universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of Master's Thesis

„Establishment of an enhanced CRISPR repressor platform in haploid mouse embryonic stem cells."

verfasst von / submitted by

Kemal Can Arat

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet	UA 066 830
Studienrichtung lt. Studienblatt / Degree programme as it appears on The student record sheet	Masterstudium Molekulare Mikrobiologie Mikrobielle Ökologie und Immunbiologie
Betreut von / Supervisor	Mag. Dr. Martin Leeb

Acknowledgments

First and foremost I would like to express my sincere gratitude to my supervisor Mag. Dr. Martin Leeb for having given me the opportunity to perform my master's thesis work in his laboratory. The fact that he was always there to help and support, made working in his laboratory a valuable and fun experience. His dedication in science, his hard work and patience are definitely one of the many things I have learnt from him.

In my daily work I have been blessed with a friendly group of Leeboratory colleagues. Therefore I would also like to express my gratitude to Julia Ramesmayer, Elena Galimberti, Michelle Huth, Laura Santini, Kitti Dora Csalyi and Luis Miguel Cerron Alvan for their support and friendship during the times spent in- and outside of the lab. Furthermore, a special note of thanks goes to the members of the Bücker laboratory, Henry Thomas and Merrit Romeike for providing me their help on my experiments.

Last but not least, I would like to thank my family and friends. A very sincere love and gratitude goes out to my parents Başak and İbrahim Orhan Arat for their never ending support all throughout my life leading up to this special occasion. Sizi çok seviyorum. A warm thank you as well to my eternal soulmate İbrahim Yılmaz and little Dr. Joshua Mardell for being there, all the time for better or worse, with their biggest love.

To my grandfather,

Habil Akıltepe

Table of Contents

Abstract

Even though most of the available genetic screens are successful, there are limitations in their approaches. Especially, most available CRISPR screening libraries focus on only protein coding genes, which make up only approximately 2 to 5% of mammalian genomes. However, these screens cannot interrogate the function of gene clusters, enhancer clusters, or most importantly, as yet unknown genetic regulatory elements. Therefore, the rest of the genomes remain undetectable.

In this project, we present an approach, which solves this particular problem of all genome-wide screening technologies and therefore, leads us to investigate genome function in a wider and reliable manner. The ultimate task of this project is to generate haploid mouse embryonic stem cell (mESC) prototypes for an innovative genome-wide screening system by combining insertional mutagenesis with an advanced CRISPR/Cas9 system based recruitment of the epigenetic silencing machinery. The second most important task is then to validate the prototype's functionality; we have to determine the range of the insertional mutagenesis. For this, we will be looking at the levels of H3K9me3, the most expected histone modification caused by the system. By analysing the silencing range, we will be able to evaluate the power of the silencing machinery within the prototype cell lines. The prototype requires three genetic components expressed simultaneously in one haploid mESC line. These components are dCas9-KRAB-MeCP2 transgene, multiple specific guide RNAs (gRNA) and the heterochromatinization-landing platform (HLP).

First, in order to observe if the silencing machinery would be established successfully, we produced a simple diploid mESC prototype. The preliminary monitoring on this prototype stated that, except for the gRNA efficiency issue, the working system could be generated in the diploid mESCs.

Next, overcoming the gRNA efficiency issue by generating a single vector for multiple gRNAs, we established the silencing machinery in haploid mESCs by generating 3 different prototypes. As the preliminary monitoring, via Western blot

and flow cytometry, we verified the system's functionality. Later broad monitoring was done on the system's silencing power. We first located the integration sites of the HLPs within the prototype's genomes. Then, we analysed the silencing range of the HLPs and observed their effects on the genomic regions nearby by looking at the H3K9me3 epigenetic modification. Moreover, we analysed the expression patterns of the genes on and near the HLP. After a series of ChIP and qPCR experiments, we were able to understand the system's capacity of silencing and so the ability of the prototypes. In the end, we concluded that one working haploid prototype was successfully produced and the silencing range of this prototype was very promising in the sense of leading the way for advanced haploid prototypes.

Later, in order to establish more controllable and secure silencing machinery, we upgraded the way of integrating dCas9-KRAB-MeCP2 transgene by generating a vector with a chemically inducible feature and a different integration method. Succeeding this, the next goal was to establish the advanced haploid mESC prototype. Unfortunately, due to some experimental issues on cells ploidy, our attempt was failed. However, for future attempts, we were able to develop and produce the genetic components that are needed for the silencing machinery.

In conclusion, even though we still have some way to go until we produce the advanced haploid mESC prototype, throughout the project, we were able to bring an understating of how to establish the silencing machinery in the test prototypes and show the power of the silencing mechanism within.

Zusammenfassung

Obwohl die meisten verfügbaren genetischen Screenings erfolgreich sind, gibt es Einschränkungen in ihren Ansätzen. Insbesondere konzentrieren sich die meisten verfügbaren CRISPR-Screening-Bibliotheken nur auf Protein-kodierende Gene, die nur etwa 2 bis 5% der Säugetiergenome ausmachen. Diese Screenings können jedoch nicht die Funktion von Genclustern, Enhancer-Clustern oder vor allem unbekannten genetischen regulatorischen Elementen abfragen. Daher bleibt der Rest des Genoms nicht nachweisbar.

In diesem Projekt stellen wir einen Ansatz vor, der dieses spezielle Problem aller genomweiten Screening-Technologien löst und uns daher dazu veranlasst, die Genomfunktion umfassender und zuverlässiger zu untersuchen. Die ultimative Aufgabe dieses Projekts besteht darin, Prototypen haploider embryonaler Mausstammzellen (mESC) für ein innovatives genomweites Screeningsystem zu generieren, indem Insertionsmutagenese mit einer fortschrittlichen CRISPR / Cas9-System-basierten Rekrutierung der epigenetischen Silencing-Maschinerie kombiniert wird. Die zweitwichtigste Aufgabe besteht dann darin, die Funktionalität des Prototyps zu validieren. Wir müssen den Bereich der Insertionsmutagenese bestimmen. Dazu werden wir uns die Levels von H3K9me3 ansehen, die am meisten erwartete Histonmodifikation, die vom System verursacht wird. Durch die Analyse des Silencing-Bereichs können wir die Leistung der Silencing-Maschinerie innerhalb der Prototyp-Zelllinien bewerten. Der Prototyp erfordert drei genetische Komponenten, die gleichzeitig in einer haploiden mESC-Linie exprimiert werden. Diese Komponenten sind dCas9-KRAB-MeCP2-transgene, multiple spezifische Guide-RNAs (gRNA) und die Heterochromatinisierungs-Landeplattform (HLP).

Um zu beobachten, ob der Silencing-Mechanismus erfolgreich eingerichtet werden kann, haben wir zunächst einen einfachen diploiden mESC-Prototyp hergestellt. Die vorläufige Überwachung dieses Prototyps ergab, dass mit Ausnahme des Problems der gRNA-Effizienz das Arbeitssystem in den diploiden mESCs generiert werden könnte.

Als nächstes haben wir das Problem der gRNA-Effizienz durch die Erzeugung eines einzelnen Vektors für mehrere gRNAs gelöst und die Silencing-Maschinerie in haploiden mESCs durch Generierung von 3 verschiedenen Prototypen etabliert. Als vorläufige Überwachung haben wir mittels Western Blot und Durchflusszytometrie die Funktionalität des Systems überprüft. Später wurde die Stummschaltung des Systems umfassend überwacht. Wir haben zuerst die Integrationsstellen der HLPs innerhalb des Genoms des Prototyps lokalisiert. Dann analysierten wir den Silencing-Bereich der HLPs und beobachteten ihre Auswirkungen auf die Genomregionen in der Nähe, indem wir die epigenetische Modifikation von H3K9me3 untersuchten. Darüber hinaus analysierten wir die Expressionsmuster der Gene auf und in der Nähe des HLP. Nach einer Reihe von ChIP- und qPCR-Experimenten konnten wir die Fähigkeit des Systems zur Stummschaltung und damit die Fähigkeit der Prototypen verstehen. Am Ende kamen wir zu dem Schluss, dass ein funktionierender haploider Prototyp erfolgreich hergestellt wurde und der Silencing-Bereich dieses Prototyps im Sinne einer Vorreiterrolle für fortschrittliche haploide Prototypen sehr vielversprechend war.

Später, um kontrollierbarere und sicherere Silencing-Maschinen zu etablieren, haben wir die Art der Integration des dCas9-KRAB-MeCP2-Transgens verbessert, indem wir einen Vektor mit einem chemisch induzierbaren Merkmal und einer anderen Integrationsmethode erzeugt haben. Danach war das nächste Ziel die Etablierung des fortschrittlichen haploiden mESC-Prototyps. Leider ist unser Versuch aufgrund einiger experimenteller Probleme mit der Zellploidie gescheitert. Für zukünftige Versuche könnten wir jedoch die genetischen Komponenten entwickeln und produzieren, die für die Schalldämpfungsmaschinerie benötigt werden.

Obwohl wir bis zur Herstellung des fortschrittlichen haploiden mESC-Prototyps noch einen weiten Weg vor uns haben, könnten wir während des gesamten Projekts ein Verständnis dafür vermitteln, wie der Silencing-Mechanismus in den Testprototypen eingerichtet werden kann, und die Leistungsfähigkeit innerhalb des Silencing-Mechanismus darlegen.

1. Introduction

1.1. Mouse embryonic stem cells (mESC) lines

Embryonic stem cells (ESCs) are the ex vivo equivalent of the epiblast lineage of the blastocyst an early-stage pre-implantation embryo, for this reason, ESCs share the same developmental capacity to differentiate into any of these three primary germ layers: definitive endoderm, ectoderm and mesoderm (Czechanski et al. 2014). The most defining features of the ESCs are the high capacity for in vitro self-renewal combined with their developmental pluripotency. Mouse embryonic stem cells (mESCs) are derived from the blastocyst (from the inner cell mass) of pre-implantation-stage embryos (Evans et al. 1981). The progenitor cells, which give rise to the mESCs, are located in the epiblast layer of the late blastocyst (4 - 4.5 days post coitum). In this state, these progenitor cells express some pluripotency-associated factors e.g. OCT4 and NANOG (Brook and Gardner 1997). Other than their capacity for stable pluripotency and self-renewal in vitro, they are crucial tools for genetic engineering (Bradley et al. 1992). The Nobel Prize–winning research, which genes could be modified in mice by using mESCs was published more than 3 decades ago, and until today, individual researchers around the globe have created almost 50,000 genetically modified alleles for every gene in the mouse genome (Eppig et al. 2012). Furthermore, mESCs are used for essential research on pluripotency as well as for the development of ES cell oriented therapies.

1.2.1. Haploid mESC lines

Haploid mESCs are a desirable model system for inspecting mammalian biology. They offer great possibilities in genetic screens and also strong models to study ploidy effects. Their high reliability for genome modification and relatively stable genome enables the production of large-scale mutant libraries that can later be used

for genetic screens on phenotype studies (Li et al. 2010).

Mammalian cells are challenging when it comes to the forward genetics application since they cannot simply be investigated for phenotypic consequences caused by recessive mutations because of their diploid genome (Leeb et al. 2015). Recently, for cell-based genome screening, CRISPR/Cas9 based methods have been developed and been used successfully (Wang et al., 2014). These methods are quite powerful and they require a great amount of knowledge to study the genes in question but unfortunately, this can lead to some mismanagement in large-scale genetic screens. At the same time, in haploid cells, it is possible to introduce hemizygous mutations for genomic regions with no prior knowledge is available. In the last couple of years, haploid mESC cultures have been also used in this direction and provided promising results (Leeb et al. 2014).

1.2. Genome-wide screenings

Genetic screens are substantial tools to characterise gene functions. Currently, they are used in research specifically to understand the genetic basis of cellular processes and therefore the function of the unknown within the whole genome.

In the last decades, there has been an important development in mammalian genome engineering and this had a substantial reflection on genetic screen studies. Also, the methods and approaches for screens have been advanced. Recently, some of the studies have used the CRISPR/Cas9 system in genome-wide screens to avoid several of the errors associated with siRNA screens (Wang et al 2014; Shalem et al. 2014). Meanwhile, RNA interference has continued to stay as the method of choice for cell culture screens. Furthermore, there have been developments on mammalian transposons and therefore their success in whole-organism screens (Pettitt et al. 2013). As it can be seen that there are plenty of approaches when it comes to genetic screen studies. Finally, it appears that among these approaches, the CRISPR/Cas9

system is more accurate and seems to produce fewer off-target effects when compared to other existing methods (Wang et al., 2014).

Although most of these approaches are very successful, there are limitations to all of them. Especially, with current methods, only protein-coding genes can be targeted, and therefore they are not capable of interrogating the whole regulatory potential of the mammalian genome. Most available CRISPR screening libraries focus on only approximately 2 to 5% of mammalian genomes. The rest of the genomes remain undetectable with common CRISPR or insertional mutagenesis based screens.

1.3. Project description & establishing prototype cells

1.3.1. Project description

As explained above, although, most of the current genetic screens are successful there are limitations in their approaches. Specifically, most available CRISPR screening libraries focus on only protein coding genes (makes up only approximately 2 to 5% of mammalian genomes). However, these screens cannot interrogate the function of gene clusters, enhancer clusters, micro RNA clusters, or most importantly, as yet unknown genetic regulatory elements. Therefore, these current methods (e.g. CRISPR or insertional mutagenesis based screens) are blind to the regulatory potential of large parts of the genome.

In this project, we try to develop an approach, which solves this problem of all genome-wide, screening technologies. The ultimate task of this project is to generate prototype mESC lines for an innovative genome-wide screening system by combining insertional mutagenesis with an advanced CRISPR/Cas9 system based recruitment of the epigenetic silencing machinery. Thereby, we hope to provide new opportunities to investigate genome function in a wider and reliable manner.

After generating the prototype mESC lines, the second most important task of this project is to test the system's ability. In order to validate the prototype's functionality, we have to determine the range of the insertional mutagenesis provided by the CRISPR/Cas9 system based recruitment of the epigenetic silencing machinery. For this, we will be looking at the levels of H3K9me3, the most expected histone modification caused by the system. By analysing the silencing range, we will be able to evaluate the power of the silencing system within the prototype cell lines.

1.3.2. Prototype mouse embryonic stem cells

In the last decade, improvements on CRISPR/Cas9 system as well as the onset of haploid mammalian cell systems have improved our ability to perform comprehensive genome-wide genetic screens. Besides, the genome-editing feature of the CRISPR/Cas9 system is improved by the use of haploid ES cells (Joung et al. 2017). For instance, in haploid ES cells, CRISPR/Cas9 system has been used for the generation of multiple knockout, knockdown and knockin mutations with high efficiency, following the application in genome-wide screens (Horii et al. 2013).

Mutant homozygous mammalian cells are complicated to generate because they carry a diploid genome. Due to the availability of the second allele, a mutant heterozygous cell might not show a disrupted phenotype. Therefore, studying the gene functions in mammalian cells is not easy and can be challenging most of the time. Haploid cells have only a single copy of each chromosome and interruption of one allele can lead to loss-of-function phenotypes directly. In the last decade, improvements have been done to generate mESCs, which can provide an excellent tool for genetic analyses (Leeb & Wutz, 2011). Besides, haploid ES cells continues to hold the most of the biological aspects of diploid ES cells, excluding their karyotype. A genome wide analysis showed that diploid and haploid ES cells have identical expression profiles (Horii et al. 2013).

By taking into consideration all advantages of haploid mouse ES cells explained above, for our prototypes, we are using haploid mESCs.

1.4. Genetic setup of prototype haploid mESC line

The prototype requires three genetic components expressed simultaneously in one haploid mESC line. These components are chemically (e.g. Doxycycline) inducible dCas9-KRAB-MeCP2 transgene, multiple specific guide RNAs (gRNA) and the heterochromatinization landing platform (HLP). Transgene and the gRNAs are identical in each prototype cell whereas, the HLP will be the variable component of the system.

1.4.1. dCas9-KRAB-MeCP fusion protein

Originally, mediating the adaptive immunity system in bacteria and archaea, the clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 system has developed as a strong tool for genome editing in the last decade (Cho et al., 2013). Cas9 is a protein with endonuclease activity and via guide RNAs (gRNA) it can be directed to specific DNA sequences. This can be achieved due to the complementarity between a special Cas9-related gRNA and the target locus. Since it is possible to alter the target locus by only changing the introduced gRNA with Cas9, the use of the system has been rapidly accepted for unbiased genome-wide screens (Shalem et al., 2014). However, cutting activity of Cas9 forms DNA double-strand breaks, which can then cause cellular toxicity. Besides modifications generated by Cas9 are irreversible. These, unfortunately, limit its applications (Mandegar et al., 2016).

As explained before, the prototype will provide a novel system with Cas9 based recruitment of the epigenetic silencing machinery by combining insertional mutagenesis in haploid ES cells. Therefore, in this project, instead of using a regular

Cas9, which causes site-specific DNA cleavage, we wanted to use an advanced Cas9 method where the endonuclease Cas9 was transformed into a programmable transcriptional repressor.

Cas9 protein can be mutated to achieve nuclease-dead Cas9 (dCas9) by the amino acids essential for DNA catalytic activity. This protein remains capable of DNA binding but lacks endonuclease activity (figure 1). When addressed to the transcriptional start site of a gene, dCas9 protein can substantially arrest RNA polymerase passing and this leads to gene silencing (Qi et al., 2016). Further advancement in transcriptional interference can be accomplished with the addition of transcriptional repression domains such as the Krüppel-associated box (KRAB) and MeCP2 to dCas9 in order to produce the dCas9–KRAB-MeCP2 fusion protein (Yeo et al., 2018). The KRAB domain inhibits transcription with the help of KAP1 protein that recruits co-repressors including SETDB1 (known as H3K9 methyltransferase), heterochromatin protein 1 (HP1) and various histone deacetylases. (Friedman et al., 1996). MeCP2, the second transcription-repression domain binds to another set of regulators including the SIN3A–histone deacetylase co-repressor complex and DNA methyltransferase DNMT1 (Jones et al., 1998).

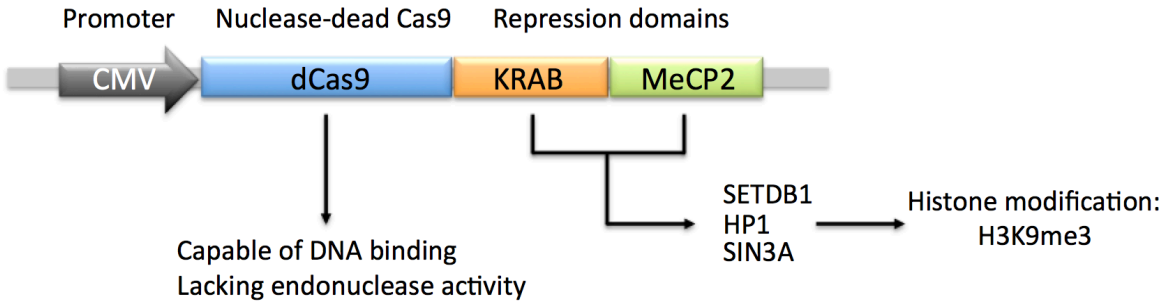


Figure 1: Schematic representation of dCas9-KRAB-MeCP2 transgene
Schematic of dCas9-KRAB-MeCP2 repressor. Nuclease-dead Cas9 remains competent for DNA binding, but unable to have endonuclease activity. Transcriptional repressors, KRAB and MeCP2 lead to H3K9me3 histone modification.

The indirect cooperation between these agents of the transcriptional repressors KRAB and MeCP2 coordinates DNA methylation and histone modifications like

H3K9me3 and H3K27me3 thereby leading to the heterochromatinisation. Among these H3K9me3 is the most expected modification (figure 1). Therefore, by looking at the H3K9me3 levels, we can evaluate the power of the fusion protein. Overall, the features of the dCas9-KRAB-MeCP2 fusion protein make it the present gold standard for dCas9 based repression research (Yeo et al., 2018).

1.4.1.1 Tetracycline-controlled transcriptional activation

In the project, we planned to deliver the dCas9-KRAB-MeCP2 fusion gene within the genome of the prototype, along with the Tetracycline-Controlled transcriptional activation system. In this method of inducible gene expression, transcription is reversibly switched on or off when the antibiotic tetracycline or one of its derivatives (e.g. doxycycline) is present (Gossen et al. 1995). Thus, we aimed to induce the transcriptional activity of the dCas9-KRAB-MeCP2 transgene at any time when Doxycycline was introduced (figure 2).

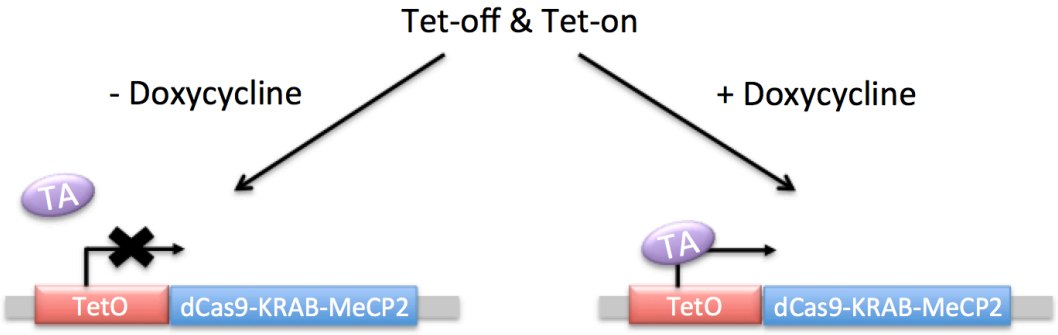


Figure 2: Tetracycline-controlled transcriptional activation
TA = transactivator, TetO = tetracycline operator. In this regulative system, TA is continuously expressed, and the transgene (e.g. dCas9-KRAB-MeCP2) is controlled by TetO. The TA contains a tet repressor (controlled by e.g. Doxycycline) fused to a transcription factor activation domain.

1.4.2. Multiple gRNA expression vector (STAgR)

The second genetic component that is identical for each prototype cell is a group of highly specific guide RNAs (gRNA). These gRNAs will recruit dCas9-KRAB-MeCP2 fusion protein to the specific genomic targets. However, these genomic targets are

distinct and do not exist in the mammalian genomes. Our guides are designed to have manufactured targets that come with the heterochromatinisation landing platform (HLP). Into the prototypes, we will introduce more than one specific genomic target. Therefore, we need to deliver several gRNAs to cover all the targets. The expression of this group of gRNAs will be achieved using available technology (Breunig et al., 2018).

Many of the several Cas9 proteins in CRISPR tools are strictly dependent on the synchronic distribution and expression of multiple gRNAs. This includes our strategy, the use of dead Cas9 with transcriptional repressors (Balboa et al., 2015). However, bringing together several individual gRNA expression vectors in order to achieve the expression of several gRNAs in single cells is complex and has its limits, in vitro, as well as in vivo. The reason for that is with the number of guides used, the ratio of cells expressing an entire set of gRNA decreases, and those that achieve the expression of a complete set of gRNAs, hardly meet stoichiometric levels (Breunig et al., 2018). Besides, other than the fact that we depend on the availability of several gRNAs in single cells, a substantial need for cost-effective and rapid vector generation arises. Thereby, in the project, we used a straightforward, rapid, and cost-effective method (string assembly gRNA cloning – STAgR) to construct working expression vectors for multiple gRNA deliveries with customizable and reliable features.

A detailed protocol on the multiple gRNA vector cloning and the information on the components of the vector can be found under the methods (see section 2.2.) and results (see section 3.2.1.) sections. In figure 10 (page 74) and figure 11 (page 76), overview of the multiple gRNA expression vectors that were produced and schematic representation of the genomic components of these vectors are shown.

1.4.3. Heterochromatinisation landing platform (HLP)

Last but not least, the variable part of the screening system will be the heterochromatinisation landing platform (HLP). The expertise on our campus VBCF in silico designed an adaptive platform for dCas9 fusion protein landing. This landing platform was synthesised to provide multiple (more than 20) landing sites for dCas9 fusion proteins. It also contains a fluorescent reporter for tracking purposes (figure 3). Our expectation from this genetic component is that it will facilitate nucleation of an off state of chromatin. This implies a more obstructed structure, where the transcriptional machinery is not accessible.

The reason why the HLP is the variable part of the screening system is that in a screen, it will be deposited at random locations within the genome. This will increase the number of integration sites of the platform. Thus, we will be able to generate and identify lots of different integration sites thereby, achieving wide-ranging genome coverage. Throughout the project, we used PiggyBac transposable element. Any other, practical and substantial integration method can be used.

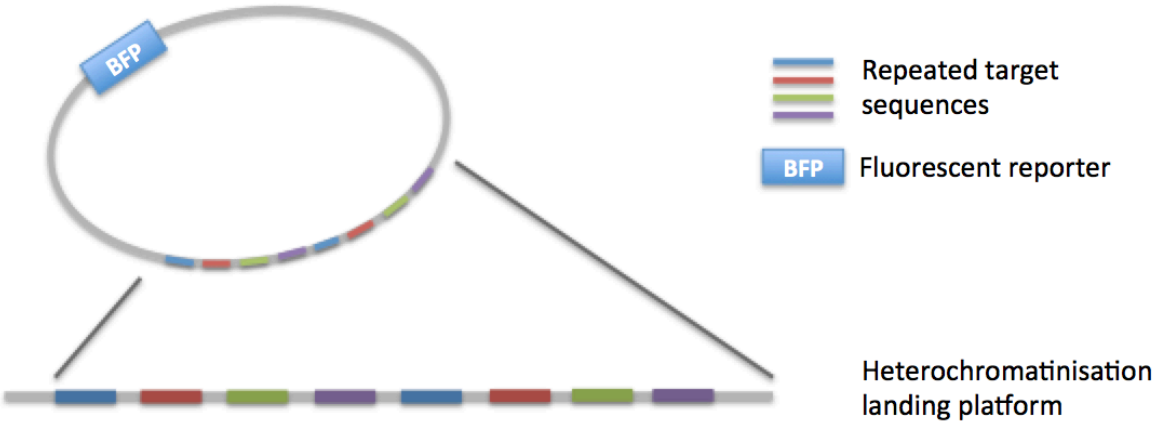


Figure 3: Schematic representation of the heterochromatinisation landing platform vector
Schematic representation of the heterochromatinisation landing platform. The HLP consist of multiple landing sites that are complementary to the dCas9-KRAB-MeCP2 fusion proteins recruited by the guide RNAs already present in the cell. It also contains a blue fluorescent protein gene.

