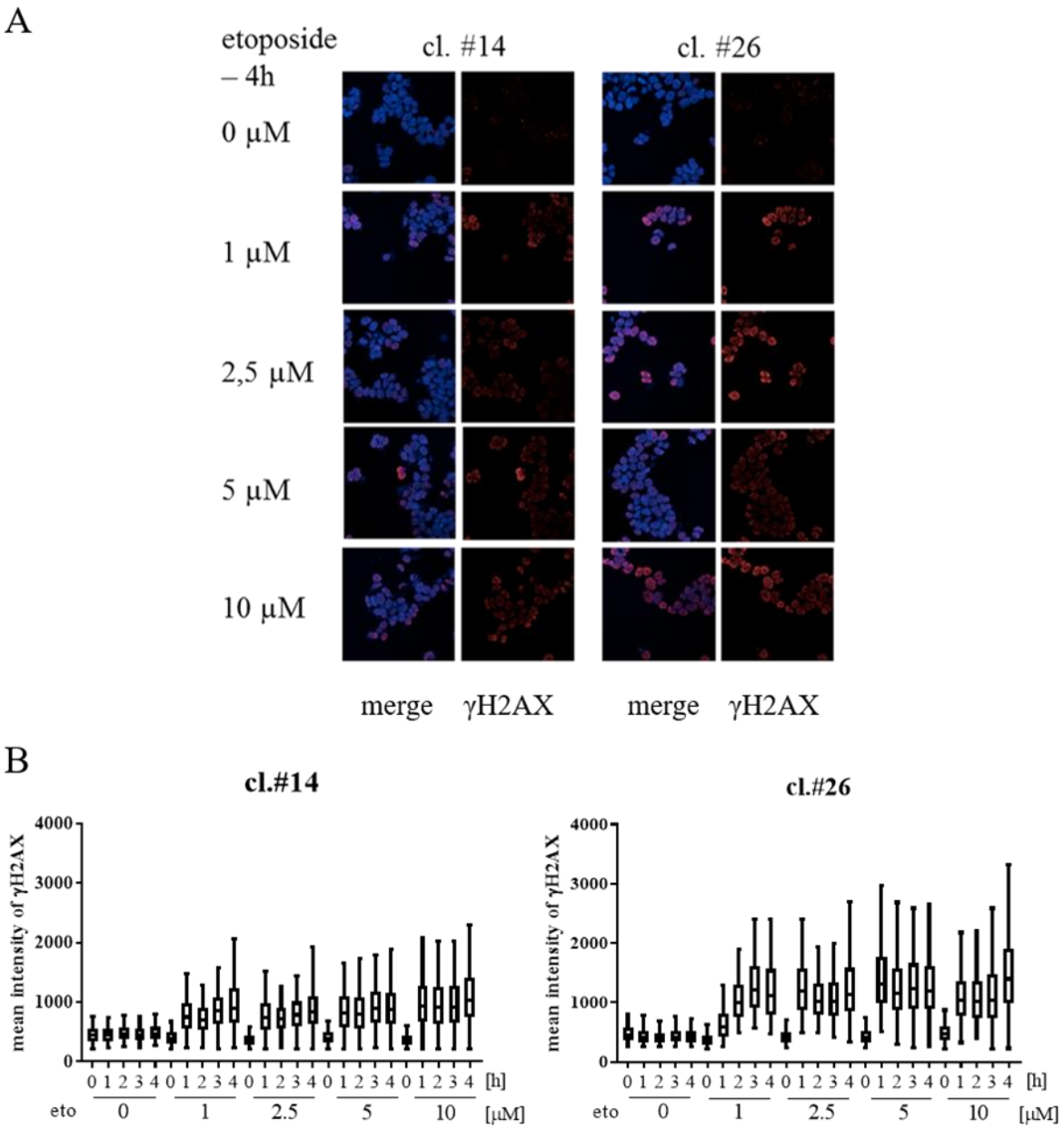


Figure 23. Optimisation of etoposide treatment conditions.

Representative images **(A)** and quantification **(B)** of γH2AX mean intensity in the nucleus using anti-γH2AX antibody after 4h of etoposide treatment at different concentrations: 0/1/2,5/5/10µM.

2.11. Induction of translocations by I-SceI and Etoposide

Those optimised conditions of DSBs induction were applied to create chromosome translocations between a promoterless Hygromycin-selection gene and active promoters. In order to do that, I induced the I-SceI expression with dox/dex/Shield and 16h later cells were treated with etoposide (1µM) for 5h. The treatment was followed by release and 48h later the selection with Hygromycin (350 µg/ml) started. In order to quantify translocations frequency,

I seeded equal numbers of cells for each condition at the beginning of the selection process. 14 days later the resistant colonies were stained with crystal violet and counted (Figure 24A-B).

Up to 0,02% (1:5000) of seeded cells were resistant to Hygromycin and formed colonies. Surprisingly the highest numbers of colonies were counted in the negative control, in cells just treated with doxycycline, where no or very little colonies were expected. It was not a case for other negative controls, such as etoposide only or no treatment control where no resistant colonies were obtained, suggesting it was not a result of insufficient selection. We hypothesized that all obtained colonies acquired a Hygromycin resistance and decided to further investigate the mechanism behind this process.

First, the I-SceI recognition sites were amplified by PCR on genomic DNA from single colonies, using primers flanking the region. In case of created translocation the primer binding site should be lost and therefore no PCR product would be amplified. Surprisingly, in each tested clone the amplification product was detected. Some of them were completely intact or smaller in size suggesting the formation of short deletions. On the other hand, many products consisted of multiple bands indicating the presence of more complex genomic rearrangements (Figure 24C).

Clearly, the mechanism of resistance acquisition differs between those outcome groups. In the case of complex rearrangements, we hypothesized that after induction of DSBs by I-SceI translocation events with random promoters could occur. However, the reason for antibiotic resistance in clones with little modifications within I-SceI recognition sites was less obvious. One possibility was that the expression of Hygromycin-resistance gene is activated by an endogenous promoter of the PPP1R12C gene, where the cassette is integrated. Secondly, it was observed that cells treated with etoposide were less viable than non-treated ones, which could lead to a false negative result, by loss of translocation positive cells due to high toxicity.

To avoid such aspecific effects the system and experimental workflow were further optimized. It included the reduction of etoposide concentration to 0,5 μM for all future experiments to decrease cell toxicity. Additionally, the promoter region of PPP1R12C was modified to eliminate the possibility of endogenous expression of Hygromycin-resistance gene.

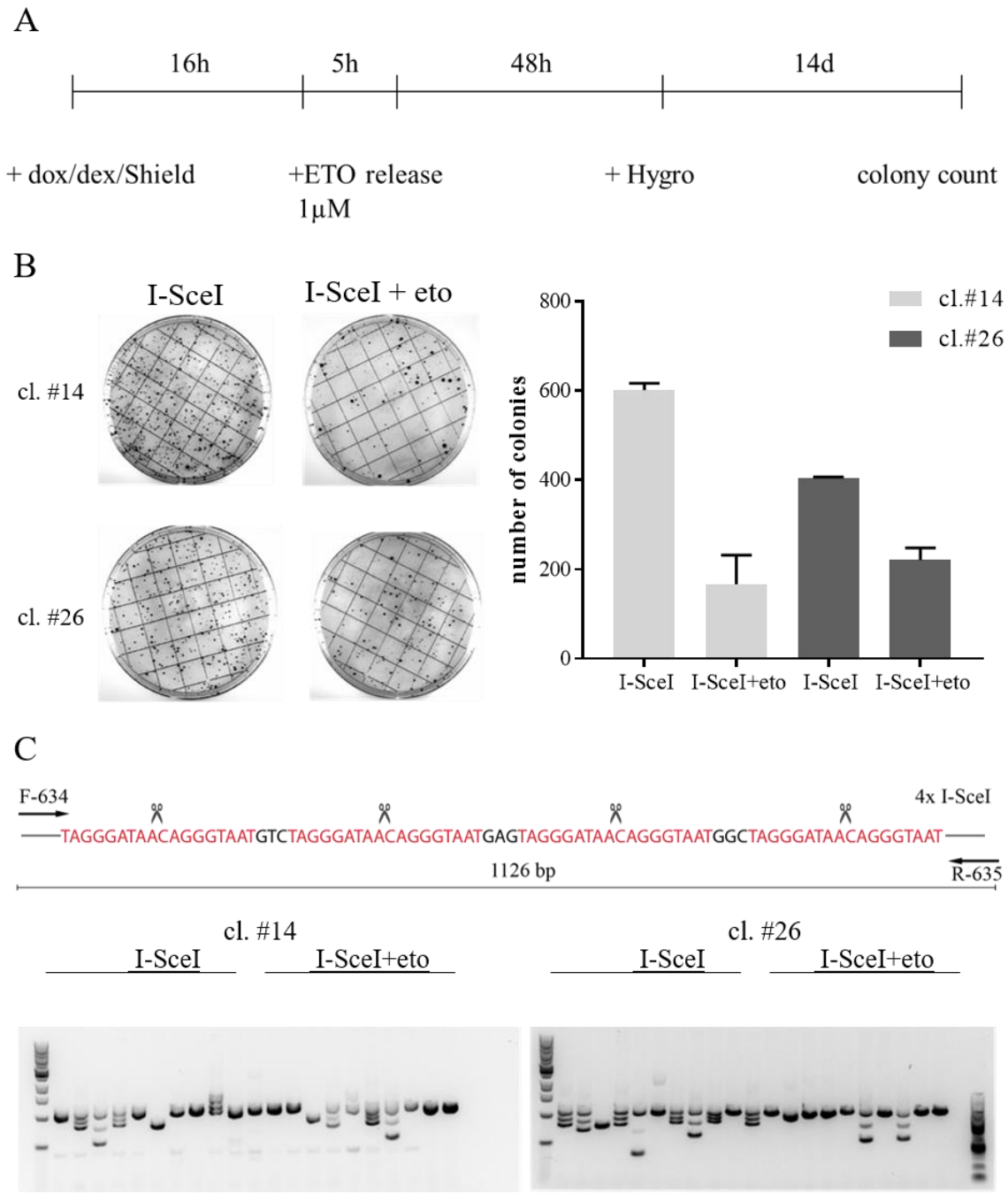


Figure 24. Induction of translocations by I-SceI and etoposide.

A. Experimental workflow. DSBs were induced by I-SceI only (16h) or combination of I-SceI and 1 µM etoposide (5h), followed by 14 days selection with Hygromycin (350 µg/ml). **B.** Resistant colonies of two example clones #14 and #26 were stained with crystal violet and counted. Error bars show standard deviation of 2 biological replicates. **C.** The region containing I-SceI recognition sites (sequence in red) was amplified by PCR on genomic DNA collected from single resistant colonies and compared to WT sequence (negative ctrl (C-) is a no template PCR reaction).

2.12. Modification in a promoter sequence of a PPP1R12C gene

The goal of this modification was to reduce the high number of background colonies formed after induction of DSBs by only I-SceI. During the procedure, a deletion of around 460 bp was induced by CRISPR/Cas9 to impair the possible binding of transcription factors. The applied sgRNAs targeted the region upstream exon 1 of PPP1R12C gene. Single clones were obtained through lentiviral infection with MULE plasmids carrying both sgRNAs and GFP (Figure 25A).

Efficient modification of promoter region was tested by qPCR measuring the expression of endogenous levels of Hygromycin-resistance gene in GFP positive clones. A Hygro+ cell line stably expressing a Hygromycin-resistance gene was used as a positive control and HAP1 cells without integrated selection cassette as a negative control. The results show that the deletion within the promoter region led to a 2-fold decrease of endogenous expression of Hygromycin-resistance gene in generated clones (Figure 25B). However, it was not completely abolished. Due to highly repetitive nature of a promoter sequence, it wasn't possible to confirm the deletion on a genomic level.

To further confirm the successful modification of a PPP1R12C promoter region, DSBs were induced by I-SceI in all tested clones. After a 14 days selection with Hygromycin, resistant colonies were stained with crystal violet and counted. The result in Figure 25C shows that the number of background colonies was greatly reduced for clones #PM-7 and #PM-14. Clone #PM-14, was chosen for additional validation, because of the lowest levels of endogenous expression and lowest number of background colonies.

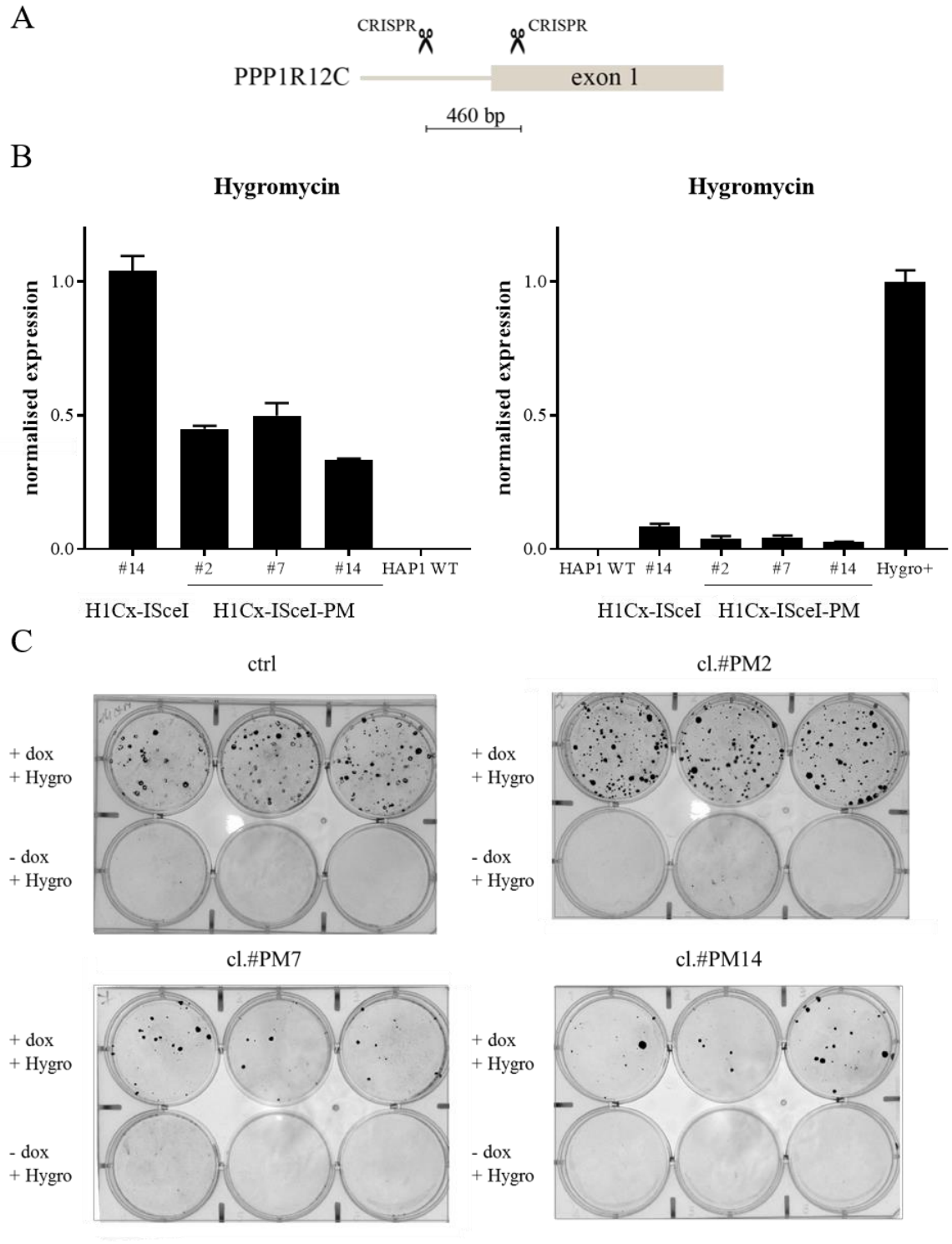


Figure 25. Inactivation of a promoter of PPP1R12C gene.

A. The design of a CRISPR/Cas9 induced deletion of 460 bp within the PPP1R12C promoter region. **B.** Hygromycin expression levels were quantified by qPCR in clones after transduction with virus containing sgRNAs targeting promoter sequence; left panel– values normalized to a clone without deletion (H1Cx-I-SceI); right panel - values normalized to positive control (cell line stably expressing Hygromycin resistance gene-Hygro+), error bars show standard deviation of 3 technical replicates. **C.** DSBs were induced by I-SceI in clones before (ctrl) and after (#14-PM) modification of promoter sequence and Hygromycin-resistant colonies were stained with crystal violet and counted.

2.13. Induction of translocations with I-SceI and etoposide after PPP1R12C promoter modification

In the following experiments I applied changes in the system, such as reduced etoposide concentration and inactivation of PPP1R12C promoter.

The induction of translocations with I-SceI and etoposide was repeated using clone #PM-14 in the presence of negative controls such as only I-SceI or only etoposide treatment and non-treated sample. The quantification of crystal violet stained resistant colonies in Figure 27A shows the efficient reduction of background colonies in negative controls - etoposide only or non-treated cells. It again proves successful modification of PPP1R12C promoter region. As expected treatment with both I-SceI and etoposide resulted in a bigger number of resistant colonies than in I-SceI only condition, which is an indication of higher translocation frequencies. It suggests that previously obtained high background levels could have been resulting from Hygromycin-resistance gene expression driven by endogenous promoter.

Additionally, I checked the changes in the region containing I-SceI recognition sites by PCR amplification on genomic DNA collected from single resistant clones. Again it showed no or only small (<100 bp) changes in the size of products. Interestingly, the complex rearrangements were hardly detected after the promoter modification (Figure 26B). The sequencing analysis of those PCR products revealed that the observed deletions correspond exactly to the processing of I-SceI recognition sites, displaying loss of fragment encompassing 1-3 sites.

Furthermore, the expression levels of Hygromycin-resistance gene in those clones were increased in comparison to non-treated cells (I-SceI-PM) in both I-SceI only and I-SceI with etoposide conditions. This suggests that induction of DSBs by I-SceI only leads to the acquisition of resistance by unknown mechanism and translocations might not be the only rearrangements involved (Figure 27).

It is possible that due to high instability of HAP1 cell line it is not haploid anymore. In this condition, only not or slightly modified allele gets amplified by PCR, whereas another one, possibly translocated and leading to resistance, is not detectable with this assay. To confirm this hypothesis, I analysed the ploidy of cells and breakage of the AAVS1 locus after DSBs induction by FISH.

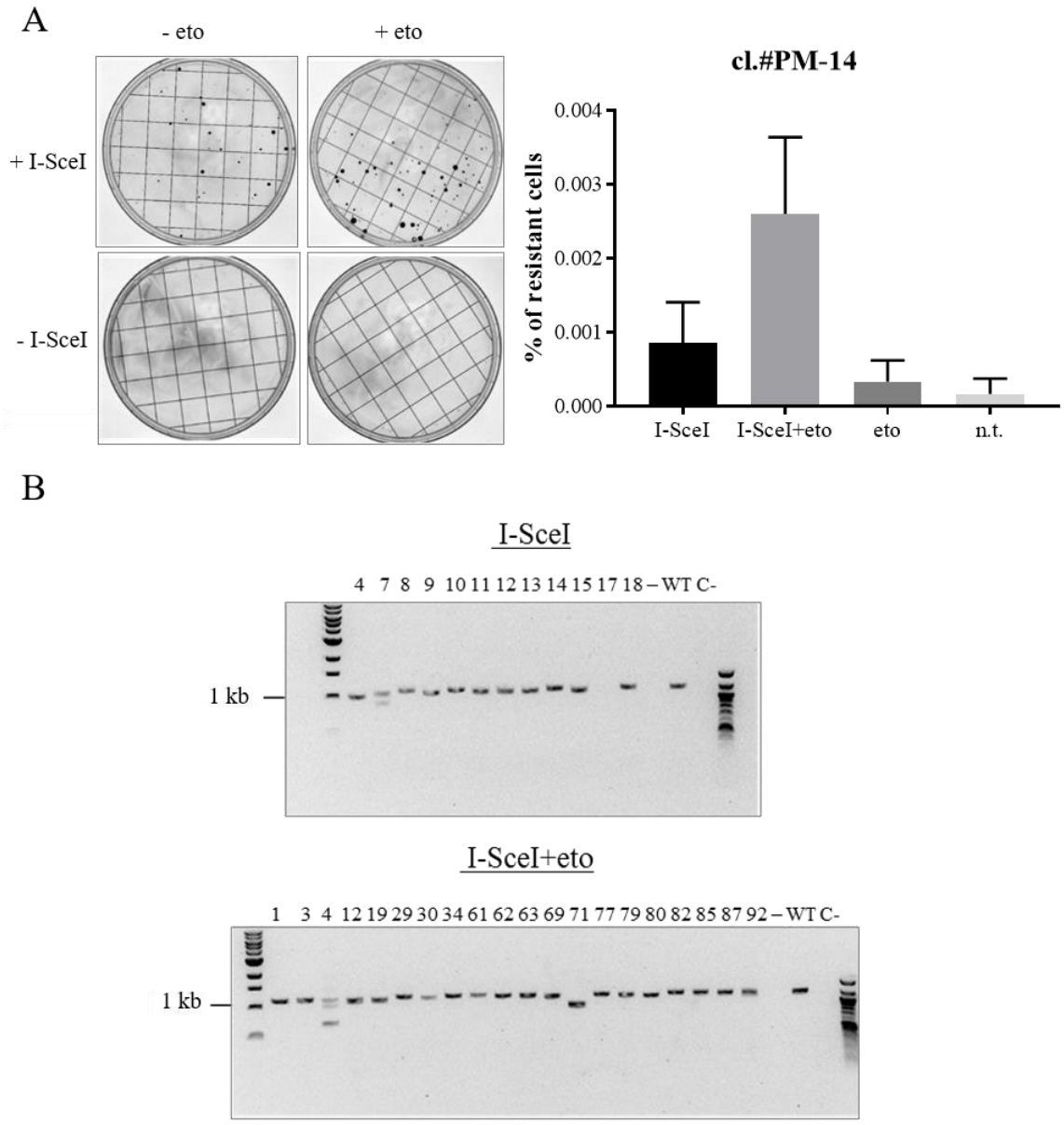


Figure 26. Induction of translocations with I-SceI and etoposide after PPP1R12C promoter modification.

A. After induction of DSBs by I-SceI or I-SceI and etoposide in clone #PM-14, Hygromycin-resistant colonies were stained with crystal violet and counted. Error bars show standard deviation of 2 biological replicates. **B.** PCR amplification of the region containing I-SceI recognition sites in the material collected from single resistant colonies of clone #PM-14 (WT- positive ctrl, negative ctrl (C-) is a no template PCR reaction).