As explained before, available screening methods focus on interrogating 2 to 5% of the mammalian genome, which only corresponds to the protein-coding genes. Gene clusters, micro RNA clusters, or yet unknown regulatory elements are left out. In a screen, the heterochromatinisation landing platform will higher our chances to interrogate better by having many landing sites for dCas9-KRAB-MeCP2 and having many different integration sites within the genome. Therefore, it will hopefully solve the problem of all genome-wide screening technologies.

1.4.4. Overview of the prototype haploid mESC line

Once the identical genetic components (dCas9-KRAB-MeCP2 and multiple gRNA expression vector) set in the haploid mESCs, the screening prototypes will be ready to be used. Later, the variable genetic component, HLP will be randomly integrated within the genomes of the prototypes. Then, upon induction of cells with Doxycycline, the working dCas9-KRAB-Mecp2 fusion protein will be recruited by the corresponding gRNAs onto the landing sites of target sequences of the HLP Thus, shut down of large genomic regions flanking the platform integration sites will occur including the BFP reporter (figure 4A).

Many genes and regulatory elements are found in clusters. Shutting off multiple at the same time is impossible with the current screening methods. With a working prototype haploid mESC line, we expect to provide simultaneous deactivation of gene clusters, enhancer clusters and/or inactivation of regulatory elements. After the system is induced with dCas9-KRAB-MeCP2 as we expect strong heterochromatinisation of the chromatin, we wish to observe the silencing of not only the gene clusters but also these elements in the close range from the HLP integration sites (figure 4B).

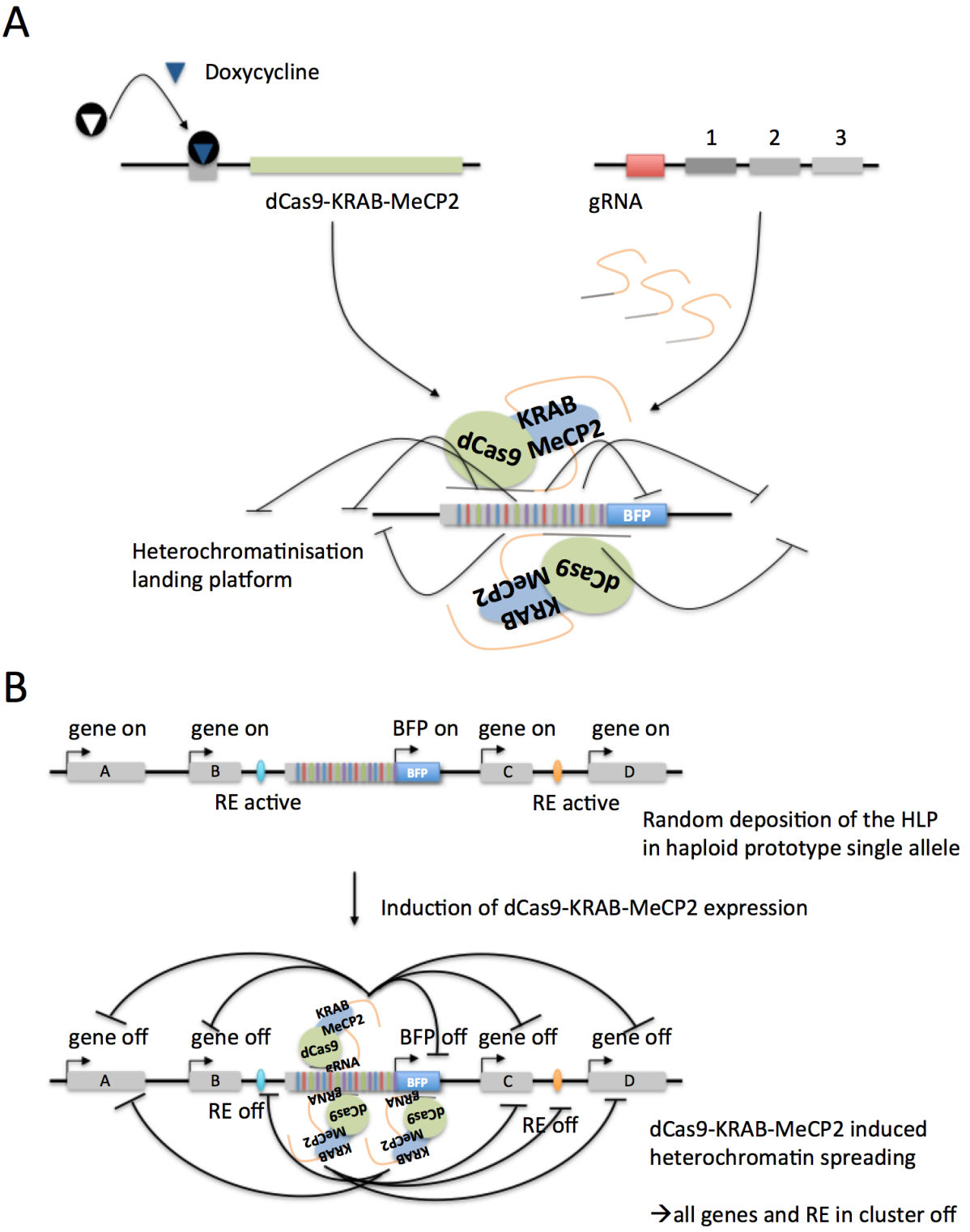


Figure 4: Proposed workflow of the prototype haploid mESC line
(A) Genetic setup of prototype haploid mESC line. It consists of 3 genetic components. Doxycycline inducible dCas9-KRAB-MeCP2 and multiple gRNA expression vector are the identical components of every prototype cell. The variable component is the HLP. (B) Random deposition of the HLP throughout the prototype genome. With the dCas9-KRAB-MeCP2 induced heterochromatin spreading, simultaneous silencing of multiple genetic elements occurs. We expect these elements to be shut down: gene clusters, microRNA clusters, enhancer clusters, and regulatory elements (RE).

1.5. Monitoring the results

The preliminary monitoring will be done to verify the generation of the prototype mESC line. After establishing the prototype with the stable genetic components (identical for each prototype cell line), dCas9-KRAB-MeCP2 fusion protein and multiple gRNA expression vector, we will have to validate the availability of the components within the cells. Therefore, following the confirmation of the fusion protein expression via Western blot, the first monitoring will be performed to analyse the expression of the reporter gene (e.g. red fluorescent protein) that comes with the gRNA vector. For this, we will analyse the cell profile of the prototype via flow cytometry. Later, we will transfect the prototype with the variable genetic component, the HLP. As we stated before, HLP contains a BFP reporter gene. Therefore, for the verification of this step, we will again perform flow cytometry to observe cell profiles with corresponding fluorescent signals. Upon the Doxycycline induction of the prototype mESC line, dCas9-KRAB-MeCP2 transgene will be active and cells will be expected to shut down the BFP reporter gene expression due to the heterochromatinisation of the landing platform. To validate the functionality of the silencing system we will again perform flow cytometry to observe cell profiles. In brief, the preliminary monitoring of the prototype mESC line will be done via flow cytometry on the expression of the different reporter genes.

Later, broad monitoring will be done to interrogate the power of the prototype mESC line by analysing the silencing system's ability. First of all, we will have to locate the integration sites of the HLP within the prototypes genome. For this, we will be using a novel method called Splinkerette PCR. Once we locate the HLP integration sites, the next step to observe the silencing range of the heterochromatinisation that is spreading from the integration sites. This will be the investigation of the H3K9me3 modification. For this, we will perform Real-time polymerase chain reaction (qPCR) following the chromatin immunoprecipitation (ChIP). Besides, via qPCR we will interrogate the expression levels of some genes nearby to the integration sites. In

conclusion, the goal of the broad monitoring will be to analyse the levels of the aforementioned epigenetic modification in a logical range within the genome, in order to be able to hypothesise the power of the repressor system so, the prototype mESC line.

2. Materials and Methods

2.1. Vectors

In the table below, all the vectors that were used in this thesis project have listed. In the results section, short names were used to refer the vectors.

Full name	Short name	Description	Antibiotic	Marker
95_pB-CAGGS-dCas9-KRAB-MeCP2	pB_dCas9	PiggyBac vector includes dCas9-KRAB-MeCP2 with CAG promoter	Ampicillin	Blasticidin
87_p1302_new_Hygro	HLP vector	PiggyBac vector includes the landing platform	Ampicillin	Hygromycin/BFP
124_pB_TetO/dCas9-KRAB-MeCP2	pB_dox_dCas9	PiggyBac vector includes dCas9-KRAB-MeCP2 with tetracycline operator	Ampicillin	Puromycin
125_Tol2_TetO/dCas9-KRAB-MeCP2	Tol2_dox_dCas9	Tol2 vector includes dCas9-KRAB-MeCP2 with tetracycline operator	Ampicillin	Puromycin
121_blue_cmv1_STAgR_4gRNAs	Blue_4gRNA_cmv	Multiple gRNA vector promoted by CMV with BFP marker	Ampicillin	BFP
122_red_cmv1_STAgR_4gRNAs	Red_4gRNA_cmv	Multiple gRNA vector promoted by CMV with RFP marker	Ampicillin	RFP
123_red_Ef1alpha_STAgR_4gRNAs	Red_4gRNA	Multiple gRNA vector promoted	Ampicillin	RFP

		by EF1alpha with RFP marker		
35_U6 BsaI gRNA	gRNA backbone	Backbone vector for all gRNA sequences that were used	Kanamycin	-
#41_PB_TRE_MCS_ Ubqc_rtta_Ires_Puro	Vector #41	Vector contains tetracycline operator sequences used for cloning pB/Tol2_dox_dCas9 vectors	Ampicillin	Puromycin
#29_pDTA-ROSA26-EF1aCas9IRESneo	Vector #29	Vector contains EF1alpha promoter used for cloning the vector Red_ mltp_gRNA	Ampicillin	Neomycin
pUPATrap-TMat	Tol2 vector	Vector contains Tol2 sequences used for cloning Tol2_dox_dCas9 vector	Ampicillin	-

2.2. String assembly guide RNA (STAgR) - multiple gRNA expression vector cloning

2.2.1. Guide RNA and promoter sequences

Four gRNA cassettes were used to establish STAgR vector. Those are the gRNA sequences from the vectors "Flyg1", "Flyg2", "Flyg3" and "Ctrlg2".

gRNA vectors	gRNA sequences (5′→3′)
Flyg1	GCCACAATTGTCGATCGTCA
Flyg2	GACCGGAACGATCTCGCGTA
Flyg3	GCCCTGCGAGTCATCGGACA
Ctrlg2	GTAGCGAACGTGTCCGGCGT

For each gRNA cassette, 1 gRNA scaffold and 1 promoter were used. In total, 3 different promoters were used. Those are human U6 (hU6), human H1 (hH1) and mouse U6 (mU6).

gRNA scaffold and promoters	Primer sequences (5′→3′)
Scaffold_forward	GTTTTAGAGCTAGAAATAGCAAGTT
hU6_reverse	CGGTGTTTCGTCCTTT
hH1_reverse	CTGGGAAAGAGTGGTCTCATACAG
mU6_reverse	CAAACAAGGCTTTTCTCCAAGG

2.2.2. Primers

In order to generate Gibson fragments for each string as well as the vector backbone, primers below were designed.

Primer name	Primer sequences (5′→3′)*
STAgR_BackbonegRNAscf_Ctrlg2_F	**GTAGCGAACGTGTCCGGCGT**GTTTTAGAGCTAGAAATAGCAAGTT
STAgR_BackbonehU6_Flyg1_R	**TGACGATCGACAATTGTGGC**CGGTGTTTCGTCCTTT
STAgR_hH1gRNAscf_Flyg1_F	**GCCACAATTGTCGATCGTCA**GTTTTAGAGCTAGAAATAGCAAGTT
STAgR_hH1hH1_Flyg2_R	**TACGCGAGATCGTTCCGGCT**CTGGGAAAGAGTGTCTCATACAGA
STAgR_mU6gRNAscf_Flyg2_F	**GACCGGAACGATCTCGCGTA**GTTTTAGAGCTAGAAATAGCAAGTT
STAgR_mU6mU6_Flyg3_R	**TGTCCGATGACTCGCAGGGC**CAAACAAGGCTTTTCTCCAAGG
STAgR_hU6gRNAscf_Flyg3_F	**GCCCTGCGAGTCATCGGACA**GTTTTAGAGCTAG AAATAGCAAGTT
STAgR_hU6hU6_Ctrlg2_R	**ACGCCGGACACGTTCGCTAC**CGGTGTTTCGTCCTTT

*Bold sequences stand for each gRNA sequence

2.2.3. PCR amplification

Promoter, scaffold and gRNA sequence containing gRNA cassettes were synthesised via PCR using homemade Phusion Polymerase. When producing each string, a vector with the corresponding promoter was used. Also, for the backbone

production, 2 different vectors containing either red or blue fluorescent protein gene were used.

Reaction for synthesis of each string and backbone as follows:

Component	Amount/Volume for 25µl reaction
Thermo Scientific™ Phusion Green HF Buffer	5 µl
dNTPs	0.5 µl of 10mM
Forward and Reverse Primers	1.25 µl of 10µM
DNA Template	10 ng
Phusion Polymerase	0.25 µl
DMSO	0.75 µl
ddH_2O	Up to 25 µl

The following PCR program was used for each reaction:

1) Preheating at 98°C for 1.30 minutes
2) Denaturing at 98°C for 10 seconds
3) Primer annealing at 59°C for 30 seconds
4) Extension at 72°C for strings 30 seconds for backbones 1.30 minutes
5) Final extension at 72°C for 10 minutes

2.2.4. DpnI restriction enzyme digestion

Before continuing with the Gibson assembly reaction, to get rid off the plasmid template, the PCR products were digested with Dpn1 (NEB #R0176). Dpn1 restriction enzyme does not cleave the PCR product, but it cleaves only *E. coli* Dam methylase-methylated plasmid DNA. All of the reactions were incubated at 37°C for one hour. After that amplified and restricted PCR products were analysed on a 1% agarose gel.

Component	Final concentration	Final volume
PCR reaction	Variable	20 µl
10X CutSmart® Buffer (NEB #B7204S)	1X	2.5 µl
Dpn1 (NEB #R0176)	Variable	0.5 µl
ddH_2O	-	2 µl
		= 25 µl

2.2.5. Gibson assembly reaction

Gibson assembly is a cloning method that allows multiple DNA fragments to join together through only one isothermal reaction. In this experiment, it was used to assemble the strings and the backbones previously generated by PCR reactions into a single vector with multiple gRNA cassettes. 3 different gRNA expression cassettes with each blue and red expressing backbone were assembled.

For this purpose, Gibson Assembly® Cloning Kit from NEB (#E5510S) was used. Also, to calculate the amount of the PCR products to be used for the reaction, NEBio Calculator was utilised. As negative controls, 2 reactions without any inserts were performed.

	3 and more fragments to be assembled
Recommended DNA ratio	vector:insert = 1:3
Total amount of DNA	0.2 – 1 pmol
2X Gibson Assembly Master Mix	10 µl
Total volume	20 µl

Samples were incubated in a thermocycler at 50 °C for 60 minutes. Following the reaction, samples were ready for the bacterial transformation.

2.2.6. *Escherichia coli* transformation

NEB® 5-alpha Competent *E. coli* cells were used from the same kit. Competent cells were thawed on ice for around 15 minutes and for each Gibson assembly reaction 2 µl of the chilled assembly product was added to the thawed competent cells. The samples were gently mixed by flicking the tube 4–5 times and then placed on ice for 30 minutes followed by a heat shock step at 42°C for 30 seconds. After the heat shock, cells were incubated on ice for 2 minutes and 950 µl of room temperature SOC media was added to each transformation batch. Upon this step, cells were regenerated by incubating them at 37°C for 60 minutes with 650 rpm rotation. As the final step, 100 µl of the cells were inoculated on warm LB plates containing Ampicillin (100 µg/ml) and plates were incubated overnight at 37°C.

2.2.7. Colony PCR

In order to identify bacterial colonies that harbour vectors with the correct number of assembled gRNA expression cassettes, Colony PCR was performed on both blue and red expressing vector plates. Two primers designed for this purpose bind outside regions of the gRNA expression cassettes. Therefore, the length of the amplicon indicates the number of cassettes within the vector. The primers:

Primer name	**Primer sequences (5'→3')**
STAgR_control_F	ACTGGATCCGGTACCAAGG
STAgR_control_R	TTACGGTTCCTGGCCTTTTG

From each positive plate 24 colonies were analysed. REDTaq® ReadyMix™ PCR Reaction Mix from Sigma-Aldrich was used for the PCR reaction. For 10 reactions:

Component	Final volume
REDTaq® ReadyMix™	62.5 µl
Forward and reverse primers (100mM)	1 µl
ddH_2O	50,5 µl
	= 115 µl

For reactions, a 96 well PCR master plate was used. Each well gets 12 µl master mix. Single colony is picked with p10 micropipette tip and carefully transferred to the master plate. Right after this, tip is quickly put into the corresponding well and swirled rigorously.

The following PCR program was used for Colony PCR:

1) Initial denaturation, 5 min at 94 °C
2) Denaturation, 30 sec at 94 °C
3) Annealing, 30 sec at 55 °C (depends on Tm)
4) Extension, 2 min at 72 °C
5) Return to denaturation step for 38 additional cycles
6) Final extension, 5 min at 72 °C

PCR amplicons are later analysed on a 1% agarose gel. From the results of the colony PCR, cultures that are harbouring the correct vectors are identified and later inoculated in a 5ml overnight LB culture (with 100µg/ml Ampicillin) at 37°C. One blue and one red expressing STAgR harbouring vector was reproduced with QIAGEN® Plasmid Maxi prep kit and verified with Sanger sequencing by using the same primers stated at section 2.2.7.

2.2.8. Further cloning of red expressing STAgR vector

This set of tasks were planned to replace CMV promoter in the constructed red fluorescent protein expressing STAgR vector with an EF1-alpha promoter. This promoter region was cloned from the vector #29_pDTA-ROSA26-EF1aCas9IRESneo that is in the main plasmid list.

2.2.8.1. Preparing red expressing STAgR backbone

In order to cut open the STAgR vector and get rid off the CMV promoter region, it was digested with Not1-HF (NEB #R3189), Nhe1-HF (NEB #R3131). Restriction digestion protocol:

Component	Final concentration	Final volume
Vector #41	5 μg	2.5 μl
10X CutSmart® Buffer (NEB #B7204S)	1X	8 μl
2 Restriction enzymes	Variable	2 μl each
ddH_2O	-	65.5 μl
		= 80 μl

The reaction was incubated at 37°C overnight. The next day, the restriction enzyme was heat-inactivated at 80°C for 30 minutes. The enzymes create sticky ends. To get rid of the sticky ends, the reaction batch was treated with 1 μl of DNA Polymerase I, Large (Klenow) Fragment (NEB #M0210) and 1 μl of 10mM dNTPs, and incubated at room temperature for 30 minutes. After this step, it is necessary to clean the DNA products and remove excess enzymes and nucleotides. For this purpose, Wizard® SV Gel and PCR Clean-Up System kit from Promega was used. The cut-open STAgR vector was eluted in 35μl nuclease free water. Furthermore, in order to avoid the vector polymerising to itself 5' phosphate was removed. Protocol:

Component	Final concentration	Final volume
Reaction batch	-	35 µl
10X CutSmart® Buffer (NEB #B7204S)	1X	5 µl
Alkaline phosphatase (CIP) (NEB # M0290)	Variable	1 µl
ddH_2O	-	9 µl
		= 50 µl

The reaction was incubated at 37°C for 60 minutes. Removal of the agents from the Alkaline phosphatase step was performed by running the reaction batch on a 0,8% agarose gel and extracting it. Protocol:

- 50 µl of the sample was added with 10 µl DNA loading dye (6X). Also, 0,7 µl of Midori Green Direct dye (#MG06) from Nippon Genetics was added to visualise bands in the absence of EtBR
- For each ladder (GeneRuler 1kb plus (SM1331) from Thermo Fisher) 1ul of Midori Green Direct dye was added
- Run at 80V for around 1.5 hours
- The extraction of the sample from the gel was done again by performing Wizard® SV Gel and PCR Clean-Up System kit from Promega. The cut-open STAgR vector was eluted in 35µl nuclease free water

2.2.8.2. Preparing EF1-alpha promoter region

EF1-alpha promoter insert was produced from the vector #29 by using PCR. The insert should have overhangs, which are complementary to the ends of the cut-open STAgR vector. Therefore, primers contained these overhang sequences. Designed primers:

Primer name	Primer sequences (5′→3′)
Ef1-a-prom#29_F	GTGCCACCTGACGTCGCTAGGGATCCACGCGTGAG
Ef1-a-prom#29_R	AGTGAGTCGTATTAGCGGCCCCAAGCTTCCCACGA

PCR for synthesis of EF1-alpha as follows:

Component	Amount/Volume for 25µl reaction
Thermo Scientific™ Phusion Green HF Buffer	5 µl
dNTPs	0.5 µl of 10mM
Forward and Reverse Primers	0.5 µl of 10µM
DNA Template	100 ng
Phusion Polymerase	0.5 µl
DMSO	2 µl
ddH_2O	Up to 25 µl

Since, melting temperature for each primer is between 60 and 64°C a gradient PCR was performed. Therefore, annealing temperature was set between 58°C in 65°C. The following PCR program was used for each reaction:

Repeat	Steps	Temperature	Time
-	Initial Denaturation	98°C	1 Minute
38X	Denaturation	98°C	10 Seconds
38X	Annealing	58-65°C	20 Seconds
38X	Extension	72°C	1.5 Minutes (30sec/kb)
-	Final extension	72°C	5 Minutes

After the PCR reaction Dpn1 restriction digestion was performed (see Section 2.2.4.).

2.2.8.3. Gibson assembly and bacterial transformation

The insert, EF1-alpha promoter that is produced by PCR and the cut-open STAgR vector containing TetO repeat sequences were assembled together by using Gibson assembly method. This time the homemade Gibson kit was used.

	3 and more fragments to be assembled
Recommended DNA ratio	vector:insert = 1:2
Total amount of DNA	0.2 – 0.5 pmol
Homemade Gibson Master Mix	3 µl
Total volume	10 µl

Samples were incubated in a thermocycler at 50 °C for 60 minutes. Following the reaction, samples were ready for the bacterial transformation. Only one negative control was used. For the bacterial transformation the same steps were followed (see Section 2.2.6.).

2.2.8.4. Colony PCR

In order to identify bacterial colonies that harbour correct vector colony PCR was performed. Two primers designed for this purpose.

Primer name	**Primer sequences (5′→3′)**
Forward ef1a_redstgr_1r	TCGCTTCATGTGACTCCACG
Reverse ef1a_redstgr_2f	TTGTAATCGGGGATGTCGGC

For the colony PCR protocol see Section 2.2.7. Constructed vector is named #123_red_EF1alpha_STAgR_4gRNAs.

2.3. dCas9-KRAB-MeCP2 vector cloning

2.3.1. Doxycycline inducible dCas9-KRAB-MeCP2

In this task, a vector that contains dCas9-KRAB-MeCP2 insert with the tetracycline operator (TetO) repeat sequences was established. For this purpose, from the general

plasmid list, vector #95_pB-CAGGS-dCas9-KRAB-MeCP2 and as the backbone #41_PB_TRE_MCS_Ubqc_rtta_Ires_Puro were used.

2.3.1.1. Preparing TetO backbone

In order to cut open the vector #41, it was digested with EcoR1-HF (NEB #R3101). Restriction digestion protocol:

Component	Final concentration	Final volume
Vector #41	5 µg	2.5 µl
10X CutSmart® Buffer (NEB #B7204S)	1X	8 µl
EcoR1-HF (NEB # R3101)	Variable	3 µl
ddH_2O	-	66.5 µl
		= 80 µl

The reaction was incubated at 37°C overnight. The next day, the restriction enzyme was heat-inactivated at 65°C for 30 minutes. The enzyme EcoR1 creates sticky ends. To get rid of the sticky ends, the reaction batch was treated with 1 µl of DNA Polymerase I, Large (Klenow) Fragment (NEB #M0210) and 1 µl of 10mM dNTPs, and incubated at room temperature for 30 minutes. After this step, it is necessary to clean the DNA products and remove excess enzymes and nucleotides. For this purpose, Wizard® SV Gel and PCR Clean-Up System kit from Promega was used. The cut-open vector #41 was eluted in 35µl nuclease free water. Furthermore, in order to avoid the vector polymerising to itself 5' phosphate was removed. Including this step, the steps from now on were same as the section 2.2.8.1.

2.3.1.2. Preparing dCas9-KRAB-MeCP2 insert

dCas9-KRAB-MeCP2 insert was produced from the vector #95 by using PCR. The insert should have overhangs, which are complementary to the ends of the cut-open

#41 vector. Therefore, primers contained these overhang sequences. Designed primers:

Primer name	**Primer sequences (5′→3′)***
dCas9-K-M_Tet-O_F	**CGGTGGTACCGACTCGAA**TTGCCACCATGGACAAGAAGTAC
dCas9-K-M_Tet-O_R	**TCGAGCCCGGGCGGAATT**GAACAAACGACCCAACACCC

*Bold sequences are complementary to the Tet-O vector ends

Since, the expected length of the insert is more than 5 kb PrimeSTAR® GXL DNA Polymerase was used. Reaction mixture:

Component	**Amount/Volume for 25µl reaction**
5X PrimeSTAR GXL Buffer	5 µl
dNTPs	2 µl of 10mM
Forward and Reverse Primers	0.75 µl of 10µM
DNA Template	10 ng
PrimeSTAR® GXL DNA Polymerase	0.5 µl
ddH_2O	Up to 25 µl

PCR conditions:

Repeat	**Steps**	**Temperature**	**Time**
-	Initial Denaturation	98°C	30 Seconds
35X	Denaturation	98°C	10 Seconds
35X	Annealing	60°C	15 Seconds
35X	Extension	68°C	5.30 Minutes (1min/kb)
-	Final extension	68°C	10 Minutes

After the PCR reaction Dpn1 restriction digestion was performed (see Section 2.2.4.).