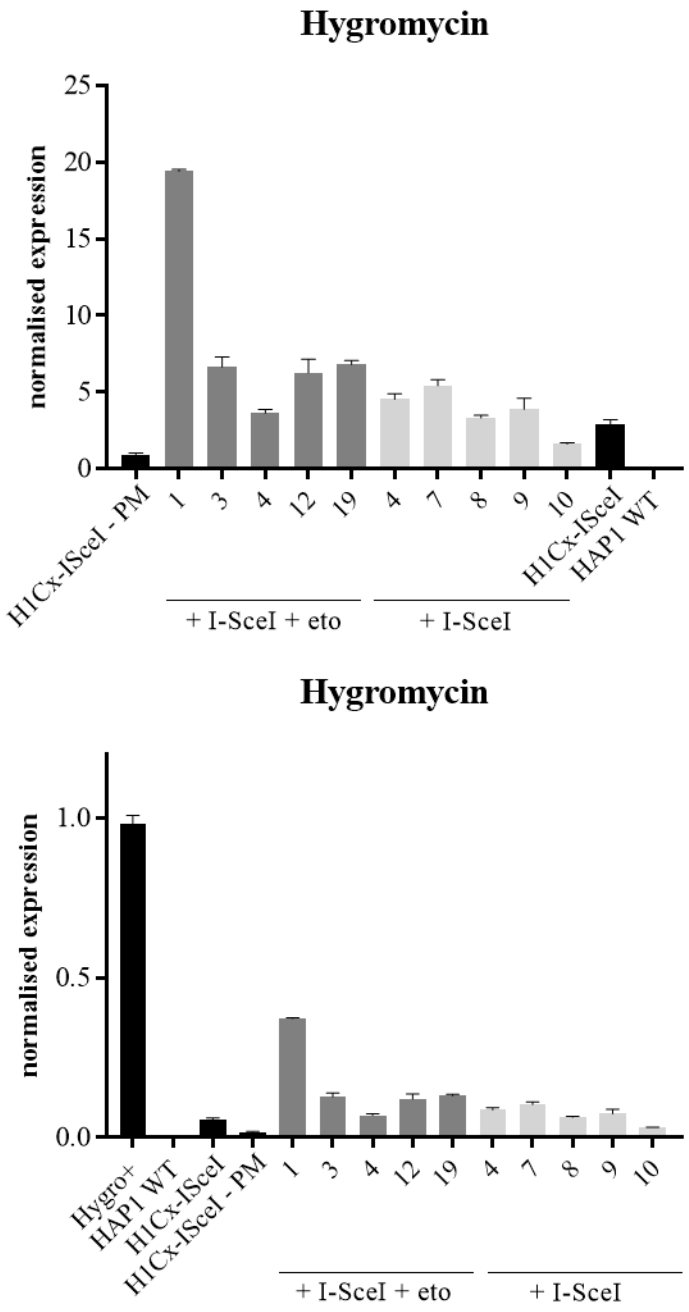


Figure 27. Quantification of translocations induced with I-SceI and etoposide after PPP1R12C promoter modification.

Hygromycin expression levels in resistant clones were quantified by qPCR. Left panel - values are normalized to the expression levels of clone #PM-14, right panel – values are normalized to positive control (Hygro+ - cell line stably expressing Hygromycin resistance gene); error bars show a standard deviation of 3 technical replicates.

2.14. Multicolour FISH analysis of AAVS1 locus

The FISH experiments were performed according to a standard protocol on fixed cells obtained from Hygromycin-resistant clones. Hybridisation probes against chromosome 19 flanking the AAVS1 locus were 500 kb apart. The 5' probe was labelled with Alexa-488 fluorophore and 3' probe with Alexa-568 fluorophore Figure 28.

The ploidy of cells was measured by counting the number of red and green probes per nucleus. The resistant clones were compared with a control non-treated clone in which all cells had an equal number of green and red spots. Interestingly, already by looking at the control sample it was obvious that HAP1 cells used in this study are not haploid anymore, but mostly diploid or even polyploid. The FISH analysis revealed that for many clones the quantification of both probes was not equal. In those cells number of red spots was often twice as much as of green ones (Figure 29).

In the control condition, both probes colocalize, which indicates the intact locus, whereas the separation or loss of spots is a sign of locus breakage and/or formation of chromosome rearrangement. To investigate this possibility I re-examined the experimental samples choosing representative clones with both equal and unequal number of spots and checked the chromosome structure by performing the FISH on metaphase spreads.

The metaphase spreads were prepared with standard protocol using Colcemid (0.02 μg/ml) and Calyculin (50nM). For FISH on metaphase spreads I used the same protocol and probes as described before for multicolour analysis of AAVS1 locus. First of all, I noticed the variability in rearrangement types within individual samples, which may be the outcome of high-instability of the HAP1 cell line during prolonged periods of culture or cross contamination of different clones (Figure 31). I observed that for some clones (I-SceI cl.#7 and I-SceI+eto cl.#4 and 19) all or majority of spreads had no detectable changes within chromosome structure indicated by equal number of green and red spots which colocalised. This was in agreement with previous quantification of spots for AAVS1 locus in cell nuclei.

Chromosome translocations, defined here by the loss of green spot coinciding with juxtaposed DNA fragment upstream of red spot, were detected in 50% of clones, but constituted less than 30% of all counted spreads. Other rearrangements that occurred included deletions and inversions. In spreads with deletions the green spot was lost as well, but there was no additional DNA attached at the chromosome end. This type of chromosomal abnormality was present in all but one of clones in which DSBs were induced with only I-SceI and in only one

clone treated with I-SceI and etoposide. Inversions, where number of spots was equal but they were present on opposite ends of chromosome, were found in only one clone.

These results prove that designed system can be used for generation of chromosome translocations. However, it has to be highlighted that other types of rearrangements are also formed and make cells resistant to Hygromycin by other unknown mechanisms. More detailed analysis including bigger number of clones is required to better understand those processes.

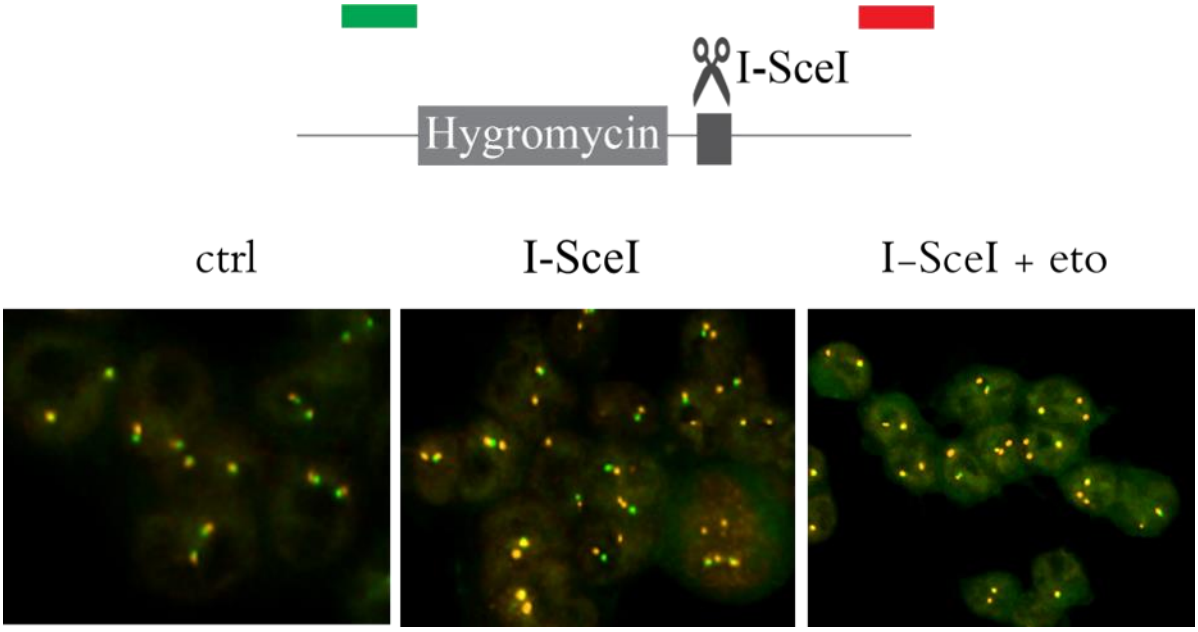


Figure 28. Multicolour FISH analysis of AAVS1 locus.

FISH was performed using green (5') and red (3') probes flanking the Hygromycin selection cassette at AAVS1 locus on chromosome 19. The integrity of the locus was tested in Hygromycin resistant clones obtained after DSBs induction with I-SceI or I-SceI and etoposide.

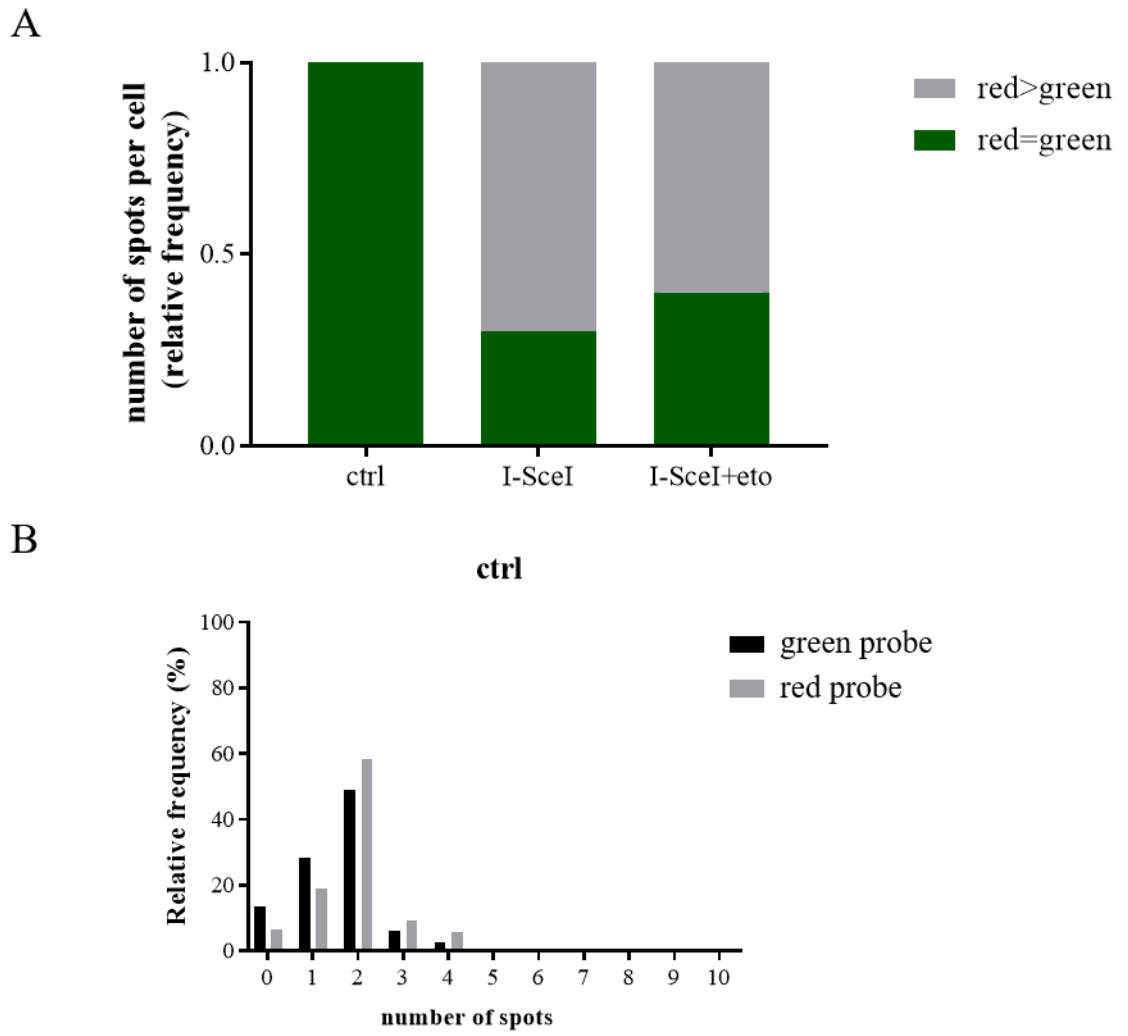


Figure 29. Quantification of number of spots labelling AAVS1 locus.

A and B. The number of red and green spots per nuclei in all tested clones was quantified by automatic software analysis.

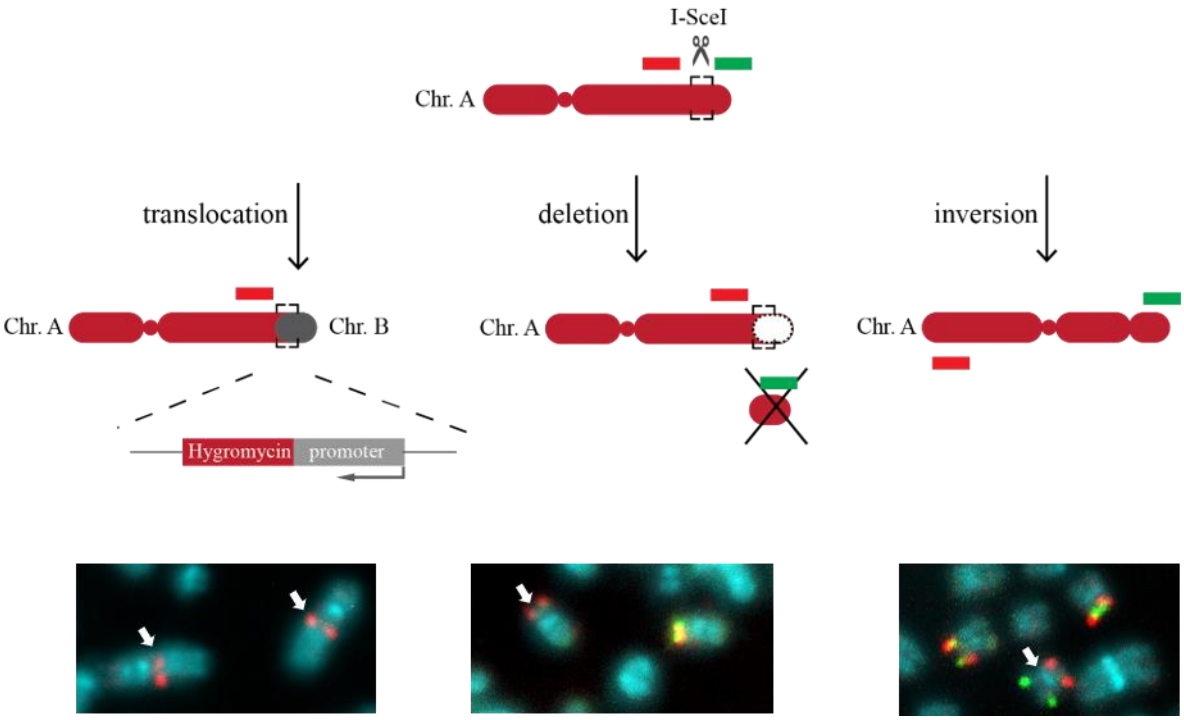


Figure 30. Detection of chromosome rearrangements by analysis of metaphase spreads.

A scheme of chromosome rearrangements formed after induction of DSBs with I-SceI endonuclease and FISH approach to detect them. Fluorescent probes label the region flanking the cutting site and DNA was stained with DAPI. Telomeric probe was labelled with Alexa-488 (green) and centromeric probe with Alexa-568 (red). Fluorescent microscope images show examples of each rearrangement type found in tested clones. White arrows show the chromosome with stated rearrangement. Formation of translocations was detected as loss of green probe binding and juxtaposition of additional chromosome piece upstream the red probe. Also in case of deletion the green probe was lost, but there was no additional piece of DNA upstream the cutting site. Observed inversions had equal number of green and red probe, but localised at opposite ends of chromosome.

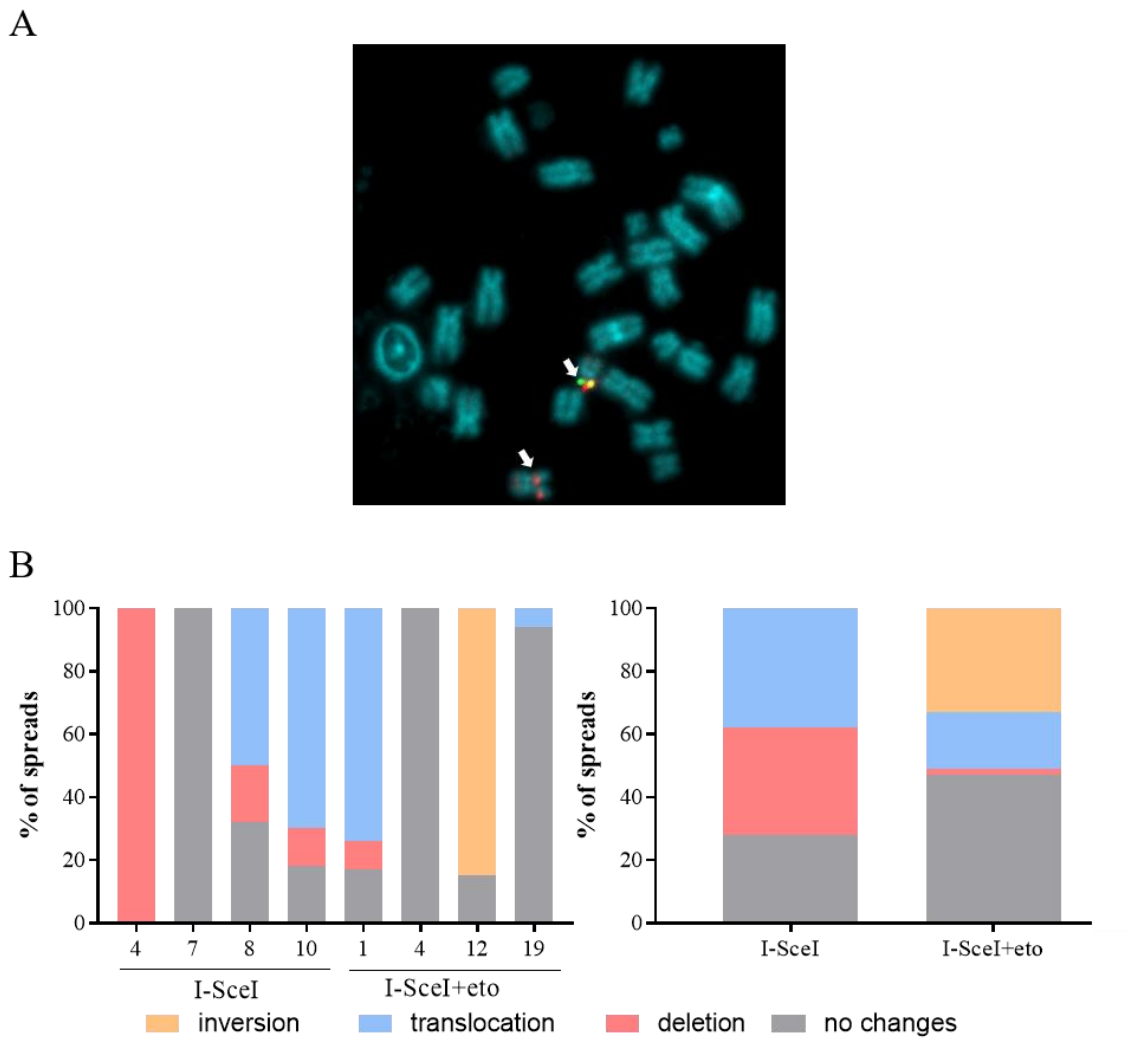


Figure 31. Quantification of chromosome rearrangements by analysis of metaphase spreads.

A. Example fluorescent microscope image of FISH on metaphase spread performed on a sample obtained from a Hygromycin-resistant colony. DSBs in those cells were induced by I-SceI and etoposide and the translocation within targeted chromosome 19 was detected and is marked with white arrow. **B.** Quantification of all detected rearrangement types: insertions, deletions and translocations within selected Hygromycin-resistant clones. Left panel shows results for individual clones and right panel a comparison between samples with I-SceI only and I-SceI and etoposide induced DSBs.

3. Discussion

Chromosome translocations are dangerous liaisons and hallmarks of several types of cancer. Their formation is triggered by endogenous or exogenous factors and can be a result of failures in DSBs repair. Different cellular aspects have been implicated to influence the frequency of translocations events. They include chromosome sequence features that dictate the propensity to breakage, the chromatin environment which influences proximity and contributes to movement, and end-joining pathways responsible for chromosome end-repair. However, only few factors have been identified to promote or suppress chromosome translocations. During my thesis, I aimed to identify novel factors involved in this process, to improve the knowledge of mechanisms behind. For this purpose, I built a novel cellular system, which is compatible with genome-wide screenings and allows enrichment of translocation positive cells in the pool. I also tested the performance of screening at control conditions and proved that the reporter assay can be applied for the induction of different chromosome rearrangements.

3.1. Advantages and limitations of the translocation reporter system

First of all, as shown in this study, the designed system is compatible with many ways of inducing damage. It allows to create rearrangements through induction of DSBs with designer nucleases and chemotherapeutics, such as etoposide. Potentially, also IR could lead to similar effects. Secondly, it gives a huge advantage of selecting translocation positive cells, which improves the performance and applications of subsequent analytical methods. Most importantly, the system is compatible with genome-wide CRISPR/Cas9 based screens and can be used in evaluating cellular processes related to genome stability. The drawback of this particular reporter design is the fact that only translocations of resistance gene with active promoters give a selective advantage to cells, therefore it is limited to specific loci. The choice of sgRNA libraries over shRNA based ones, although increasing the efficiency of depleting factors, narrow down the targets to only non-essential genes. Finally, the HAP1 cell line where the reporter system is integrated is not stable and becomes polyploid over time. It is possible to insert the same resistance cassette into other cell lines, however, predicted efficiency of both integration, but also gene knock-outs would be much lower.

3.2. HAP1 cell line stability

The designed translocation assay was integrated into the haploid human HAP1 cell line, to increase the efficiency of generating complete single-gene knockouts during the screening.

The cell line is widely used in genetic screenings. The ploidy of cell line was controlled by FACS or FISH at initial steps of creating the system: while making the HAP1-Cas9 cell line and after the excision of cassette fragment with Cre recombinase (data not shown). Final FISH experiments performed to prove the formation of rearrangements, showed that cells are not stable and with time turn diploid or even polyploid. Similar observations were described in several publications. The suggested reason for losing haploid features is lower growth rates of haploid cells due to longer chromosome segregation, which results in diploid cells overtaking the culture (Olbrich et al. 2017). Therefore it is important to control the ploidy of cells and make sure that at the time of the screening, most importantly at the step of generating knock-outs with the library, cells should be for example pre-sorted to use only the haploid pool.

Another drawback of using this cell line was the low efficiency of transfection. Reported by the Horizon Discovery (the supplier of the cell line) the efficiency of HAP1 transfection with standard transfection reagents is rather low <30%. Therefore other techniques of DNA delivery were applied in this project: electroporation and viral transduction. The laborious optimisation of those conditions let to stably integrate a selection cassette within the genome of HAP1 cells and to induce DSBs and chromosome rearrangements at high frequencies.

Further analysis of screening hits may require a generation of cancer-relevant translocation models in other cell lines as well. However, HAP1 cells have been successfully used to induce by CRISPR/Cas9 and quantify by PCR *MLL-AF4* translocations in a parallel project (data not shown). Studies on chromosome translocations involve the combination of various techniques. HAP1 cells are suitable for genetic engineering and molecular biology based methods. However, due to their nature they are not very adherent and grow forming cell clusters, therefore are not suitable for high-throughput FISH based approaches.

Since HAP1 cells originate from KBM7 cell line from CML patient, they harbour the Philadelphia chromosome and are positive for the BCR-ABL fusion protein. It has been shown that the expression of this oncogenic protein has an impact on genome stability. BCR/ABL cells have elevated expression levels of polymerase β and higher rates of spontaneous DNA damage. This fusion protein can modulate repair pathways, e.g. by enhancing the expression levels of RAD51 which is associated with drug resistance in leukaemia cells. It also causes strong downregulation of BRCA1 and a delay of mitotic checkpoint activation (Wolanin et al. 2010; Slupianek et al. 2001; Skorski 2002). All of it should be considered when analysing results obtained in this cell line. Further analysis should be performed in other cell lines with well characterized repair pathways, e.g. U2OS, HeLa or HEK293T.