2.3.1.3. Gibson assembly and bacterial transformation

The insert, dCas9-KRAB-MeCP2 that is produced by PCR and the cut-open vector #41 containing TetO repeat sequences were assembled together by using Gibson

assembly method (see Section 2.2.5.). Only one negative control was used. Also, for the bacterial transformation the same steps were followed (see Section 2.2.6.).

2.3.1.4. Colony PCR

In order to identify bacterial colonies that harbour correct vector colony PCR was performed. Two primers designed for this purpose. Forward primer is in the backbone vector #41 upstream to the expected insertion site and reverse primer is in the dCas9-KRAB-MeCp2 insert downstream to the expected insertion site. The primers:

Primer name	**Primer sequences (5′→3′)**
Forward 59pBtetOHy_F2	CCAAGGCCAGACTAGACGAG
Reverse qPCR_Sgk3_F	GTGCGATGAGGTTTTCGAGC

For the colony PCR protocol see Section 2.2.7. Constructed vector is named #124_pB_TetO/dCas9-KRAB-MeCP2_puro.

2.3.2. Tol2 Transposon system with dCas9-KRAB-MeCP2

In this task, a vector that contains dCas9-KRAB-MeCP2 insert with the Tol2 Transposon system was established. For this purpose, from the general plasmid list, vector #124_pB_TetO/dCas9-KRAB-MeCP2_puro and pUPATrap-TMat (Tol2 vector) were used. Assembly was performed as 3 fragment Gibbson:

- Backbone is the cut out w/ Not1 and Xho1 Tol2 vector
- 1st fragment is 1st half the pB_Tet-O/dCas9-KRAB-MeCP2_puro
- 2nd fragment is 2nd half the pB_Tet-O/dCas9-KRAB-MeCP2_puro

2.3.2.1. Preparing Tol2 backbone

In order to cut out the Tol2 region off the vector, it was digested with Xho1 (NEB #R0146), Not1-HF (NEB # R3189), Sma1 (NEB # R0141). Restriction digestion protocol:

Component	Final concentration	Final volume
Vector #41	5 µg	2.5 µl
10X CutSmart® Buffer (NEB #B7204S)	1X	8 µl
3 Restriction enzymes	Variable	1.5 µl each
ddH_2O	-	65 µl
		= 80 µl

The reaction was incubated at 37°C overnight. The next day, the restriction enzyme was heat-inactivated at 65°C for 30 minutes. To get rid of the sticky ends, the reaction batch was treated with 1 µl of DNA Polymerase I, Large (Klenow) Fragment (NEB #M0210) and 1 µl of 10mM dNTPs, and incubated at room temperature for 60 minutes. After this step, it is necessary to clean the DNA products and remove excess enzymes and nucleotides. For this purpose, Wizard® SV Gel and PCR Clean-Up System kit from Promega was used. The Tol2 backbone was eluted in 35µl nuclease free water. Furthermore, in order to avoid the vector polymerising to itself 5' phosphate was removed. Including this step, the steps from now on were same as the section 2.2.8.1.

2.3.2.2. Preparing fragments produced from pB_Tet-O/dCas9-KRAB-MeCP2_puro

The fragments were produced from the vector #124_pB_TetO/dCas9-KRAB-MeCP2_puro by using PCR. The insert should have overhangs, which are complementary to the ends of the cut-out Tol2 vector. Therefore, primers contained these overhang sequences. Designed primers:

Primer name	Primer sequences (5′→3′)
Tol2/Tet-o-dCas9_F	CCGGCTCGATCTCCGCGGCCGGCCAGACTAGACG AGTTTAC
Tol2/dCas9_3Gibbson_ 1st_R	GACGAGGTTCTTGTAATTCTCAAGCATGAC
Tol2/dCas9_3Gibbson_ 2nd_F	AGAATTACAAGAACCTCGTCAGTCTCGGAT
Tol2/Tet-o-dCas9_R	TGGATATGATGCATGCTCGAGGGATACGGGGAAA AAGCCA

Since, the expected length of the insert is more than 5 kb PrimeSTAR® GXL DNA Polymerase was used (see Section 2.3.1.2).

2.3.2.3. 3 fragment Gibson assembly and bacterial transformation

Two fragments from the vector #124_pB_TetO/dCas9-KRAB-MeCP2_puro that is produced by PCR and the cut-out Tol2 vector containing Tol2 sequences were assembled together by using Gibson assembly method (see Section 2.2.5.). Only one negative control was used. Also, for the bacterial transformation the same steps were followed (see Section 2.2.6.).

2.3.2.4. Colony PCR

In order to identify bacterial colonies that harbour correct vector colony PCR was performed. Two primers designed for this purpose. Forward primer is in the backbone Tol2 vector and reverse primer is in the first fragment of #124_pB_TetO/dCas9-KRAB-MeCP2_puro. The primers:

Primer name	Primer sequences (5′→3′)
Forward Tol2/dcas9_seq_F	GGAACAAAAGCTGGAGCTCC
Reverse qPCR_Sgk3_F	GTGCGATGAGGTTTTCGAGC

For the colony PCR protocol see Section 2.2.7. Constructed vector is named #125_Tol2_TetO/dCas9-KRAB-MeCP2_puro.

2.4. Working with mouse embryonic stem cells (mESC)

2.4.1. Used mouse ESC lines

The table below shows all the cell lines used during the project.

Cell line	Ploidy	Source
C2	Diploid	Gertsenstein et al., 2010
RC9	Diploid	Li et al., 2018
H129	Haploid	Leeb et al. 2012
HOCT4	Haploid	Leeb et al. 2012
HREX	Haploid	Leeb et al. 2012

2.4.2. Maintenance of the cells

Mouse ES cell lines were cultured in 2i/LIF medium on plates coated with 0.2% gelatine. This is the most efficient way of keeping them pluripotent in culture.

2.4.2.1. Solutions and growth medias

Chiron, 10mM, 60µL aliquots:

- Molecular weight of CHIR = 465.3 g/mol
- CHIR-glass bottle: 25mg (or 50mg)

5,3724mL DMSO is pipetted in a 15mL conical centrifuge tube and 1mL in CHIR-glass bottle to get everything. Then close and invert. After that, it is pipetted back

into the tube and repeated until all CHIR is used. Later, conical centrifuge tube is put into 37°C water bath until yellowish and clear. Aliquot a 60µL in 0.5mL tubes.

PD03, 10mM, 25µL aliquots:

- Molecular Weight of PD03: 482.19 g/mol

It is diluted using DMSO. For example, 10mg PD03, dissolve it in 2.1mL DMSO. Aliquot a 25µL in 0.5mL tubes.

LIF (Leukemia inhibitory factor), 10µg/mL, 600 µL aliquots:

It is diluted to 10µg/mL with PBS. To sterile it is filtered using 0.22µm filters. Aliquots are stored at -80°C, and some as working stocks are stored at -20°C.

L-Glutamine (LG):

Thaw stock bottle at 4°C, for about 24 hours. If necessary, it is put very shortly in a 37°C water bath. It is diluted in room temperature. Make 6 or 4.5 mL aliquots. Aliquots can be stored at -20°C.

0.2% Gelatine:

50mL 2% Gelatine stock solution is poured into fresh 500mL PBS and stored at room temperature.

Trypsin, 0.5%, EDTA, 1 mM solution:

Thaw 100 mL of 2.5% trypsin (Life Technologies) at 4°C overnight. Add this to 500 mL of PBS and add 1 mL of 0.5 M EDTA (pH=8). After mixing well to sterile, it is filtered using 0.22µm filters. Make aliquots in 50-ml conical centrifuge tubes. Aliquots can be stored up to 1 year at −20°C.

ES-DMEM (serum medium+LIF):

Recipe:

- 500 ml high-glucose DMEM – Sigma Aldrich

- 85 ml fetal bovine serum - Sigma Aldrich
- 6 ml 100× L-glutamine
- 6 ml 100× penicillin/streptomycin - Gibco
- 6 ml 100× Non-Essential amino acids - Gibco
- 6 ml 100× sodium pyruvate - Gibco
- 600 µl 2-mercaptoethanol (55mM) – Sigma Aldrich
- 550 µL LIF

After mixing, the media is flow filtered (0.22µm filter). This media can be stored up to 1 month at 4ºC.

ES DMEM with 2i:

Two inhibitors are added to an aliquot of ES DMEM media. This aliquot is generally 50 mL.

- 10mMChiron: 1:3333 → 15 µL in 50mL
- 10mM PD03 1:10000 → 5 µL in 50mL

This media can be stored up to 1 week at 4ºC.

2.4.2.2. Plate sizes and volumes

96well plate ~0.35cm2 Trypsin: 30µL + stop medium 100 µL New medium: 200µL	**6 cm dish** 21cm2 Trypsin: 500µL + stop medium 1mL New medium: 4mL
24 well plate 2cm2 Trypsin: 70 µL + stop medium 250 µL New medium: 500µL	**T25** 25cm2 Trypsin: 500µL + stop medium 2mL New medium: 8mL
12 well plate 4cm2 Trypsin: 100µL + stop medium 500 µL New medium: 1mL	**10 cm dish** 55cm2 Trypsin: 1mL + stop medium 4mL New medium: 10mL
6 well plate	**15 cm dish**

10cm2 Trypsin: 250µL+ stop medium 1mL New medium: 2mL	152 cm2 Trypsin: 3 mL + stop medium 12mL New medium: 20mL

2.4.2.3. Passage conditions

The cells are passaged when they reach a sizeable ICM outgrowth (usually on day 2 to 3). The growth media is removed by using a vacuum and a Pasteur pipet. After washing the cells with PBS, depending on the size of the plate/well x µl of 0.5% trypsin/EDTA is added. Incubation time is maximum of 3 minutes at 37°C in the cell culture incubator. To further induce the cell detachment carefully tap the plates. This step should be carefully done, in order to avoid cell suspension splashing onto the lid. Again, depending on the size of the plate/well add x µl ES-DMEM + 2i (containing 15% FBS) to inactivate the trypsin. Pipet up and down several times to distribute the media in the cell suspension. Then, the cell suspension is transferred into a 15 mL conical centrifuge tube. In order to collect the cells and get rid off, the trypsin cells are span down at 1000x g for 5 minutes. The remaining media/trypsin mixture is removed by using a vacuum and a Pasteur pipet. Later, cells are resuspended in the ES-DMEM medium before seeding into a plate/well coated with 0.2% gelatine. ES-DMEM +2i medium is added in corresponding volumes into the plate/well. The medium is replaced on the next day with fresh ES-DMEM +2i. Consequently, fresh medium is added every other day.

2.4.2.4. Cell numbers according dish sizes

In order to have confluent cultures, for maintenance, cells always are re-plated in 1:10 ratio to the surface area of the dish. This corresponds to $8x10^4$ cells/cm^2. However, cell doublings must be taken into consideration. Every 12 to 24 hours Mouse ES cells double the amount. Therefore, the half amount is used. This corresponds to $4x10^4$ cells/cm^2. Sometimes even less to $3\text{-}2x10^4$.

2.4.3. Transfection conditions

2.4.3.1. Chemical transfection

For chemical transfections Lipofectamine 2000 from Thermo fisher (cat #11668019) was used.

In a 6 well plate it's need:

- 0.2% gelatine: coat the wells
- 3µL lipofectamine 2000 (L2k) per well
- 500µL serum-free medium (Opti-MEM) per well
- In total 1-2µg plasmid DNA
- ES-DMEM +2i medium for the cells
- Negative control: most cases empty vector

Protocol:

Prepare Lipofectamine 2000 master mix:

- 250µL serum-free medium (OptiMEM) per well
- 3µL L2k per well

Let them for 5 minutes incubation. Meanwhile prepare DNA master mix:

- Make dilutions of DNA to distribute evenly between all wells
- Pipet 250µL OptiMEM in each well
- Pipet DNA in each well and mix by pipetting up and down
- Plasmid (eg. 100-500ng) and transposase (eg. 500-1000ng)
- Add L2k solution to wells (250µL per well)
- Incubate at RT -> complex formation at least 30 minutes (can go up to 6 hours)

In the meantime prepare the cells:

- Cell suspension (trypsinize) and count
- Add $2x10^5$ cells per well
- Prepare cell suspension in a tube to split cells equally
- Add ES-DMEM + 2i medium to a total volume of 2mL per well

Change medium next day if possible and if selection needed start after approximately 24 hours.

2.4.3.1. Electroporation

For electroporation two different gadgets were used. One of them is Gene Pulser Xcell™ Electroporation System from Bio-Rad (Cat #16-2660). The second one is Neon™ Transfection System from Thermo-Fisher (Cat #MPK5000). The protocol steps are not that different from each other, but Neon™ Transfection System is more sophisticated in its terms.

Protocol for Gene Pulser Xcell™ Electroporation System from Bio-Rad (For a 10 cm dish):

- 1-2x10^6 cells with 10- 15 μg/μL of DNA
- Cell pellets are washed in 10ml PBS, and span down 1100xg for 5 minutes
- Cell suspension is resuspended in 800ul PBS then transfered into 4mm cuvettes from Gene Pulser Xcell™ Electroporation System
- Program for electroporation: 270V, 500uf, unlimited Ω, 4mm
- Cells are carefully collected from the cuvettes by using Pasteur pipettes and pipette controller
- Later cells are directly transferred into the 10cm dishes with ES-DMEM+2i medium
- Change medium next day if possible and if selection needed start after approximately 24 hours.

Protocol for Neon™ Transfection System from Thermo-Fisher:

In this transfection method 10 cm dishes were used as well. The instruction manual and the protocol can be found at Thermo-Fisher's website under the catalogue number MPK5000.

2.4.4. Selection conditions

Selection was done with antibiotics with conditions below:

Name	Stock concentration	Working concentration	Use
Hygromycin	50 mg/mL	125 μg/mL	1:300
Blasticidin	10 mg/mL	10 μg /mL	1:2000
Puromycin	1 mg/mL	1 μg/mL	1:2000
Doxycyclin	1 mg/mL	100 ng - 1 μg/mL	1:2000
Normocin	50 mg/mL	100 μg/mL	1:500

2.4.5. Picking up the colonies

Transfected cells are selected with the corresponding antibiotics. After sorting the desired cell population with FACS, these cells are plated in 6 cm dishes in different numbers. These numbers are mostly 200, 500, 1000 and 2000 thousand cells in 6 cm dishes. Later, single colonies are collected one by one and seeded in 96 well dishes. Protocol:

- Prepare a 96-well plate with 10 μL of PBS in each well that is needed
- Prepare a 96-well plate with necessary amount of 0.2% gelatine in each well that is needed
- PBS wash the plates where the colonies are picked and then leave cells in PBS. Colonies are collected in PBS
- With P200 micropipette pick colonies by sucking them with strongly but carefully and transfer them into the wells with PBS. Pipette up and down few times
- After, the needed number of colonies is picked up and put in PBS wells deattach them by adding 25-30 μL of 0.5% trypsin. Incubate at 37C for ~3 minutess
- Meanwhile, discard 0.2% gelatine from the wells of the other 96-well plate and add 100 μl medium (ES-DMEM+2i)

- Tyrpsinized colonies get 75 µl media and are mixed by pipetting up and down before they are transferred into the 96 well plate with the medium

2.4.7. Fluorescence-activated cell sorting (FACS)

2.4.7.1. Höchst staining and sorting

Höchst is toxic and induces some cell death. As a variation on this procedure, haploid ES cell cultures of sufficient purity (>10% haploid) can simply be obtained by sorting based on physical FS/SS (forward/side scatter) characteristics. Gates are set using a diploid control ES cell line. This is the method of choice in case a sorter equipped with a UV laser is unavailable and for routine purifications from a largely haploid cell culture. Materials:

- 10 mg/ml Höchst 33342 in H2O (Life Technologies, cat. no. H3570) ES-DMEM medium with 15% FBS.
- Haploid ES cell cultures
- 15-ml conical centrifuge tubes (e.g., BD Falcon)
- Falcon 70 µm cell strainer (Corning, cat. no. 352350)
- Polypropylene FACS tube (Corning, cat. no. 352063)
- Cell sorter: e.g., Beckman Coulter Moflo or Becton Dickinson FACS Aria equipped with an UV laser
- Cell culture plates and flasks of appropriate size, gelatine coated
- Additional reagents and equipment for dissociating haploid ES cell cultures (e.g. Tyripsin)

Protocol:

- Prepare a Höchst 33342 1:100 working dilution by adding 15 µl of 10 mg/ml Höchst 33342 dye to 1 ml ES-DMEM with 15% FBS. Keep Höchst 33342–containing solutions and media away from direct light, as the dye is light sensitive.

- Dissociate haploid ES cell cultures to prepare a single-cell suspension in 2i/LIF medium. Prepare a cell suspension of at most 2 × 107 cells/ml.
- At higher concentrations, Höchst staining will not reach saturation, and the cell-cycle profile will not show sufficient resolution for purification of haploid ES cells. Cells should be resuspended in a total volume of at least 300 μl.
- Add the diluted Hoechst 33342 from the first step to the cells at 1:10 ratio (the final concentration of Hoechst is 15 μg/ml) and incubate in a 15-ml conical centrifuge tube for 20 to 30 min at 37°C in a tissue culture incubator. Avoid longer incubations, as this can increase cell death.
- Pass the stained cell suspension through a 30-μm cell strainer into a polypropylene FACS tube to remove clumps before sorting on a cell sorter.
- Place FACS tubes on ice immediately and proceed to flow purification. Once on ice, finish sorting quickly, ideally within 1 hour.
- Collect the cells gated to the Höchst 33342 intensity corresponding to 1*n* DNA content in a FACS collection tube containing 0.5 ml ES–DMEM with 15% FBS.
- After sorting, transfer the cell suspension from the sorting tube to a 15-ml conical centrifuge tube with 10 ml ES-DMEM with 15% FBS using a centrifuge (5 min at 250 × *g*, room temperature).
- Remove the medium, carefully avoiding any loss of the cell pellet.
- Resuspend the cell pellet in 2i/LIF medium and plate on a fresh cell culture flask or plate coated with gelatin.
- Passage the cells every 2 or 3 days. Continue with Propidium iodide (PI) staining of the cells

2.4.7.2. Getting prepared for sorting without Höchst staining

Steps are similar to the Höchst sorting. Again as a variation on this procedure, haploid ES cell cultures of sufficient purity (>10% haploid) can simply be obtained by sorting based on physical FS/SS (forward/side scatter) characteristics. Gates are set using a diploid control ES cell line. This is the method of choice in case a sorter

equipped with a UV laser is unavailable and for routine purifications from a largely haploid cell culture. For materials check section 2.4.7.1. Protocol:

- Media preparation with 1:500 Normocin (In order to prevent cells being contaminated while cells are sorted in the cell sorter.).
- Prepare FACS tubes with ES-DMEM+2i and 1:500 Normocin media of 5-ml in 15ml Falcon tubes.
- After sorting, transfer the cell suspension from the sorting tube to a 15-ml conical centrifuge tube with 10 ml ES-DMEM with 15% FBS using a centrifuge (5 min at 250 × *g*, room temperature).
- Remove the medium, carefully avoiding any loss of the cell pellet.
- Resuspend the cell pellet in 2i/LIF medium and plate on a fresh cell culture flask or plate coated with gelatin.
- Passage the cells every 2 or 3 days. Continue with Propidium iodide (PI) staining of the cells

2.4.7.3. Propidium iodide (PI) staining of the cells

Cell cycle analysis based on Propidium iodide staining is a simple way to assess ploidy of haploid ESC cultures, and should be performed regularly (ideally every or every second passage). Materials:

- Haploid ES cell cultures
- 70% ethanol, cold
- Phosphate-buffered saline (Life Technologies, cat. no. 10010-015)
- 100 μg/ml RNase A (e.g., Sigma, cat. no. R4642)
- 50 mg/L Propidium iodide in PBS (store at 4°C protected from light)
- Flow cytometer (e.g., Cyan ADP, Beckman Coulter)
- Benchtop centrifuge

Protocol:

- Set aside and pellet a small proportion of the total cells (ideally >105 cells) at the time of passaging.
- Fix in 1 ml (or 500mL) cold 70% ethanol by drop-wise addition while vortexing at low speed.
- Incubate at 4°C for at least 30 minutes to over night
- Centrifuge 5 min at 500 × g, room temperature, in a benchtop centrifuge, wash once in PBS, and centrifuge again as before.
- Remove supernatant and resuspend pellet in 50 µl of 100 µg/ml RNase A.
- Add 250 µl of 50 mg/L PI and record cell cycle profile on a flow cytometer.
- Gate out cell aggregates using forward and side scatter parameters and record cell cycle in the PI channel.

2.5. Protein methods

2.5.1. Protein extraction

Collected cells are washed with PBS. After this step, cell pellets can be stored at -80°C for up to 1 week. Preparation of the protein lysis buffer:

- 10x RIPA buffer (Thermo-Fisher cat #10017003)
- 10x Protease inhibitor (Roche cat #04693159001) corresponds to 1tablet in 1mL H_2O

Cell pellet is resuspended in lysis buffer (10-100µL depending on size of pellet) and incubated on ice for 15 to 20 minutes. Protein suspension is span down for 25-30 minutes at 13000 rpm at 4°C. The supernatant is transferred to new microcentrifuge tube. Protein lysate can be stored at -20°C.

2.5.2. Bradford assay

Bradford reagent (Sigma-Aldrich cat #B6916) is diluted in H_2O with 1:5 ratio. For the measurement 96 well tissue culture plate is used. Every well gets 200µL Bradford

reagent and 1µL lysate and incubated at room temperature for 5 minutes. Measurements are done with the microplate reader.

2.5.3. Western blot

2.5.3.1. Acrylamide gel electrophoresis

Proteins were separated according to their lengths by polyacrylamide electrophoresis. Since, dCas9-KRAB-MeCp2 protein is just over 200 kDa, mostly 8 to 10% gels were used. Preparation of the gels:

- put TEMED last and use it 1000x (eg. 10µL in 10mL) !
- ~8mL per 1st separating gel
- ~4mL for 2nd stacking gel
- Pour 1st separating gel, fill up with Isopropanol, wait 40 minutes
- After the 1st gel is polymerised wash off the Isopropanol with tap water
- Pour 2nd stacking gel, wait

Assemble electrophoresis chamber by putting glass with the gel as much down as possible and close chamber tightly minding the green edges. Preparation of the 10x SDS-PAGE running buffer:

- 30 gr Tris Base
- 144 gr Glycine
- Fill up to 1L with H_2O and autoclave
- Add 10 gr SDS

Assemble the electroph system and fill the chamber up until buffer overflows, then put buffer into outer chamber.

2.5.3.2.Sample preparation and gel run

To prepare 10 mL Lämmli buffer 6x:

- 4,7mL Glycerol
- 1,2mL Tris 0,5M pH 6.8

- 1,2g SDS pellets
- 6mg Bromphenol blue
- Add H_2O up to 10mL and then heat to 65°C to dissolve the SDS pellets. Finally add 0,93mL DTT
- Aliquot a 500µL, can be stored at -20° C

To prepare the protein samples ready to load, thaw 6x Lämmli buffer at room temperature. Then add the corresponding amount of water. At last add the needed amount of the protein lysate. In order to denature the protein boil the samples at 95 °C before loading them. Load also 5 µL of broad range protein marker. Running time is almost 2 hours. 80V until the protein samples reach the border of separating gel and then 120V.

2.5.3.3. Transfer and blocking

Prepare 1L Transfer buffer a while in advance, store at 4 ° C:

- Fill up with 800mL MQ and 100mL Methanol
- In the cold room add 100mL of 10x transfer buffer stock

Prepare sandwich:

- Remove stacking gel
- Every piece & gloves must be soaked with transfer buffer
- Cut membrane (1 per gel) in sponge size and remove the protective foil
- Set up sandwich: (**+) White face of the apparatus** – Sponge – Paper – NC membrane – Gel – Paper – Sponge – **Black face of the apparatus (-)**
- Roll with pipet tip
- Put sandwich in chamber (black to black, white to red)
- Put ice pack into chamber, fill up completely with buffer and close the lid
- Put whole chamber into tray with full of ice as possible
- Run 2h at 200 mA (0.4 A) for 1 apparatus

Blocking:

- 10% milk: milk powder 5g and fill with PBST up to 50 mL

- Put membrane in a tray with 10-20mL 10% milk
- Block for 45min - 2h at room temperature

2.5.3.4. Hybridization and detection

The membrane is left with primary antibody overnight at 4°C. Primary antibody for dCas9 is a homemade antibody and it is used 1:100 in 5% milk blocking solution. The name is "Cas9 8C". For the Tubulin control, antibody is used 1:5000. The next day membrane is washed for 3 times, each time for 7 minutes with PBST. The secondary antibody is an anti-mouse HRP and it is used 1:10000/1:12000 in again 5% milk blocking solution. The membrane is left with secondary antibody for at least 1 hour at room temperature. Later, the membrane is washed for 3 times, each time 10 minutes with PBST.

For detection, dry membrane with Kimtech® Science™ Kimwipes™ Delicate Task Wipers (cat #34133). Then place it on transparent foil. Incubate the membrane for 5 minutes with Amersham ECL Select Western Blotting Detection Reagent from GE Healthcare Life Sciences (cat #RPN2235). Again dry it with Kimwipes™. Put membrane in Bio-Rad ChemiDoc™ Imagers within the foil. Detection settings:

- 5 seconds exposure time for 10 minutes and 11 pictures
- For marker image use the colorimetric option

2.6. DNA and RNA methods

2.6.1. DNA extraction

Throughout this project, for every DNA extraction "Puregene-Core-A" kit from Qiagen was used. Protocol can be found in instructions book provided with the kit (Pg. 34-36). There were some modifications on done the protocol:

- Start from step 6. Just resuspend your pellets in required amount of cell lysis solution

- Step 8 and 9 → everything on ice
- Step 9 and 12 → centrifuge full speed
- Step 16 → 65 ° C with 1000 rpm
- Step 17 → 50 μL DNA Hydration solution

2.6.2. Total RNA extraction

Throughout this project, for every total RNA extraction "EXTRACTME TOTAL RNA KIT (EM09.1)" from Blirt was performed. After extraction, RNA samples should be stored at -80 ° C.

2.6.3. Reverse transcription polymerase chain reaction (RT-PCR)

Calculate how much extracted total RNA would be needed for 500 ng of DNA sample. For the reverse transcription reaction "SensiFAST™ cDNA Synthesis Kit" from Bioline is used. Protocol:

- The total reaction volume is 20 μL.
- 1 μL of Reverse transcriptase enzyme
- Reaction Buffer solution provided by the kit should always be 5X of the total volume. Therefore, 4 μL.
- The rest up to 20 μL is water.

Thermocycler settings:

Cycles	**Temperatures**	**Times**	**Notes**
1X	95° C	2 min	Polymerase activation
40X	95° C 60-65° C 72° C	5s 10s 5-20s	Denaturation Annealing Extension

2.6.4. Real-time polymerase chain reaction (qPCR)

2.6.4.1. qPCR primers for haploid mESC prototypes H129-dCas9/HLP-1, 2 and 3

These primers designed for both qPCR and ChIP experiments. There are primers for gene expression qPCR experiments for the genes nearby the integration sites of HLPs. Also, there are primers for ChIP qPCR to investigate H3K9me3 modification.

qPCR primers: Primers designed for several genes for every integration of each subclone. The properties of the primers are as follows:

- Lengths are between 19-23 bp
- Tm s are between 60-62.5
- Produced amplicon size is between 100-150bp
- All primers are within the mRNA coding sides of the gene
- The primers were designed for the genes that have relatively high expression. Genes that are already lowly expressed were ignored.

Also for each gene expression qPCR experiments we used 2 pluripotency and 2 housekeeping markers. Those are:

Pluripotency markers	
ML269 real Oct4 F	GTGGACCTCAGGTTGGACTGG
ML270 real Oct4 R	CTTCTGCAGGGCTTTCATGTC
ML267 real Nanog F	TCGAATTCTGGGAACGCCTC
ML268 real Nanog R	GTCTTCAGAGGAAGGGCGAG

Housekeeping genes	
ML398 real bActin F	CTCCTAGCACCATGAAGATCAAGA
ML399 real bActin R	TCAAAGAAAGGGTGTAAAACGCAG
835 real ribosprotein L32_FW2	GGCTTTTCGGTTCTTAGAGGA
836 real ribosprotein L32_RW2	TTCCTGGTCCACAATGTCAA

ChIP qPCR primers: For ChIP qPCR primers, I designed one set upstream and one set downstream of the integration site. I look for the regions between 1 to 150kb from the HLP integration. The properties of primers are as follows:

- Lengths are between 20-22 bp
- Tm s are between 59-61C
- Produced amplicon size is between 100-150bp

Also for each gene expression ChIP qPCR experiments we used 2 following positive and negative primers:

Positive controls	
1448_MLV 5'LTR FW	CAGGATATCGGTGGTCAAGC
1449_MLV 5'LTR RV	TTCCGCTTTATCTGAGAACC
1450_MLV Gag FW	GTACCATCCTGAGGCCATCC
1451_MLV Gag RV	AGTCTAACCTTGCAGCACTGG
1452_IAP probe FW	CCCAGTTGGGTTGTGAAAGAAAACT
1453_IAP probe RV	TTTACTAGTTAGGCACAGCGGAGGC
Negative controls	
1454_Gapdh promoter FW	AGCATCCCTAGACCCGTACAGT
1455_Gapdh promoter RV	GGGTTCCTATAAATACGGACTGC
ML271 real Sox2F	CAGGAGTTGTCAAGGCAGAGA
ML272 real Sox2 R	CTTAAGCCTCGGGCTCCAAA

2.6.4.2. Protocol

For the qPCR "SensiFAST™ SYBR® No-ROX Kit" from Bioline is used. Protocol for 1 single reaction:

- The total reaction volume is 12 µL
- 6 µL of Sybr-Green Buffer (2X) including the DNA polymerase
- 0.05 µL of each primer (forward and reverse)
- For the cDNA 1/10 dilution should be made. From this stock use 3 µL.
- After distributing 9 µL of the master mix (buffer + primers + water) into wells add the 3 µL of cDNA 1/10 diluted stock.

Thermocycler settings:

Cycles	Temperatures	Times	Notes
1X	95° C	2 min	Polymerase activation
40X	95° C 60-65° C 72° C	5s 10s 5-20s	Denaturation Annealing Extension

2.6.5. Splinkerette PCR (spPCR)

2.6.5.1. Primers for haploid mESC prototypes H129-dCas9/HLP-1, 2 and 3

	Adaptor oligonucleotides		Sequences
	Splinkerette-TruSeq bottom A		/5phos/g*atcggaagagcacacgctgaagatttttttttcaaaaaa*a
	Splinkerette-TruSeq top A overhang		g*ttcccatggtactactcatagtgactggagttcagacgtgtgctcttccgatc*t
	Primers for the 1st PCR		**Sequences**
	SplAP1v2		g*ttcccatggtactactcatagtgact*g
	NGS_PB5pr_1		g*atatacagaccgataaaacacatgcgtc*a
	NGS_PB3pr_1		g*acggattcgcgctatttagaaagaga*g
	Primers for the 2nd PCR		**Sequences**
	qPCR2.2		c*aagcagaagacggcatacgagatcgtgatgtgactggagttcagacgtgtg*c
For 5' end of PB integrations		NGS_PB5pr_2	a*atgatacggcgaccaccgagatctacaccacgcatgattatctttaacgtacgtca*c
		NGS_TSAS_i27:	c*aagcagaagacggcatacgagataggaatgtgactggagttcagacgtgtg*c
		NGS_TSAS_i28:	c*aagcagaagacggcatacgagatcttttggtgactggagttcagacgtgtg*c
		NGS_TSAS_i29:	c*aagcagaagacggcatacgagattagttggtgactggagttcagacgtgtg*c
For 3' end of PB integrations		NGS_PB3pr_2	a*atgatacggcgaccaccgagatctacacatgcgtcaattttacgcagactat*c
		NGS_TSAS_i30:	c*aagcagaagacggcatacgagatccggtggtgactggagttcagacgtgtg*c
		NGS_TSAS_i31:	c*aagcagaagacggcatacgagatatcgtggtgactggagttcagacgtgtg*c
		NGS_TSAS_i32:	c*aagcagaagacggcatacgagattgagtggtgactggagttcagacgtgtg*c

2.6.5.2. Protocol

Splinkerette PCR (spPCR) is useful to locate the exact genomic insertion site for the transgenes. This technique was originally generated to amplify the genomic DNA between a known target gene and a restriction site (Horn et al., 2007).

Shearing the genomic DNA:

- Digest 2–3 μg of genomic DNA with 4 units of Sau3AI (NEB# R0169L) or BamHI (NEB# R3136L) restriction enzymes in a 30 μL reaction overnight.
- Next day, before starting with the PCR reactions, heat-inactivate the enzyme at 65 ° C for 20 minutes.

Adaptor annealing:

- Prepare the Splinkerette adaptors mix 150 pmol of each oligonucleotide in 100 μL water with 0.5X NEB Buffer 2 (NEB #B7202) and heat to 95 ° C for 10 minutes to denature.
- Anneal the oligonucleotides (see section 2.6.5.1) by cooling the mixture slowly to room temperature.
- Annealed Splinkerette adaptors can be stored at 20 ° C.

Adaptor ligation with the digested genomic DNA:

- Set up ligation reactions as follows: 5 μL of digested genomic DNA, 3 μL Splinkerette adaptor, 2 μL T4 ligase buffer (NEB #M0202L), 1 μL T4 ligase (NEB #M0202L), 9 μL water. Incubate at 16 ° C overnight.
- Next day, inactivate the ligase by heating the reaction to 65 ° C for 10 min.

1st and 2nd PCR:

- Use 1 μL of ligation reaction as the template for the first PCR. Any standard polymerase can be used (e.g. Platinum Taq). Set up a standard 25 μL reaction using the primers (see section 2.6.5.1) and make 30 cycles as follows: 94 ° C for 30 seconds, 62 ° C for 30 seconds, 72 ° C for 90 seconds, followed by a final extension at 72 ° C for 5 minutes.
- Set up the 2nd PCR, using 1 μL of the 1st PCR sample as the DNA template and follow the same PCR conditions as above.

2.7. Chromatin immunoprecipitation (ChIP)

2.7.1 ChIP buffer recipes

Block Solution (50 mL) [store at 4°C]

250 mg	BSA powder	0.5% BSA (w/v)
50 mL	PBS (1x)	1x PBS

Lysis Buffer 1 – LB1 (50 mL) [store at RT]

2.5 mL	1 M Hepes-KOH pH 7.5	50 mM Hepes
1.4 mL	5 M NaCl	140 mM NaCl
100 uL	0.5 M EDTA	1 mM EDTA
10 mL	50% glycerol	10% glycerol
2.5 mL	10% NP-40	0.5% NP-40
1.25 mL	10% Triton X-100	0.25% TX-100
32.25 mL	dH20	

Lysis Buffer 2 – LB2 (50 mL) [store at RT]

0.5 mL	1 M Tris-HCl pH 8.0	10 mM Tris
2.0 mL	5 M NaCl	200 mM NaCl
100 uL	0.5 M EDTA	1 mM EDTA
50 uL	0.5 M EGTA	0.5 mM EGTA
47.35 mL	dH20	

Lysis Buffer 3 – LB3 (50 mL) [store at RT]

0.5 mL	1 M Tris-HCl pH 8.0	10 mM Tris
1.0 mL	5 M NaCl	100 mM NaCl
100 uL	0.5 M EDTA	1 mM EDTA
50 uL	0.5 M EGTA	0.5 mM EGTA
0.5 mL	10% Na-Deoxycholate	0.1% Na-Deoxycholate
1.25 mL	20% N-lauroylsarcosine	0.5% N-lauroylsarcosine
46.60 mL	dH20	

RIPA Wash Buffer (250 mL) [store at RT]

12.5 mL	1 M Hepes-KOH pH 7.5	50 mM Hepes
25.0 mL	5 M LiCl	500 mM LiCl
0.5 mL	0.5 M EDTA	1 mM EDTA
25.0 mL	10% NP-40	1% NP-40
17.5 mL	10% Na-Deoxycholate	0.7% Na-Deoxycholate
169.5 mL	dH2O	

Elution Buffer (50 mL) [store at RT]

2.5 mL	1 M Tris-HCl pH 8.0	50 mM Tris
1.0 mL	0.5 M EDTA	10 mM EDTA

5.0 mL	10% SDS	1% SDS
41.5 mL	dH2O	

2.7.2. Protocol

For ChIP high numbers of cells are needed. Therefore, cells are cultured in 15 cm dishes. The first step is crosslinking of the cells. It follows like this. Remove the media and add 10 mL PBS to the cells. Crosslink cells with 1% formaldehyde for 10 minutes at room temperature. Quench with 0.125 M glycine for 10 minutes at again room temperature. Remove PBS/Formaldehyde/Glycine and wash the cells 2x on the plate with ice-cold PBS + 0.01% Triton mixture. Harvest with cell scraper and spin down at 500x g for 5 minutes at 4°C. Flash freeze pellets in liquid nitrogen and store at -80°C.

When ready to start with the experiment thaw frozen pellets on ice for 30-60 minutes. Prepare Lysis Buffers (LB1/2/3) by adding protease inhibitors (Complete Protease Inhibitor Cocktail tablets (Roche cat #04693159001), prepare a 50x or 100x solution prior to use) to 1X and chilling on ice. Turn on the Sonicator and cool the waterbath to 4°C. Resuspend pellet in 5 mL LB1 by pipetting and rotate vertically at 4°C for 10 minutes. Spin 5 minutes at 1350 x g at 4°C (GH-3.8 rotor = 2400 rpm) and aspirate supernatant with gel loading tip. Resuspend pellet in 5 mL LB2 by pipetting and rotate vertically at room temperature for 10 minutes. Spin 5 minutes at 1350 x g at 4°C and aspirate supernatant with gel loading tip. Resuspend pellet in 3 mL LB3 by pipetting. Sonicate the samples for 17 cycles with 30s on and 45s off. This has to be optimized for each cell line and for each cell pellet. Transfer each sonicated sample to a 1.5 mL microfuge tube and spin 10 min at 16000 x g (max) at 4°C to pellet cellular debris ("pellet"). There should be almost no pellet left over. Add 300 µL 10% Triton X-100 to ~3 mL sonicated lysate (1% final conc.). Aliquot ~3300 µL into 6 microfuge tubes: 3 x 1 mL and 3 x 100 µL.

- 1 mL aliquots for ChIP (for up to 3 different ChIP reactions per sample)
- 100 µL aliquots as "input" (are 10% of input for each ChIP reaction)

For histone modifications we go down to 500 µl of extract and 5ug of antibody. For H3K9me3 modification this antibody from Active-Motif (cat #39161) was used. Add 5 µg of antibody to 1 mL sonicated chromatin aliquot for each ChIP reaction and invert to mix. Rotate the samples vertically at 4°C overnight (12-16 hours) to make antibody to chromatin. Store "input" samples at -20°C. Flash freeze any unused sonicated chromatin aliquots (1 mL) in liquid nitrogen and store at -80°C

For this specific antibody Protein G magnetic Dynabeads was used. Aliquot 100 µL Dynabead slurry (50 µL beads) into a microfuge tube on ice for each ChIP reaction. Wash Dynabeads 3 times in 1.5 mL cold Block Solution:

- Add 1 mL cold Block Solution and mix by invert/flick
- Put in magnetic holder and wait for beads to settle on side
- Pour off liquid and repeat
- Last wash: remove all liquid with gel loading tip

Add all ~1 mL of antibody bound chromatin to washed beads and invert to mix. Here it is often a good idea to quickly spin down the extract to remove any unwanted debris that might have formed. Rotate the samples vertically at 4°C for at least 2-4 hours to make antibody bind to beads. Chill RIPA Wash Buffer on ice and set water bath to 65°C while rotating. Wash bound beads at least 5 times in 1 mL cold RIPA and keep on ice:

- Add 1 mL cold RIPA and mix by invert/flick
- Put in magnetic holder and wait for beads to settle on side
- Pour off liquid and repeat
- Last wash: Spin 3min at 950 x g and remove all liquid with gel loading tip

Wash beads 1 time in 1 mL TE + 50 mM NaCl on ice (as described above). Spin 3 minutes at 950 x g at 4°C and remove all remaining TE with gel loading tip. Add 210 µL Elution Buffer at room temperature and vortex briefly to mix. Elute for 15 minutes at 65°C in water bath with brief vortexing every few minutes. Spin down beads 1 min at 16000 x g at room temperature and add to magnetic holder. Transfer supernatant (~200 µL) to fresh microfuge tube once beads settled. Thaw "input"

samples from -20°C and add 300 µL (3 volumes) Elution Buffer and mix by brief vortexing. Incubate ChIP/input samples in 65°C water bath overnight (12-16 hours) to reverse crosslinks.

On the next day, add 1 volume TE buffer/tube at RT (to dilute SDS) and invert to mix:

- input = 400 µL -> 400 µL TE
- ChIP = 200 µL -> 200 µL TE

Add RNase A to 0.2 mg/mL final concentration and invert to mix:

- input = 800 µL -> 8 µL 20 mg/mL RNase A
- ChIP = 400 µL -> 4 µL 20 mg/mL RNase A

Incubate in 37°C water bath for 2 hours to digest RNA. Add $CaCl_2$ to 5.25 mM using 300 mM $CaCl_2$ + 10 mM Tris pH 8.0 and invert to mix:

- input = 800 µL -> 14 µL 300 mM $CaCl_2$
- ChIP = 400 µL -> 7 µL 300 mM $CaCl_2$

Add Proteinase K to 0.2 mg/mL final concentration and invert to mix:

- input = 800 µL -> 8 µL 20 mg/mL Proteinase K
- ChIP = 400 µL -> 4 µL 20 mg/mL Proteinase K

Incubate in 55°C water bath for 30 minutes to digest protein. Phenol-Chloroform extraction is performed to DNA at room temperature. Add 1 volume (input = 800 µL and ChIP = 400 µL) 25:24:1 Phenol-Chloroform-Isoamyl alcohol and mix by inverting. Spin 5 minutes at 13200 rpm (max) at room temperature to separate layers. Add 1 volume Chloroform to the tubes and mix by inverting. Then, again spin for 5 minutes at 13200 rpm to separate layers. Transfer the aqueous phase above the gel to a new 1.5ml tube. Ethanol precipitate DNA. This step must be done on ice. Add the following in order and mix by inversion:

- 2 volumes ice cold 100% ethanol
- 1/10th volume 3M NaOAc
- 1.5 µL 20 mg/mL glycogen

Incubate 30 minutes at -20°C to precipitate DNA. Then spin 30 minutes at 13200 rpm (max) at 4°C to pellet DNA (small white pellet). Pour off EtOH and remove remainder with gel loading tip to leave pellet. Later add 0.5 mL ice cold 70% ethanol and invert/flick to resuspend. Spin 15 minutes at 13200 rpm (max) at 4°C to pellet DNA (small white pellet) and pour off EtOH and remove remainder with gel loading tip to leave pellet. Air dry pellet at room temperature on bench for ~10 minutes (until pellet turning clear). Finally, resuspend pellet in desired volume of dH_20 at room temperature by vortexing:

- input = 50 µL (make sure that you combine both pellets at this point in 50 µL total volume) and ChIP = 50 µL

DNA can be stored at -20°C until needed for subsequent steps.

3. Results

As explained in the introduction, in order to address a significant problem of all genome-wide screening technologies, the ultimate task of this Master's thesis project was to produce screening haploid mouse embryonic stem cell (mESC) prototypes by using a silencing platform that launches epigenetic modifications with the help of an enhanced CRISPR repressor system. The system within the prototypes was planned consist of three genetic components. Those were, a silencing platform that induces heterochromatinisation, which is called the heterochromatinisation landing platform (HLP), the dCas9-KRAB-MeCP2 transcriptional repressor and the guide RNA (gRNA) sequences that is recruited to the HLP in order to achieve the H3K9me3 epigenetic repression. After establishing the prototype cells, functional tests were planned conducted in order to validate the functionality of the system.

In the aim of establishing the advanced haploid mESC prototype, the components had to be tested together in different cell lines, some of them had to be adjusted and some had to be advanced or completely altered. Overall, during the whole project, the main goal was to bring the system to its best shape in order to be able to

construct the advanced haploid mESC prototype. Throughout the project, one simple diploid prototype and some working haploid prototypes (from the same mESC line) were produced. Later on, the attempt on making the advanced haploid prototype, unfortunately, was failed. In the table 1 and the figure 5 below, the prototypes are shown with their different components and the schematic representations explain the system for each prototype.

Prototypes	Cell name	Ploidy	Genetic components	Results/Notes
Simple diploid prototype	C2	Diploid	pB_dCas9 vector, the HLP vector and individual gRNA vectors	Upon transfection of the gRNAs system works
Working haploid prototype	H129	Haploid	pB_dCas9 vector, the HLP vector and Red_4gRNA vector	Upon transfection of the multiple gRNA expression vector system works
Advanced haploid prototype	HREX	Haploid	Tol2_dox_dCas9 vector, the HLP vector and Red_4gRNA vector	Prototype cells were never be able to kept haploid until the multiple gRNA vector transfection

Table 1: Produced mESC prototypes during the project
Table shows produced cell prototypes throughout whole project with the information on their ploidy and the genetic components, which they contain. Also, in the results/notes column for each prototype, the systems functionality is explained.

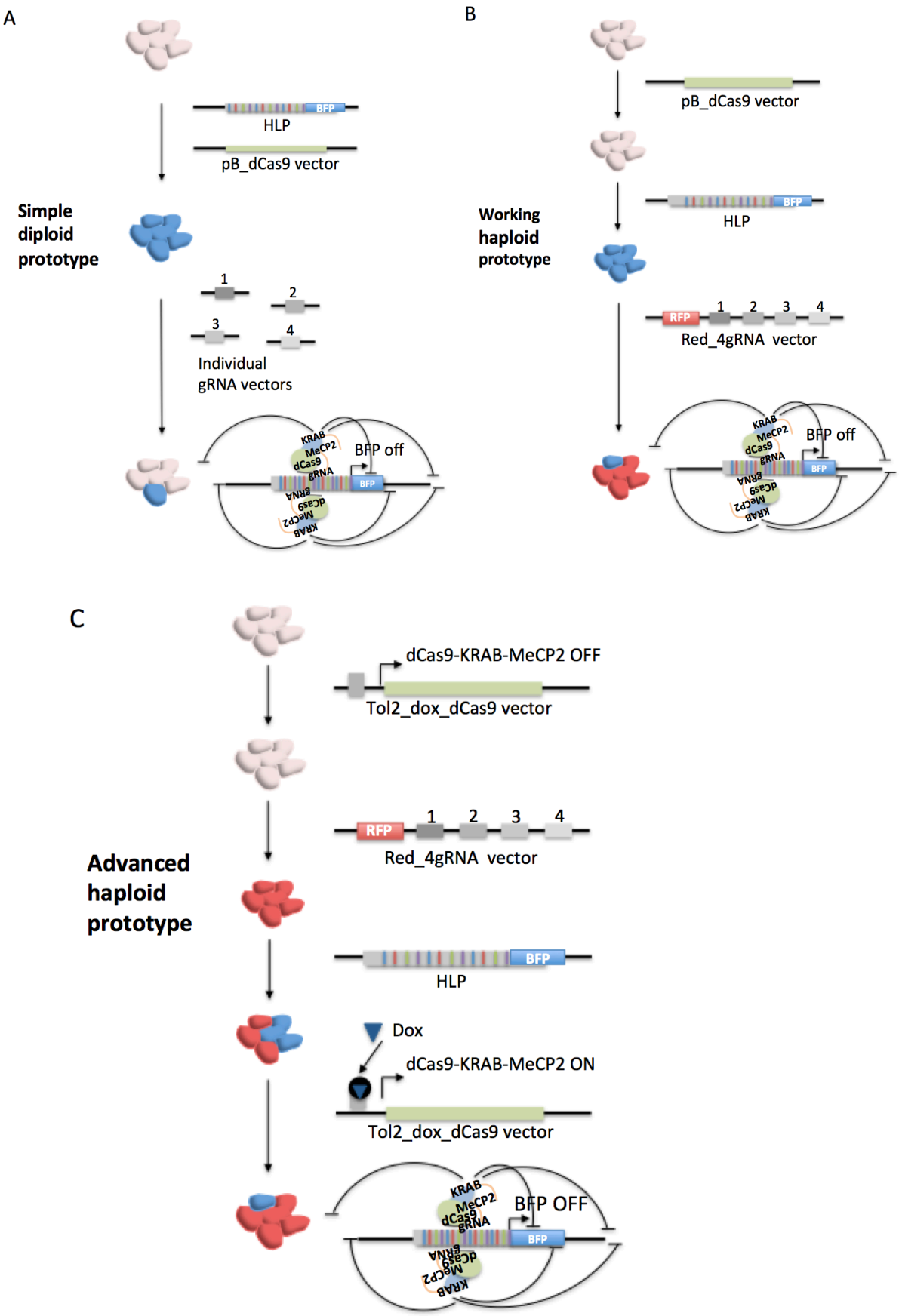
First, the simple diploid prototype was produced. As mentioned, it consisted of three genetic components expressed concurrently. These were dCas9-KRAB-MeCP2 transgene, which was integrated with the transposon system PiggyBac (pB_dCas9 vector), the silencing platform HLP (the HLP vector integrated with PiggyBac as well) and individual gRNA vectors. Upon transfection of the gRNAs system was working.

Secondly, the working haploid prototype was produced. Again, the system contained three genetic components. pB_dCas9 vector was used once more to introduce the dCas9-KRAB-MeCP2 transgene and also the HLP was integrated into the genome with the same transposon system PiggyBac. This time, instead of

individual gRNA vectors, one vector containing multiple gRNA expression cassettes (Red_4gRNA vector) was used. The system's functionality was validated. Once we were sure about the prototype, we looked for H3K9me3 modification on the HLP and sites nearby the integration within the genome and we were able to observe the silencing range of the system.

Later on, after advancing and altering some components, we attempted to construct one advanced haploid prototype. Unfortunately, due to some experimental issues on cells ploidy, our attempt was failed. However, the genetic components for this prototype were known and ready to use. This time, the dCas9-KRAB-MeCP2 transgene was planned to be introduced via Tol2 transposon system in order to distinguish the integration sites of the HLP, which was integrated into the genome with PiggyBac. Also, by using Doxycycline inducible system, the transcriptional activity of the transgene would be controlled at anytime. Thereby, the system in the advanced haploid prototype was aimed to be more unique and secure. Eventually, our task is to establish this prototype.

Finally, in the next parts of the results chapter, the detailed experiments on the production of each prototype and each genetic component required for the system are found.



Legend on the next page →

Figure 5: Schematic representation of the produced/planned mESC prototypes during the project
Overview of the mESC prototypes. simple diploid and haploid lines were established during the project. Advanced haploid prototype is still planned to be constructed. The box below is the legend for the schematic display. (A) Produced simple diploid prototype is shown. Upon the transfection of individual gRNA vectors into the dCas9-KRAB-MeCP2 and the HLP established cells, functional tests showed that the system is working. (B) Working haploid prototype was established by transfecting each component into the subclones. This time, multiple gRNA expression vector "Red_4gRNA" was used. The system is validated. (C) Advanced haploid prototype is yet to be produced. The system is steady and controllable. Again every component is placed subclonally. Benefitting the Doxycycline inducible transcriptional activation, it's planned to be the switch of the prototype.