3.3. Improvements and limitations of the screening setup.

Preparation of the screening required optimisation of its individual steps: cell transfection, DSB induction, selection of resistant colonies and detection of translocations.

Delivering the sgRNAs to cells was obtained through transduction, which gave the highest efficiency. Also as confirmed by T7 assay it gave high rates of cut. However, it resulted in high toxicity and death of cells. Therefore, it should be further optimised, probably compromising between efficiency of cut and toxicity by using lower virus titers. Selection with Hygromycin lasts 14 days, which is much longer than the time required for translocations to form. Difficulties in establishing the correct antibiotic concentration may also come from the fact, that changing the promoter that drives the expression of this gene may influence the transcription levels. In other words, stronger promoters may lead to higher and weaker ones to lower expression of Hygromycin-resistance gene. This means that some cells although harbouring translocation giving them resistance may still die during selection due to an excessive amount of antibiotic.

The experimental set-up of screening included the choice of the sgRNA library for single-gene knock-outs. The selected library provided by Sabatini and Lander is commercially available as one- or two-vector system. Since the preparation of translocation reporter assay already required Cas9 expression, only the sgRNA carrying plasmids from the library are required. The targeting sgRNAs are cloned into the pLX-sgRNA plasmid with Blasticidin as a selectable marker. The fact that this library is available with specific antibiotic resistance gene dictated the whole design of the screening conditions. The advantage of using this specific library over others is first of all novel approach to sgRNA design, which increases cleavage efficiency and decreases off-target effects. Second of all, it is possible to purchase it as smaller sub-pool libraries, which are divided based on function and include sgRNAs targeting kinases, ribosomal, cell-cycle and nuclear proteins (Tim Wang et al. 2014; Miles, Garippa, and Poirier 2016). This library has been successfully used in other studies for example for identification of essential genes in the human genome (T. Wang et al. 2015). The disadvantage of using this library is the lack of flexibility in terms of using a selectable marker, which influence the choice of antibiotic resistance genes at other stages of the screening.

Initial experiments which aimed to generate translocations by inducing DNA breaks by I-SceI and etoposide resulted in a high number of background resistant colonies. They were occurring mainly in cells with only I-SceI cut and sporadically in untreated condition, without

any cut. One hypothesis behind this phenomenon was that probably expression of promoterless-Hygromycin resistance gene is activated by the endogenous promoter of the PPP1R12C gene, where selection cassette was integrated and is further enhanced by additional unknown mechanism.

There are other examples of endogenous promoter activating the expression of inserts described in the literature. For example, in a study by (DeKelver et al. 2010) ZFN based targeted transgenesis into AAVS1 locus resulted in consistent expression of different types of transgenes, including reporter constructs. Those expression levels were relatively low but sufficient for FACS and isolation of pools based on selection. In other publications, promoterless puromycin-polyA cassette and promoterless GFP gene integrated downstream the PPP1R12C promoter or within intron 1 respectively, were expressed endogenously (Zou et al. 2011; L. Smith et al. 2016). Interestingly, as presented in this thesis, deleting fragment of PPP1R12C promoter resulted in decreased expression levels of Hygromycin resistance gene in the absence of any damage. This suggests that indeed the endogenous expression of the insert could take place.

In order to avoid a possible effect of endogenous expression and minimise the number of background colonies, the promoter of the PPP1R12C gene was modified. The CRISPR/Cas9 induced deletion contained 460 bp region upstream exon 1 of the gene. In subsequent qPCR analysis, it was proven that, after the modification, the obtained clones had reduced levels of resistance gene expression. However, it was not completely abrogated. All attempts to confirm the deletion on the DNA level by PCR or inverse-PCR failed. The most probable reasons for difficulties in amplification of this region was its high GC content >70% and repetitive nature of promoter sequence. As an alternative approach, the whole-genome sequencing could be applied to both confirm the integration of a selection cassette, prove the deletion in promoter sequence and exclude the possibility of non-target integration in other positions in the genome.

It is noteworthy that the highest levels of background colonies were observed with single I-SceI cut. Induction of DSBs by I-SceI required integration of recognition sites to the genome and expression of the enzyme. The integrated cassette contained an array of four neighbouring I-SceI recognition sites, which increases cutting rates. Faithfull repair of such breaks results in reconstitution of recognition sequence and recurrent cleavage. Imperfect religation can modify the sequence, which enables further cleavage. This may include short insertions and deletions or translocations in the presence of a second random break on another chromosome (Qiu et al. 2017). The current notion in the field is that the formation of translocation requires

simultaneous breaks at two non-homologous chromosomes. Experiments in which I induced DSBs by I-SceI and etoposide showed that cutting one chromosome with I-SceI only already leads to translocation as observed by FISH. The number of tested clones was not sufficient to make meaningful conclusions, but these observations give a new perspective to the possible mechanisms of translocation formation.

First possibility is that one DSB is sufficient to trigger the formation of chromosome translocation. Previous studies investigating the outcomes of single DSB misrepair don't present a clear view on this phenomenon. It has been shown that after the ZFN induction of single break within *MLL* locus insertions and deletions were the most common rearrangements, which I also observed in this project (Figure 12). Interestingly, in the same study, translocations were detected only in the presence of DNA-PK inhibitor, which was accompanied by an increase in SSA usage, suggesting that the repair of this DSB is coupled to NHEJ (Do et al. 2012). Also, the impairment of γH2AX expression enhanced the generation of balanced translocations and megabase-pair inversions after a single engineered DNA DSB (Qiu et al. 2017)**.** The frequency of rearrangements formed after single DSB varies based on experimental setup. On one hand, some studies report very low frequencies of translocation events after the single DNA cut, also in case of DSBs induced by both I-SceI and etoposide, which authors explain as result of suboptimal conditions (Varga and Aplan 2005; Aplan 2011). On the other hand, genome-wide studies combined with high-throughput sequencing use Cas9 or I-SceI induced breaks as a bait to identify multiple translocation partners (Marnef, Cohen, and Legube 2017).

Another interesting mechanism was observed in the study of Piazza et al. They showed that a single lesion can lead to translocation of two other chromosomes in the process of multi-invasion-induced rearrangement (MIR) (Piazza, Wright, and Heyer 2017). Altogether, those studies and results obtained in this thesis implicate the possibility that only one broken chromosome can lead to chromosomal rearrangements. Although this hypothesis demands much more extensive research, it is possible that mechanisms behind chromosome translocations are much more complex than it is known up to date and defining its basic principles requires revisiting.

It cannot be excluded that observed rearrangements are the effect of off-target cleavage at I-SceI cryptic sites. They restriction site differs from the canonical one 1-5 bp and it has been shown that they were not only efficiently cut in the presence of an enzyme, but also contributed to the generation of translocations. Therefore, it is possible that some of the detected

rearrangements were formed because of DSBs at two chromosomes – one with integrated I-SceI cutting site and second random one with cryptic I-SceI site. However, it is just a hypothesis which would require further experiments to be validated, e.g. by applying HTGTS which was previously used to reveal such genomic sites (Chiarle et al. 2011).

Formation of chromosome rearrangements from single DSBs is an interesting notion but it does not explain the cases where no structural changes were observed in Hygromycin-resistant clones. The possible explanation of this phenomenon might come from the process of break-induced transcription. It is known that DNA damage inhibits the pre-existing transcription initiation at active sites, which is usually recovered after religation of broken ends (Vítor et al. 2019). Additionally, RNA polymerase II has been shown to be enriched at broken sites and its activity leads to the production of transcripts forming DNA-RNA hybrids. Those hybrids regulate resection process and recruitment of RPA (Ohle et al. 2016). Some studies suggest that silencing of coding transcription may coexist with the formation of non-coding RNA at the break sites. This DNA damage response RNA (DDRNAs) are generated in a DROSHA- and DICER-dependent manner (W. T. Lu et al. 2018). Active antisense transcription was also detected at sites of endogenous DSBs by analysis of DSBCapture and transcriptome data from NHEK cells. Taking into consideration the fact that break-induced transcription facilitates DNA repair, investigating its mechanisms is also intriguing in the context of translocation biogenesis (D'Adda di Fagagna 2014; Vítor et al. 2019; Michelini, et al. 2017).

3.4. Alternative applications of the translocation reporter system

Although different types of chromosome rearrangements were induced with the designed translocation reporter, performance of the screening according to the initial experimental plan was not possible. The reason was sub-optimal conditions for the growth of translocation positive – Hygromycin-resistant cells after induction of DSBs with CRISPR/Cas9. The revisiting of individual steps of the screening set-up and their further optimisation are necessary before proceeding to the screening step to achieve informative and reliable results. Changes in the method of DSBs induction towards I-SceI and etoposide improved the viability of cells and as shown by FISH analysis also allowed to induce translocations. This approach can be alternatively applied in a screening.

The main goal of generating this particular translocation reporter system was to perform the screening to identify suppressors of translocations formation. However, it is also possible to utilise the system for other purposes. As shown in the analysis of breakpoint junctions intra-

and inter-chromosomal repair differ in terms of processing the broken ends, e.g. in case of the length of resected fragments. It is also locus-specific, since there is a striking difference in the repair of LDHA and Hygromycin cuts. The length of deletion was increased in the absence of cNHEJ component - Lig4 (Figure 12). It is in agreement with previous studies showing that in the absence of this pathway, a-EJ takes place, which relies on microhomologies and therefore requires more extensive resection (Ghezraoui, et al. 2014). It would be interesting to investigate this process further and identify novel factors influencing resection, as they potentially play a role in the formation of translocations.

CRISPR/Cas9 induced translocations are easily detectable with PCR techniques because of usage of specific primers flanking known breakpoints. The limitation of this technique is usually the low sensitivity, which does not allow to detect low-frequency events. Additionally, the quantification results might be biased by the primer design. Especially in case of short (<500bp) PCR products, there is a risk of losing a primer binding site due to extensive resection, which results in false negative results.

The huge advantage of the reporter system created in this thesis is the fact that cells positive for translocations are preselected, which enriches the pool. It could improve the sensitivity of the assay and allow to work with more samples without drastically increasing the costs. The combination of this novel translocation reporter, LAM-HTGTS and Hi-C data available for HAP1 cell line could be used for correlative studies on the influence of genome architecture and transcriptional activity on the frequency of translocations.

Another interesting approach would be to screen for factors promoting/suppressing therapy-related translocations, which occur in patients treated with etoposide. The treatment with this drug is often associated with secondary leukemia, linked to rearrangements within the MLL gene. Identification of such factors could contribute to the development of new drugs to be used in combinatorial therapies to decrease side effects of initial treatment (Libura et al. 2005).

3.5. Conclusions

In conclusion, the results presented in this study show that the developed novel translocation reporter system can be used to induce different types of chromosome rearrangements and could be compatible with a genome-wide screening approach. However, currently, high levels of background colonies after single I-SceI cut and difficulties to maintain healthy resistant cells, preclude its use for screening approaches. Additionally, the system can be utilised in studies on other biological processes, such as DNA repair or break-induced transcription, for example to investigate repair pathway choices, end-resection and interplay between DNA repair and transcription. Taken together, this study provides a novel tool to study processes related to genome maintenance.

4. Materials and methods

4.1. Molecular biology methods

4.1.1. Isolation of plasmid DNA from bacteria

Isolation of plasmid DNA from bacteria was performed with PureLink Quick Plasmid Miniprep Kit and QIAfilter Plasmid Midi Kit according to manufacturer's protocol.

4.1.2. Isolation of BAC DNA

Isolation of BAC DNA was performed with NucleoBond® Xtra BAC and ZR BAC DNA Miniprep Kit according to manufacturer's protocol.

4.1.3. Transformation of competent cells

The following standard protocol of cells transformation was applied to all bacteria strains used in this study. 50 µl of competent cells were used for each reaction. First, bacteria and plasmid were mixed and incubated on ice for 30 min. Heat-shock of cells was carried at 42°C for 45 s in a water bath, followed by 2 min incubation on ice. Next, 250 µl of SOC was added to cells and incubated for 1h at 37°C in an incubator shaking at 180 rpm. Finally, cells were spread on the agar plate with the appropriate antibiotic (Kanamycin/Ampicillin) and incubated overnight at 37°C.

4.1.4. Extraction of genomic DNA from cell lines

Total genomic DNA was isolated from cells using the Qiagen DNA Blood&Tissue Kit according to manufacturers' protocol. The concentration of extracted DNA was measured by NanoDrop, with the exception of samples used for ddPCR, which were measured with Qubit dsDNA HS Assay Kit.

4.1.5. Extraction of total RNA from cell lines

Total RNA was isolated from cells using the RNeasy Plus Mini Kit. The concentration was determined by NanoDrop.

4.1.6. Purification of PCR products for sequencing

PCR products were purified using the QIAquick PCR Purification Kit or gel-excised and processed with the QIAquick Gel Extraction Kit according to manufacturers' protocol.

4.1.7. DNA sequencing

All sequence analysis was performed by GATC Biotech. Sequences in Fasta and ABI format were evaluated with ApE and CLC Sequence Viewer 7 software.

4.1.8. Analysis of breakpoint junctions' sequences

This analysis was performed to characterize breakpoint junctions on the sequence level. For analysis of intrachromosomal repair products of T7 PCRs were used. For interchromosomal repair products of PCRs for detection of translocations, inversions and deletions were analysed. All samples were collected from WT and Lig4$^{-/-}$ cells.

First, all PCR products were purified with the PCR Purification Kit. All PCRs were performed using Phusion or Q5 polymerases, which have 3′→ 5′ exonuclease activity and remove 3'-A overhangs required for cloning into pGEM-T Easy vector containing 3′-T overhangs. The 3' A-tailing reaction was performed at 72°C for 20 minutes (Table 1.).

Table 1. Adding A overhangs to PCR products.

Component	**Amount**
10x Standard Taq Reaction Buffer	1x
PCR product	20 µl
dATPs (10 mM)	1 µl
Taq polymerase	0,2 µl
Water	up to 50 µl
Total	50 µl

The ligation of PCR products into the pGEM-T Easy vector was performed at RT for 1h (Table 2.). The amount of each insert was calculated based on the following formula, with a molar ratio of insert:vector 3:1:

$$ng\ of\ insert = \frac{\text{ng of vector x kb size of insert}}{\text{kb size of vector}} * (insert{:}\ vector\ molar\ ratio)$$

Table 2. pGEM-T Easy system – parameters of ligation reaction.

Component	**Amount**
2x Quick Ligase Buffer	1x
pGEM-T Easy vector	50 ng
T4 DNA ligase	1 µl
PCR product	~3 µl
Total	10 µl

The 5 µl of ligation product were used for transformation of 50 µl of XL1-Blue bacteria. Cells were spread on the agar plate with the appropriate antibiotic, IPTG (40 µl of 0,1M) and Xgal (40 µl) and incubated overnight at 37°C. Next day the white colonies were picked, seeded on 96-well agar plate and sent for sequencing.

4.1.9. Amplification of genomic or plasmid DNA by PCR

Parameters of distinct PCR reactions were optimized individually for a specific set of primers. Table 3-7 present general PCR protocols according to the polymerase type used for the reaction. All PCRs were performed on C1000 Touch Thermal Cycler (Bio-Rad). Detailed PCR protocols can be found in Supplementary part. Amplified products were validated on 1% agarose gel, stained with Gel Red nucleic acid dye and imaged with GelDoc XR+.

Table 3. Protocol of PCRs using Q5 polymerase.

Reaction component	**Final amount**
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer F	0.5 µM
Primer R	0.5 µM
DNA	50-200 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	**Time [s]**	
Initial denaturation	98	30-60	
Denaturation	98	10-30	x30-35
Annealing	64.3-72	20-30	
Extension	72	30-300	
Final extension	72	120-600	
Hold	12		

Table 4. Protocol of PCRs using Phusion polymerase.

Reaction component	**Final amount**
5x Phusion HF Buffer	1x
10 mM dNTPs	200 µM
Primer F	0.5 µM
Primer R	0.5 µM
DNA	100-125 ng
Phusion Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	**Time [s]**	
Initial denaturation	98	30-180	
Denaturation	98	10-30	x35
Annealing	64.3-67.9	20-30	
Extension	72	30-60	
Final extension	72	300-600	
Hold	12		

Table 5. Protocol of PCRs using Quick-Load® Taq polymerase.

Reaction component	Final amount
2x Quick-Load® Taq Master Mix	1x
Primer F	0.5 µM
Primer R	0.5 µM
DNA	40-100 ng
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	95	30	
Denaturation	95	30	x30
Annealing	58-63	30-60	
Extension	68	60-90	
Final extension	68	300	
Hold	12		

Table 6. Protocol of PCRs using Pfu Ultra II polymerase.

Reaction component	Final amount
10x PfuUltra II Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer F	0.5 µM
Primer R	0.5 µM
DNA	50 ng
Pfu Ultra II	1 µl
Water	up to 50 µl
Total	**50 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	95	120	
Denaturation	95	20	x30
Annealing	60	20	
Extension	72	20	
Final extension	72	180	
Hold	12		

Table 7. Protocol of digital droplet PCR (ddPCR).

Reaction component	Final amount
2x ddPCR Supermix for Probes (no dUTP)	1x
Primer 927	900 nM
Primer 162	900 nM
Primer 1161	900 nM
Primer 1162	900 nM
Probe 1367	250 nM
Probe 1166	250 nM
DNA	100 ng
Water	up to 20 µl
Total	**20 µl**

	Temp. [°C]	Ramp rate	Time [s]	
Initial denaturation	95	2°C/s	600	
Denaturation	94		30	x40
Annealing	58.8		60	
Extension	98		600	
Hold	12			

4.1.10. Reverse transcriptase reaction

The RT reactions were performed using SuperScript II Reverse Transcriptase (Invitrogen) with random hexamers and 500 ng of RNA, according to manufacturers' protocol.

4.1.11. qPCR for quantification of Hygromycin-resistance gene expression.

All qPCRs were carried out using a Real-Time PCR Detection System CFX96 (Bio-Rad). 1 µl of cDNA template was used for quantification of gene expression levels. Parameters of reaction are presented in Table 8. The results were analysed with Bio-Rad CFX Manager software and data was normalized to GAPDH signal.

Table 8. Protocol of qPCR for quantification of Hygromycin-resistance gene expression.

Reaction component	Final amount
SSOFast EvaGreen supermix	1x
Primer 348	750 nM
Primer 956	750 nM
Primer 1163	750 nM
Primer 1164	750 nM
cDNA	1 µl
Water	up to 20 µl
Total	20 µl

	Temp. [°C]	Time [s]	
Initial denaturation	98	120	
Denaturation	98	10	x40
Annealing/extension	60	25	
Melt curve	65-95	continuous	

4.1.12. T7 assay

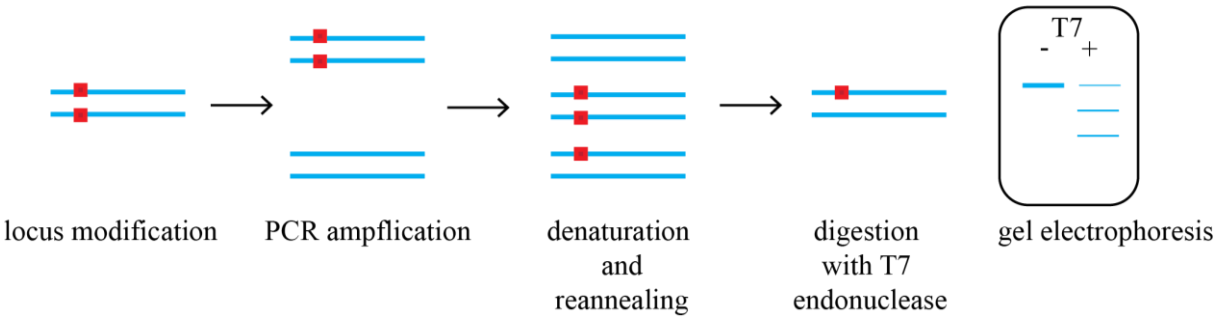


Figure 32. Workflow of T7 mismatch assay.

The genomic locus targeted with CRISPR/Cas9 is cut and processed by DNA repair mechanisms, which may lead to the introduction of insertions or deletions within the target sequence. The modified locus is amplified by PCR using primers flanking the region of interest. Afterwards, PCR products are denaturated and reannealed, which triggers the formation of homo- and heterodimers. T7 endonuclease can recognize and cleave mismatches in heterodimers. Finally, the digestion products are analyzed on an agarose gel, where multiple bands are the indication of successful cut and repair of targeted locus.

The protocol for T7 mismatch assay was modified from (Renouf et al. 2014).

First, the fragment of interest was amplified by PCR on genomic DNA using a polymerase with proofreading activity. Each specific PCR reaction required individual optimization. The quality of amplification was verified on a 1% agarose gel. Correct PCR products were purified with Qiagen PCR Purification Kit.

For treatment with T7 endonuclease, 70 ng of each purified product were mixed with an equal volume of 2X NEB Buffer 2.1 in two different PCR tubes. Then, the DNA was denatured and re-annealed by incubating it in a thermal cycler with following set up: 95°C for 5 min, 95°C to 25°C at −0.5°C/sec, and 15 min at 4°C.

Afterwards, 5 U of T7 endonuclease was added in one of the two tubes and the same volume of 2X NEB Buffer 2.1 in the second one as a control. The reaction was incubated at 37°C for 20 min. The enzyme was inactivated with 1.5 µL of 0,25M EDTA. Final reaction products were verified on a 1% agarose gel.

4.1.13. Digital droplet PCR (ddPCR)

Prior to amplification, the concentration of genomic DNA was measured with Qubit and it was digested with Hind III enzyme to improve the accessibility of the template during PCR reaction. Digestion parameters are listed in Table 9.

Table 9. Digestion of gDNA for ddPCR.

Reaction component	**Amount**
Hind III	10 U/µg
10x Buffer Cut Smart	1x
DNA	1.5 µg
Water	Up to 50 µl
Total	50 µl

The PCR reaction components were assembled as presented in Table 7. The PCR mix was partitioned into water-oil droplets by loading 20 µl of reaction and 70 µl of droplet generation oil into an 8-well cartridge and placing it in a QX200 Droplet Generator. 40 µl of obtained droplets were transferred to a PCR plate and amplified on the T100 Thermal Cycler (Bio-Rad) with reaction parameters as in Table 7. After PCR, individual droplets were analyzed for the presence of a fluorescent signal with Droplet Reader. The final quantification of positive and negative droplets was performed with QuantaSoft software with “absolute quantification” parameter and manually adjusted thresholds.

4.2. Cell culture and generation of stable cell lines

4.2.1. Culturing conditions

HAP1 (Horizon Discovery) is a human near-haploid fibroblast-like adherent cell line derived from the male chronic myelogenous leukemia (CML) cell line KBM-7. Cells were cultured in Iscove's Modified Dulbecco's Media (IMDM) supplemented with 10% fetal bovine serum

(Gibco), 2mM L-glutamine (Gibco), and 100 U/mL penicillin/streptomycin (Gibco). Medium for cell lines with doxycycline-inducible systems contained tetracycline-free FBS (Biochrome). HEK293T cells were cultured in (Dulbecco's Modified Eagle Medium) DMEM supplemented with 10% fetal bovine serum (Gibco), 2mM L-glutamine (Gibco), and 100 U/mL penicillin/streptomycin (Gibco). All cells were cultured at 37°C and 5% CO2.

4.2.2. Generation of dox-inducible Cas9 cell line

HEK293T packaging cell line was split a day prior to transfection to achieve 80% confluency. Transfection mix 1 included 5 µg of each packaging vector: pMDLg/pRRE, pRSV-REV, pMD2.G and 8 µg of pCW_Cas9 together with 450 µl of Optimem. Transfection mix 2 included 30 µl of XtremeGene HP and 450 µl of Optimem. Mixes 1 and 2 were incubated separately for 5 minutes and then together for another 20 minutes. Afterwards, the mix was added dropwise to the medium of HEK293T cells. The day after transfection the medium of HEK293T was changed to IMDM containing Tet-approved FBS and HAP1 cell line was prepared to reach 60% confluency the next day.