3.1. Producing and testing diploid mESC prototype

3.1.1. Task and experimental setup

As mentioned above, the ultimate task of this master thesis project is using the enhanced CRISPR repressor system, the dCas9-KRAB-MeCP2 transcriptional repressor, with the target heterochromatinisation landing platform (HLP) in order to build a screening haploid mouse ES cell line prototype and run functional tests. That being said, the starting point of the project was to validate the system's functionality in diploid mouse ES cells. This prototype should have consisted of three genetic components expressed concurrently. These were dCas9-KRAB-MeCP2 transgene, multiple specific guide RNAs (see section 2.2.1.) and the HLP that contains a blue fluorescent protein (BFP) expressing gene. As diploid mouse ES cell line, C2 was used to construct the system. The first step in this experiment involved the simultaneous chemical transfection of the pB_dCas9 vector, containing only dCas9-KRAB-MeCP2 transgene that is promoted with CAG promoter and HLP vector, containing the HLP with the transposon system PiggyBac. Cells, later selected with ES-DMEM 2i containing Hygromycin and Blasticidin and surviving cells were preserved. The next step was to transfect multiple specific gRNA vectors into the cells containing both vectors with electroporation. The cells transfected with gRNA vectors in different ratios, meaning, either only single one or all of them at the same time. Upon transfection of the gRNAs, translated dCas9-KRAB-Mecp2 fusion protein would be recruited by the gRNAs to the heterochromatin dimer of the HLP that has the target sequences of the matching gRNA sequences. Hence, shut down genomic

regions on and near the platform integration sites was expected to occur including the BFP reporter gene. Subsequently, this produced diploid mouse ES cell line prototype could be verified upon checking the expression levels of BFP with flow cytometry.

3.1.2. Results

After transfection of the pB_dCas9 and HLP vectors in the C2 diploid cell line, the prototype with dCas9-KRAB-MeCP2 transgene and the HLP was aimed to be generated. In order to validate this, two strategies were followed. Firstly, the cell cultures were checked under the florescent microscope by using the blue laser aiming to see the expression of the BFP reporter gene within the HLP vector. Two cultures of the same cell line with the same transfection conditions and one negative control (C2 cell line with empty vector transfection) were checked simultaneously. After the verification of the BFP expression of these positive cells (data not shown), the second step was to detect the presence of the fusion protein dCas9-KRAB-MeCP2 within the cells via Western blot. Therefore, whole-cell extracts were developed from the C2 cell culture variants 1 and 2. Protein production profiles of these two variants of the C2 cell line are presented in Figure 6. The expected dCas9-KRAB-MeCP2 fusion protein is 201 kDa (Yea et al. 2018) For the positive control, RC9 mouse ES cell line was used. This cell line contains a non-engineered, novel cas9 gene, which transcribes cas9 protein in the size of 164 kDa. Western blot results showed that in both variants, fusion protein could be detected whereas the negative control confirms the absence of the dCas9-KRAB-MeCP2.

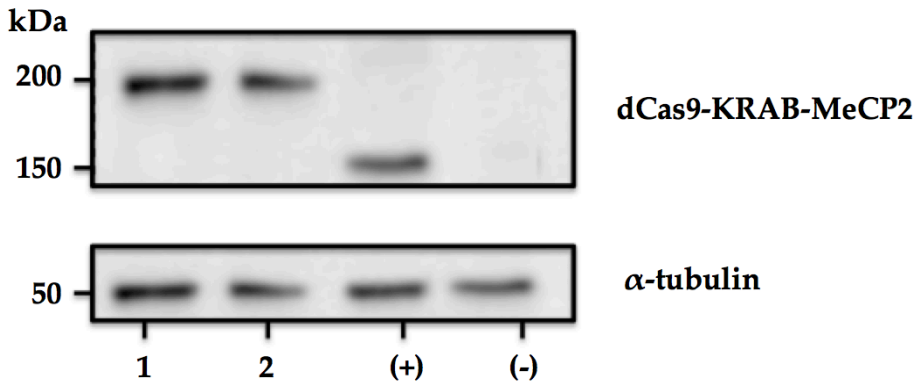


Figure 6: Western blot analysis of fusion protein dCas9-KRAB-MeCP2
Whole-cell extracts prepared from C2 cell line transfected with vector #95 and #87 and these cell lines were analysed for the fusion protein dCas9-KRAB-MeCP2 production via Western blot. Whole-cell extracts prepared from RC9 and wild type C2 (with empty vector transfection) cell lines were used as positive and negative control, respectively. Lanes 1-2: C2 cell line (transfected with vector #95 and #87) culture variants 1 and 2.

Following the confirmation of the presence of two necessary genetic components in the diploid C2 cell line prototype, the next step was to test the system's functionality by transfecting multiple, individual, specific gRNA vectors. For this experiment, only one variant from the diploid prototype was used. After gRNAs were produced in the cells, these gRNAs were expected to recruit the translated dCas9-KRAB-Mecp2 fusion protein onto the heterochromatin dimer of the HLP. Thereby, BFP reporter gene expression was expected to be silenced due to the heterochromatinisation of the gene region. The diploid C2 prototype cell line was transfected with individual gRNA (electroporation with Gene Pulser Xcell™) vectors in two different scenarios (figure 7) In the first scenario, only one of the gRNA vectors, containing only one specific gRNA expression sequence, was introduced to the cells. For the second scenario, all 4 of the individual gRNA vectors were transfected into the cells. As a negative control, one vector with a non-working gRNA sequence (ctrlg1) was used. This gRNA vector was transfected into the prototype several times before, but it never showed a decrease in BFP expression. Therefore, vector was decided to use as the negative control.

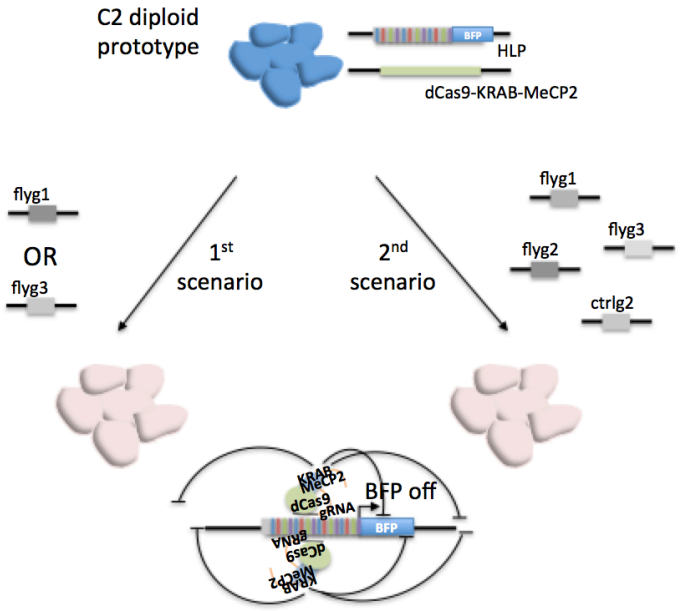


Figure 7: The diploid C2 prototype cell line transfection with individual gRNA vectors
The diploid C2 prototype cell line was transfected with individual gRNA vectors in two different scenarios. In the first scenario, either flyg1 or flyg3 gRNA vector was introduced to the cells. For the second scenario, all 4 of the individual gRNA vectors (flyg1, 2, 3 and ctrlg2) were transfected into the cells.

Upon 48 hours to the transfection of the vectors, cell profiles were monitored with flow cytometry (figure 8 and 8). The comparison of the expression of the BFP reporter gene in wild-type and prototype cells without gRNAs can be seen in figure 8A. As expected in wild-type cells there is almost no blue signal detected while the prototype cells that have the HLP express a high number of blue signals. In both cases, more than %75 of whole cell population expressed blue or did not express blue. The negative control continued to keep the BFP expression after 48 hours upon the transfection with a false gRNA vector (figure 8B).

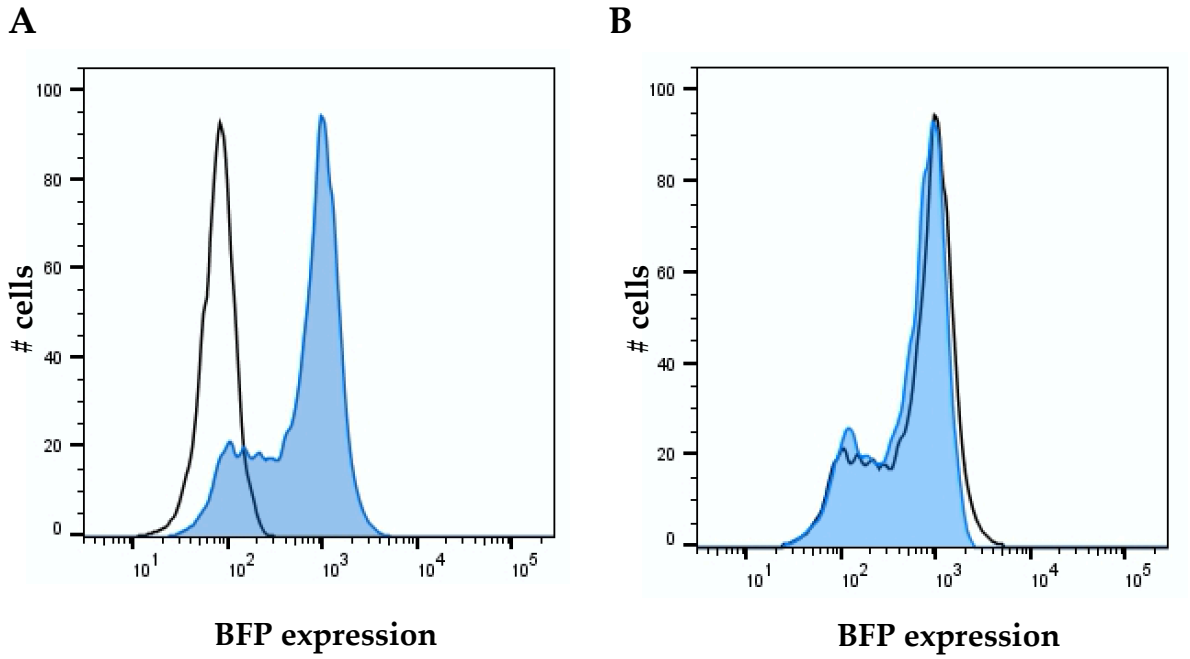


Figure 8: Reduction of BFP reporter gene expression in wild-type and prototype cells with the negative control (ctrlg1) transfection
(A) Flow cytometry analysis of the BFP reporter gene expression in wild-type diploid C2 ES cell line (Blank) and diploid C2 prototype without any gRNA transfection (Blue) and (B) the comparison of BFP reporter gene expression of again the diploid C2 prototype without any gRNA transfection (Blue) and the negative control (ctrlg1) transfection.

Prototype cells transfected with only one different gRNA showed a clear shift on the BFP expression profiles (figure 9A). In both cases, there was a decrease in the number of blue cells slightly higher than a half-fold change when compared to the prototype cells without gRNAs. Cell profiles of the prototype that was transfected with all 4 different gRNA vectors were similar (figure 9B). There was still a decrease in the BFP expression level. However, the level of decline was evidently lower, when compared to the scenario where the expression of one individual gRNA vector was introduced in the prototype.

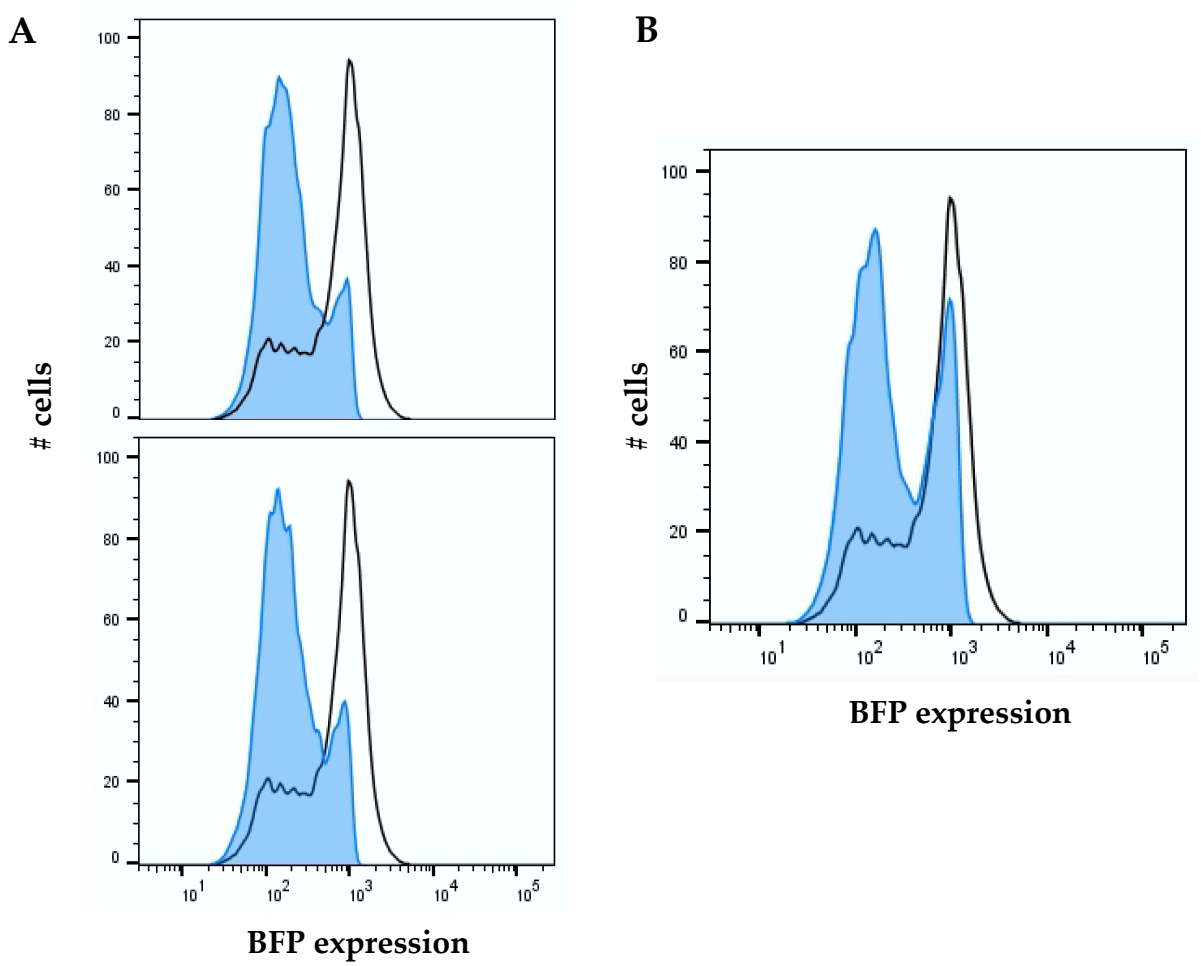


Figure 9: Comparison on BFP expression levels in prototype cells with a single gRNA vector transfection and with 4 different gRNA vector transfection
(A) Flow cytometry analysis of BFP reporter gene expression of the diploid C2 prototype with single gRNA vector (flyg1 or flyg3 gRNA expression sequence) transfection (Blue) and diploid C2 prototype without any gRNA transfection (Blank). (B) Comparison of BFP reporter gene expression of the diploid C2 prototype with 4 different gRNA vector (flyg1, flyg2, flyg3, ctrlg2 gRNA expression sequences) transfection (Blue) and diploid C2 prototype without any gRNA transfection (Blank).

3.1.3. Discussion

The first step of the task, producing the diploid C2 mouse ES cell prototype, which consists of three genetic components (dCas9-KRAB-MeCP2 transgene, multiple specific guide RNAs and the HLP that contains a blue fluorescent protein (BFP) expressing gene), was achieved. The first validation of the system's functionality was done via Western blot. Cell prototypes were shown to obtain and able to produce dCas9-KRAB-MeCP2 fusion protein. Besides, cells were checked under a fluorescent microscope in order to show the BFP reporter gene expression. Later, as the second

validation, via flow cytometry, the system's silencing on BFP expression was questioned. Before this, the prototype was transfected with different gRNA vectors (each of them containing different gRNA expression sequences). In flow cytometry results we could see a clear shift on the BFP expression after the gRNA transfection of the cells. In the first scenario where the prototype was transfected with only one gRNA vector (either flyg1 or flgy3), more than half of the whole cell population expressing blue dropped off to 24%, indicating the silencing of the BFP reporter gene. However, when diploid C2 prototype transfected simultaneously with 4 different gRNA vector (flyg1, flyg2, flyg3, ctrlg2 gRNA expression sequences) silenced cell population was only reduced to 40%. In other words, when more than one individual gRNA expression vector was transfected into the prototype cells, the shutting down efficiency of the system got lower. The reason for this could be due to the poor transfection efficiency of simultaneously introduced gRNA vectors into the prototype, availability of the each individual gRNAs during the recruitment of the dCas9-KRAB-MeCP2 fusion protein to the HLP or some unforeseen factors related with the transcription interference of more than one gRNA genes at the same time. This conclusion makes it obvious that the solutions for the transfection of gRNA vectors should be changed and enhanced. Other than the gRNA efficiency issue, this experiment proved that the working system could be established in the diploid C2 mouse ES cell line. Therefore, the next step was to construct the haploid mouse ES cell prototype after improving the efficiency of concurrent transfection of multiple, specific gRNA vectors.

3.2. Multiple gRNA expression vector – STAgR cloning

3.2.1. Task and experimental setup

After the successful production of the diploid prototype, there was one crucial point to be solved prior to the production of haploid prototypes. Since the ratio of cells expressing a complete set of 4 gRNAs was decreasing with the number of gRNAs

used, it was obvious that combining multiple individual gRNA expression vectors to achieve expression of all 4 in the diploid prototype has its limits. Therefore, the task was to generate a functional multiple gRNA expression vector containing the complete set of 4 gRNA sequences.

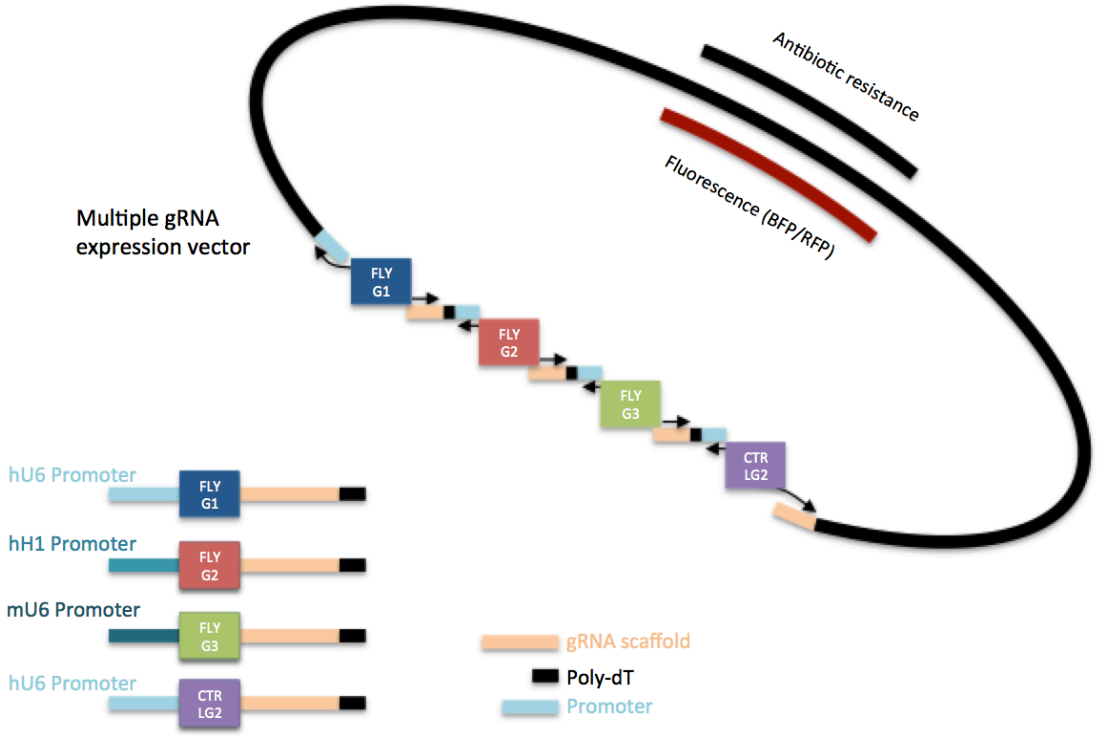


Figure 10: Protocol for multiple gRNA expression vector with 4 individual gRNA cassettes
Overview of the multiple gRNA vector construction procedure, 4 different strings and vectors were used to assemble gRNA expression cassettes with different promoters and gRNA scaffolds.

For the multiple gRNA expression vector, 4 gRNA sequences were used to establish. As they were used in the diploid prototype, these vectors were "Flyg1", "Flyg2", "Flyg3" and "Ctrlg2". After deciding on promoters and gRNA scaffolds were going to be used for each of the gRNA sequences, PCR primers were designed. The primers consisted of the gRNA sequences and overhangs that are complementary with the promoters and gRNA scaffolds. Sense gRNA sequences were joined to the forward primer, which binds to the scaffold part. On the other hand, the antisense gRNA sequences were joined to the reverse primer sequences of the specific promoters (figure 10). The sense gRNA sequence (fig. 10 purple) from the last gRNA-Ctrlg2

and the antisense from the first gRNA-Flyg1 (fig. 10 blue) separately were added to the primers used for the production of multiple gRNA expression vector backbone. In the same way, the rest of the primers were attached with the sense and antisense gRNA sequences, thereby using them as overhangs in the aimed order. For the backbones, two identical vectors except with either BFP or RFP markers were used. In the next step, these primers were used for PCR amplification to produce the Gibson fragments (containing the gRNA sequence, promoter, and scaffold) for assembly of the vectors. Later, Gibson assembly was performed to construct the multiple gRNA expression vectors either with BFP or RFP. These multiple vectors were named "blue_4gRNA_cmv" and "red_4gRNA_cmv". Moreover, due to the expected shutdown of the human cytomegalovirus (CMV) immediate-early enhancer and promoter within ES cells, the promoter in "red_4gRNA_cmv" vector was replaced with a human elongation factor-1 alpha (EF1 alpha) promoter. The resulting vector was named "red_4gRNA".

3.2.2. Result and discussion

For the production of the multiple gRNA expression vectors, short overhang primers provided with individual gRNA sequences were used to amplify the strings containing only the promoters (hU6, hH1, mU6), a transcriptional stop signal (poly dT) and a gRNA hairpin (figure 10). In total, 3 different multiple gRNA expression vectors were cloned (figure 10). In every vector, the numbers and the combination of gRNA expression cassettes are identical whereas the fluorescence reporters and the promoters of them are different. The most efficient vector for the validation of transfection experiments in ES cells was expected to be "red_4gRNA" since the others contain a CMV promoter for the fluorescence reporter gene. Vectors containing more than one gRNA expression sequences nearby essentially rely on an efficient promoter and terminator sequences in order to cope with the transcriptional interference that holds up the performance of individual gRNA sequences (Breunig

et al. 2018). Therefore, for each gRNA sequence, a unique promoter sequence was used (figure 10 and 11).

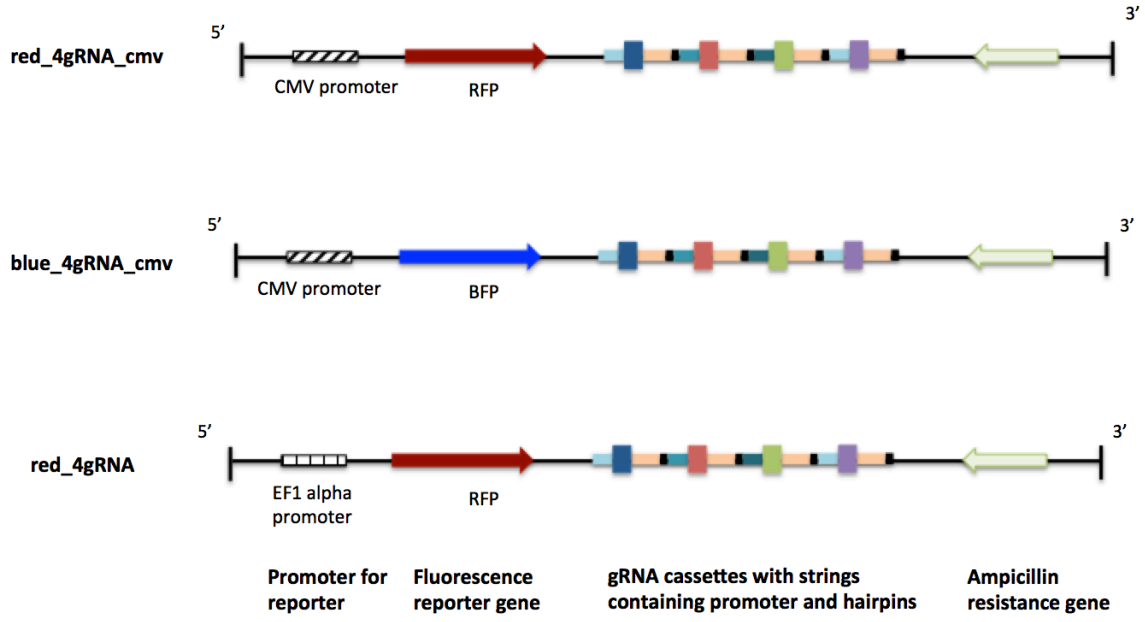


Figure 11: Schematic representation of the genomic components of 3 different multiple gRNA expression vectors
Overview of the multiple gRNA expression vectors. They are consisting of fluorescence reporter genes (either RFP or BFP with different promoters) as validators of transfection experiments, identical set of 4 different gRNA cassette combined together and the Ampicillin resistance gene.

It's shown that the STAgR method provided some advantages that existing cloning strategies don't. For example, any chosen combination or the specific number of gRNA sequences could be cloned into the vector, since every gRNA expression cassette was amplified from the same string. Moreover, it was convenient because individual gRNA sequences provide the required homology for Gibson assembly. Also, with the combination of different strings, different Cas9 or expression systems could be customised (Breunig et al. 2018). By completing this task, the quest for fixing the poor shutting down efficiency of simultaneously introduced gRNA vectors was hopefully enlightened. Therefore, the next step was to construct the haploid mouse ES cell prototype with these multiple gRNA expressing vectors.

3.3. Producing and testing haploid mESC prototype

3.3.1 Establishing and testing the system in haploid H129 and Hoct4 lines

3.3.1.1. Task and experimental setup

After, the successful construction of the multiple gRNA expression vectors, It was expected to have higher transfection efficiency for the further transfections. The next step was to establish and validate the system in haploid mouse ES cells. This haploid prototype was planned to be a preliminary and would light the way for a better, substantial haploid prototype in the next steps. The system was to consist of three genetic components expressed simultaneously like before. Only this time, instead of single gRNA vectors, the prototype was going to be introduced with the multiple gRNA vector that contains the sequences of 4 specific gRNA (flyg1, 2, 3 and ctrlg2). As before, the other components were going to be dCas9-KRAB-MeCP2 transgene and the HLP that contains a blue fluorescent protein (BFP) expressing gene. In addition to this, as haploid mouse ES cell line, H129 and Hoct4 were used to construct the system.

The task started with the chemical transfection of the pB_dCas9 vector, containing only dCas9-KRAB-MeCP2 transgene, which is promoted with CAG promoter and by using the transposon system PiggyBac. Since it's necessary to keep the cells haploid, the following tissue culture work was done carefully (see section 2.4.6.). After 2 days of the transfection, cells were passaged into a 10 cm dish. Both of the cell lines later selected with ES-DMEM 2i containing Blasticidin only for 3 days. Longer selection times were avoided in order not to create a stressful environment for the cells, ultimately to keep as many of them still haploid. Prior to the fluorescence-activated cell sorting (FACS), surviving cells were preserved. During FACS, the Hoechst staining method was used to detect haploid population and sorting was only applied

with the haploid gate. Sorted cells were seeded in different densities into 6cm dishes before the colonies picked up after 7 days. These colonies from each H129 and Hoct4 cell lines were seeded in 96 well plates and later passaged into larger dishes until their ploidy checked with Propidium Iodide (PI) staining. Later, the ones that are still haploid were to check to have dCas9-KRAB-MeCP2 fusion protein via Western Blot.

For the next step, the clones that are still haploid and containing the fusion protein were chosen to continue the task with. In this step, 3 different clones that were chosen from the previous step were used for chemical transfection of HLP vector, containing the HLP by using the transposon system PiggyBac. Until it was achieved to have the clones with the fusion protein and the HLP, the next steps were the same as previously. Only this time cells were selected for 3 days with Hygromycin. Prior to the multiple gRNA expression vector transfection positive clones were frozen for the further steps and for the next stage, 3 of them (H129-dCas9/HLP-1,2 and 3) were chosen to continue with. Prior to the introduction of the vector red_4gRNA_cmv in these prototype clones; vector was linearized in order to advance the chances of productive genomic integration. Also because, there is any available genomic integration model, e.g. transposable elements present in the vector. The transfection of vector red_4gRNA_cmv was done chemically. Since this vector contains a RFP reporter gene, upon transfection of the vector, prototype cells were expected to express red and the blue colours simultaneously for a short time. The blue expression was expected to be reduced by time due to the silencing of the BFP reporter gene. Initial investigation was done with florescent microscopy. Later, to validate these haploid mouse ES cell line prototypes, the expression levels of BFP and RFP expression within the cells were checked via flow cytometry.

3.3.1.2. Results

The experiment started with the transfection of the pB_dCas9 vectors into H129 and Hoct4 haploid cell lines, to generate the parental clones with the dCas9-KRAB-

MeCP2 transgene. As a result, 5 clones seemed to be successfully produced. In order to validate this, one strategy was followed. It was detecting the presence of the fusion protein dCas9-KRAB-MeCP2 within the clones via Western blot. Therefore, whole-cell extracts were developed from these 5 different clones and proteins were extracted. For detecting the expression profiles of Cas9 protein the same antibodies were used. The profiles of 5 parental clones are presented in figure 12. The expected dCas9-KRAB-MeCP2 fusion protein is 201 kDa. For the positive control, RC9 mESC line was used. This cell line contains a non-engineered, novel cas9 gene, which transcribes cas9 protein in the size of 164 kDa (Yeo et al. 2018). As a negative control, the wild type of H129 cell line was used. Image quantification software results showed that in 3 clones (2 from H129 cell line, 1 from Hoct4), fusion protein could be detected whereas for the other 2 clones could not (figure 12). Also, both positive and negative controls confirm the validity of the procedure.

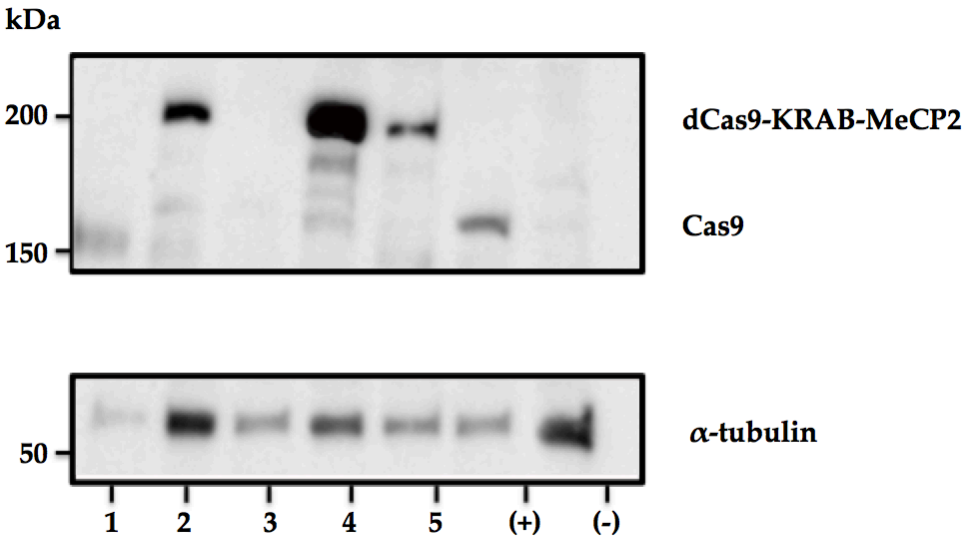


Figure 12: Western blot analysis of fusion protein dCas9-KRAB-MeCP2 in 5 different haploid clones

Protein samples from 5 different parental clones of H129 and Hoct4 mESC lines (transfected with vector pB_dCas9 vector) were analysed for the fusion protein dCas9-KRAB-MeCP2 production via Western blot. Protein extracts prepared from RC9 and wild type H129 (with empty vector transfection) cell lines were used as positive and negative control, respectively. Lanes 1-5 in order: the clones H129-dCas9/1, 2, 3, 4 and Hoct4-dCas9/1 clone E7.

After confirming the presence of one of the genetic components in 3 clones from haploid H129 and Hoct4 cell lines, the next step was to further establish the system by producing new subclones with the HLP and later test the system's functionality by transfecting the multiple gRNA expression vector "Red_4gRNA_cmv". Firstly, for making the subclones H129-dCas9/1 (from now on referred as H129-dCas9) parental clone was used. In total three prototypes with the fusion gene and the HLP were produced. These were subclones of H129-dCas9/HLP-1, 2, 3. Upon 48 to 72 hours to production of these subclones, the comparison on cell profiles of the wild type H129 and the subclone H129-dCas9/HLP-1 was monitored with flow cytometry. The cell profile of wild type H129 cell line on the expression of the BFP and RFP reporter gene can be seen in figure 13A. As expected in wild type cells there were neither blue nor red signal detected. While at the same time, the clone H129-dCas9/HLP-1 contains the fusion gene and the HLP was expected to express the BFP but not RFP and this was the case when the cell profile of the clone was observed (figure 13B).

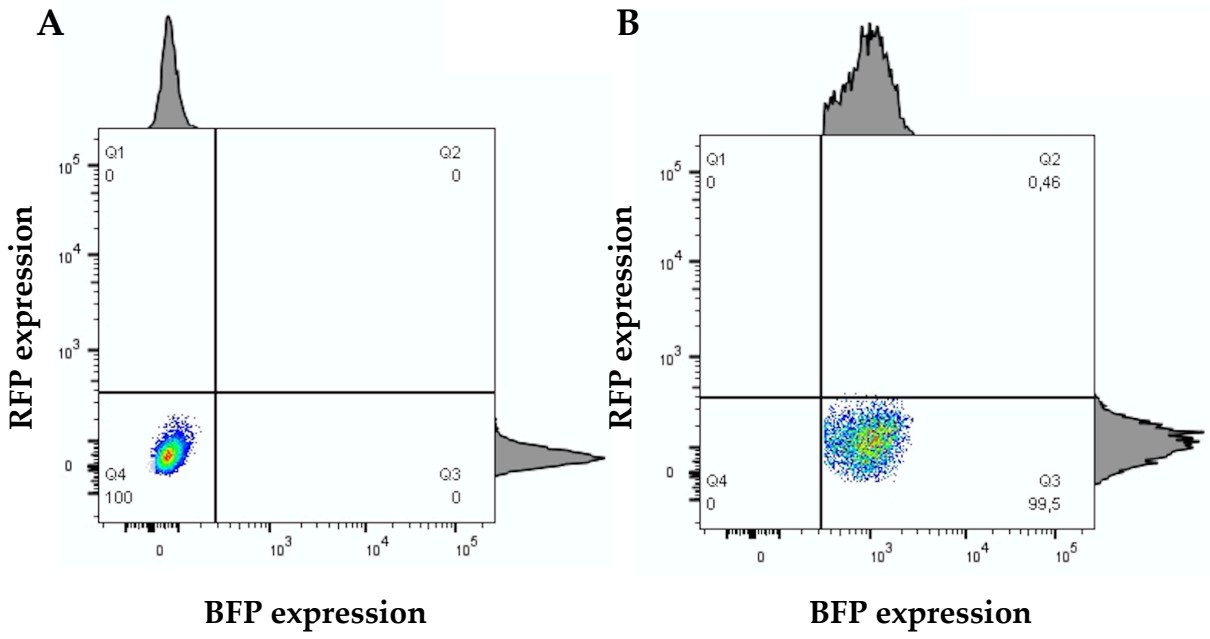


Figure 13: Comparison of the BFP and RFP expression levels on cell profiles of the wild type H129 and the clone H129-dCas9/HLP-1

(A) Flow cytometry analysis of wild type H129 on BFP and RFP reporter gene expression. The population is in Q4 where there is neither BFP nor RFP expression could be detected. (B) Same analysis for the clone containing the HLP (H129-dCas9/HLP-1). This time, the cell population belongs to Q3 where BFP expression is present with no RFP signal.

Later, to complete the required genetic components of the prototype and test the system, red_4gRNA_cmv was transfected to these subclones. Thereby, 3 different haploid prototypes were produced. After the transfection of "Red_4gRNA_cmv" vector, prototype cells were expected to express red and blue colours simultaneously for a short time because the gRNA vector contains an RFP reporter gene. At the molecular level, following the transcription of gRNAs, they were expected to recruit the translated dCas9-KRAB-Mecp2 fusion protein onto the heterochromatin dimer of the HLP. Therefore, while the RFP reporter gene was expressed, BFP expression had to be silenced due to the heterochromatinisation of the gene region. Upon 24 hours to the transfection of the gRNA vector, cell profiles were monitored with flow cytometry (figure 14). Observations for the subclone H129-dCas9/HLP-1 showed that almost 85% of the cell population don't express the BFP gene whereas before the gRNA vector transfection the whole population was blue. The shift has seen in the direction from Q2 to mostly Q1 and Q4 (figure 14A). In this case, the potential haploid prototype would be selected from Q1 (still RFP expressing but not BFP), 33% of the whole cell population. For the subclones H129-dCas9/HLP-2 and 3, the analysis showed that, in both cases, more than 68% of the whole population didn't express the BFP reporter gene anymore while approximately 40% of this population (in Q1) kept the red signal (figure 14B and 14C). This meant that, in both cases, 40% of the whole population could be used as potential haploid prototypes. However, for these subclones, there wasn't a clear cell aggregation in Q1, which would indicate the correct cell population (figure 14B and 14C). Overall, for all the subclones, test results were promising due to the clear shifts from BFP to RFP expression. On average, 70% of the whole cell population for each subclone managed to suppress the BFP expression while approximately 40% of this population sustained the RFP expression.

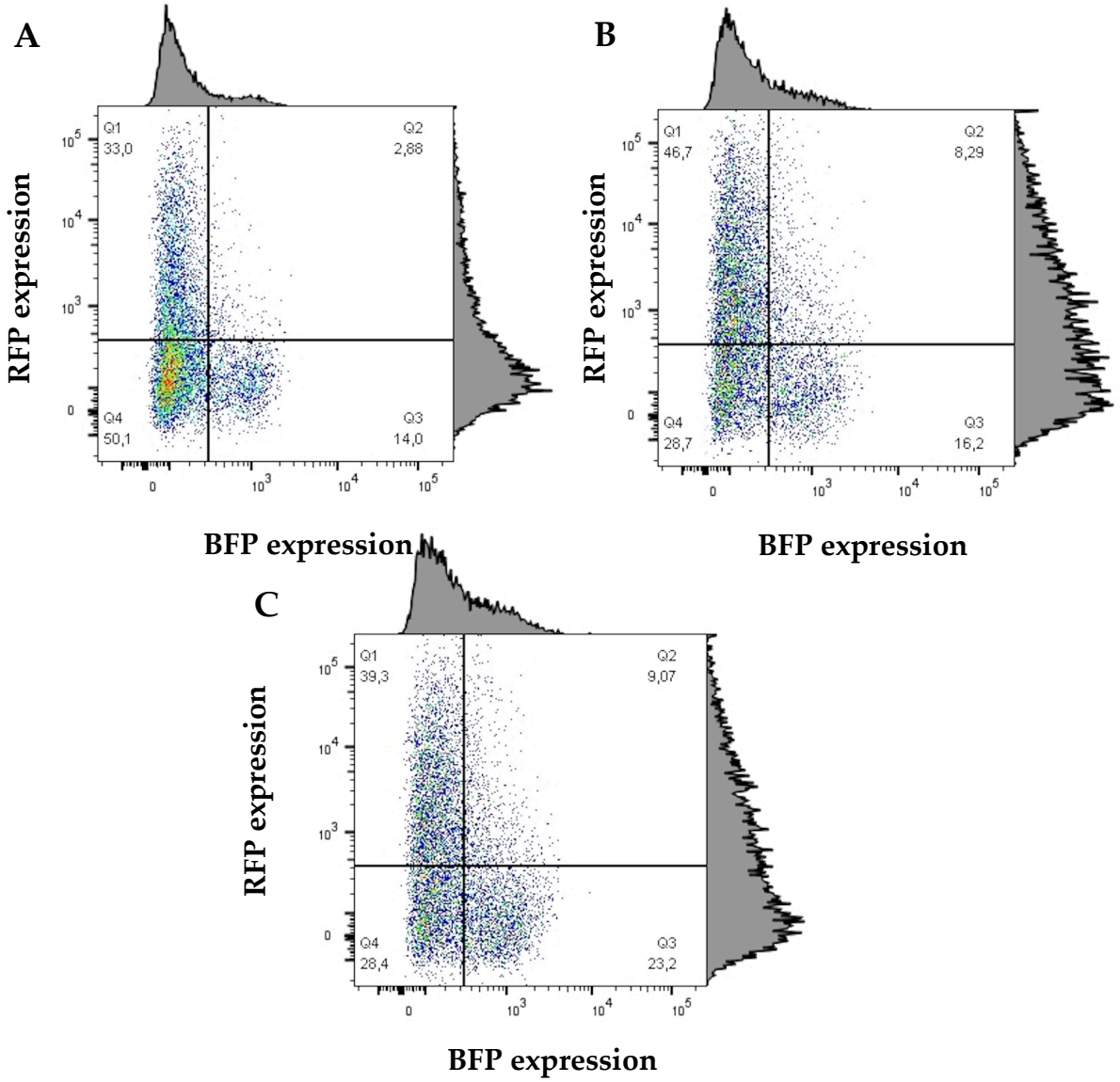


Figure 14: Comparison of the BFP and RFP expression levels on cell profiles of the subclones H129-dCas9/HLP 1, 2 and 3 upon 24 hours to the transfection of "Red_4gRNA_cmv" vector
(A) Flow cytometry analysis of the subclone H129-dCas9/HLP on BFP and RFP reporter gene expression upon 24 hours to the transfection of "Red_4gRNA_cmv" vector. The clear shift was detected from Q2 to Q1 and Q4 where is no BFP expression could be detected. (B) Same analysis for the subclone H129-dCas9/HLP-2. This time, the cell population belongs to Q1 is more populated more than Q4 where neither BFP nor RFP expression is present. (C) Finally, the same analysis for the subclone H129-dCas9/HLP-3. The cell profiles look similar like in (B), but this time with more population in Q3, the cells with the blue signal.

3.3.1.2. Discussion

With this step of the project, producing the working haploid mouse ESC prototypes those consist of three genetic components (dCas9-KRAB-MeCP2 transgene, the HLP that contains a blue fluorescent protein (BFP) expressing gene and gRNA vector

"red_4gRNA_cmv") was achieved. Firstly, via Western blot, parental clones of H129 and Hoct4 validated to obtain and able to produce dCas9-KRAB-MeCP2 fusion protein. Later, the subclones of H129-dCas9 (HLP-1, 2 and 3) containing the HLP were produced. In the next step, in order to establish prototypes and test them, these subclones were transfected with multiple gRNA expression vector "red_4gRNA_cmv". Afterwards, for testing the prototypes via flow cytometry, the system's silencing on BFP expression while sustaining the RFP expression was questioned. Results indicate that, upon 24 hours to the introduction of gRNA vector into the system, for all three subclones, 70 to 85% of the whole cell population were able to shut down the BFP expression indicating the heterochromatinisation of the HLP and possible sites nearby. However, when we observe the red signal from the same subclones the average was only 40%. If it was the case, what happened to the rest of 30 to 35% of the cells with no blue signal? They were probably still transcribing the gRNAs meaning that the gRNA vector is integrated into the genome and working so, the BFP gene was getting silenced. The high possible answer to this; the expected shutdown of the human cytomegalovirus (CMV) immediate-early enhancer and promoter within mouse ESC, the promoter for the RFP gene in "red_4gRNA_cmv" vector. In other words, 70 to 85% of the subclones were able to process the system successfully, but we could only be sure of 40% among them. Therefore, for monitoring a stable RFP expression we could not use this vector. Instead, it was planned to use the vector ("red_4gRNA"), which was replaced with a human elongation factor-1 alpha (EF1 alpha) promoter.

Overall, the results were promising. From these 3 subclones of H129-dCas9/HLP-1, 2 and 3 the majority of the cell populations showed the functionality of the system. Therefore, these subclones could be used as the working haploid mouse ESC prototypes and should be tested further on the system's silencing and heterochromatinisation mechanisms. For this, firstly, these subclones were going to be transfected with "red_4gRNA" vector and with new prototypes, further tests were going to be conducted.

3.3.2. Locating the integration sites of the landing platform

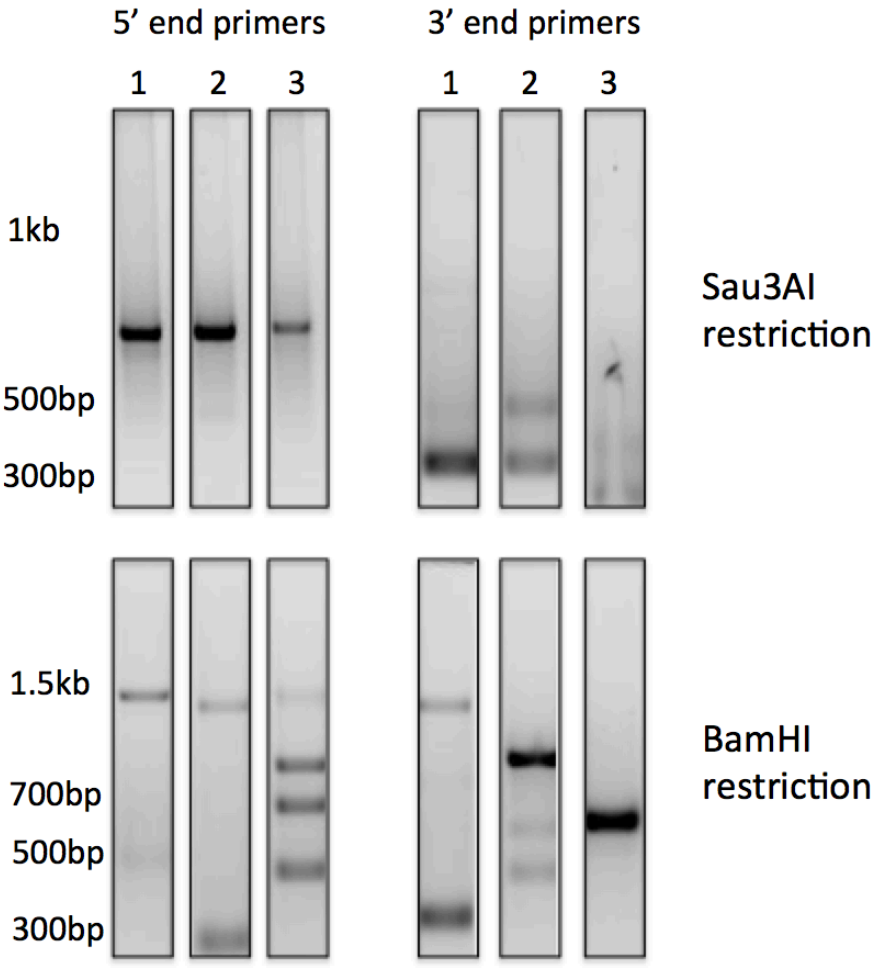
3.3.2.1. Task and experimental setup

Working haploid prototypes showed that the system was able to function for all the subclones. The next step, as explained above, had to be on making prototypes more stable. Therefore, subclones (w/o gRNA vector) needed to be transfected with the "red_4gRNA" vector in order to secure the RFP signal within the cells. Upon 24 and 48 hours to the transfection, the subclones H129-dCas9/HLP-1, 2 and 3 were validated under the florescent microscope by looking at the BFP and RFP signals. Once more, after being positive about the prototypes, eventually, we wanted to locate the integration sites of the silencing platform HLP. The method "Splinkerette PCR" was used to locate the integration sites of the silencing platform HLP within the genome of some of the haploid prototype clones that were produced. For detailed information about the method see section 2.6.3. and introduction. After collecting DNA samples and ligating adaptor sequences, the H129-dCas9 parental clone and the subclones H129-dCas9/HLP-1 and 3 were monitored several times. In total, two Splinkerette PCR experiments were performed. For each experiment, different restriction enzymes ("Sau3AI" and "BamHI") with according adaptor sequences were used. Besides, to wider the interrogation, both 3′ and 5′ end PiggyBac primers were used. After two rounds of PCR for each experiment, as the first step of monitoring the results, the samples were run on 2% Agarose gel. From the gel results, the bands those had the same size for the parental clone and the subclones stated the dCas9-KRAB-MeCP2 transgene integration. At the same time, the bands with different sizes for the subclones indicated the HLP integrations within the genome of the prototypes. The second step on monitoring results was sending all these different bands from all clones to Sanger sequencing. The primers for sequencing were chosen among the adaptor and PiggyBac sequences. Later, all these sequencing results were analysed with genome browser software from UCSC and each band on the Agarose gel from each clone was associated with a locus within the

genome of these prototypes. Thereby, including the dCas9 transgene, the integration sites of the HLP were detected for each working haploid prototype.

3.3.2.2. Results and discussion

DNAs from all the clones were restricted either with Sau3AI or BamHI and ligated with the corresponding adaptor sequences. These DNA/adaptor sequences were used as the template DNA for the first reaction of Splinkerette PCR. For each PCR reaction, either 5′ end or 3′ end PiggyBac integration primers were used. Relative Agarose gel results are represented in figure 15. Each band from each clone indicated an integration site of the HLP or the transgene dCas9-KRAB-MeCP2. When the parental clone H129-dCas9 had bands with the subclones on the same size (figure 15A and 15B), these bands indicted only the transgene integration.



Legend on the next page →

Figure 15: Relative Agarose gel representations of both Splinkerette PCR results
DNAs from all clones either restricted with Sau3AI or BamHI then ligated with the corresponding adaptor sequences. These DNA-adaptor sequences used as DNA template for the first reaction of Splinkerette PCR. For each PCR reaction either 5′ end or 3′ end PiggyBac integration primers were used. Each band from each clone represents an integration site of the HLP or the transgene dCas9-KRAB-MeCP2. The bands with the same sizes from each clone including the parental clone H129-dCas9 stands for only the transgene integration. The rest of the bands expected to represent the HLP integration sites. 1 stands for the clone H129-dCas9, 2 for H129-dCas9/HLP-1 and 3 for H129-dCas9/HLP-3.

The reason why is, the only PiggyBac integration in common for both parental and the subclones could be the dCas9-KRAB-MeCP2. Furthermore, the rest of the bands from the subclones were expected to represent possible HLP integration sites (figure 15A and 15B). The initial observations were done. Then all bands were cut out from the gel in order to purify the DNA fragments. Later on, all these bands were sent to Sanger sequencing. The sequencing results were analysed in order to validate if they harbour the PiggyBac repeat sequences according to the 3' or 5' end repeats. After confirming, they were further checked with the UCSC Genome Browser.