Two days after transfection the medium of HEK293T cells containing virus was passed through the 0,45 µM filter and together with Polybrene (8 µg/ml) was used to exchange medium of HAP1 cells. On the following day, the medium of HAP1 cells was changed again. The selection with Puromycin (2 µg/ml) was started 2 days after infection.

4.2.3. Generation of stably expressing Cas9 cell line

HEK293T packaging cell line was split a day prior to transfection to achieve 80% confluency. Transfection mix 1 included 5 µg of each packaging vector: pMDLg/pRRE, pRSV-REV, pMD2.G and 8 µg of LentiCRISPR v2 together with 500 µl of DMEM. Transfection mix 2 included 60 µl of Pei and 500 µl of DMEM. Mixes 1 and 2 were incubated separately for 5 minutes and then together for another 15 minutes and added to HEK293T medium dropwise. On the following day, the medium was changed and HAP1 cell line was prepared to reach 60% confluency day after.

Two days after transfection the medium of HEK293T cells containing virus was passed through the 0,45 µM filter and together with Polybrene (8 µg/ml) was used to exchange medium of HAP1 cells. On the following day the medium of HAP1 cells was changed again. The selection with Puromycin (2 µg/ml) was started 2 days after infection.

4.2.4. Integration of a selection cassette into the AAVS1 locus

HAP1-Cas9 cells stably expressing Cas9 were cotransfected with pLX-sgRNA (8 µg) and AAVS1hygroNEOP2aGFP (16 µg) using Neon Transfection System with buffer R. Electroporation parameters are presented in Table 10.

Table 10. Neon Transfection System parameters.

Voltage	1575 V
Width	10 ms
Number of pulses	3
Number of cells	6 mln

2 days later selection with G418 (800 µg/ml) was started and GFP positive clones were selected for further validation.

4.2.5. Excision of Neomycin resistance gene and GFP by Cre recombinase

In order to excise the fragment of selection cassette including Neomycin resistance gene and GFP, Cre recombinase recognizing loxP sites was used. To express Cre recombinase, first a packaging cell line HEK293T was transfected with 5 µg of each packaging vectors: pMDLg/pRRE, pRSV-REV, pMD2.G and 8 µg of pCMV-R-Cre using 60 µl of Pei transfection reagent. 2 days later, the virus soup was collected, filtered with 0,4 µM filter and concentrated 25 times with Lenti-X Concentrator. Cells were incubated with virus in the presence of Polybrene (8 µg/ml) for 24h. Positive clones were selected based on expression of mCherry, loss of GFP expression and further confirmed by PCR on genomic DNA using primers flanking the loxP sites (Table 33.).

4.2.6. Generation of H1Cx mutant cell lines

4.2.6.1. sgRNA cloning into pLX-sgRNA

In order to create knock-out of repair factors such as Lig4, nucLig3 and Parp1.

The knock-out was achieved by applying CRISPR/Cas9 technology. H1Cx cells are already stably expressing Cas9 protein. To deliver sgRNAs targeting the sequence of interest they were cloned into the pLX-sgRNA plasmid. MULE system couldn’t be used for this purpose, because it does not contain a destination vector with Blasticidin resistance.

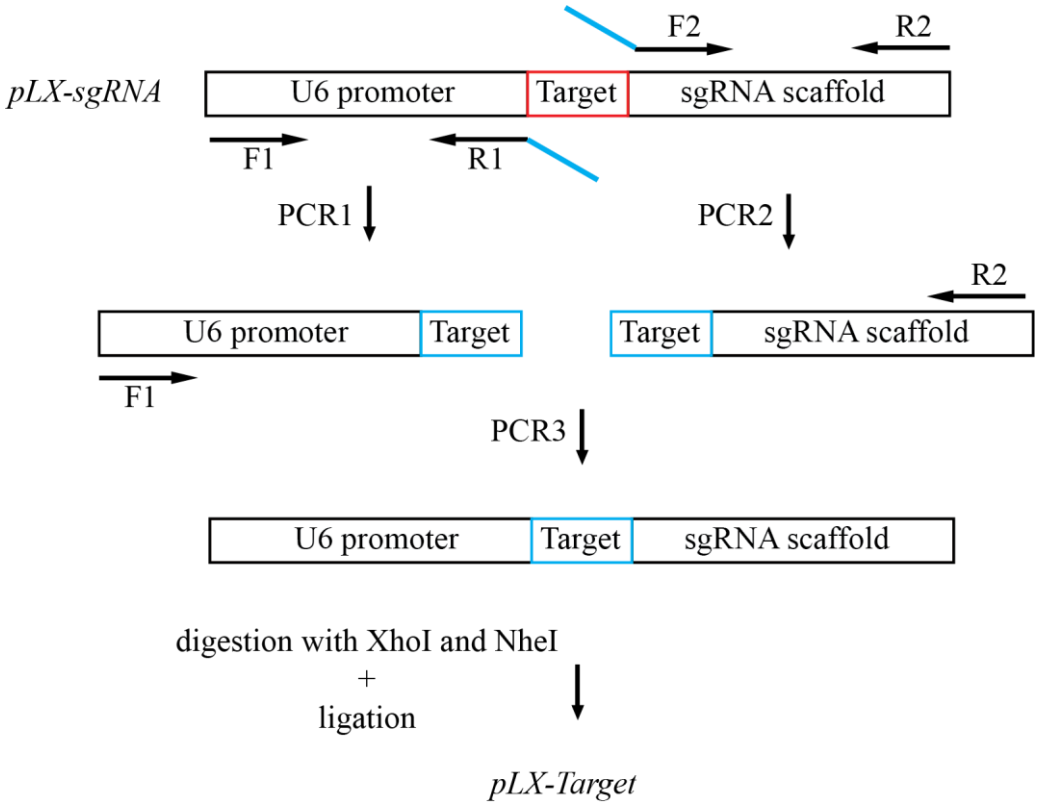


Figure 33. Workflow of sgRNA cloning into pLX-sgRNA.

The procedure of cloning sgRNAs into pLX-sgRNA is depicted on Figure 33. Target sgRNA sequences are cloned between the XhoI and NheI sites of the pLX-sgRNA. First, the target sequence was designed using CRISPR Design Tool crispr.mit.edu developed by Zhang lab. The goal was to create a frame-shift mutation within the coding sequence of the gene of interest.

For this cloning procedure, 4 primers were required: 2 specific to vector and 2 sgRNA-specific. In PCR 1 and PCR 2 pLX-sgRNA was used as a template with primers F1+R1 and F2+R2 respectively. Products of PCR1 and 2 were extracted from the gel using the Gel Extraction Kit and used as a template for PCR 3 with primers F1+R2.

After purification with the PCR Purification Kit, products and pLX-sgRNA were digested with XhoI and NheI enzymes (Table 11.).

Table 11. Digestion of pLX-sgRNA.

Component	**Amount**	
	pLX-sgRNA	PCR products
10x Buffer Cut Smart	1x	1x
XhoI	20 U	20 U
NheI	20 U	20 U
DNA	3 µg	800 ng
Water	up to 50 µl	up to 50 µl
Total	50 µl	50 µl

After digestion, all products were purified from the gel using the QIAquick Gel Extraction Kit. Ends of pLX-sgRNA were dephosphorylated with Antarctic Phosphatase Table 12.

Table 12. Dephosporylation of pLX sgRNA.

Component	**Amount**
Antarctic Phosphatase Reaction Buffer 10x	1x
Digested Vector	50 µl
Antarctic Phosphatase	2 µl
Water	up to 60 µl
Total	60 µl

Ligation of the insert with pLX-sgRNA was performed with Quick Ligation Kit for 5 minutes at RT (Table 13.) and followed by transformation of Stbl3 bacteria with the standard protocol using 4 µl of ligation mix.

Table 13. pLX-sgRNA-parameters of ligation reaction.

Component	**Amount**
Insert	8 ng
Vector	40 ng
2x Quick Ligase Buffer	1x
Quick Ligase	1 µl
Water	up to 21 µl
Total	21 µl

Resistant colonies were grown as a liquid culture and bacteria DNA was extracted with PureLink Quick Plasmid Miniprep Kit. The successful cloning was confirmed by sequencing with primer 239.

4.2.6.2. Generation of stable cell lines

HEK293T packaging cell line was split a day prior to transfection to achieve 80% confluency. Transfection mix included 0,25 µg of each packaging vector: pMDLg/pRRE, pRSV-REV, pMD2.G and 0,5 µg of specific vector (Table 14.) together with 3,75 µl of Xtreme Gene HP and 100 µl of Optimem. It was incubated for 30 minutes and added to HEK293T medium dropwise. The medium of HEK293T cells was changed after 6h and H1Cx cell line was prepared to reach 60% confluency on the following day.

Two days after transfection the medium of HEK293T cells containing virus was passed through the 0,45 µM filter and together with Polybrene (8 µg/ml) was used to exchange medium of H1Cx cells. On the following day, the medium of H1Cx cells was changed again. The selection with Blasticidin (5 µg/ml) was started 3 days after infection.

Table 14. Plasmids used for knock-outs of repair factors' genes.

Plasmid	Depleted factor
pLX_Lig4_sgRNA	Lig4
pLX_Parp1_sgRNA	Parp1
pLX_Lig3_1_sgRNA	Lig3_1
pLX_Lig3_2_sgRNA	Lig3_2
pLX_Lig3_3_sgRNA	Lig3_3

4.2.7. Generation of MULE expression plasmids

Multiple Lentiviral Expression System Kit (MULE) was used for simultaneous expression of multiple genetic elements from one plasmid. The goal was to introduce two sgRNAs and a reporter gene into an expression plasmid.

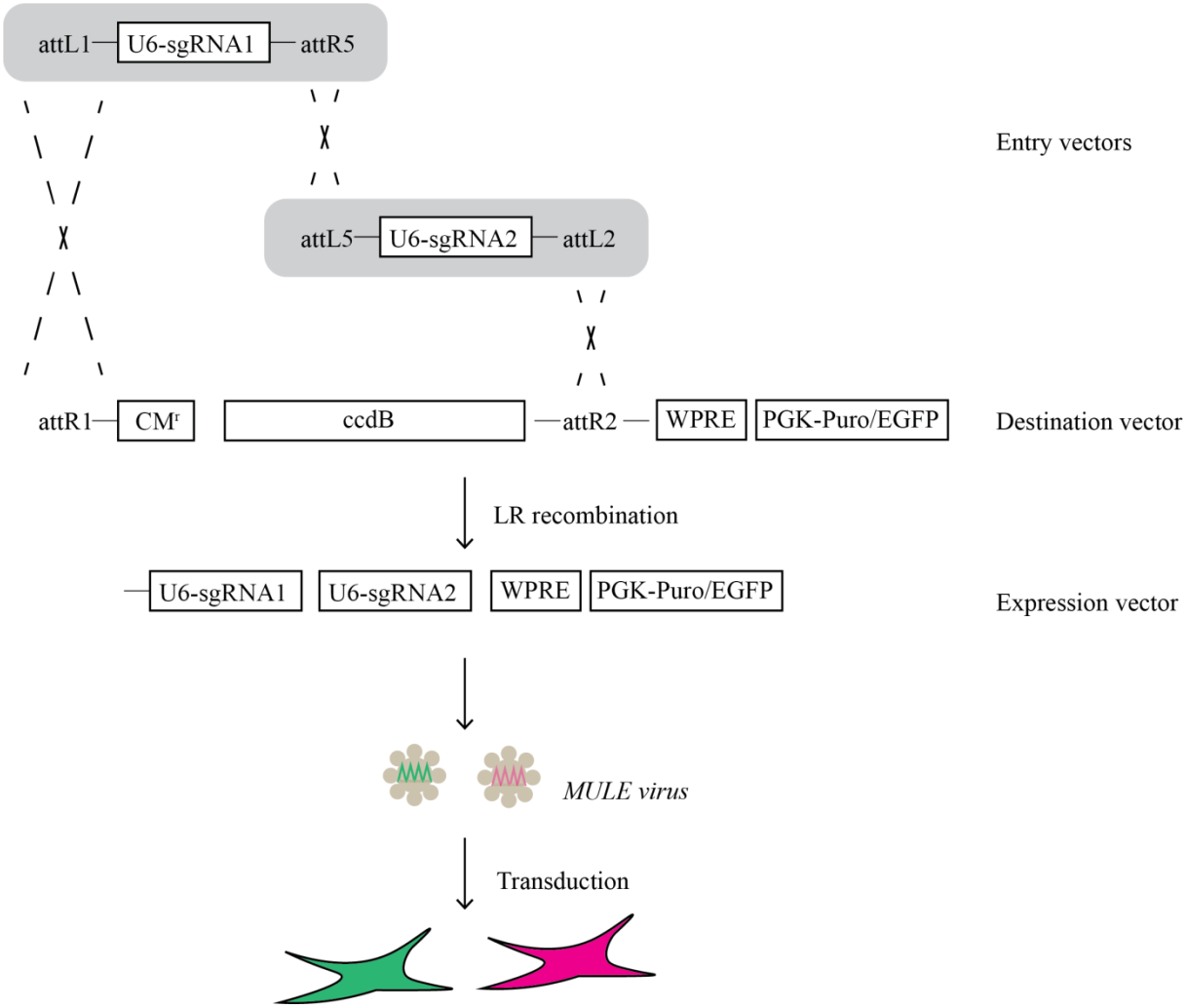


Figure 34. Workflow of Multiple Lentiviral Expression System Kit cloning.

4.2.7.1. Cloning into entry vectors

The design of sgRNA was performed with CRISPR Design Tool crispr.mit.edu developed by Zhang lab. sgRNAs with the highest scores (above 90) and targeting the genes in the suitable position were chosen. Cloning procedure was performed based on (Albers et al. 2015) and general cloning protocol "GeCKO – lentiviral CRISPR toolbox" from Zhang Lab. First, 5' overhangs were added to oligonucleotides – ACCG for the forward oligo and AAAC for the reverse complement one. Pairs of oligonucleotides were phosphorylated and annealed with parameters presented in Table 15.

Table 15. Annealing of oligonucleotides.

Component	Volume [µl]
oligo A (100 µM)	1
oligo B (100 µM)	1
10x T4 Ligation Buffer (NEB)	1
T4 PNK (NEB)	0.5
Water	up to 10
Total	10

Temperature [°C]	Time
37	30 min
95	5 min and then ramp down to 25°C at 5°C/min

Entry plasmids pMuLE ENTR U6 stuffer sgRNA scaffold L1-R5 and pMuLE ENTR U6 stuffer sgRNA scaffold L5-L2 were linearized for 1h at 50°C (Table 16). Vector ends were dephosphorylated using Antarctic Phosphatase for 30 min at 37°C followed by 5 min at 70°C (Table 17). Ligation of annealed oligonucleotides and digested entry vector was performed with the Quick Ligation™ Kit for 10 min at RT (Table 18.).

Table 16. Digestion of MULE entry vectors.

Component	Amount
Buffer 3.1 (NEB)	5 µl
Vector	5 µg
BfuAI	5U
Water	up to 50 µl
Total	50 µl

Table 17. Dephosphorylation of pMULE ENTR U6 plasmids with Antarctic Phosphatase.

Component	Amount
Antarctic Phosphatase Reaction Buffer 10x	1x
Digested Vector	50 µl
Antarctic Phosphatase	1 µl
Water	up to 60 µl
Total	60 µl

Table 18. Components of ligation reaction.

Component	Amount
Digested entry vector	50 ng
Oligonucleotides duplex (1:200)	1 µl
2x Quick Ligase Buffer	1x
Quick Ligase	1 µl
Water	up to 11 µl
Total	11 µl

4 µl of ligation products were used for transformation of 50 µl of One Shot® Stbl3™ bacteria. The reaction was incubated on ice for 30 min. Heat-shock of cells was carried at 42°C for 45 s in a water bath and followed by 2 min incubation on ice. Next, 250 µl of SOC was added to cells and incubated for 1h at 37°C in an incubator shaking at 180 rpm. Finally, cells were spread on the agar plate with the appropriate antibiotic and incubated overnight at 37°C. Resistant colonies were picked and grown as a liquid culture, which was used to extract bacteria DNA with PureLink Quick Plasmid Miniprep Kit on the following day. The completeness of the sgRNA cloning into the entry vector was confirmed by sequencing using primer 91.

4.2.7.2. Cloning into the destination vector

MULE destination vectors used in this study include pMuLE Lenti Dest eGFP and pMuLE Lenti Dest Neo. The recombination reaction between entry and destination vectors was performed using Gateway™ LR Clonase™ II Enzyme Mix.

LR reaction (Table 19) was performed at 25°C for 1 hour and inactivated with 1 µL of Proteinase K solution for 10 min at 37°C, which was followed by transformation of Stbl3 bacteria as described before. Finally, cells were spread on the agar plate with the appropriate antibiotic and incubated overnight at 37°C. Resistant colonies were picked and grown as a liquid culture, which was used to extract bacteria DNA on the following day. The successful cloning was confirmed by digestion and sequencing with primer 50 (Table 20).

Table 19. Components of LR reaction.

Component	Amount
Destination vector	75 ng
Entry 1	37.5 ng
Entry 2	37.5 ng
Clonase	1 µl
1xTE	up to 5 µl
Total	5 µl

Table 20. Digestion of MULE expression vector.

Component	Amount
10x Buffer Cut Smart	1x
Expression vector	100 ng
AgeI	10 U
EcoRV	10 U
Water	up to 10 µl
Total	10 µl

4.2.8. Generation of HAP1 cell line stably expressing HA-DD-I-SceI-GR

4.2.8.1. Cloning of HA-DD-ISceI-GR construct

The strategy was to create the DD-ISceI-GR fusion protein expressed under the doxycycline-inducible promoter in lentiviral plasmid pLentiCMV Tre 3G Neo Dest.

All components of the fusion protein were assembled together using GeneArt® Seamless Cloning and Assembly Kit. The first step included a design of oligonucleotides containing sequence specific to one of the components and overhangs allowing their recombination (Figure 20). All oligonucleotides are listed in Table 33. The designed fragments were amplified by PCR using pLVX-PTuner and ISceI-GR-RFP plasmids as a template (PCR Table 6) and purified with the PCR Purification Kit.

Prior to recombination reaction the entry plasmid pENTR1A was digested with KpnI and EcoRV and purified with the PCR Purification Kit (Table 21.).

Table 21. Digestion of pENTR1A vector.

Component	Amount
10x Buffer Cut Smart	1x
KpnI	30U
EcoRV	30U
pENTR1A vector	2.5 µg
Water	up to 50 µl
Total	50 µl

The recombination reaction was performed at RT for 30 min using 60 ng of entry plasmid and 70-100 ng of inserts (Table 22.). Directly afterwards 3 µl of the reaction mix were used to transform the DH10B™ $T1^R$ SA competent cells according to standard protocol.

Table 22. Seamless cloning – parameters of recombination reaction.

Component	**Amount**
Insert 1 - DD	70 ng
Insert 2 - I-SceI	78 ng
Insert3 - GR	107 ng
pENTR1A	60 ng
2x GeneArt® Enzyme Mix	5 µl
Water	up to 10 µl
Total	10 µl

Resistant colonies were grown as a liquid culture and bacteria DNA was extracted with the PureLink Quick Plasmid Miniprep Kit. To confirm successful recombination the DNA was digested with KpnI, EcoRI and SacII (Table 23.).

Table 23. Digestion of recombination products.

Component	**Amount**
10x Buffer Cut Smart	1x
KpnI	5U
EcoRI	5U
SacII	5U
DNA	50 ng
Water	up to 10 µl
Total	10 µl

In the next step, the HA tag was added to the DD-I-SceI-GR construct. Pairs of oligonucleotides were phosphorylated and annealed, according to "GeCKO – lentiviral CRISPR toolbox" from Zhang Lab (Table 24).

Table 24. Annealing of oligonucleotides for HA-tag.

Component	**Volume [µl]**
oligo 47 (100µM)	1
oligo 48 (100µM)	1
10x T4 Ligation Buffer (NEB)	1
T4 PNK (NEB)	0.5
Water	up to 10
Total	10

Temp. [°C]	**Time**
37	30 min
95	5 min and then ramp down to 25°C at 5°C/min

The entry vector pENTR1A-DD-I-SceI-GR was digested with KpnI Table 25. and prior to ligation with HA-tag, its ends were dephosphorylated for 30 minutes at 37°C and 5 minutes at 70°C (Table 26.).

Table 25. Digestion of pENTR1A-DD-I-SceI-GR.

Component	**Amount**
10x Buffer Cut Smart	1x
KpnI	10U
DNA	1.5 µg
Water	up to 50 µl
Total	50 µl

Table 26. Dephosphorylation of pENTR1A-DD-I-SceI-GR with Antarctic Phosphatase.

Component	**Amount**
Antarctic Phosphatase Reaction Buffer 10x	1x
Digested Vector	50 µl
Antarctic Phosphatase	1 µl
Water	up to 60 µl
Total	60 µl

The final ligation reaction took place at RT for 5 minutes with the Quick Ligation Kit (Table 27). Directly after, 4 µl of ligation mix were used for transformation of Stbl3 bacteria according to standard protocol. The accurate cloning was confirmed by sequencing using bacteria DNA extracted from resistant colonies and primer 49.

Table 27. Parameters of ligation reaction.

Component	**Amount**
Digested entry vector	30 ng
HA oligonucleotides duplex (1:200)	1 µl
2x Quick Ligase Buffer	1x
Quick Ligase	1 µl
Water	up to 11 µl
Total	11 µl

The sequencing results revealed that were was a mismatch within a sequence of DD component of the fusion protein, so it was modified using QuikChange II Site-Directed Mutagenesis Kit (Table 28).

Table 28. Parameters of site-directed mutagenesis reaction.

Component	**Amount**
10x Reaction Buffer	2,5 µl
pENTR1A-DD-I-SceI-GR	24 ng
Primer F (P57)	6 pM
Primer R (P58)	6 pM
10 mM dNTPs	0.5 µl
Polymerase Pfu Ultra	1.25U
Water	up to 25 µl
Total	25 µl

Cycles	**Temp. [°C]**	**Time [s]**
1x	95	30
12x	95	30
	55	60
	68	780
	68	120

To eliminate parental methylated DNA from the sample, products of mutagenesis reaction were digested with 5U of DpnI for 1h at 37°C. This was followed by the transformation of Stbl3 bacteria using 2 µl of reaction. Successful mutagenesis was tested by sequencing. The final step of creating the plasmid was the LR recombination reaction performed according to the protocol of Gateway™ LR Clonase™ II Enzyme mix (Table 29).

Table 29. Parameters of LR recombination reaction.

Component	**Amount**
pENTR1A-DD-I-SceI-GR	150 ng
Destination vector pLenti CMV Tre3G Neo	150 ng
5x LR Clonase™ II	1x
TE Buffer	up to 10 µl
Total	**10** µl

1 µl of the reaction was used to transform Stbl3 bacteria with a standard protocol. Successful cloning was confirmed by digestion with KpnI and sequencing with primer 50 (Table 30).

Table 30. Digestion of pENTR1A-HA-DD-I-SceI-GR.