In both Splinkerette PCR results, it was observed that all the clones had bands that are the same size (figure 15) and when checked with the genome browser, the hits are in the same loci for all the clones. The parental clone H129-dCas9 and the subclones H129-dCas9/HLP-1 and 3 had the same hits on chromosome 14. Since all the bands had the same size and all the hits are in the same loci for these clones, it was decided that this indicated the dCas9-KRAB-MeCP2 transgene integration. Other results showed that each subclone had unique hits other than the chromosome 14 hit which stands for the transgene integration. These hits were indicating the silencing platform HLP integrations. After analysing them with the genome browser, we came up with the table 2 below. In this table, integration sites of the HLP of the working haploid prototypes H129-dCas9/HLP-1 and 3 are shown. Furthermore, with the help of the genome browser, the regions nearby the integration sites of the HLP were detected and if applicable, closer genes were revealed. The goal was accomplished. Since the integration sites were located, the broad monitoring was going to be done in the next step in order to observe the silencing range of the HLP

and effect on the genomic regions nearby. Thus, we would answer the question of how far the silencing goes from the landing platform.

Clones/Proto-types	Integration sites	Close regions to the integration site	
		Downstream	Upstream
H129-dCas9/HLP-1	Chr. 15 (EFR3A gene)	-	OC90 gene-50kb
	Chr. 3 (BMPR1B gene)	UNC5C gene-100kb	-
H129-dCas9/HLP-3	Chr. 5 (Non-coding region)	RPL21 gene-50kb	GTF31 gene-50kb
	Chr. 13 (PRSS16 gene)	ZFP184 gene-50kb	-

Table 2: Expected integration sites of the heterochromatinisation landing platform within the genome of the working haploid prototypes
The integration sites of the HLP and the nearby regions from the integration sites are shown in the table for each working haploid prototype.

3.3.3. Heterochromatinisation of the landing platform and the sides nearby

3.3.3.1. Task and experimental setup

We got closer to the step of having the haploid mESC prototypes that would be using the silencing platform HLP, which launches epigenetic modifications with the help of the genetic components of the repressor system. After finding out the integration sites of the HLP for working haploid prototypes H129-dCas9/HLP-1 and 3 we were ready to analyse the silencing range of the HLP and effect on the genomic regions nearby. Therefore, our plan was to observe the heterochromatinisation of the integration sites and the sites nearby by looking at the H3K9me3 epigenetic modification. Furthermore, we also wanted to analyse the expression patterns of the genes on and near the HLP. Thus, we would answer the question of how far the silencing goes from the landing platform. Therefore, our strategy was first to perform quantitative polymerase chain reaction (qPCR) in order to see the expression levels of the genes on and near the integration sites and later chromatin immunoprecipitation (ChIP) to investigate H3K9me3 methylation range again on and near the integration sites. To help with the experimental planning, we constructed maps (figure 16). In

each of these maps, there is the schematic representation of the integration sites of the HLP and the genes nearby from both haploid prototypes.

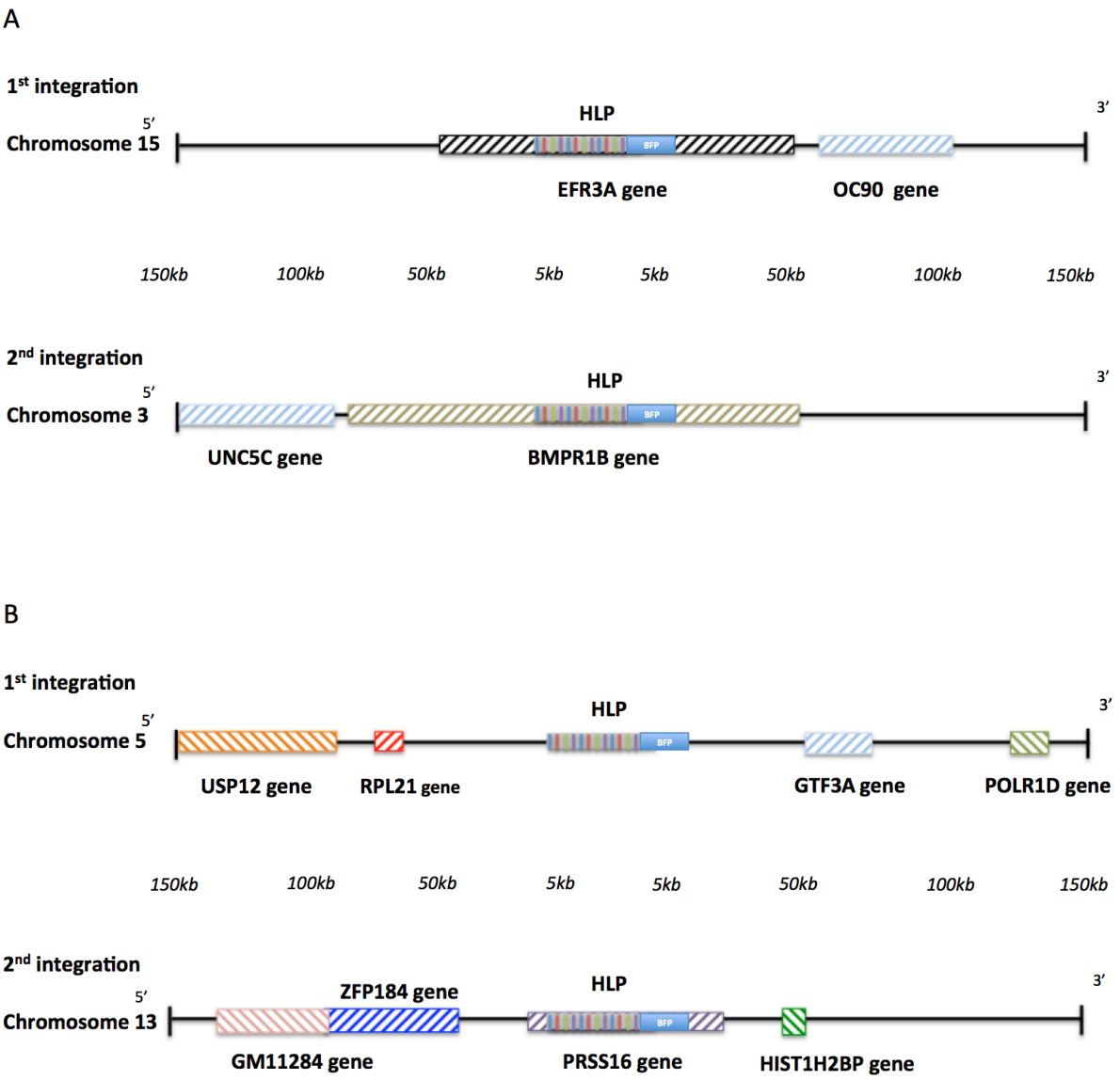


Figure 16: Schematic representation of the integration sites of the HLP and the genes nearby from working haploid prototypes H129-dCas9/HLP-1 and 3

The expected integration sites of the HLP and the nearby genes from the integration sites are shown in the table for each working haploid prototype. Blue and the green box within the HLPs stand for BFP reporter gene and Hygromycin resistance gene. (A) For H129-dCas9/HLP-1 prototype 1st expected integration of the HLP was within the EFR3A gene located in the chromosome 3. 2nd found integration was located in BMPR1B gene, chr. 15. (B) For H129-dCas9/HLP-3 prototype 1st expected integration of the HLP was located in a non-coding region of the chromosome 3. 2nd found integration was within PRSS16 gene located in the chr. 13.

For qPCR experiments, at first, we wanted to analyse the genes on the landing platform where the silencing is expected the most and the genes in the range of maximum 100 kb. The genes we analysed within the HLP were the BFP reporter gene and the antibiotic gene Hygromycin. For the genes nearby, we chose to analyse the ones already with relatively high expression. All these genes for each case are stated in figure 16. Also for prototype H129-dCas9/HLP-3 we further investigated the expression levels of the genes approximately 150-200 kb away from the HLP integration sites (figure 16).

The next step was the investigation of the H3K9me3 modification in order to answer the question of how far the heterochromatinisation goes from the silencing platform. For this, on each prototype, we performed qPCR following the chromatin immunoprecipitation (ChIP). The goal was to analyse the levels of the aforementioned epigenetic modification in a logical range within the genome, in order to be able to hypothesise the power of the repressor system so, the prototypes. Once more, at first, we wanted to analyse the H3K9me3 modification on the silencing platform where the heterochromatinisation is expected the most and the regions in the range of between 1 to 100 kb from the HLP. For this, we used the same qPCR primers (BFP and Hygromycin genes) for the HLP and also we designed one set upstream (1-100kb) and one set downstream (1-100kb) of the integration site of each case for both haploid prototypes. Some of these primers were designed from the coding regions of the genes that are shown to be in nearby locations (figure 16). After we covered the region up to 100kb from the integration sites for each case of the prototypes, only for the cases of H129-dCas9/HLP-3 we further wanted to investigate the epigenetic modification in the range up to 200 kb.

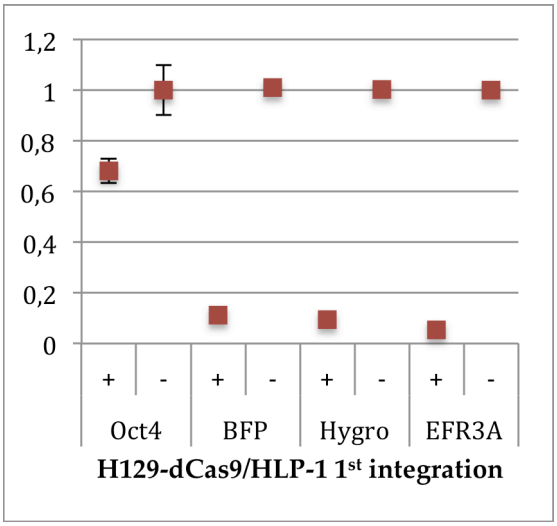
3.3.3.2. Results and discussion

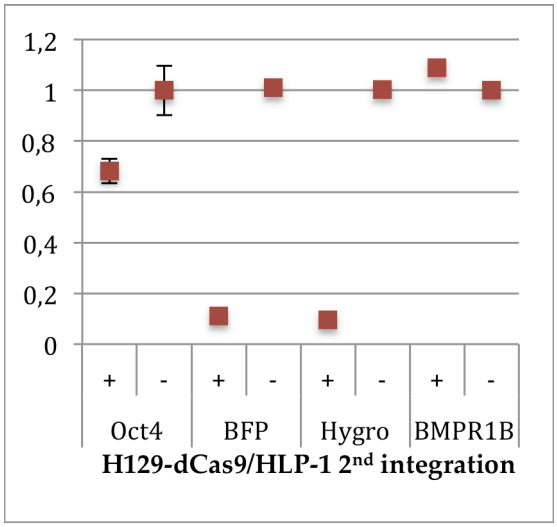
For both qPCR experiments (expression and ChIP) the prototypes were observed with their versions that contain no multiple gRNA expression vector. Therefore, we

were able to compare the expression levels of the questioned genes and the levels of H3K9me3 modifications with and without the working repressor system.

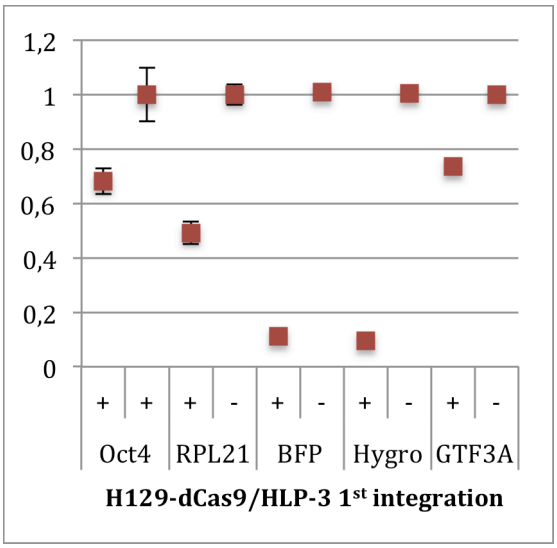
At very first, for gene expression qPCR experiments, we wanted to analyse the genes on the HLP where the silencing is expected the most and the genes in the range of maximum 100 kb (figure 17). The genes within the HLP were BFP and Hygromycin and the expression levels of these genes were observed for every case. Besides, the expression levels of the genes that are integrated by the HLPs like ERF3A, BMPR1B (H129-dCas9/HLP-1) and PRSS16 (H129-dCas9/HLP-3) were analysed alongside with the genes nearby to these integration sites like RPL21, GTF3A and so on (figure 16). All samples were normalised to the housekeeping Actb (Beta-actin) gene. Oct4 gene was used as a positive control since it was active as a pluripotency control.

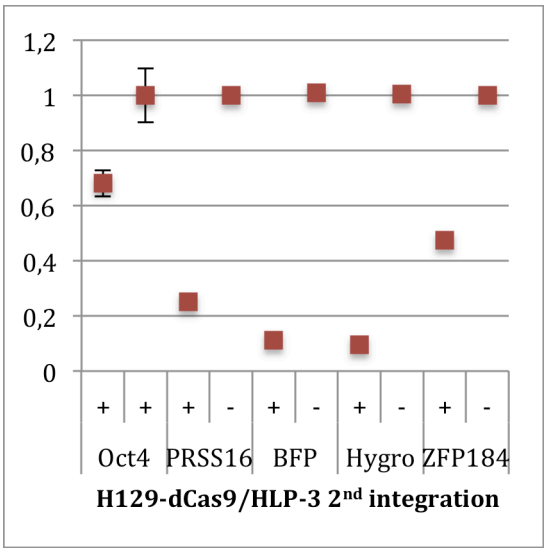
A





B



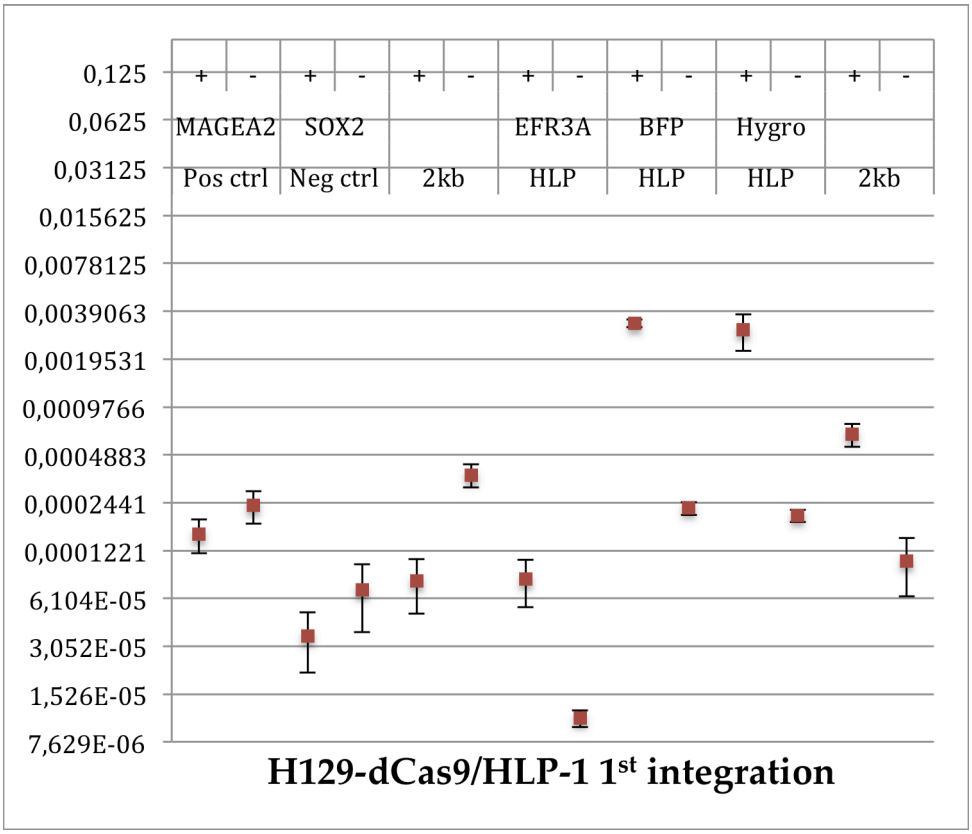


Legend on the next page →

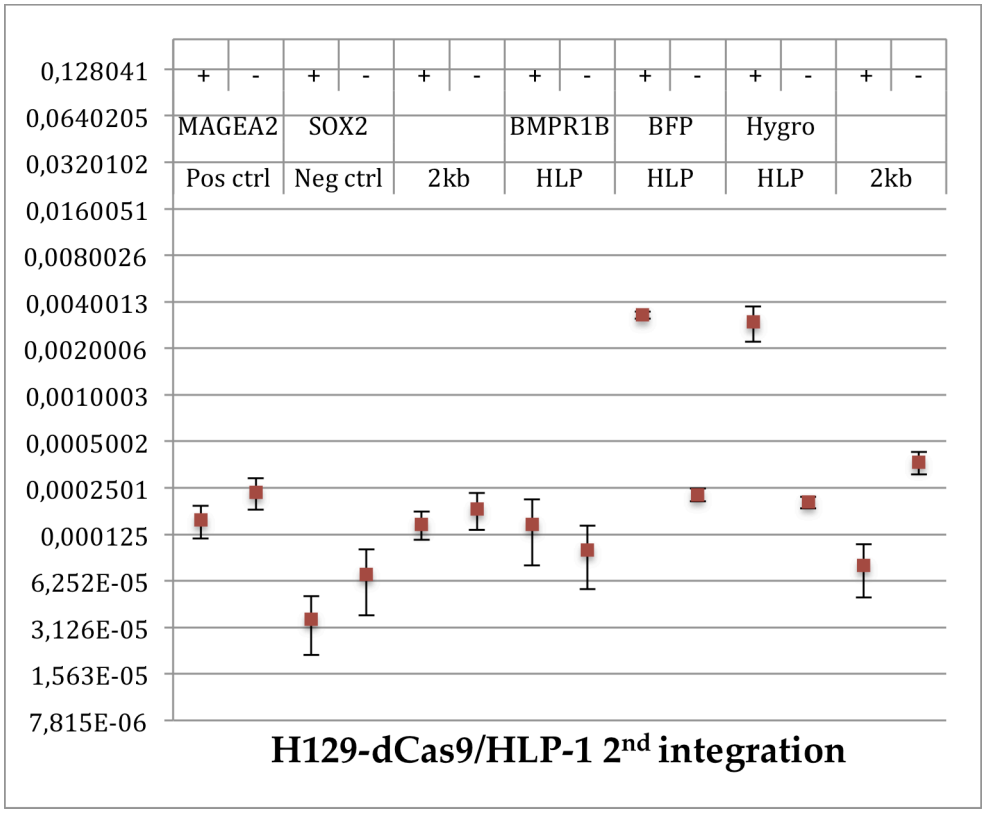
Figure 17: Expression levels of the genes on and nearby the integration sites of the HLPs from both prototypes H129-dCas9/HLP-1 and 3 analysed with quantitative polymerase chain reaction (qPCR)
Analysing the genes on the HLP where the silencing is expected the most and the genes in the range of maximum 100 kb. The genes within the HLP were BFP and Hygromycin and the expression levels of these genes were observed for every case. All samples were normalised to the housekeeping Actb (Beta-actin) gene. Normalisation was done by having the square of the difference between quantitation cycle (Cq) levels of the samples and the housekeeping gene. Error bars were calculated by observing the standard deviation levels of both the sample and the housekeeping gene. Oct4 gene was used as a positive control since it was active as a pluripotency control. (+) samples stand for the prototypes with gRNAs and (–) samples stand for the prototypes with gRNAs. Normalised results of the (-) samples then fixed to 1 in order to make a better comparison between (+) and (-) samples. (A) Results for the prototype H129-dCas9/HLP-1 and (B) for the prototype H129-dCas9/HLP-3.

Gene expression qPCR results showed that for the OCT4 control there was a slight decrease when compared to the control samples (figure 17). Therefore, we interpreted the results considering this fact. In both prototypes, as expected, for BFP and Hygromycin genes, a 1-fold decrease is observed when compared to the control samples (figure 17). This suggested that we were able to observe the silencing caused by the heterochromatinisation on the platform. In prototype H129-dCas9/HLP-3, for both integration sites, the nearby genes (RPL21, GTF3A, PRSS16 and ZPF184) could be suppressed, because they showed less gene expression levels compared to their control samples (figure 17B). On the other hand, in prototype H129-dCas9/HLP-1, for nearby gene BMPR1B, we could not observe a difference in the expression levels whereas EFR3A gene, which the HLP was integrated into, had a 1-fold decrease (figure 17A). Overall, these results indicate that the repressor system works successfully on the HLP for both prototypes and silencing seemed to spread the nearby genes. This gave us the confidence to expect the silencing of the system in further regions. After all, one of the tasks of this project is to understand the system's silencing capacity. In order to question the range of the heterochromatinisation that spreads from the silencing platform, the next step was the investigation of the H3K9me3 modification. For this, on each prototype, we performed qPCR following the ChIP. Therefore we analysed the levels of the epigenetic modification in a logical range within the results didn't show substantial modification levels other than on the silencing platform.

A



B



Legend on the next page →

Figure 18: H3K9me3 modification levels of the regions up to 5 kb from the integration sites of the HLPs from the prototype H129-dCas9/HLP-1 analysed with quantitative polymerase chain reaction (qPCR) following the chromatin immunoprecipitation (ChIP)
For the working haploid prototype H129-dCas9/HLP-1, analysing the regions on the HLP where the silencing is expected the most and the regions nearby in the range of maximum 5 kb. All samples were normalised to their input samples (no detection of the H3K9me3 modification). Normalisation was done by having the square of the difference between quantitation cycle (Cq) levels of the samples and the housekeeping gene (The volume differences between ChIP and input samples were taken into the consideration). Error bars were calculated by observing the standard deviation levels of both the ChIP and the input samples. MAGEA2 gene was used as a positive control. SOX2 gene was used as a negative control. (+) samples stand for the prototypes with gRNAs and (–) samples stand for the prototypes with gRNAs. Results showed in logarithmic scale with the base 2. In (A) 1st integration case and in (B) 2nd integration case of the prototype H129-dCas9/HLP-1are shown.

Compared to the control samples, in the HLP region, we observed high levels of H3K9me3 modification. Also for the 1st integration case of the H129-dCas9/HLP-1 prototype, in the region 2 kb away, a slight difference was observed (figure 18A). Other than these results, nothing remarkable for the nearby regions was analysed. Consequently, these results indicate that the repressor system works successfully on the HLP for the prototype H129-dCas9/HLP-1. However, the silencing seemed to remain limited only with the landing platform. This is what we observed from the levels of H3K9me3 modification. Since the modification levels didn't increase with the range, we decided not to observe further regions for this prototype.

For the prototype H129-dCas9/HLP-3, since we observed a promising silencing range on the qPCR experiment for gene expression, we decided to question the H3K9me3 levels for this prototype in a range up to 150 kb from the integration sites. In addition to the regions where the nearby genes located, we also analysed the regions exactly 100 kb and 150 kb far away for both of the cases of the prototype H129-dCas9/HLP-3 (figure 19). Firstly, for both cases, the results showed substantial modification on the silencing platform. For the 1st integration case, we still observed high levels of the modification 2 kb far away for both up- and downstream when compared to the control samples. Besides, in the region of RPL21 gene (85kb upstream), we analysed a slight increase in the modification level (figure 19A). The second HLP integration case, the situation looked even better. We were able to observe high levels of the

H3K9me3 modification still in the regions 150 and 100 kb upstream from the integration site.