Component	**Amount**
10x Buffer Cut Smart	1x
KpnI	5U
DNA	200 ng
Water	up to 10 µl
Total	10 µl

4.2.8.2. Generation of stable cell line expressing DD-I-SceI-GR

HEK293T packaging cell line was split a day prior to transfection to achieve 80% confluency. Transfection mix included 0,25 µg of each packaging vector: pMDLg/pRRE, pRSV-REV, pMD2.G and 0,5 µg of HA_DD_ISceI_GR_inpLentiCMVTre3GNeo or pLenti CMV rtTA3 Blast together with 3,75 µl of Xtreme Gene HP and 100 µl of Optimem. It was incubated for 30 minutes and added to HEK293T medium dropwise. The medium of HEK293T cells was changed after 6h and H1Cx cell line was prepared to reach 60% confluency on the following day.

Two days after transfection the medium of HEK293T cells containing virus was passed through the 0,45 µM filter and concentrated 5 times with Lenti-X Concentrator. Together with Polybrene (8 µg/ml), it was used to exchange the medium of H1Cx cells. On the following day, the medium of H1Cx cells was changed again. The selection with Blasticidin (10 µg/ml) and G418 (800 µg/ml) was started 6 days after infection.

4.2.9. Modification of PPP1R12C gene promoter

4.2.9.1. Preparation of expression plasmid

To inactivate the promoter of PPP1R12C gene a fragment of 460 bp upstream exon 1 was deleted using CRISPR/Cas9 system. sgRNAs targeting this region were cloned into pMULE Lenti Dest eGFP plasmid (according to the protocol described before) using oligonucleotides 789-792.

4.2.9.2. Generation of stable cell line expressing sgRNA targeting the promoter of PPP1R12C

HEK293T packaging cell line was split a day prior to transfection to achieve 80% confluency. Transfection mix included 0,25 µg of each packaging vector: pMDLg/pRRE, pRSV-REV, pMD2.G and 0,5 µg of pMULE-PPP1R12C-promoter-LentiDestGFP vector together with 3,75 µl of Xtreme Gene HP and 100 µl of Optimem. It was incubated for 30 minutes and added to HEK293T medium dropwise. The medium of HEK293T cells was changed after 6h and H1Cx-I-SceI (cl.#14 and cl.#26) cell line was prepared to reach 60% confluency on the following day.

Two days after transfection the medium of HEK293T cells containing virus was passed through the 0,45 µM, concentrated 5 times with Lenti-X Concentrator and together with Polybrene (8 µg/ml) was used to exchange medium of H1Cx-I-SceI cells. On the following day the medium of H1Cx-I-SceI cells was changed again. The selection of positive clones was based on GFP expression.

4.2.10. Generation of HAP1 cell line stably expressing Hygromycin-resistance gene

HEK293T packaging cell line was split a day prior to transfection to achieve 80% confluency. Transfection mix included 0,25 µg of each packaging vector: pMDLg/pRRE, pRSV-REV, pMD2.G and 0,5 µg of pLenti_CMV_rtta3_Hygro vector together with 3,75 µl of Xtreme Gene HP and 100 µl of Optimem. It was incubated for 30 minutes and added to HEK293T medium dropwise. The medium of HEK293T cells was changed after 6h and HAP1 cell line was prepared to reach 60% on the following day.

Two days after transfection the medium of HEK293T cells containing virus was passed through the 0,45 µM filter, concentrated 5 times with Lenti-X Concentrator and together with Polybrene (8 µg/ml) was used to exchange medium of HAP1 cells. On the following day, the medium of HAP1 cells was changed again. The selection of positive clones was based on Hygromycin (200-500 µg/ml) resistance and PCR analysis.

4.2.11. Optimization of cell culture conditions in soft agar

The goal of this experiments was to optimize conditions of HAP1 cell line cultivation in soft agar to ensure stable growth of Hygromycin-resistant colonies. Culture dishes contained 2 layers of agar: bottom layer 1% agar in 1:1 ratio with 2x IMDM and upper layer 0,6% in 1:1 ratio with 2xIMDM containing cells. Both layers contained antibiotic (Hygromycin) at 350/500 µg/ml. Pouring each layer was followed by 30 minutes incubation at RT to allow its solidification. On top of 2 agar layers, there was a 1x IMDM with the antibiotic.

4.3. Microscopy-based techniques

4.3.1. Immunofluorescence (IF) staining

Cells for IF were fixed on PLL coated coverslips/wells with 4% PFA for 15 min and washed 3 times with PBS. Then cells were permeabilized with PBS-0.3% Triton X for 5 min, followed by 3 washings with PBS. In the next step, the blocking solution of 3% BSA was applied for 1h. Incubation with primary antibodies was performed overnight at 4°C.

On the second day of the procedure, the sample was washed 3 times with PBS-Tween and incubated with secondary antibody for 1h. It was followed by another 3 washings with PBS-Tween and labelling with DNA stain Hoechst/DAPI (1:2000;1:1000) for 15 min. After final washings with PBS, the samples were ready for imaging.

4.3.2. Fluorescent in situ hybridization (FISH)

Cells were fixed on PLL coated coverslips with 4% PFA for 15 min and washed 3 times with PBS. It was followed by the 20 min incubation with 0.5% saponin/0.5% Triton X-100 in PBS and again washed 3 times with PBS. Then cells were treated with 0,1M HCL for 15 min, washed with 2xSSC for 10 min and incubated with 50% formamide in 2xSSC for 30 min.

FISH probes were fluorescently labelled by nick translocation using Nick Translation Kit according to manufacturers' protocol.

In the probe preparation the components were mixed as shown in Table 31. Afterwards the reaction was spun-down at 14700 rpm at 4°C for 20 min. The obtained pellet was dried in SpeedVac for 3 min and resuspended in 7 µl of hybridization buffer.

Table 31. Components of FISH probe reaction mix.

Reaction component	Amount per cover slip
probe DNA	80 ng
COT-1	3 µg
tRNA	20 µg
3M sodium acetate pH 5.2	1/10 Vol
100% ice-cold EtOH	2 Vol

To denature the probe it was hybridized on coverslip at 85°C for 5 min and incubated overnight at 37°C in a humid chamber. Next day the rubber, used to seal coverslip to the slide, was removed by 2xSSC solution. Then, the sample was washed 3 times with 1xSSC (preheated to 45°C) for 5 min, followed by another 3 washes of 5 min each with 0.1xSSC (preheated to 45°C). Afterwards, the coverslip was washed with water and mounted on the slide with 3 µl of Vectashield containing DAPI. Coverslips with cells were imaged with a high-content microscope (Opera Phenix) and quantification of red and green dots was performed with Harmony - automatic analysis software.

4.3.3. Preparation of metaphase spreads

Cells for preparation of metaphase spreads were prepared the day before so that on the day of treatment they were still in the phase of exponential growth.

To arrest mitotic cells in metaphase the Colcemid at a final concentration of 0.02 µg/ml was added to the culture medium and incubated for 90 min at 37°C. 30 min before the end of incubation, Calyculin (50nM) was added.

Afterwards, the medium was collected and mixed with PBS used to wash cells. Cells were trypsynised and spun down together with the collected medium for 7 min at 1000 rpm. The cell pellet was gently resuspended in 10 ml of 0.075 M KCl (preheated to 37°C) and incubated for 15 min at 37°C. In the following step, few drops of fixative (MeOH: Acetic acid – 3:1) were added and again cells were collected by 8 min centrifugation at 800 rpm. The cell pellet was resuspended in 1 ml of fixative and incubated for 60 min on ice. Final fixation included 3 repetitions of resuspending cell pellet in 1 ml of fixative and spinning down for 5 min at 500 rpm. The final volume had to be adjusted individually to each sample.

Coverslips for metaphase spreads were coated with PLL, washed with 70% EtOH and incubated at -20°C for 10 min before the procedure. A small volume of cells in fixative was dropped on coverslip from the height of 20 cm. Immediately after, the coverslip was transferred to a heating block prewarmed to 50°C and dried for 2 min.

Coverslips with cells were imaged with AF-7000 wide field microscope (Leica Biosystems) and acquired images were processed with ImageJ software.

4.3.4. Image acquisition and automatic image analysis

For microscopy imaging followed by signal quantification the Opera Phenix High Content Screening System (PerkinElmer) was used. Images of all HAP1 derived cell lines were acquired with 40x water objective NA 1.1 in spinning disc confocal mode. 4-8 planes in a distance of 1 μm were obtained with a camera binning of 2, which yields the pixel size of 300 nm. Nuclei were detected within a Hoechst channel for 60 ms at 50% laser power (LP). Other variable targets were detected within channels: Alexa 488 for 100-400 ms, Alexa 568 for 100-500 ms and Alexa 647 at 400 ms, all at 50% LP.

FISH experiments on AAVS1 locus were imaged obtaining 11 planes in a distance of 0,8 μm and over 100 fields per sample. Nuclei were detected within a DAPI channel for 240 ms at 100% LP. Fluorescent probes were detected within channels: Alexa 488 and Alexa 568 for 500 ms at 100% LP.

The following excitation/emission wavelengths were applied: Hoechst/DAPI (405/435-480), Alexa 488 (488/500-550), Alexa 568 (561/570-630), Alexa 647 (640/650-760).

All imaging data obtained with Opera Phenix was analysed with Harmony® Software (PerkinElmer). The segmentation of nuclei within Hoechst/DAPI channel was performed using a building block “Find Nuclei” Method B with common threshold 0,3 and area >15 μm^2. The quantification of mean intensity of each channel within the population “Nuclei selected” was calculated with a “Calculate intensity properties” function. Detection of the γH2AX foci was obtained with “Find spots” Method B with detection sensitivity of 0,48 and splitting sensitivity of 0,5.

Specifically for FISH analysis the segmentation of nuclei within DAPI channel was performed using a building block “Find Nuclei” Method M with common threshold 0,46 and splitting sensitivity of 0,3. Probe signal was detected as spots with “Find spots” Method C with radius ≤ 3,5, contrast 0,25 and uncorrected spot to region intensity > 0,25.

Images from FISH on metaphase spreads were examined under AF-7000 wide field microscope (Leica Biosystems) with a 100×/1.4 oil immersion objective and the ORCA Flash 4.0 camera (Hamamatsu). For each sample 10 planes in the distance of 2,46 μm were acquired. Images

were visualized with the LAS AF software (Leica Biosystems) and analysis was performed with ImageJ software.

4.3.5. Quantification of Hygromycin-resistant colonies

Cells were incubated with a staining solution (0,4% crystal violet/20% EtOH) for 30 min at RT. Afterwards, they were washed with water several times. Stained colonies were imaged with GelDoc XR+.

4.4. Materials

Table 32. List of plasmids used in this study.

Dat. number	Plasmid name	Source	Purpose
NA	pLVX-PTuner	Clontech #632173	amplify DD
NA	ISceI-GR-RFP	Addgene #17654	amplify GR and I-SceI
1	pENTR1A	Addgene #17398	entry plasmid
6	pLenti CMV rtTA3 Blast	Addgene #26429	expression of third-generation reverse tetracycline-regulated transactivator gene
8	pCW-Cas9	Addgene #50661	inducible expression of Cas9 protein
9	pLX-sgRNA	Addgene #50662	expression of sgRNA
10	LentiCRISPR v2	Addgene #52961	stable expression of Cas9 protein
21	pMDL_pRRE	gift from Richly lab	packaging plasmid
22	pRSV-REV	gift from Richly lab	packaging plasmid
23	pMD2.G	gift from Richly lab	packaging plasmid
39	HA_DD_ISceI_GR_inpLentiCMVTre3GNeo	Roukos lab	expression of HA-DD-I-SceI-GR fusion protein
42	pLenti_CMV_rtta3_Hygro	Addgene #26730	generation of cells stably expressing Hygromycin-resistance gene
49	LentiCRISPRv2_AVV1_sgRNA	Roukos lab	CRISPR/Cas9 mediated break

			within AAVS1 locus
50	pCMV_R_Cre	Addgene #27546	for expression of Cre recombinase
71	AAVS1hygroNEOP2aGFP	Invitrogen	contains selection cassette
187	pMULE_CS3_ISOC2_Lenti Dest Neo	Roukos lab	simultanous induction of DSBs by 2 sgRNAs
201	pMuleLentiDesteGFP	Roukos lab	
224	pMULE_CS3_ctrl_Lenti Dest eGFP	Roukos lab	
225	pMULE_CS3_ctrl_Lenti Dest_Neo	Roukos lab	
226	pMULE_CS3_LDHA1_Lenti Dest eGFP	Roukos lab	
227	pMULE_ctrl_LDHA1_Lenti Dest eGFP	Roukos lab	
229	pMULE_ctrl_PPP5C_Lenti Dest eGFP	Roukos lab	
230	pMULE_CS3_PPP5C_Lenti Dest eGFP	Roukos lab	
231	pMULE_ctrl_ctrl_Lenti Dest eGFP	Roukos lab	
313	pMULE-PPP1R12c-promoter-LentiDestGFP	Roukos lab	
244	pLX_Lig4_sgRNA	Roukos lab	generation of knock-out cell lines
245	pLX_Parp1_sgRNA	Roukos lab	
246	pLX_Lig3_1_sgRNA	Roukos lab	
247	pLX_Lig3_2_sgRNA	Roukos lab	
248	pLX_Lig3_3_sgRNA	Roukos lab	

Table 33. List of primers used in this study (all primers were purchased from Sigma).

Dat. number	Name	Sequence	Purpose
412	F_P71_cassette_pcr_end5	TGCTATACGAAGTTATGCGG	confirmation of selection cassette integration
413	R_P71_cassette_pcr_end5	CCCTGTGAAAGATGCCTG	
348	F_neo_cassette71_inside3	GACAATGGCCGCATAACAGC	
349	R_neo_cassette71_inside3	CCCAGTCATAGCCGAATAGCC	
400	F_P71_cassette_pcr_beg3	CTCTTCCGATGTTGAGCCCC	
401	R_P71_cassette_pcr_beg3	CCAGCACTAGCTGAATTACCCTG	
433	F_P71_Cre_out_PCR2	CCGTGGTTGGCTTGTATGGA	confirmation of excision by Cre recombinase
434	R_P71_Cre_out_PCR2	CTGGCAAGGAGAGAGATGGC	
449	F_PPP5C2_CS3_trans1	AAGGTTGGACCCTTTTTGGC	detection of chromosome rearrangements
450	R_PPP5C2_CS3_trans1	AGACTAGCTGAGCTCTCGGA	
634	F_AAVS1_Trans_PCR_3	CCTCCTTCAGGTTCCGTCTTC	

635	R_AAVS1_Trans_PCR_3	GGGAGATGCAAT AGGTCAGGC	
150	For CRISPR_lig4 2nd seq	ACGAGAAGATTC ATCACCGCTT	confirmation of frame shift mutations
151	Rev CRISPR_lig4 2nd seq	CGTTCACCATCTA GCTTGGTTT	
152	For CRISPR_Parp1 seq	TTCTAAAGTGTGG GAGGGGC	
153	Rev CRISPR_Parp1 seq	AGAACTGGTGGG AAAGCCTG	
892	F_Lig3_sgRNA1_seq1	GCCCAGTGACTTC AGCATTC	
893	R_Lig3_sgRNA1_seq1	AGGGATTGGGCA CCACTTTG	
1161	F_GAPDH_intron	CCCCACACACAT GCACTTACC	normalisation of ddPCR results
1162	R_GAPDH_intron	CCTAGTCCCAGG GCTTTGATT	
1163	F_GAPDH_exon	CTGCACCACCAA CTGCTTAG	normalisation of qPCR results
1164	R_GAPDH_exon	GTCTTCTGGGTGG CAGTGAT	
643	F_LDHA_T7PCR_1	AAAGCAGCGTCG AGTTTTGG	T7 PCR/detection of chromosome rearrangements
644	R_LDHA_T7PCR_1	AAGCCATGGTTCC CGCTTAG	
398	F__PPP5C_T7PCR_3	CAGCGACATTCA CAACCCAC	
399	R__PPP5C_T7PCR_3	CAGCACACTCAG TCCTCTCG	
161	F_AAVS1_CS2/3III_T7PCR	CCTGAGTCCGGA CCACTT	
162	R_AAVS1_CS2/3III_T7PCR	CCCGAATCCACA GGAGAACG	
956	R_Hygromycin_seq	CCT TTG CCC TCG GAC GAG T	checking expression levels of Hygromycin-resistance gene
927	F_CS3_LDHA_droplet_PCR	AACTGAAGGTCG TCCTGACT	detection of translocations with ddPCR
572	F_ISOC2_T7PCR_1	GGCGGGGAACAG ACAATAAAAG	detection of chromosome rearrangements
17	F_2a_VecDD	caattcagtcgactggatccg gtacCATGGGAGTG CAGGTGGAAACC ATC	cloning of HA-DD-ISceI-GR
18	R_2a_IsceIDD	GTT TTT CAT GCC AGC ACT AGC TGA TTC CGG TTT TAG AAG CTC CAC ATC	
19	F_2b_DDISceI	CCGGAATCAGCT AGTGCTGGCATG	

		AAAAACATCAAA AAAAAC	
20	R_2b_GRISceI	CCC GCG GTA GCC AGC ACT AGC TGA TTT CAG GAA AGT TTC GGA GG	
21	F_2c_ISceIGR	CCTGAAATCAGCT AGTGCTGGCTACC GCGGGTATCGGA AATGTCTTC	
22	R_2c_VecGR	AGA AAG CTG GGT CTA GAT TCA TCT AGA TCC GGT GGA TCC AAA TTT TTG	
47	For_KpnI-HA-KpnI	CATGtacccatacgatgtt cctgactatgcgGGTAC	
48	Rev_KpnI-HA-KpnI	CCGCATAGTCAG GAACATCGTATG GGTACATGGTAC	
49	F_pENTR1A_Seq	TAA ACT GCC AGG CAT CAA ACT	sequencing of ENTR1A
57	F_C2_pENTR1A_extraG_mut	GAC TAT GCG GGT ACC ATG GGA GTG CAG GTG GAA ACC	mutagenesis
58	R_C2_pENTR1A_extraG_mut	GGT TTC CAC CTG CAC TCC CAT GGT ACC CGC ATA GTC	
50	F_pLenti_HA_seq	TCCAGTGTGGTGG AATTC	sequencing of destination plasmids
258	F_mule_entry_sgRNA__AAVS1_CS3	ACCGAGG GAG ACA TCC GTC GGA GA	MULE cloning
259	R_mule_entry_sgRNA__AAVS1_CS3	AAACtctccgacggatgt ctccctC	
278	F-control-sgRNA-MULE	ACCG GTCATGTCACTTA TCAAGTC	
279	R-control-sgRNA-MULE	AAAC GAC TTG ATA AGT GAC ATG AC	
313	F_mule_entry_sgRNA_PPP5C1	ACCGCCCTATTGG CTCTCTCTACC	
314	R_mule_entry_sgRNA_PPP5C1	AAACGGTAGAGA GAGCCAATAGGG	
495	F_mule_entry_sgRNA_ISOC2	ACCGAAGGCCGA CCATGAATTGGG	
496	R_mule_entry_sgRNA_ISOC2	AAACCCC AAT TCA TGG TCG GCC TT	

628	F_LDHA_sgRNA1	ACCGACTTAGACT CCCAGCGCACG	
629	R_LDHA_sgRNA1	AAACCGT GCG CTG GGA GTC TAA GT	
789	F_mule_entry_sgRNA_PPP1R12C _1	ACCGACG CGG CTG TCC AGT CGA AT	
790	R_mule_entry_sgRNA_PPP1R12C _1	AAACATT CGA CTG GAC AGC CGC GT	
791	F_mule_entry_sgRNA_PPP1R12C _2	ACCGCGGGCGGG CGGTGCGATGTC	
792	R_mule_entry_sgRNA_PPP1R12C _2	AAACGAC ATC GCA CCG CCC GCC CG	
239	F_pLX-sgRNA_seq	cgggtttattacagggacag cag	pLX-sgRNA cloning
658	F1_pLXsgRNA_universal	AAACTCGAGTGT ACAAAAAAGCAG GCTTTAAAG	
659	R2_pLXsgRNA_universal	AAAGCTAGCTAA TGCCAACTTTGTA CAAGAAAGCTG	
660	F2_pLXsgRNA_lig4	GCTTATACGGATG ATCATAAGTTTTA GAGCTAGAAATA GCAA	
661	R1_pLXsgRNA_lig4	TTA TGA TCA TCC GTA TAA GC GGTGTTTCGTCCT TTCC	
662	F2_pLXsgRNA_parp1	GAAGTACGTGCA AGGGGTGTAGTTT TAGAGCTAGAAA TAGCAA	
663	R1_pLXsgRNA_parp1	TAC ACC CCT TGC ACG TAC TTCGGTGTTTCGT CCTTTCC	
319	F_P71_cassette_pcr_beg	ATGCCGTGTTCAC TCGCTG	T7 for I-SceI
956	R_Hygromycin_seq	CCT TTG CCC TCG GAC GAG T	

Table 34. List of Taq-Man probes used for ddPCR (all purchased from Integrated DNA Technologies IDT).

Database number	Name	Sequence	Purpose
1367	FAM-Hygro	5'-/56-FAM/TGC GTC CCG/ZEN/ CCT CCC CTT CTT GTA GGC CT/3IABkFQ/-3'	ddPCR
1166	HEX_GAPDH	5'-/5HEX/CTCCCACTC/ZEN/CTG ATTTCTGGAAAAGAGC/3IABkFQ/3'	

Table 35. List of primary antibodies used in this study.

Name	Product number	Supplier	Host organism	Conc.
Flag M2	20047221	Agilent Technologies	Mouse	1:1000
anti-phospho-histone H2A.X (Ser139), clone JBW301	05-636	Merck Millipore	Mouse monoclonal	1:2000
DNA Ligase IV	ab193353	Abcam	Rabbit	1:1000
PARP (46D11)	9532	Cell Signaling	Rabbit monoclonal	1:1000
DNA ligase III	GTX70145	GeneTex	Mouse monoclonal	1:1000
HA 3F10	11867423001	Sigma /Roche	Rat monoclonal	1:1000

Table 36. List of secondary antibodies used in this study.

Name	Product number	Supplier	Host organism	Concentration
Alexa Fluor 488 anti-rabbit IgG (H+L)	A11034	Mol. Probes	goat	1:1000
Alexa Fluor 488 anti-rat IgG (H + L)	A21208	Mol. Probes	donkey	1:1000
Alexa Fluor 568 anti-mouse IgG (H + L)	A10037	Mol. Probes	donkey	1:1000
Alexa Fluor 647 anti-mouse IgG (H + L)	A31571	Mol. Probes	donkey	1:1000

Table 37. List of BAC probes used in this study.

Name	Target	Purpose
RP11-434L20	C1S	to determine the proximity of chromosomes in HAP1 cells
RP11-1078C15	DSP	
RP11-384G4	AAVS1 locus	
RP11-1105H6	AF4-3'	
RP11-607L24	AF9-3'	
RP11-114A7	ENL-3'	
CTD-2159M9	MLL-5‘	
RP11-142P13	AAVS1 locus	multicolour analysis of AAVS1 locus
RP11-119A6	AAVS1 locus	

Table 38. List of kits used in this study.

Name	Supplier	Catalog number
DNeasy Blood & Tissue Kit	Qiagen	69506
QIAquick PCR Purification Kit	Qiagen	28106
Monarch® PCR & DNA Cleanup Kit	New England Biolabs	T1030S
QIAquick Gel Extraction Kit	Qiagen	28706
PureLink Quick Plasmid Miniprep Kit	Thermo Fisher Scientific	K210011
QIAfilter Plasmid Midi Kit	Qiagen	12245
Quick Ligation™ Kit	New England Biolabs	M2200
NucleoBond® Xtra BAC	Macherey-Nagel	740436.25
ZR BAC DNA Miniprep Kit	Zymo Research	D4048
RNeasy Plus Mini Kit	Qiagen	74134
GeneArt® Seamless Cloning and Assembly Kit	Invitrogen	A14603
QuikChange II Site-Directed Mutagenesis Kit	Agilent	200523
Nick Translation Kit	Abbott	7J0001
Neon™ Transfection System 100 µL Kit	Life technologies	MPK10025
pGEM®-T Easy Vector System I	Promega	A1360
Mix & Go! E. coli Transformation Buffer Set	Zymo Research	T3002
Qubit dsDNA HS Assay Kit	Thermo Fisher	Q32851

Table 39. List of bacteria strains used in this study.

Name	Supplier	Catalog number
DH10B™ $T1^R$ SA	Thermo Fisher Scientific	12331013
One Shot® Stbl3™	Invitrogen	C737303
XL1-Blue	Agilent Technologies	200523

Table 40. List of other reagents used in this study.

Name	Supplier	Catalog number
10x T4 Ligation Buffer	NEB	B0202S
10x NEB Buffer 2.1	NEB	B7202S
3M sodium acetate pH 5.2	Sigma-Aldrich	S7670
Accutase	Sigma-Aldrich	A6964-100ML
Acetic acid	Sigma-Aldrich	33209
Agarose	Biozym	840006
Ampicilin	IMB Core Facility	NA
Antarctic Phosphatase +buffer	NEB	B0289S
Blasticidin	Invivogen	ant-bl-1
BSA albumin sigma	Sigma	A3294-100g
10x Buffer 3.1	NEB	B7203S
Calyculin A	Santa Cruz Biotechnology	sc-24000
CellTiter-Blue Cell Viability Assay	Promega GmbH	G8080
Chloramphenicol	IMB Core Facility	NA

COT Human DNA	Sigma	11581074001
Crystal violet	Sigma-Aldrich	C6158-50G
ddPCR™ Supermix for Probes (No dUTP)	Biorad	#1863024
Dexamethasone	Sigma	D4902-25MG
Dextran sulfate	Sigma	67578
DMEM	Thermo Fisher Scientific	21969-035
DMSO	Sigma-Aldrich	D8418-500ml
DNA ladder 1kb	NEB	#N0468
DNA ladder 100 bp	NEB	#N0467
dNTPS	Invitrogen	10297-018
Doxycycline	Sigma	D9891-5G
EDTA	IMB Core Facility	NA
EtOH	Sigma	32205-4x2.5L
EtOH denatured	Sigma	24194-4x5L-R
Etoposide	Sigma	E1383-25mg
FBS	Gibco	16050-122
FBS-Tet approved	Biochrom	S0115
Formamide	Honeywell	47670-1L-F
G418	Invivogen	ant-gn-1
Gateway® LR Clonase® II Enzyme mix	Thermo Fisher Scientific	11791100
Gelatin	Sigma-Aldrich	G1393-20ml
GelRed	GeneON	#S420
Glucose	IMB Core Facility	NA
Hoechst	Molecular Probes	H3570
Hybridization buffer	10% dextran sulfate/ 50% formamide (deionized pH7)/ 2xSSC / 1% Tween20	NA
Hydrochloric acid	Sigma	30721-1L-GL
Hygromycin	Invivogen	ant-hg-5
IMDM (powder)	ThermoFisher	42200014
IPTG	Sigma	I6758-10G
Iscove's Modified Dulbecco's Medium (IMDM)	Fisher Scientific	21980065
Kanamycin	IMB Core Facility	NA
KaryoMAX™ Colcemid™ Solution in HBSS	ThermoFisher	15210-040
KCl	IMB Core Facility	NA
LB agar	IMB Core Facility	NA
Lenti-X Concentrator	Takara/Clontech	631232
L-Glutamine	Thermo Fisher Scientific	25030-024
Loading buffer	NEB	B7024S
Methanol	Sigma	32213-2.5L
Noble agar ultrapure (J10907 100GM)	Affymetrix USB	9002-18-0
Optimem	Invitrogen	31985047
Paraformaldehyde	Sigma	158127-3KG
PBS	Thermo Fisher Scientific	14190-094

Pei	Polysciences, Inc.	24885
Penicilin/Streptomycin	Thermo Fisher Scientific	15140-122
Pfu Ultra	Agilent	600670
Phusion	NEB	M0530
PLL	Sigma	P1274
Polybrene	Merck	TR-1003-G
Proteinase K	Active Motif	104492
Puromycin	Invivogen	ant-pr-1
Q5	NEB	M0491
Quick-Load® Taq 2X Master Mix	NEB	M0271L
random hexamers	ThermoFisher	N8080127
Saponin	Sigma	47036-50G-F
Shield-1	Takara/Clontech	632189
SOB Medium	Roth	AE27.1
SOC Medium	SOB+20 mM Glucose	NA
SSC	IMB Core Facility	NA
SSoFast	Biorad	172 5204
SsoFast™ EvaGreen® PCR Supermix	BioRad	172 5204
SuperScript™ II Reverse Transcriptase	Thermo Fisher Scientific	18064014
SybrSafe	ThermoFisher	S33102
T4 DNA ligase	NEB	M0202
T4 PNK	NEB	M0201
T4 Polynucleotide Kinase (PNK)	New England BioLabs	M0201 S
T7 endo	New England BioLabs	M0302 S
Triton X100	Sigma	T8787 -250 ml
Trypsin	Thermo Fisher Scientific	25200-056
Tween -20	Sigma	P9416-100 ml
VECTASHIELD Antifade Mounting Medium	Vector Laboratories	H-1000
Water	Fresenius Kabi	40676.00.00
X-Gal	Sigma	3117073001
Xtreme Gene HP	Sigma-Aldrich	06366236001
Yeast RNA 10tubes in 1 box AM7118	FiSher Scientific	10702487

Supplementary

Table S 1. Protocol of T7 PCR for checking efficiency of a cut within AAVS1 locus (A).

Reaction component	Final amount
5x Phusion HF Buffer	1x
10 mM dNTPs	200 µM
Primer 215	0.5 µM
Primer 216	0.5 µM
DNA	100 ng
Phusion Polymerase	0.02 U/µl
DMSO	1.5 µl
Water	up to 50 µl
Total	**50 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x35
Annealing	64.3	20	
Extension	72	30	
Final extension	72	600	
Hold	12		

Table S 2. Protocol of T7 PCR for checking efficiency of a cut within AAVS1 locus (B).

Reaction component	Final amount
5x Phusion HF Buffer	1x
10 mM dNTPs	200 µM
Primer 244	0.5 µM
Primer 245	0.5 µM
DNA	100 ng
Phusion Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	180	
Denaturation	9	10	x32
Annealing	6	20	
Extension	72	30	
Final extension	72	600	
Hold	12		

Table S 3. Protocol of T7 PCR for checking efficiency of a cut upstream Hygromycin-resistance gene.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 161	0.5 µM
Primer 162	0.5 µM
DNA	80 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x35
Annealing	64.3	20	
Extension	72	30	
Final extension	72	600	
Hold	12		

Table S 4. Protocol of T7 PCR for LDHA gene.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 643	0.5 µM
Primer 644	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x35
Annealing	72	20	
Extension	72	60	
Final extension	72	300	
Hold	12		

Table S 5. Protocol of T7 PCR for PPP5C gene.

Reaction component	Final amount
5x Phusion HF Buffer	1x
10 mM dNTPs	200 µM
Primer 398	0.5 µM
Primer 399	0.5 µM
DNA	100 ng
Phusion Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	60	
Denaturation	98	10	x35
Annealing	67.9	30	
Extension	72	60	
Final extension	72	300	
Hold	12		

Table S 6. T7 PCR for cutting efficiency of I-SceI endonuclease.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 319	0.5 µM
Primer 956	0.5 µM
DNA	150 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x32
Annealing	68.2	20	
Extension	72	300	
Final extension	72	300	
Hold	12		

Table S 7. Protocol of PCR to confirm mutation in Lig4 gene.

Reaction component	Final amount
5x Phusion HF Buffer	1x
10 mM dNTPs	200 µM
DMSO	3%
2.5mM $MgCl_2$	0.4 µl
Primer 150	0.5 µM
Primer 151	0.5 µM
DNA	125 ng
Phusion Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	180	
Denaturation	98	10	x35
Annealing	65	30	
Extension	72	30	
Final extension	72	300	
Hold	12		

Table S 8. Protocol of PCR to confirm mutation in Parp1 gene.

Reaction component	Final amount
5x Phusion HF Buffer	1x
10 mM dNTPs	200 µM
Primer 152	0.5 µM
Primer 153	0.5 µM
DNA	100 ng
Phusion Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	180	
Denaturation	98	30	x35
Annealing	67	30	
Extension	72	60	
Final extension	72	300	
Hold	12		

Table S 9. Protocol of PCR to confirm mutation in Lig3 gene.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 892	0.5 µM
Primer 893	0.5 µM
DNA	200 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	30	
Annealing	69.7	30	x30
Extension	72	240	
Final extension	72	300	
Hold	12		

Table S 10. Protocol of PCR for detection of deletion.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 572	0.5 µM
Primer 162	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x32
Annealing	67.9	20	
Extension	72	120	
Final extension	72	300	
Hold	12		

Table S 11. Protocol of PCR for detection of inversion.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 449	0.5 µM
Primer 450	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x32
Annealing	64.6	20	
Extension	72	90	
Final extension	72	300	
Hold	12		

Table S 12. Protocol of PCR for detection of translocation.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 643	0.5 µM
Primer 162	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x32
Annealing	66.6	20	
Extension	72	90	
Final extension	72	300	
Hold	12		

Table S 13. Protocol of PCR to confirm the presence of selection cassette at AAVS1 locus – 5' end.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 400	0.5 µM
Primer 401	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x30
Annealing	70	20	
Extension	72	240	
Final extension	72	300	
Hold	12		

Table S 14. Protocol of PCR to confirm the presence of selection cassette at AAVS1 locus – within the cassette.

Reaction component	Final amount
2x Quick-Load® Taq Master Mix	1x
Primer 348	0.5 µM
Primer 349	0.5 µM
DNA	40 ng
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	95	30	
Denaturation	95	30	x30
Annealing	59	60	
Extension	68	90	
Final extension	68	300	
Hold	12		

Table S 15. Protocol of PCR to confirm the presence of selection cassette at AAVS1 locus – 3' end.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 412	0.5 µM
Primer 413	0.5 µM
DNA	150 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x33
Annealing	65.5	20	
Extension	72	240	
Final extension	72	300	
Hold	12		

Table S 16. Protocol of PCR to check the efficiency of excision after expression of Cre recombinase.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 433	0.5 µM
Primer 434	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x30
Annealing	72	20	
Extension	72	150	
Final extension	72	120	
Hold	12		

Table S 17. Protocol of conventional PCR for the region containing I-SceI recognition sites.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 634	0.5 µM
Primer 635	0.5 µM
DNA	100 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	10	x35
Annealing	69.8	20	
Extension	72	60	
Final extension	72	300	
Hold	12		

Table S 18. Protocol of conventional PCR to test primers for ddPCR.

Reaction component	Final amount
5x Q5 Reaction Buffer	1x
10 mM dNTPs	200 µM
Primer 927	0.5 µM
Primer 162	0.5 µM
DNA	50 ng
Q5 Polymerase	0.02 U/µl
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	98	30	
Denaturation	98	20	x35
Annealing	61.2	20	
Extension	72	90	
Final extension	72	300	
Hold	12		

Table S 19. Protocol of PCR for cloning sgRNAs into pLX-sgRNA.

Reaction component	Final amount
2x Quick-Load® Taq Master Mix	1x
Primer F	0.5 µM
Primer R	0.5 µM
DNA	100 ng
Water	up to 25 µl
Total	**25 µl**

	Temp. [°C]	Time [s]	
Initial denaturation	95	30	
Denaturation	95	30	x30
Annealing	58-63	30	
Extension	68	60	
Final extension	68	300	
Hold	12		

Bibliography

Abel, Haley, John Pfeifer, and Eric Duncavage. 2015. *Translocation Detection Using*. *Clinical Genomics*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-404748-8.00010-1.

Abildinova, Gulshara, Zhanara Abdrakhmanova, Helena Tuchinsky, Elimelech Nesher, Albert Pinhasov, and Leon Raskin. 2016. "Fast Detection of Deletion Breakpoints Using Quantitative PCR." *Genetics and Molecular Biology* 39 (3): 365–69. https://doi.org/10.1590/1678-4685-GMB-2015-0159.

Albers, Joachim, Claudia Danzer, Markus Rechsteiner, Holger Lehmann, Laura P Brandt, Tomas Hejhal, Antonella Catalano, et al. 2015. "A Versatile Modular Vector System for Rapid Combinatorial Mammalian Genetics." *The Journal of Clinical Investigation* 125 (4): 1603–19. https://doi.org/10.1172/JCI79743DS1.

Allan, James M., and Lois B. Travis. 2005. "Mechanisms of Therapy-Related Carcinogenesis." *Nature Reviews Cancer* 5 (12): 943–55. https://doi.org/10.1038/nrc1749.

Aplan, D. 2011. "Efficient Repair of DNA Double-Strand Breaks in Malignant Cells with Structural Instability." *Cancer* 683: 115–22. https://doi.org/10.1016/j.mrfmmm.2009.10.016.

Aten, Jacob a, Jan Stap, Przemek M Krawczyk, Carel H van Oven, Ron a Hoebe, Jeroen Essers, and Roland Kanaar. 2004. "Dynamics of DNA Double-Strand Breaks Revealed by Clustering of Damaged Chromosome Domains." *Science (New York, N.Y.)* 303 (5654): 92–95. https://doi.org/10.1126/science.1088845.

Bemark, Mats, and Michael S. Neuberger. 2000. "The C-MYC Allele That Is Translocated into the IgH Locus Undergoes Constitutive Hypermutation in a Burkitt's Lymphoma Line." *Oncogene* 19 (30): 3404–10. https://doi.org/10.1038/sj.onc.1203686.

Blomen, Vincent A, Lucas T Jae, Johannes W Bigenzahn, Joppe Nieuwenhuis, Jacqueline Staring, Roberto Sacco, Ferdy R Van Diemen, et al. 2015. "Gene Essentiality and Synthetic Lethality in Haploid Human Cells." *Science* 350 (62).

Boch, Jens, and Ulla Bonas. 2010. "*Xanthomonas* AvrBs3 Family-Type III Effectors: Discovery and Function." *Annual Review of Phytopathology* 48 (1): 419–36. https://doi.org/10.1146/annurev-phyto-080508-081936.

Bont, Rinne De, and Nik van Larebeke. 2004. "Endogenous DNA Damage in Humans: A Review of Quantitative Data." *Mutagenesis* 19 (3): 169–85. https://doi.org/10.1093/mutage/geh025.

Brandsma, Inger, and Dik C. Gent. 2012. "Pathway Choice in DNA Double Strand Break Repair: Observations of a Balancing Act." *Genome Integrity* 3 (1): 1. https://doi.org/10.1186/2041-9414-3-9.

Brown, Jessica S, Stephen P Jackson, and Stephen P Jackson. 2015. "Ubiquitylation , Neddylation and the DNA Damage Response." *Open Biology*.

Brunet, E., D. Simsek, M. Tomishima, R. DeKelver, V. M. Choi, P. Gregory, F. Urnov, D. M. Weinstock, and M. Jasin. 2009. "Chromosomal Translocations Induced at Specified Loci in Human Stem Cells." *Proceedings of the National Academy of Sciences* 106 (26): 10620–25. https://doi.org/10.1073/pnas.0902076106.

Brunet, Erika, and Maria Jasin. 2018. “Induction of Chromosomal Translocations with CRISPR-Cas9 and Other Nucleases: Understanding the Repair Mechanisms That Give Rise to Translocations.” In , 15–25. https://doi.org/10.1007/978-981-13-0593-1_2.

Bunting, Samuel F., Elsa Callén, Nancy Wong, Hua-Tang Chen, Federica Polato, Amanda Gunn, Anne Bothmer, et al. 2010. “53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks.” *Cell* 141 (2): 243–54. https://doi.org/10.1016/j.cell.2010.03.012.

Bunting, Samuel F, and Andre Nussenzweig. 2013. “End-Joining, Translocations and Cancer.” *Nature Reviews. Cancer* 13 (7): 443–54. https://doi.org/10.1038/nrc3537.

Burman, Bharat, Tom Misteli, and Gianluca Pegoraro. 2015. “Quantitative Detection of Rare Interphase Chromosome Breaks and Translocations by High-Throughput Imaging.” *Genome Biology* 16 (1): 146. https://doi.org/10.1186/s13059-015-0718-x.

Burman, Bharat, Zhuzhu Z Zhang, Gianluca Pegoraro, Jason D Lieb, and Tom Misteli. 2015. “Histone Modifications Predispose Genome Regions to Breakage and Translocation.” *Genes & Development*, 1–10. https://doi.org/10.1101/gad.262170.115.GENES.

Cannan, Wendy J., Ishtiaque Rashid, Alan E. Tomkinson, Susan S. Wallace, and David S. Pederson. 2017. “The Human Ligase IIIα-XRCC1 Protein Complex Performs DNA Nick Repair after Transient Unwrapping of Nucleosomal DNA.” *Journal of Biological Chemistry* 292 (13): 5227–38. https://doi.org/10.1074/jbc.M116.736728.

Cao, Jennifer Yinuo, Aunoy Poddar, Leslie Magtanong, Jennifer H Lumb, Trevor R. Mileur, Michael A Reid, Cole M Dovey, et al. 2019. “A Genome-Wide Haploid Genetic Screen Identifies Regulators of Glutathione Abundance and Ferroptosis Sensitivity.” *Cell Reports* 26 (6): 1544-1556.e8. https://doi.org/10.1016/j.celrep.2019.01.043.

Casper, Anne M., Paul Nghiem, Martin F. Arlt, and Thomas W. Glover. 2002. “ATR Regulates Fragile Site Stability.” *Cell* 111 (6): 779–89. https://doi.org/10.1016/S0092-8674(02)01113-3.

Chandrasegaran, Srinivasan, and Dana Carroll. 2016. “Origins of Programmable Nucleases for Genome Engineering.” *Journal of Molecular Biology* 428 (5): 963–89. https://doi.org/10.1016/j.jmb.2015.10.014.

Chiarle, Roberto, Yu Zhang, Richard L. Frock, Susanna M. Lewis, Benoit Molinie, Yu Jui Ho, Darienne R. Myers, et al. 2011. “Genome-Wide Translocation Sequencing Reveals Mechanisms of Chromosome Breaks and Rearrangements in B Cells.” *Cell* 147 (1): 107–19. https://doi.org/10.1016/j.cell.2011.07.049.

Choi, Peter S, and Matthew Meyerson. 2014. “Targeted Genomic Rearrangements Using CRISPR/Cas Technology.” *Nature Communications* 5 (1): 3728. https://doi.org/10.1038/ncomms4728.

Chu, Gilbert, Lisa G. DeFazio, Rachel M. Stansel, and Jack D. Griffith. 2002. “Synapsis of DNA Ends by DNA-Dependent Protein Kinase.” *EMBO Journal* 21 (12): 3192–3200. https://doi.org/10.1093/emboj/cdf299.

Clark, Steven S, William M. Crist, and Owen N. Witte. 1989. “Molecular Pathogenesis of Ph-Positive Leukemias.” *Annual Review of Medicine* 40 (1): 113–22. https://doi.org/10.1146/annurev.me.40.020189.000553.

Cohen, Sarah, Nadine Puget, Yea-Lih Lin, Thomas Clouaire, Marion Aguirrebengoa, Vincent Rocher, Philippe Pasero, Yvan Canitrot, and Gaëlle Legube. 2018. "Senataxin Resolves RNA:DNA Hybrids Forming at DNA Double-Strand Breaks to Prevent Translocations." *Nature Communications* 9 (1): 533. https://doi.org/10.1038/s41467-018-02894-w.

Cremer, T., and C. Cremer. 2001. "Chromosome Territories, Nuclear Architecture and Gene Regulation in Mammalian Cells." *Nature Reviews Genetics* 2 (4): 292–301. https://doi.org/10.1038/35066075.

D'Adda di Fagagna, Fabrizio. 2014. "A Direct Role for Small Non-Coding RNAs in DNA Damage Response." *Trends in Cell Biology* 24 (3): 171–78. https://doi.org/10.1016/j.tcb.2013.09.008.

Day, Tovah A., Jacob V. Layer, J. Patrick Cleary, Srijoy Guha, Kristen E. Stevenson, Trevor Tivey, Sunhee Kim, et al. 2017. "PARP3 Is a Promoter of Chromosomal Rearrangements and Limits G4 DNA." *Nature Communications* 8. https://doi.org/10.1038/ncomms15110.

DeKelver, Russell C., Vivian M. Choi, Erica A. Moehle, David E. Paschon, Dirk Hockemeyer, Sebastiaan H. Meijsing, Yasemin Sancak, et al. 2010. "Functional Genomics, Proteomics, and Regulatory DNA Analysis in Isogenic Settings Using Zinc Finger Nuclease-Driven Transgenesis into a Safe Harbor Locus in the Human Genome." *Genome Research* 20 (8): 1133–42. https://doi.org/10.1101/gr.106773.110.

Difilippantonio, Michael J., Jie Zhu, Hua Tang Chen, Eric Meffre, Michel C. Nussenzweig, Edward E. Max, Thomas Ried, and André Nussenzweig. 2000. "DNA Repair Protein Ku80 Suppresses Chromosomal Aberrations and Malignant Transformation." *Nature* 404 (6777): 510–14. https://doi.org/10.1038/35006670.

Difilippantonio, Simone, Eric Gapud, Nancy Wong, Ching Yu Huang, Grace Mahowald, Hua Tang Chen, Michael J. Kruhlak, et al. 2008. "53BP1 Facilitates Long-Range DNA End-Joining during V(D)J Recombination." *Nature* 456 (7221): 529–33. https://doi.org/10.1038/nature07476.

Do, To Uyen, Bay Ho, Shyh Jen Shih, and Andrew Vaughan. 2012. "Zinc Finger Nuclease Induced DNA Double Stranded Breaks and Rearrangements in MLL." *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* 740 (1–2): 34–42. https://doi.org/10.1016/j.mrfmmm.2012.12.006.

Dobbs, Tracey A., John A. Tainer, and Susan P. Lees-Miller. 2010. "A Structural Model for Regulation of NHEJ by DNA-PKcs Autophosphorylation." *DNA Repair* 9 (12): 1307–14. https://doi.org/10.1016/j.dnarep.2010.09.019.

Dong, Y., R-C Du, Y-T Jiang, J. Wu, L-L Li, and R-Z Liu. 2012. "Impact of Chromosomal Translocations on Male Infertility, Semen Quality, Testicular Volume and Reproductive Hormone Levels." *Journal of International Medical Research* 40 (6): 2274–83. https://doi.org/10.1177/030006051204000625.

Dong, Zirui, Lingfei Ye, Zhenjun Yang, Haixiao Chen, Jianying Yuan, Huilin Wang, Xiaosen Guo, et al. 2018. "Balanced Chromosomal Rearrangement Detection by Low-Pass Whole-Genome Sequencing." *Current Protocols in Human Genetics* 96 (1): 8.18.1-8.18.16. https://doi.org/10.1002/cphg.51.

Doudna, Jennifer A., and Emmanuelle Charpentier. 2014. “The New Frontier of Genome Engineering with CRISPR-Cas9.” *Science* 346 (6213). https://doi.org/10.1126/science.1258096.

Druker, Brian J. 2008. “Translation of the Philadelphia Chromosome into Therapy for CML.” *Blood* 112 (13): 4808–17. https://doi.org/10.1182/blood-2008-07-077958.

Dueva, Rositsa, and George Iliakis. 2013. “Alternative Pathways of Non-Homologous End Joining (NHEJ) in Genomic Instability and Cancer.” *Translational Cancer Research* 2 (3): 163–77. https://doi.org/10.3978/j.issn.2218-676X.2013.05.02.

Elliott, Beth, Christine Richardson, and Maria Jasin. 2005. “Chromosomal Translocation Mechanisms at Intronic Alu Elements in Mammalian Cells.” *Molecular Cell* 17 (6): 885–94. https://doi.org/10.1016/j.molcel.2005.02.028.

Essletzbichler, Patrick, Tomasz Konopka, Federica Santoro, Doris Chen, Bianca V Gapp, Robert Kralovics, Thijn R Brummelkamp, and Sebastian M B Nijman. 2014. “Megabase-Scale Deletion Using CRISPR / Cas9 to Generate a Fully Haploid Human Cell Line.” *Genome Research*, 2059–65. https://doi.org/10.1101/gr.177220.114.

Eykelenboom, Jennifer E., Gareth J. Briggs, Nicholas J. Bradshaw, Dinesh C. Soares, Fumiaki Ogawa, Sheila Christie, Elise L.V. Malavasi, et al. 2012. “A t(1;11) Translocation Linked to Schizophrenia and Affective Disorders Gives Rise to Aberrant Chimeric DISC1 Transcripts That Encode Structurally Altered, Deleterious Mitochondrial Proteins.” *Human Molecular Genetics* 21 (15): 3374–86. https://doi.org/10.1093/hmg/dds169.

Fessler, Evelyn, and Lucas T Jae. 2018. “Haploid Screening for the Identification of Host Factors in Virus Infection.” In , 1836:121–37. https://doi.org/10.1007/978-1-4939-8678-1_6.

Flavia Michelini, Sethuramasundaram Pitchiaya, Valerio Vitelli, Sheetal Sharma, Ubaldo Gioia, Fabio Pessina, Matteo Cabrini, Yejun Wang, Ilaria Capozzo, Fabio Iannelli, Valentina Matti, Sofia Francia, G.V. Shivashankar, Nils G. Walter, and Fabrizio d’Adda di Fagagna. 2017. “Damage-Induced LncRNAs Control the DNA Damage Response through Interaction with DDRNAs at Individual Double-Strand Breaks.” *Nature Cell Biology* 19 (12): 1400–1411. https://doi.org/10.1038/ncb3643.

Fonslow, Bryan R, Benjamin D Stein, Kristofor J Webb, Tao Xu, Jeong Choi, Sung Kyu, and John R Yates Iii. 2015. “DNA Ligase IV Regulates XRCC4 Nuclear Localisation.” *DNA Repair* 21: 36–42. https://doi.org/10.1038/nmeth.2250.Digestion.

Frit, Philippe, Nadia Barboule, Ying Yuan, Dennis Gomez, and Patrick Calsou. 2014. “Alternative End-Joining Pathway(s): Bricolage at DNA Breaks.” *DNA Repair* 17: 81–97. https://doi.org/10.1016/j.dnarep.2014.02.007.

Gabriel, Richard, Ina Kutschera, Cynthia C Bartholomae, Christof von Kalle, and Manfred Schmidt. 2014. “Linear Amplification Mediated PCR - Localization of Genetic Elements and Characterization of Unknown Flanking DNA.” *Journal of Visualized Experiments*, no. 88 (June): 1–8. https://doi.org/10.3791/51543.

Gao, Qingsong, Wen-wei Liang, Steven M Foltz, Matti Nykter, Qingsong Gao, Wen-wei Liang, Steven M Foltz, Gnanavel Mutharasu, and Reyka G Jayasinghe. 2018. “Driver Fusions and Their Implications in the Development and Treatment of Human Cancers Resource Driver Fusions and Their Implications in the Development and Treatment of Human Cancers,” 227–38. https://doi.org/10.1016/j.celrep.2018.03.050.

Gao, Yankun, Sachin Katyal, Youngsoo Lee, Jingfeng Zhao, Jerold E. Rehg, Helen R. Russell, and Peter J. McKinnon. 2011. “DNA Ligase III Is Critical for MtDNA Integrity but Not Xrcc1-Mediated Nuclear DNA Repair.” *Nature* 471 (7337): 240–46. https://doi.org/10.1038/nature09773.

Gelot, Camille, Josée Guirouilh-Barbat, Tangui Le Guen, Elodie Dardillac, Catherine Chailleux, Yvan Canitrot, and Bernard S. Lopez. 2016. “The Cohesin Complex Prevents the End Joining of Distant DNA Double-Strand Ends.” *Molecular Cell* 61 (1): 15–26. https://doi.org/10.1016/j.molcel.2015.11.002.

Germini, Diego, Yara Bou Saada, Tatiana Tsfasman, Kristina Osina, Chloé Robin, Nikolay Lomov, Mikhail Rubtsov, Nikolajs Sjakste, Mar≳ Lipinski, and Yegor Vassetzky. 2017. “A One-Step PCR-Based Assay to Evaluate the Efficiency and Precision of Genomic DNA-Editing Tools.” *Molecular Therapy - Methods and Clinical Development* 5 (June): 43–50. https://doi.org/10.1016/j.omtm.2017.03.001.

Ghezraoui, Hind, Marion Piganeau, Benjamin Renouf, Jean-baptiste Renaud, Annahita Sallmyr, Brian Ruis, Sehyun Oh, et al. 2014. “Chromosomal Translocations in Human Cells Are Generated by Canonical Nonhomologous End-Joining.” *Molecular Cell* 55 (6): 829–42. https://doi.org/10.1016/j.molcel.2014.08.002.

Giglia-Mari, Giuseppina, Angelika Zotter, and Wim Vermeulen. 2011. “DNA Damage Response.” *Cold Spring Harbor Perspectives in Biology* 3 (1): 1–19. https://doi.org/10.1101/cshperspect.a000745.

Glover, Thomas W., Martin F. Arlt, Anne M. Casper, and Sandra G. Durkin. 2005. “Mechanisms of Common Fragile Site Instability.” *Human Molecular Genetics* 14 (SUPPL. 2): 197–205. https://doi.org/10.1093/hmg/ddi265.

Gunji, Hisako, Kazuo Waga, Fumihiko Nakamura, Kazuhiro Maki, Ko Sasaki, Yuichi Nakamura, and Kinuko Mitani. 2004. “TEL/AML1 Shows Dominant-Negative Effects over TEL as Well as AML1.” *Biochemical and Biophysical Research Communications* 322 (2): 623–30. https://doi.org/10.1016/j.bbrc.2004.07.169.

Gupta, Ankit, Ryan G. Christensen, Amy L. Rayla, Abirami Lakshmanan, Gary D. Stormo, and Scot A. Wolfe. 2012. “An Optimized Two-Finger Archive for ZFN-Mediated Gene Targeting.” *Nature Methods* 9 (6): 588–90. https://doi.org/10.1038/nmeth.1994.

Hakim, Ofir, Wolfgang Resch, Arito Yamane, Isaac Klein, Kyong-Rim Kieffer-Kwon, Mila Jankovic, Thiago Oliveira, et al. 2012. “DNA Damage Defines Sites of Recurrent Chromosomal Translocations in B Lymphocytes.” *Nature* 484 (7392): 69–74. https://doi.org/10.1038/nature10909.

Hangas, Anu, and Jaakko L O Pohjoismäki. 2019. “Twist and Turn — Topoisomerase Functions in Mitochondrial DNA Maintenance.”

Harewood, Louise, Kamal Kishore, Matthew D. Eldridge, Steven Wingett, Danita Pearson, Stefan Schoenfelder, V. Peter Collins, and Peter Fraser. 2017. “Hi-C as a Tool for Precise Detection and Characterisation of Chromosomal Rearrangements and Copy Number Variation in Human Tumours.” *Genome Biology* 18 (1): 1–11. https://doi.org/10.1186/s13059-017-1253-8.

Helmrich, Anne, Monica Ballarino, and Laszlo Tora. 2011. “Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes.” *Molecular Cell* 44 (6): 966–77. https://doi.org/10.1016/j.molcel.2011.10.013.

Hermetz, Karen E., Urvashi Surti, Jannine D. Cody, and M. Katharine Rudd. 2012. “A Recurrent Translocation Is Mediated by Homologous Recombination between HERV-H Elements.” *Molecular Cytogenetics* 5 (1): 6. https://doi.org/10.1186/1755-8166-5-6.

Hu, Jiazhi, Robin M Meyers, Junchao Dong, Rohit A Panchakshari, Frederick W Alt, and Richard L Frock. 2016. “Detecting DNA Double-Stranded Breaks in Mammalian Genomes by Linear Amplification – Mediated High-Throughput Genome-Wide Translocation Sequencing.” *Nature Protocols* 11 (5): 853–71. https://doi.org/10.1038/nprot.2016.043.

Indiviglio, S. M., and A. A. Bertuch. 2009. “Ku’s Essential Role in Keeping Telomeres Intact.” *Proceedings of the National Academy of Sciences* 106 (30): 12217–18. https://doi.org/10.1073/pnas.0906427106.

Jang, Ye Eun, Insu Jang, Sunkyu Kim, Subin Cho, Daehan Kim, Keonwoo Kim, Jaewon Kim, et al. 2020. “ChimerDB 4 . 0 : An Updated and Expanded Database of Fusion Genes” 48 (November 2019): 817–24. https://doi.org/10.1093/nar/gkz1013.

Jankovic, Mila, Niklas Feldhahn, Thiago Y. Oliveira, Israel T. Silva, Kyong-Rim Kieffer-Kwon, Arito Yamane, Wolfgang Resch, et al. 2013. “53BP1 Alters the Landscape of DNA Rearrangements and Suppresses AID-Induced B Cell Lymphoma.” *Molecular Cell* 49 (4): 623–31. https://doi.org/10.1016/j.molcel.2012.11.029.

Jiang, Yanwen, Isabelle Lucas, and Michelle M Le Beau. 2015. *Chromosomal Translocations and Genome Rearrangements in Cancer*. https://doi.org/10.1007/978-3-319-19983-2.

Joung, J Keith, and Jeffry D Sander. 2013. “TALENs: A Widely Applicable Technology for Targeted Genome Editing.” *Nature Reviews Molecular Cell Biology* 14 (1): 49–55. https://doi.org/10.1038/nrm3486.

Kim, Pora, and Xiaobo Zhou. 2019. “FusionGDB : Fusion Gene Annotation DataBase” 47 (November 2018). https://doi.org/10.1093/nar/gky1067.

Kim, Y G, J Cha, and S Chandrasegaran. 1996. “Hybrid Restriction Enzymes: Zinc Finger Fusions to Fok I Cleavage Domain.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (3): 1156–60. http://www.ncbi.nlm.nih.gov/pubmed/8577732%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC40048.

Klein, Isaac a, Wolfgang Resch, Mila Jankovic, Thiago Oliveira, Arito Yamane, Hirotaka Nakahashi, Michela Di Virgilio, et al. 2011. “Translocation-Capture Sequencing Reveals the Extent and Nature of Chromosomal Rearrangements in B Lymphocytes.” *Cell* 147 (1): 95–106. https://doi.org/10.1016/j.cell.2011.07.048.

Krawczyk, P. M., T. Borovski, J. Stap, T. Cijsouw, R. t. Cate, J. P. Medema, R. Kanaar, N. A. P. Franken, and J. A. Aten. 2012. “Chromatin Mobility Is Increased at Sites of DNA Double-Strand Breaks.” *Journal of Cell Science* 125 (9): 2127–33. https://doi.org/10.1242/jcs.089847.

Kruhlak, Michael J., Arkady Celeste, Graham Dellaire, Oscar Fernandez-Capetillo, Waltraud G. Müller, James G. McNally, David P. Bazett-Jones, and André Nussenzweig. 2006. “Changes in Chromatin Structure and Mobility in Living Cells at Sites of DNA Double-Strand Breaks.” *Journal of Cell Biology* 172 (6): 823–34. https://doi.org/10.1083/jcb.200510015.

Lee, Myunggyo, Kyubum Lee, Namhee Yu, Insu Jang, Ikjung Choi, Pora Kim, Eun Jang, et al. 2017. “ChimerDB 3 . 0 : An Enhanced Database for Fusion Genes from Cancer Transcriptome and Literature Data Mining” 45 (November 2016): 784–89. https://doi.org/10.1093/nar/gkw1083.

Liang, Feng, and Maria Jasin. 1996. “Ku80-Deficient Cells Exhibit Excess Degradation of Extrachromosomal DNA.” *Journal of Biological Chemistry* 271 (24): 14405–11. https://doi.org/10.1074/jbc.271.24.14405.

Libura, Jolanta, Diana J. Slater, Carolyn A. Felix, and Christine Richardson. 2005. “Therapy-Related Acute Myeloid Leukemia-like MLL Rearrangements Are Induced by Etoposide in Primary Human CD34 + Cells and Remain Stable after Clonal Expansion.” *Blood* 105 (5): 2124–31. https://doi.org/10.1182/blood-2004-07-2683.

Lieber, Michael R. 2008. “The Mechanism of Human Nonhomologous DNA End Joining.” *Journal of Biological Chemistry* 283 (1): 1–5. https://doi.org/10.1074/jbc.R700039200.

Lin, Cheng-Yu, Ankit Shukla, John Grady, J. Fink, Eloise Dray, and Pascal Duijf. 2018. “Translocation Breakpoints Preferentially Occur in Euchromatin and Acrocentric Chromosomes.” *Cancers* 10 (1): 13. https://doi.org/10.3390/cancers10010013.

Lin, Ming Tseh, Li Hui Tseng, Roy G. Rich, Michael J. Hafez, Shuko Harada, Kathleen M. Murphy, James R. Eshleman, and Christopher D. Gocke. 2011. “Δ-PCR, a Simple Method to Detect Translocations and Insertion/Deletion Mutations.” *Journal of Molecular Diagnostics* 13 (1): 85–92. https://doi.org/10.1016/j.jmoldx.2010.11.004.

Lu, Guangqing, Jinzhi Duan, Sheng Shu, Xuxiang Wang, Linlin Gao, Jing Guo, and Yu Zhang. 2016. “Ligase I and Ligase III Mediate the DNA Double-Strand Break Ligation in Alternative End-Joining.” *Proceedings of the National Academy of Sciences* 113 (5): 1256–60. https://doi.org/10.1073/pnas.1521597113.

Lu, Wei Ting, Ben R. Hawley, George L. Skalka, Robert A. Baldock, Ewan M. Smith, Aldo S. Bader, Michal Malewicz, Felicity Z. Watts, Ania Wilczynska, and Martin Bushell. 2018. “Drosha Drives the Formation of DNA:RNA Hybrids around DNA Break Sites to Facilitate DNA Repair.” *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-02893-x.

Ma, Chu Jian, Bryan Gibb, Youngho Kwon, Patrick Sung, and Eric C. Greene. 2017. “Protein Dynamics of Human RPA and RAD51 on SsDNA during Assembly and Disassembly of the RAD51 Filament.” *Nucleic Acids Research* 45 (2): 749–61. https://doi.org/10.1093/nar/gkw1125.

Ma, Yunmei, Ulrich Pannicke, Klaus Schwarz, and Michael R Lieber. 2002. “Hairpin Opening and Overhang Processing by an Artemis/DNA-Dependent Protein Kinase Complex in Nonhomologous End Joining and V(D)J Recombination.” *Cell* 108 (6): 781–94. http://www.ncbi.nlm.nih.gov/pubmed/11955432.

Mani, R.-S., Scott A Tomlins, Kaitlin Callahan, Aparna Ghosh, M. K. Nyati, Sooryanarayana Varambally, Nallasivam Palanisamy, and A. M. Chinnaiyan. 2009. “Induced Chromosomal Proximity and Gene Fusions in Prostate Cancer.” *Science* 326 (5957): 1230–1230. https://doi.org/10.1126/science.1178124.

Marnef, Aline, Sarah Cohen, and Gaëlle Legube. 2017. “Transcription-Coupled DNA Double-Strand Break Repair: Active Genes Need Special Care.” *Journal of Molecular Biology* 429 (9): 1277–88. https://doi.org/10.1016/j.jmb.2017.03.024.

Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. “A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science* 337 (6096): 816–21. https://doi.org/10.1126/science.1138140.

Mateos-Gomez, Pedro A., Fade Gong, Nidhi Nair, Kyle M. Miller, Eros Lazzerini-Denchi, and Agnel Sfeir. 2015. “Mammalian Polymerase θ Promotes Alternative NHEJ and Suppresses Recombination.” *Nature* 518 (7538): 254–57. https://doi.org/10.1038/nature14157.

Mathas, S., S. Kreher, K. J. Meaburn, K. Johrens, B. Lamprecht, C. Assaf, W. Sterry, et al. 2009. “Gene Deregulation and Spatial Genome Reorganization near Breakpoints Prior to Formation of Translocations in Anaplastic Large Cell Lymphoma.” *Proceedings of the National Academy of Sciences* 106 (14): 5831–36. https://doi.org/10.1073/pnas.0900912106.

Meaburn, Karen J., Tom Misteli, and Evi Soutoglou. 2007. “Spatial Genome Organization in the Formation of Chromosomal Translocations.” *Seminars in Cancer Biology* 17 (1): 80–90. https://doi.org/10.1016/j.semcancer.2006.10.008.

Meisenberg, Cornelia, Sarah I Pinder, Suzanna R Hopkins, Frances M G Pearl, Penny A Jeggo, Jessica A Downs, Frances M G Pearl, Penny A Jeggo, and Jessica A Downs. 2019. “Repression of Transcription at DNA Breaks Requires Cohesin throughout Interphase and Prevents Article Repression of Transcription at DNA Breaks Requires Cohesin throughout Interphase and Prevents Genome Instability.” *Molecular Cell* 73 (2): 212-223.e7. https://doi.org/10.1016/j.molcel.2018.11.001.

Mertens, Fredrik, Bertil Johansson, Thoas Fioretos, and Felix Mitelman. 2015. “The Emerging Complexity of Gene Fusions in Cancer.” *Nature Reviews Cancer* 15 (6): 371–81. https://doi.org/10.1038/nrc3947.

Miles, Linde A., Ralph J. Garippa, and John T. Poirier. 2016. “Design, Execution, and Analysis of Pooled in Vitro CRISPR/Cas9 Screens.” *FEBS Journal* 283: 3170–80. https://doi.org/10.1111/febs.13770.

Mitelman, Felix, Bertil Johansson, and Fredrik Mertens. 2007. “The Impact of Translocations and Gene Fusions on Cancer Causation.” *Nature Reviews Cancer* 7 (4): 233–45. https://doi.org/10.1038/nrc2091.

Neal, Jessica A., and Katheryn Meek. 2011. “Choosing the Right Path: Does DNA-PK Help Make the Decision?” *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* 711 (1–2): 73–86. https://doi.org/10.1016/j.mrfmmm.2011.02.010.

Neumann, Frank R., Vincent Dion, Lutz R. Gehlen, Monika Tsai-Pflugfelder, Roger Schmid, Angela Taddei, and Susan M. Gasser. 2012. “Targeted INO80 Enhances Subnuclear Chromatin Movement and Ectopic Homologous Recombination.” *Genes and Development* 26 (4): 369–83. https://doi.org/10.1101/gad.176156.111.

Neves, H, C Ramos, M G da Silva, A Parreira, L Parreira, C. Chomienne, Kathy Howe, et al. 1999. “The Nuclear Topography of ABL, BCR, PML, and RARalpha Genes: Evidence for Gene Proximity in Specific Phases of the Cell Cycle and Stages of Hematopoietic Differentiation.” *Blood* 93 (4): 1197–1207. http://www.ncbi.nlm.nih.gov/pubmed/9949162.

Nickoloff, J A C. 2016. “Preventing the Chromosomal Translocations That Cause Cancer” 127: 176–95.

Nikiforova, M. N., J. R. Stringer, R. Blough, M. Medvedovic, J. A. Fagin, and Y. E. Nikiforov. 2000. “Proximity of Chromosomal Loci That Participate in Radiation-Induced Rearrangements in Human Cells.” *Science* 290 (5489): 138–41. https://doi.org/10.1126/science.290.5489.138.

Ohle, Corina, Rafael Tesorero, Géza Schermann, Nikolay Dobrev, Irmgard Sinning, and Tamás Fischer. 2016. “Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair.” *Cell* 167 (4): 1001-1013.e7. https://doi.org/10.1016/j.cell.2016.10.001.

Olbrich, Teresa, Cristina Mayor-Ruiz, Maria Vega-Sendino, Carmen Gomez, Sagrario Ortega, Sergio Ruiz, and Oscar Fernandez-Capetillo. 2017. “A P53-Dependent Response Limits the Viability of Mammalian Haploid Cells.” *Proceedings of the National Academy of Sciences* 114 (35): 9367–72. https://doi.org/10.1073/pnas.1705133114.

Osborne, Cameron S. 2014. “Molecular Pathways : Transcription Factories and Chromosomal Translocations” 20 (2): 296–301. https://doi.org/10.1158/1078-0432.CCR-12-3667.

Petrini, John H.J., and Travis H. Stracker. 2003. “The Cellular Response to DNA Double-Strand Breaks: Defining the Sensors and Mediators.” *Trends in Cell Biology* 13 (9): 458–62. https://doi.org/10.1016/S0962-8924(03)00170-3.

Piazza, Aurèle, William Douglass Wright, and Wolf Dietrich Heyer. 2017. “Multi-Invasions Are Recombination Byproducts That Induce Chromosomal Rearrangements.” *Cell* 170 (4): 760-773.e15. https://doi.org/10.1016/j.cell.2017.06.052.

Pierce, Andrew J., Roger D. Johnson, Larry H. Thompson, and Maria Jasin. 1999. “XRCC3 Promotes Homology-Directed Repair of DNA Damage in Mammalian Cells.” *Genes and Development* 13 (20): 2633–38. https://doi.org/10.1101/gad.13.20.2633.

Piganeau, Marion, Hind Ghezraoui, Anne De Cian, Lionel Guittat, Mark Tomishima, Loic Perrouault, Oliver René, et al. 2013. “Cancer Translocations in Human Cells Induced by Zinc Finger and TALE Nucleases.” *Genome Research* 23 (7): 1182–93. https://doi.org/10.1101/gr.147314.112.

Qiu, Zhijun, Zhenhua Zhang, Anna Roschke, Tamas Varga, and Peter D. Aplan. 2017. "Generation of Gross Chromosomal Rearrangements by a Single Engineered DNA Double Strand Break." *Scientific Reports* 7 (February): 15–17. https://doi.org/10.1038/srep43156.

Renouf, Benjamin, Marion Piganeau, Hind Ghezraoui, Maria Jasin, and Erika Brunet. 2014. "Creating Cancer Translocations in Human Cells Using Cas9 DSBs and NCas9 Paired Nicks." In *Methods Enzymology*, 251–71. https://doi.org/10.1016/B978-0-12-801185-0.00012-X.

Rieder, Dietmar, Zlatko Trajanoski, and James G. McNally. 2012. "Transcription Factories." *Frontiers in Genetics* 3 (October): 469–89. https://doi.org/10.3389/fgene.2012.00221.

Rogakou, Emmy P., Duane R. Pilch, Ann H. Orr, Vessela S. Ivanova, and William M. Bonner. 1998. "DNA Double-Stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139." *Journal of Biological Chemistry* 273 (10): 5858–68. https://doi.org/10.1074/jbc.273.10.5858.

Roix, Jeffrey J., Philip G. McQueen, Peter J. Munson, Luis A. Parada, and Tom Misteli. 2003. "Spatial Proximity of Translocation-Prone Gene Loci in Human Lymphomas." *Nature Genetics* 34 (3): 287–91. https://doi.org/10.1038/ng1177.

Rong, Yao, Shota Nakamura, Tetsuya Hirata, Daisuke Motooka, and Yi-shi Liu. 2015. "Genome-Wide Screening of Genes Required for Glycosylphosphatidylinositol Biosynthesis," 1–18. https://doi.org/10.1371/journal.pone.0138553.

Roukos, Vassilis, Rebecca C Burgess, and Tom Misteli. 2014. "Generation of Cell-Based Systems to Visualize Chromosome Damage and Translocations in Living Cells." *Nature Protocols* 9 (10): 2476–92. https://doi.org/10.1038/nprot.2014.167.

Roukos, Vassilis, and Stephan Mathas. 2015. "The Origins of ALK Translocations." *Frontiers in Bioscience (Scholar Edition)* 7 (June): 260–68. http://www.ncbi.nlm.nih.gov/pubmed/25961701.

Roukos, Vassilis, and Tom Misteli. 2014. "The Biogenesis of Chromosome Translocations." *Nature Cell Biology* 16 (4): 293–300.

Roukos, Vassilis, Ty C Voss, Christine K Schmidt, Seungtaek Lee, Darawalee Wangsa, and Tom Misteli. 2013. "Spatial Dynamics of Chromosome Translocations in Living Cells." *Science* 341.

Rulten, Stuart L., Anna E.O. Fisher, Isabelle Robert, Maria C. Zuma, Michele Rouleau, Limei Ju, Guy Poirier, Bernardo Reina-San-Martin, and Keith W. Caldecott. 2011. "PARP-3 and APLF Function Together to Accelerate Nonhomologous End-Joining." *Molecular Cell* 41 (1): 33–45. https://doi.org/10.1016/j.molcel.2010.12.006.

Salto-Tellez, Manuel, Suresh G. Shelat, Bernice Benoit, Hanna Rennert, Martin Carroll, Debra G.B. Leonard, Peter Nowell, and Adam Bagg. 2003. "Multiplex RT-PCR for the Detection of Leukemia-Associated Translocations: Validation and Application to Routine Molecular Diagnostic Practice." *Journal of Molecular Diagnostics* 5 (4): 231–36. https://doi.org/10.1016/S1525-1578(10)60479-5.

San Filippo, Joseph, Patrick Sung, and Hannah Klein. 2008. "Mechanism of Eukaryotic Homologous Recombination." *Annual Review of Biochemistry* 77 (1): 229–57. https://doi.org/10.1146/annurev.biochem.77.061306.125255.

Sander, Jeffry D, and J Keith Joung. 2014. “CRISPR-Cas Systems for Editing, Regulating and Targeting Genomes.” *Nature Biotechnology* 32 (4): 347–55. https://doi.org/10.1038/nbt.2842.

Sartori, Alessandro A, Claudia Lukas, Julia Coates, Martin Mistrik, Shuang Fu, Jiri Bartek, Richard Baer, Jiri Lukas, and Stephen P Jackson. 2007. “Human CtIP Promotes DNA End Resection.” *Nature* 450 (7169): 509–14. https://doi.org/10.1038/nature06337.

Sharma, Sumana. 2018. “Application of CRISPR-Cas9 Based Genome-Wide Screening Approaches to Study Cellular Signalling Mechanisms.” https://doi.org/10.3390/ijms19040933.

Shibata, Atsushi, Monika Steinlage, Olivia Barton, Szilvia Juhasz, Julia Kuenzel, Julian Spies, Atsushi Shibata, Penny A. Jeggo, and Markus Lobrich. 2017. “DNA Double-Strand Break Resection Occurs during Non-Homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination.” *Molecular Cell* 65 (4): 671-684.e5. https://doi.org/10.1016/j.molcel.2016.12.016.

Shuga, Joe, Yong Zeng, Richard Novak, Qing Lan, Xiaojiang Tang, Nathaniel Rothman, Roel Vermeulen, et al. 2013. “Single Molecule Quantitation and Sequencing of Rare Translocations Using Microfluidic Nested Digital PCR.” *Nucleic Acids Research* 41 (16): 1–11. https://doi.org/10.1093/nar/gkt613.

Simsek, Deniz, Erika Brunet, Sunnie Yan-Wai Wong, Sachin Katyal, Yankun Gao, Peter J McKinnon, Jacqueline Lou, et al. 2011. “DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation.” *PLoS Genetics* 7 (6): e1002080. https://doi.org/10.1371/journal.pgen.1002080.

Simsek, Deniz, Amy Furda, Yankun Gao, Jérôme Artus, Erika Brunet, Anna-Katerina Hadjantonakis, Bennett Van Houten, Stewart Shuman, Peter J. McKinnon, and Maria Jasin. 2011. “Crucial Roles for DNA Ligase III in Mitochondria but Not in Xrcc1-Dependent Repair.” *Nature* 471 (7337): 245–48. https://doi.org/10.1038/nature09794.

Simsek, Deniz, and Maria Jasin. 2010. “Alternative End-Joining Is Suppressed by the Canonical NHEJ Component Xrcc4-Ligase IV during Chromosomal Translocation Formation.” *Nature Structural & Molecular Biology* 17 (4): 410–16. https://doi.org/10.1038/nsmb.1773.

Skorski, Tomasz. 2002. “BCR/ABL Regulates Response to DNA Damage: The Role in Resistance to Genotoxic Treatment and in Genomic Instability.” *Oncogene* 21 (56 REV. ISS. 7): 8591–8604. https://doi.org/10.1038/sj.onc.1206087.

Slupianek, Artur, Christoph Schmutte, Gregory Tombline, Malgorzata Nieborowska-Skorska, Grazyna Hoser, Michal O. Nowicki, Andrew J. Pierce, Richard Fishel, and Tomasz Skorski. 2001. “BCR/ABL Regulates Mammalian RecA Homologs, Resulting in Drug Resistance.” *Molecular Cell* 8 (4): 795–806. https://doi.org/10.1016/S1097-2765(01)00357-4.

Smith, Joseph R., Sean Maguire, Lesley A. Davis, Morgan Alexander, Fentang Yang, Siddharthan Chandran, Charles Ffrench-Constant, and Roger A. Pedersen. 2008. “Robust, Persistent Transgene Expression in Human Embryonic Stem Cells Is Achieved with AAVS1-Targeted Integration.” *Stem Cells* 26 (2): 496–504. https://doi.org/10.1634/stemcells.2007-0039.

Smith, Laura, Manasa Chandra, Taihra Ul-Hasan, Gregory Cost, Michael Holmes, Philip Gregory, K.K. Wong, and Saswati Chatterjee. 2016. “AAVHSC Vectors Encoding Zinc-Finger Nucleases Mediate Efficient Targeted Integration at the Human AAVS1 Locus in CD34+ Human Hematopoietic Stem Cells.” *Molecular Therapy* 23 (May): S24. https://doi.org/10.1016/s1525-0016(16)33660-7.

Soni, Aashish, Maria Siemann, Martha Grabos, Tamara Murmann, Gabriel E. Pantelias, and George Iliakis. 2014. “Requirement for Parp-1 and DNA Ligases 1 or 3 but Not of Xrcc1 in Chromosomal Translocation Formation by Backup End Joining.” *Nucleic Acids Research* 42 (10): 6380–92. https://doi.org/10.1093/nar/gku298.

Soutoglou, Evi, Jonas F. Dorn, Kundan Sengupta, Maria Jasin, Andre Nussenzweig, Thomas Ried, Gaudenz Danuser, and Tom Misteli. 2007. “Positional Stability of Single Double-Strand Breaks in Mammalian Cells.” *Nature Cell Biology* 9 (6): 675–82. https://doi.org/10.1038/ncb1591.

Strick, R., P. L. Strissel, S. Borgers, S. L. Smith, and J. D. Rowley. 2000. “Dietary Bioflavonoids Induce Cleavage in the MLL Gene and May Contribute to Infant Leukemia.” *Proceedings of the National Academy of Sciences* 97 (9): 4790–95. https://doi.org/10.1073/pnas.070061297.

Sun, Y., X. Jiang, S. Chen, N. Fernandes, and B. D. Price. 2005. “A Role for the Tip60 Histone Acetyltransferase in the Acetylation and Activation of ATM.” *Proceedings of the National Academy of Sciences* 102 (37): 13182–87. https://doi.org/10.1073/pnas.0504211102.

Tadi, Satish K., Carine Tellier-Lebègue, Clément Nemoz, Pascal Drevet, Stéphane Audebert, Sunetra Roy, Katheryn Meek, Jean Baptiste Charbonnier, and Mauro Modesti. 2016. “PAXX Is an Accessory C-NHEJ Factor That Associates with Ku70 and Has Overlapping Functions with XLF.” *Cell Reports* 17 (2): 541–55. https://doi.org/10.1016/j.celrep.2016.09.026.

Tiwari, Ankana, Owen Addis Jones, and Kok Lung Chan. 2018. “53BP1 Can Limit Sister-Chromatid Rupture and Rearrangements Driven by a Distinct Ultrafine DNA Bridging-Breakage Process.” *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-03098-y.

Tomlins, Scott A., Bharathi Laxman, Sooryanarayana Varambally, Xuhong Cao, Jindan Yu, Beth E. Helgeson, Qi Cao, et al. 2008. “Role of the TMPRSS2-ERG Gene Fusion in Prostate Cancer.” *Neoplasia* 10 (2): 177-IN9. https://doi.org/10.1593/neo.07822.

Torres, R., M. C. Martin, A. Garcia, Juan C. Cigudosa, J. C. Ramirez, and S. Rodriguez-Perales. 2014. “Engineering Human Tumour-Associated Chromosomal Translocations with the RNA-Guided CRISPR-Cas9 System.” *Nature Communications* 5: 1–8. https://doi.org/10.1038/ncomms4964.

Tran, Anh Nhi, Fulya Taylan, Vasilios Zachariadis, Ingegerd Ivanov Öfverholm, Anna Lindstrand, Francesco Vezzi, Britta Lötstedt, et al. 2018. “High-Resolution Detection of Chromosomal Rearrangements in Leukemias through Mate Pair Whole Genome Sequencing.” *PLoS ONE* 13 (3): 1–10. https://doi.org/10.1371/journal.pone.0193928.

Tsai, Albert G., Haihui Lu, Sathees C. Raghavan, Markus Muschen, Chih Lin Hsieh, and Michael R. Lieber. 2008. “Human Chromosomal Translocations at CpG Sites and a Theoretical Basis for Their Lineage and Stage Specificity.” *Cell* 135 (6): 1130–42. https://doi.org/10.1016/j.cell.2008.10.035.

Uematsu, Naoya, Eric Weterings, Ken Ichi Yano, Keiko Morotomi-Yano, Burkhard Jakob, Gisela Taucher-Scholz, Pierre Olivier Mari, Dik C. Van Gent, Benjamin P.C. Chen, and David J. Chen. 2007. “Autophosphorylation of DNA-PKCS Regulates Its Dynamics at DNA Double-Strand Breaks.” *Journal of Cell Biology* 177 (2): 219–29. https://doi.org/10.1083/jcb.200608077.

Vanoli, Fabio, and Maria Jasin. 2017. “Generation of Chromosomal Translocations That Lead to Conditional Fusion Protein Expression Using CRISPR-Cas9 and Homology-Directed Repair,” 138–45. https://doi.org/10.1016/j.ymeth.2017.05.006.

Varga, Tamas, and Peter D. Aplan. 2005. “Chromosomal Aberrations Induced by Double Strand DNA Breaks.” *DNA Repair* 4 (9): 1038–46. https://doi.org/10.1016/j.dnarep.2005.05.004.

Vilenchik, M. M., and A. G. Knudson. 2003. “Endogenous DNA Double-Strand Breaks: Production, Fidelity of Repair, and Induction of Cancer.” *Proceedings of the National Academy of Sciences* 100 (22): 12871–76. https://doi.org/10.1073/pnas.2135498100.

Vitelli, Valerio, Alessandro Galbiati, Fabio Iannelli, Fabio Pessina, Sheetal Sharma, and Fabrizio d’Adda di Fagagna. 2017. “Recent Advancements in DNA Damage–Transcription Crosstalk and High-Resolution Mapping of DNA Breaks.” *Annual Review of Genomics and Human Genetics* 18 (1): 87–113. https://doi.org/10.1146/annurev-genom-091416-035314.

Vítor, Alexandra C., Sreerama C. Sridhara, João C. Sabino, Ana I. Afonso, Ana R. Grosso, Robert M. Martin, and Sérgio F. De Almeida. 2019. “Single-Molecule Imaging of Transcription at Damaged Chromatin.” *Science Advances* 5 (1). https://doi.org/10.1126/sciadv.aau1249.

Walter, Joachim, Lothar Schermelleh, Marion Cremer, Satoshi Tashiro, and Thomas Cremer. 2003. “Chromosome Order in HeLa Cells Changes during Mitosis and Early G1, but Is Stably Maintained during Subsequent Interphase Stages.” *Journal of Cell Biology* 160 (5): 685–97. https://doi.org/10.1083/jcb.200211103.

Wang, Guliang, Steve Carbajal, Jan Vijg, John DiGiovanni, and Karen M. Vasquez. 2008. “DNA Structure-Induced Genomic Instability in Vivo.” *Journal of the National Cancer Institute* 100 (24): 1815–17. https://doi.org/10.1093/jnci/djn385.

Wang, Hailong, and Xingzhi Xu. 2017. “Microhomology-Mediated End Joining: New Players Join the Team.” *Cell and Bioscience* 7 (1): 4–9. https://doi.org/10.1186/s13578-017-0136-8.

Wang, Minli, Weizhong Wu, Wenqi Wu, Bustanur Rosidi, Lihua Zhang, Huichen Wang, and George Iliakis. 2006. “PARP-1 and Ku Compete for Repair of DNA Double Strand Breaks by Distinct NHEJ Pathways.” *Nucleic Acids Research* 34 (21): 6170–82. https://doi.org/10.1093/nar/gkl840.

Wang, P., R.-H. Zhou, Y. Zou, C. K. Jackson-Cook, and L. F. Povirk. 1997. “Highly Conservative Reciprocal Translocations Formed by Apparent Joining of Exchanged DNA Double-Strand Break Ends.” *Proceedings of the National Academy of Sciences* 94 (22): 12018–23. https://doi.org/10.1073/pnas.94.22.12018.

Wang, T., K. Birsoy, N. W. Hughes, K. M. Krupczak, Y. Post, J. J. Wei, E. S. Lander, and D. M. Sabatini. 2015. “Identification and Characterization of Essential Genes in the Human Genome.” *Science* 350 (6264): 1096–1101. https://doi.org/10.1126/science.aac7041.

Wang, Tim, Jenny J. Wei, David M. Sabatini, and Eric S. Lander. 2014. “Genetic Screens in Human Cells Using the CRISPR-Cas9 System.” *Science* 343 (6166): 80–84. https://doi.org/10.1126/science.1246981.

Warrell, Raymond P., Stanley R. Frankel, Wilson H. Miller, David A. Scheinberg, Loretta M. Itri, Walter N. Hittelman, Rohini Vyas, et al. 1991. “Differentiation Therapy of Acute Promyelocytic Leukemia with Tretinoin (All-Trans-Retinoic Acid).” *New England Journal of Medicine* 324 (20): 1385–93. https://doi.org/10.1056/NEJM199105163242002.

Weinstock, David M, Erika Brunet, and Maria Jasin. 2007. “Formation of NHEJ-Derived Reciprocal Chromosomal Translocations Does Not Require Ku70.” *Nature Cell Biology* 9 (8): 978–81. https://doi.org/10.1038/ncb1624.

White, Ryan R, and Jan Vijg. 2017. “Do DNA Double-Strand Breaks Drive Aging?” 63 (5): 729–38. https://doi.org/10.1016/j.molcel.2016.08.004.

Wiedenheft, Blake, Samuel H Sternberg, and Jennifer A Doudna. 2012. “RNA-Guided Genetic Silencing Systems in Bacteria and Archaea” 46: 44–46. https://doi.org/10.1038/nature10886.

Williams, M, I R Rainville, and J A Nicklas. 2002. “Use of Inverse PCR to Amplify and Sequence Breakpoints of HPRT Deletion and Translocation Mutations.” *Environmental and Molecular Mutagenesis* 39 (1): 22–32. https://doi.org/10.1002/em.10040.

Wolanin, K., A. Magalska, M. Kusio-Kobialka, P. Podszywalow-Bartnicka, S. Vejda, S. L. McKenna, G. Mosieniak, E. Sikora, and K. Piwocka. 2010. “Expression of Oncogenic Kinase Bcr-Abl Impairs Mitotic Checkpoint and Promotes Aberrant Divisions and Resistance to Microtubule-Targeting Agents.” *Molecular Cancer Therapeutics* 9 (5): 1328–38. https://doi.org/10.1158/1535-7163.mct-09-0936.

Wood, Richard D, and Sylvie Doublié. 2016. “DNA Polymerase θ (POLQ), Double-Strand Break Repair, and Cancer.” *DNA Repair* 44 (512): 22–32. https://doi.org/10.1016/j.dnarep.2016.05.003.

Wray, Justin, Elizabeth A. Williamson, Sudha B. Singh, Yuehan Wu, Christopher R. Cogle, David M. Weinstock, Yu Zhang, et al. 2013. “PARP1 Is Required for Chromosomal Translocations.” *Blood* 121 (21): 4359–65. https://doi.org/10.1182/blood-2012-10-460527.

Wu, Xiaotian, Shiqi Mao, Yachen Ying, Christopher J. Krueger, and Antony K. Chen. 2019. “Progress and Challenges for Live-Cell Imaging of Genomic Loci Using CRISPR-Based Platforms.” *Genomics, Proteomics and Bioinformatics*, no. xxxx. https://doi.org/10.1016/j.gpb.2018.10.001.

Wyatt, David W., Wanjuan Feng, Michael P. Conlin, Matthew J. Yousefzadeh, Steven A. Roberts, Piotr Mieczkowski, Richard D. Wood, Gaorav P. Gupta, and Dale A. Ramsden. 2016. “Essential Roles for Polymerase θ-Mediated End Joining in the Repair of Chromosome Breaks.” *Molecular Cell* 63 (4): 662–73. https://doi.org/10.1016/j.molcel.2016.06.020.

Yan, Catherine T., Cristian Boboila, Ellen Kris Souza, Sonia Franco, Thomas R. Hickernell, Michael Murphy, Sunil Gumaste, et al. 2007. “IgH Class Switching and Translocations Use a Robust Non-Classical End-Joining Pathway.” *Nature* 449 (7161): 478–82. https://doi.org/10.1038/nature06020.

Yang, Fan, Christopher J Kemp, and Steven Henikoff. 2015. “Anthracyclines Induce Double-Strand DNA Breaks at Active Gene Promoters.” *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 773 (March): 9–15. https://doi.org/10.1016/j.mrfmmm.2015.01.007.

Yousefzadeh, Matthew J., David W. Wyatt, Kei ichi Takata, Yunxiang Mu, Sean C. Hensley, Junya Tomida, Göran O. Bylund, et al. 2014. “Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ.” *PLoS Genetics* 10 (10). https://doi.org/10.1371/journal.pgen.1004654.

Zha, S., F. W. Alt, H.-L. Cheng, J. W. Brush, and G. Li. 2007. “Defective DNA Repair and Increased Genomic Instability in Cernunnos-XLF-Deficient Murine ES Cells.” *Proceedings of the National Academy of Sciences* 104 (11): 4518–23. https://doi.org/10.1073/pnas.0611734104.

Zhang, Yu, and Maria Jasin. 2011. “An Essential Role for CtIP in Chromosomal Translocation Formation through an Alternative End-Joining Pathway.” *Nature Structural and Molecular Biology* 18 (1): 80–85. https://doi.org/10.1038/nsmb.1940.

Zhang, Yu, Rachel Patton McCord, Yu-Jui Ho, Bryan R Lajoie, Dominic G Hildebrand, Aline C Simon, Michael S Becker, Frederick W Alt, and Job Dekker. 2012. “Chromosomal Translocations Are Guided by the Spatial Organization of the Genome.” *Cell* 148 (5): 908–21. https://doi.org/10.1016/j.cell.2012.02.002.Chromosomal.

Zou, Jizhong, Colin L. Sweeney, Bin Kuan Chou, Uimook Choi, Jason Pan, Hongmei Wang, Sarah N. Dowey, Linzhao Cheng, and Harry L. Malech. 2011. “Oxidase-Deficient Neutrophils from X-Linked Chronic Granulomatous Disease IPS Cells: Functional Correction by Zinc Finger Nuclease-Mediated Safe Harbor Targeting.” *Blood* 117 (21): 5561–72. https://doi.org/10.1182/blood-2010-12-328161.

List of abbreviations

3D	three dimensional
AAVS1	adeno-associated virus integration site 1
a-EJ	alternative end-joining
ALCL	anaplastic large cell lymphoma
ALL	acute lymphocytic leukemia
AML	acute myeloid leukemia
APL	promyelocytic leukemia
BAC	bacterial artificial chromosomes
Blast	blasticidin
bp	base pair
BSA	bovine serum albumin
C	control
cDNA	complementary DNA
CFS	common fragile sites
Chr.	chromosome
cl.	clone
CML	chronic myeloid leukemia
cNHEJ	classical non-homologous end-joining
CNV	copy number variation
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CSR	class switch recombination
Ctrl	control
dCas9	deactivated Cas9
ddPCR	digital droplet PCR
DDRNAs	damage response RNA
Dex	dexamethasone
DHT	dihydrotestosteron
DISC1	Disrupted in Schizophrenia 1
DNA-PKcs	DNA-PK catalytic subunit
dNTPs	deoxynucleotides
Dox	doxycycline
DSBs	double-strand breaks

EMEA	European Medicines Agency
ETO	etoposide
EtOH	ethanol
FACS	fluorescence-activated cell sorter
FBS	fetal bovine serum
FDA	Food and Drug Administration
F	forward
FISH	fluorescence in situ hybridization
GFP	green fluorescent protein
H1C	HAP1 with a cassette
H1Cx	HAP1 with an excised aprt of cassette
HA-DD-I-SceI-GR	HA-degron-I-SceI-glucocorticoid receptor
HR	homologous recombination
Hygro	Hygromycin
IF	immunofluorescence
Ind.	inducible
IR	ionizing radiation
IRIFs	ionizing radiation-induced foci
LAM-HTGTS	linear amplification-mediated high-throughput genome-wide sequencing
Lig	ligase
LINEs	long interspersed nuclear elements
LP	laser power
MeOH	methanol
MIR	multi-invasion-induced rearrangement
MOI	multiplicity of infection
mtDNA	mitochondrial DNA
MULE	Multiple Lentiviral Expression
n.t.	not treated
NAHR	non-allelic homologous recombination

NER	nucleotide excision repair
NHEJ	non-homologous end-joining
P2A	porcine teschovirus-1 2A
pA	polyadenylation sequence
PAM	protospacer adjacent motif
PCA	prostate cancer
PIKK	phosphatidylinositol 3-kinase-related kinase
PLL	poly-L-lysine
PM	promoter modification
qPCR	quantitative PCR
RNA Pol II	RNA polymerase II
ROS	reactive oxygen species
R	reverse
RT	room temperature

RT-PCR	Reverse Transcriptase PCR
Seq	sequencing
sgRNA	single-guided RNA
shRNA	small hairpin RNA
SSA	single-stranded annealing
ssDNA	single-stranded DNA
TALEN	transcription activator-like effector nuclease
Temp.	temperature
Top2	Topoisomerase II
tracrRNA	trans-activating crRNA
UV	ultraviolet
WT	wild type
ZFN	Zinc-finger nucleases

List of figures

List of tables