A

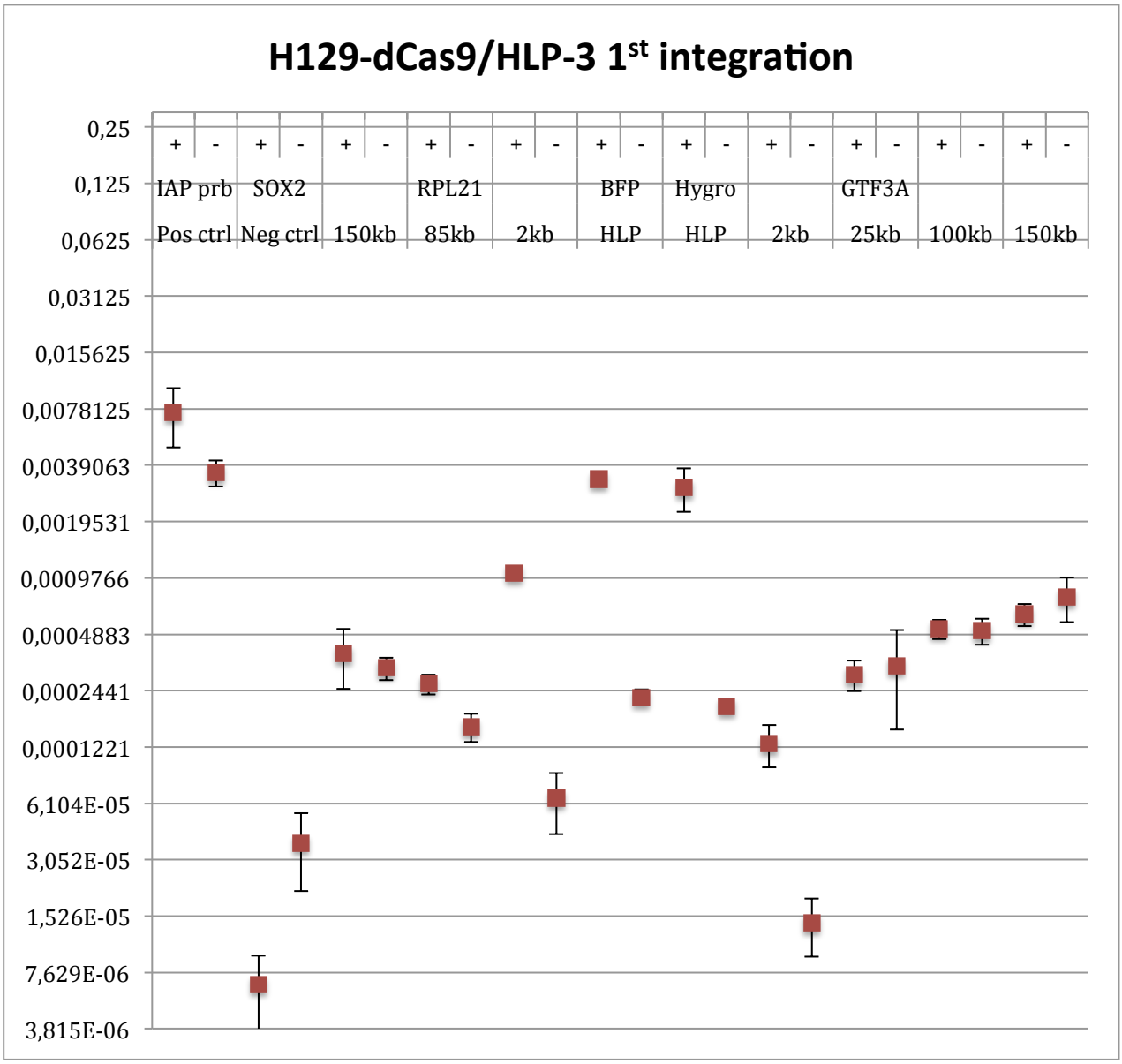


Figure continues on the next page →

B

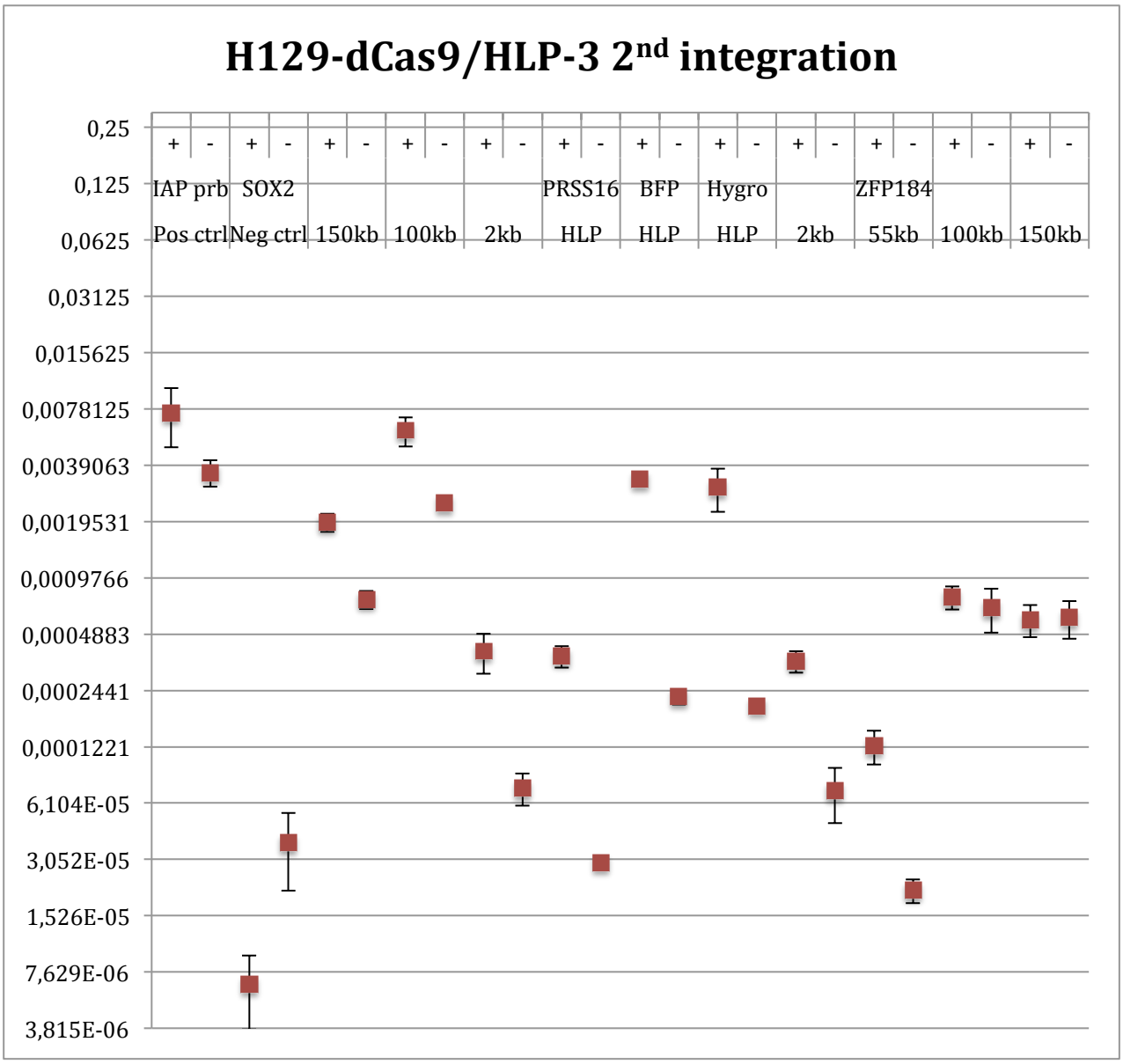


Figure 19: H3K9me3 modification levels of the regions up to 150 kb from the integration sites of the HLPs from the prototype H129-dCas9/HLP-3 analysed with quantitative polymerase chain reaction (qPCR) following the chromatin immunoprecipitation (ChIP)

For the working haploid prototype H129-dCas9/HLP-3, analysing the regions on the HLP where the silencing is expected the most and the regions up to the range of maximum 150 kb. All samples were normalised to their input samples (no detection of the H3K9me3 modification). Normalisation was done by having the square of the difference between quantitation cycle (Cq) levels of the samples and the housekeeping gene (The volume differences between ChIP and input samples were taken into the consideration). Error bars were calculated by observing the standard deviation levels of both the ChIP and the input samples. IAP gene was used as a positive control. SOX2 gene was used as a negative control. (+) samples stand for the prototypes with gRNAs and (–) samples stand for the prototypes with gRNAs. Results showed in logarithmic scale with the base 2. In (A) 1st integration case and in (B) 2nd integration case of the prototype H129-dCas9/HLP-3 are shown.

On the other hand, we were still able to observe the modification up to 55 kb downstream and from that point we couldn't observe any difference on the H3K9me3 levels (figure 19B). Therefore, concerning understanding the system's silencing capacity, the results from the prototype H129-dCas9/HLP-3 were critical and lighted the way. In figure 20, there is a schematic representation of each HLP integration of the prototype H129-dCas9/HLP-3 with the H3K9me3 levels as well as gene expression profiles.

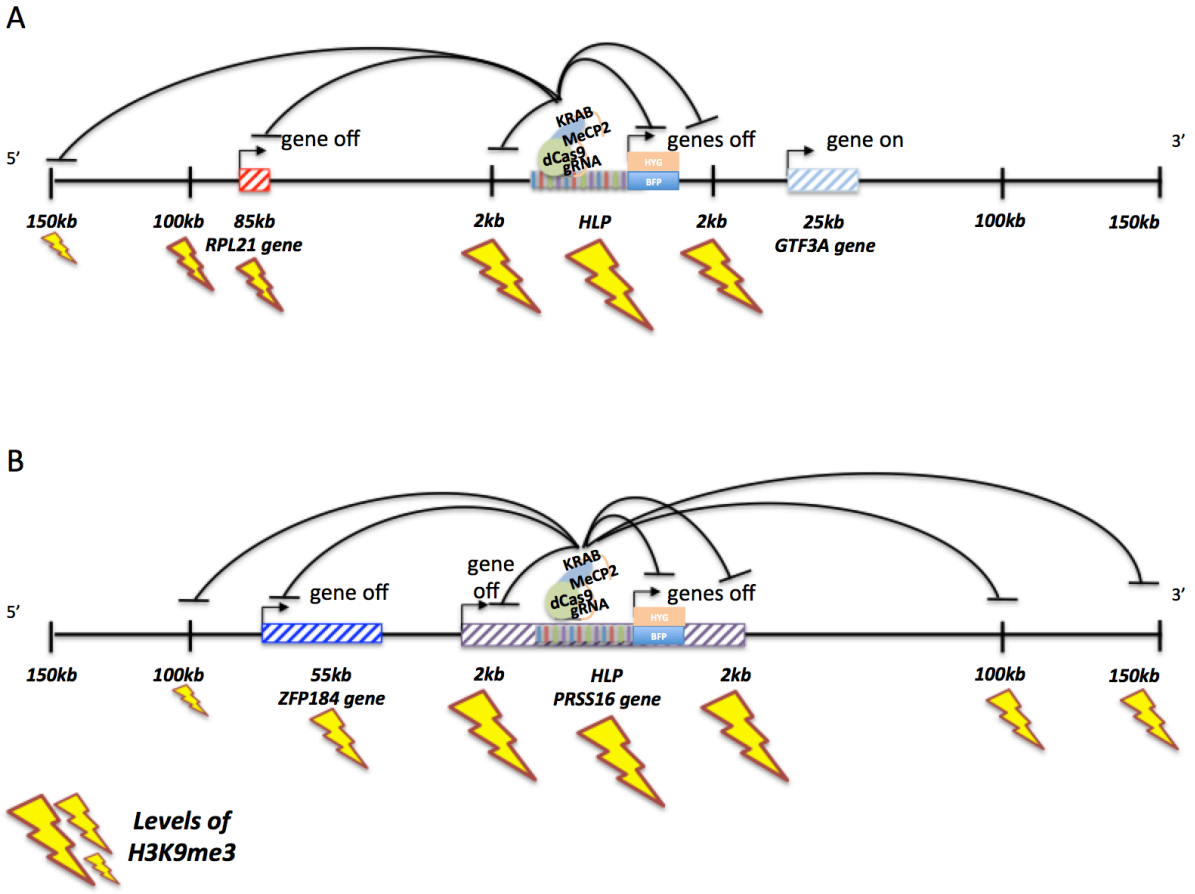


Figure 20: Schematic representation of H3K9me3 modification levels of the regions up to 150 kb from the integration sites of the HLPs from the prototype H129-dCas9/HLP-3 analysed with quantitative polymerase chain reaction (qPCR) following the chromatin immunoprecipitation (ChIP)

Upon the recruitment of dCas9-KRAB-MeCP2 fusion protein on the HLP integration sites of the working haploid prototype H129-dCas9/HLP-3, this figure shows observations on the regions nearby, where the silencing is expected the most. The analysed range is up to 150 kb. Lightning symbol represents the H3K9me3 epigenetic modification. The bigger the lightning symbol gets the higher H3K9me3 modification relatively (check figure 19). In the figure, gene expression profiles are also shown. In (A) 1st integration case and in (B) 2nd integration case of the prototype H129-dCas9/HLP-3 are shown.

Comparing the results on the gene expression and H3K9me3 modification levels showed us that the working haploid prototype H129-dCas9/HLP-3 was useful to understand the system's capacity of silencing and revealed the ability of the desired prototypes. Furthermore, these results were very helpful when it comes to advancing the prototypes in the next steps.

3.3.4. Discussion

With the multiple gRNA expression vector we generated 3 different haploid mESC prototypes. The preliminary monitoring via Western blot and flow cytometry on these 3 subclones of H129-dCas9/HLP-1, 2 and 3 verified the system's functionality. Overall, the results were promising. Therefore, these subclones could be used as the working haploid mESC prototypes and should be tested further on the system's silencing and heterochromatinisation mechanisms.

In the next step, we have shown that with the method Splinkerette PCR the exact integration sites of the HLP and the regions nearby could be detected. Therefore, a map of the whole region could be constructed. In our experiments, we used primer sequences for PiggyBac integrations. That was the case because; in the prototypes, the HLPs and dCas9-KRAb-MeCP2 transgene were integrated with the PiggyBac transposon system. For the advanced prototypes, in order to have a practical and faster determination of the integration sites of only HLPs, the introduction of the method of the transgene was aimed to be altered. Therefore, for the integration of the transgene, instead of the current transposon system, methods like Tol2, Rosa26 and etc. were aimed to be used.

After locating the integration sites of the HLPs, in order to determine the silencing range of the system, broad monitoring experiments were performed. The evaluation of the qPCR results on expression levels of potential target genes and on H3K9me3 profiles of the locations nearby and within the HLP validate the expected shutting

down efficiency of the system. By looking at these results we were able to understand the system's capacity of silencing and the ability of the desired prototypes. Besides, we were successful to produce the one working haploid prototype and this was H129-dCas9/HLP-3. When we observed the range of the heterochromatinisation from the integration sites of the HLPs, this prototype was very promising in the sense of leading the way for advanced haploid prototypes that would have a larger range of heterochromatinisation.

3.4. Cloning of the Doxycycline inducible dCas9-KRAB-MeCP2 vector

3.4.1. Task and experimental setup

Already in the previous task, the functional tests certified the system's performance in haploid mouse ES prototype cells. This working haploid prototype had to be advanced until it reaches an adequate level of prototype that is aimed to used in genome wide screening. As the first step of the improvement, in order to secure the stability of the system, the activity of dCas9-KRAB-MeCP2 transgene must have been able to be controlled. For this purpose, the plan was to use Tetracycline-controlled transcriptional activation method. In this method of controllable gene expression, transcription is reversibly switched on or off when the antibiotic tetracycline or one of its derivatives (e.g. doxycycline) is present (Gossen et al. 1995). Thus, the transcriptional activity of the transgene could be altered at anytime when all of the three genetic components expressed within the prototype genome simultaneously. Secondly as the next improvement of the system, the integration sites of dCas9-KRAB-MeCP2 transgene within the genome must have been assured and had to be distinct than the integration sites of the HLP. In the last attempts, both for the transgene and the landing platform integration we always used the vectors, containing the transposon system PiggyBac. As we showed in the last experiment, working haploid prototype production, it was challenging to distinguish the

integration sites of the transgene and the HLP with the same PiggyBac primers used for Splinkerette PCR. Therefore, the next plan was to construct a vector for dCas9-KRAB-MeCP2 transgene but this time contains a different and more stable integration system like Tol2 or Rosa26 transposon system. Thus, it would not be challenging anymore to only observe the integration sites of the HLP when PiggyBac primers were used for the Splinkerette PCR. Consequently, the ultimate task was to clone vectors that contain dCas9-KRAB-MeCP2 transgene with the Doxycycline-inducible transcriptional activation system, which is between the unique arms of the Tol2 or Rosa26 transposon system.

The first step of the task was to clone the vector with Dox-inducible system with the dCas9 fusion gene together. For this, the strategy was cutting open the vector with Tet-operator sequences and concurrently, producing dCas9-KRAB-MeCP2 insert via PCR. Later, by performing Gibson assembly, the aimed vector could be constructed. The resulting vector was named "pB_dox_dCas9". The following was going to be the cloning of the Tol2 and Rosa26 transposon systems into this vector. Again with Gibson assembly the vectors were successfully cloned.

3.4.2. Results and Discussion

Upon following the experimental setup above, two vectors were successfully cloned (figure 21). The resulting vectors were named "Tol2_dox_dCas9" and "Rosa26_dox_dCas9". The fundamental parts of these vectors were identical except for the integration frameworks. In both cases, the important genomic elements including the fusion gene were fixed between the two essential sequences of the corresponding transposon system (figure 21).

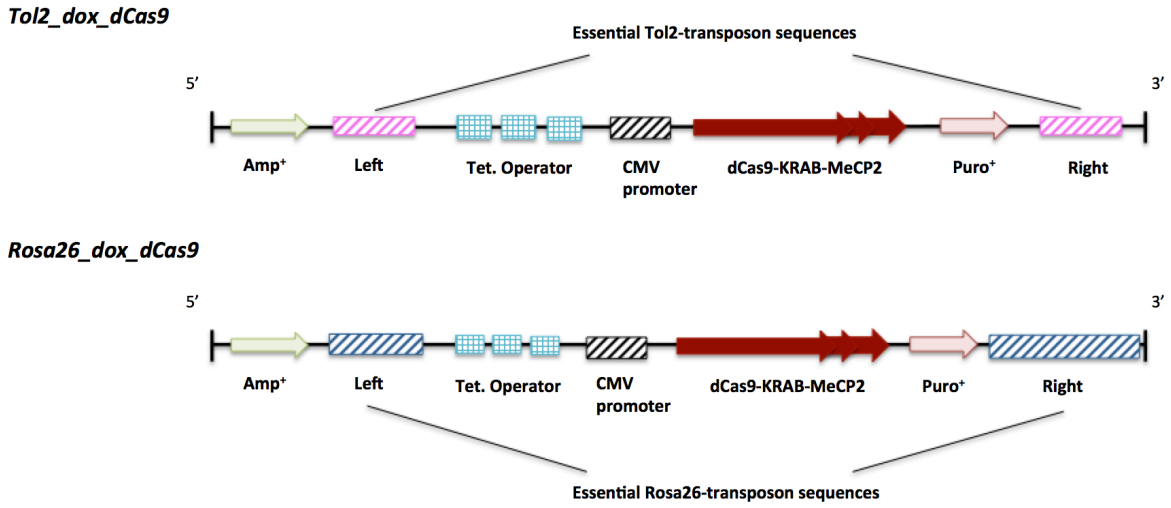


Figure 21: Schematic representation of the genomic components of 2 different Doxycycline inducible dCas9-KRAB-MeCP2 vectors
Overview of both Tol2 and Rosa26 Doxycycline-inducible dCas9-KRAB-MeCP2 vectors. They are consisting of antibiotic marker genes (both Ampicillin and Puromycin) as selective markers of transformation/transfection experiments, essential left and right sequences of the relevant transposon system (Tol2 or Rosa26), Tet-operator repeat sequences of the Doxycycline-inducible system and dCas9-KRAB-MeCP2 transgene sequence.

With these vectors, we could now integrate the dCas9-KRAB-MeCP2 fusion gene securely, meaning that by using the Doxycycline-inducible transcriptional activation, at any time, the expression of the fusion gene could be controlled. Furthermore, for the integration of the fusion gene, using different integration systems other than PiggyBac, it would take one step less to locate the HLP, which would then be integrated with PiggyBac into the genome of the prototype. Thereby, on our way to reach the goal of advancing the haploid mESC prototype, by cloning these vectors we took one further step.

3.5. Producing an advanced haploid mESC prototype

Until this point of the project, there have been many experimental tasks done in order to advance and if necessary to alter the components that were going to be used in the construction of the ultimate haploid mESC prototype. In the last task, by making the new "Tol2_dox_dCas9" vector, we planned to make the system

controllable and secure. Therefore, the next goal was to try and establish the advanced haploid mESC prototype.

This time, the dCas9-KRAB-MeCP2 transgene was going to be integrated via Tol2 transposon system in order to distinguish the integration sites of the HLP, which was going to be integrated into the genome with PiggyBac. Besides, by using Doxycycline inducible system, the transcriptional activity of the dCas9-KRAB-MeCP2 transgene was going to be controlled at anytime. Finally, again the multiple gRNA expression vectors "Red_4gRNA" was going to be used. While establishing this prototype, the placing order of the components was going to start from the transgene, later the gRNA vector and as final the HLP.

We started choosing the cell lines for the experiment. The haploid mESC lines H129 and HREX were going to be used to make the prototypes. After the transfection (attempts with chemical and electroporation) of the "Tol2_dox_dCas9" vector into these cell lines, the selection was done Puromycin for a maximum of 3 days. Puromycin is an aggressive antibiotic that also affects the pluripotency of ES cell lines. Therefore, since it causes stress, keeping cells haploid was very challenging. Unfortunately, we were not able to keep the cells haploid until the next step where we were planning to transfect the haploid clones with dCas9-KRAB-MeCP2 with the multiple gRNA vector. Therefore, we were not able to proceed with the next step.

4. Conclusion

In this project, we tried to develop an approach, which solves an important problem of all genome-wide screening technologies. The ultimate tasks of this project were to generate screening prototype mESC lines with an advanced CRISPR/Cas9 system based recruitment of the silencing platform and later, to evaluate its silencing power by analysing the H3K9me3 epigenetic modification. At the beginning of the project, the main goal was to bring the silencing system to its best shape to construct the

advanced haploid mESC prototype. Therefore, throughout the project, one simple diploid prototype and some working haploid prototypes were produced. Later on, the attempt on making the advanced haploid prototype, unfortunately, was failed.

Firstly, the preliminary monitoring on the simple diploid prototype stated that other then the gRNA efficiency issue; the working system could be generated in the diploid C2 mESCs. The HLP and fusion protein seemed to work successfully and simultaneously. Therefore, the next step was to construct the haploid mouse ES cell prototype after improving the efficiency of concurrent transfection of multiple, specific gRNA vectors.

Next, succeeding the gRNA efficiency issue by generating a single vector for multiple gRNA expression cassettes, we were ready to establish the system in the haploid mESCs and we generated 3 different prototypes. The preliminary monitoring via Western blot and flow cytometry verified the system's functionality. After, we concluded that these cell lines could be used as working haploid mESC prototypes and should be tested further on the system's silencing power due to heterochromatinisation. For the broad monitoring, for all these prototypes, we first located the integration sites of the HLPs within their genomes via Splinkerette PCR. Then, we were ready to analyse the silencing range of the HLPs and observe their effects on the genomic regions nearby. Therefore, we looked for the H3K9me3 epigenetic modification. Moreover, we also analysed the expression patterns of the genes on and near the HLP. After a series of ChIP and qPCR experiments, we were able to understand the system's capacity of silencing and so the ability of the prototypes. In the end, we came to the conclusion that we successfully produced one working haploid prototype and when we looked at the range of silencing from the integration sites of its HLPs, this prototype was very promising in the sense of leading the way for advanced haploid prototypes.

Later on, we upgraded our way of introducing dCas9-KRAB-MeCP2 transgene to the silencing system. For this, we generated a vector with the Doxycycline inducible transgene coming with a different integration method. Thus, the system became more controllable and secure. Succeeding this, the next goal was to establish the advanced haploid mESC prototype. Unfortunately, due to some experimental issues on cells ploidy, our attempt was failed. However, the genetic components for the advanced prototype were known and some of them were ready to use. For future attempts, we know that we shall introduce dCas9-KRAB-MeCP2 and the multiple gRNA expression vector with a secure integration method (e.g. Tol2 and Rosa26) which is different than the integration method of the HLP (e.g. PiggyBac). Therefore, it will be easier and faster to locate the HLPs within the genome. Besides, by controlling the silencing system with a substantial mechanism (e.g. Doxycycline) will make the system desirable for ideal screenings.

In conclusion, with this project, we proposed to fill a technological gap and provide new opportunities to investigate mammalian genome function in a wider and reliable manner. We tried to achieve this by establishing haploid mESC prototypes. Meanwhile, we performed many experimental tasks to advance and if necessary to alter the components that were going to be used in the construction of the ultimate haploid mESC prototype. Even though we might still have some way to go until we establish the advanced prototype, throughout the project, we were able to bring an understating of how to establish the silencing system in the test prototypes and show the power of the silencing mechanism that comes with.

5. References

Bradley, A., Hasty, P., Davis, A. & Ramirez-Solis, R. (1992) **Modifying the mouse: design and desire.** Biotechnology 10, 534–539.

Balboa D., Weltner J., Eurola S., Trokovic R., Wartiovaara K, Otonkoski T. (2015) **Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation.** Stem Cell Reports 5(3):448–59.

Breunig C. T., Durovic T., Neuner A. M., Baumann V., Wiesbeck M. F., Köferle A. (2018) **One step generation of customizable gRNA vectors for multiplex CRISPR approaches through string assembly gRNA cloning (STAgR).** PLoS One 13, e0196015.

Brook, F.A. & Gardner, R.L. (1997) **The origin and efficient derivation of embryonic stem cells in the mouse.** Proc. Natl. Acad. Sci. USA 94, 5709–5712.

Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. (2013) **Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease**. Nat. Biotechnology 31, 230–232.

Czechanski, A., Byers, C., Greenstein, I. et al. (2014) **Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains.** Nat Protoc 9, 559–574.

Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A. & Richardson, J.E. (2012) **The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse.** Nucleic Acids Res. 40, D881–D886.

Evans, M.J. & Kaufman, M.H. (1981) **Establishment in culture of pluripotential cells from mouse embryos.** Nature 292, 154–156

Friedman, J. R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., Rauscher, F.J. (1996) **KAP-1, a novel corepressor for the highly conserved KRAB repression domain.** Genes Dev. 10, 2067–2078.

Gertsenstein, M, Nutter, L.M.J., Reid, T, Pereira, M., Stanford, W.L., Rossant J. (2010) **Efficient Generation of Germ Line Transmitting Chimeras from C57BL/6N ES Cells by Aggregation with Outbred Host Embryos.** PLoS ONE 5(6): e11260.

Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., Bujard, H. (1995). **Transcriptional activation by tetracyclines in mammalian cells.** Science 268 (5218).

Horii, T., Hatada, I., (2013) **Genome engineering using haploid embryonic stem cells using the Cas9/RNA system.** PeerJ 15(5),83-94.

Horn, C., Hansen, J., Schnütgen, F., Seisenberger, C., Floss, T., Irgang, M., De-Zolt, S., Jones, P. L. (1998) **Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription**. Nat. Genetics 19, 187–191.

Joung, J., Konermann, S., Gootenberg, J. S., Abudayyeh, O. O., Platt, R. J., Brigham, M. D., Sanjana, N. E., & Zhang, F. (2017) **Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening.** Nature protocols 12(4), 828–863.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014) **Genetic exploration of the exit from self-renewal using haploid embryonic stem cells.** Cell Stem Cell 14:385-393.

Leeb, M., Perry, A. C. F. and Wutz, A. (2015) **Establishment and use of mouse haploid ES cells.** Curr. Protoc. Mouse Biol. 5:155-185.

Leeb, M., Wutz, A. (2011) **Derivation of haploid embryonic stem cells from mouse embryos.** Nature 479:U131–U164.

Leeb, M., Wutz, A. (2012) **Establishment of epigenetic patterns in development.** Chromosoma 121, 251–262.

Li, M. A., Jason, S.L., Tilgner, K., Ong, S. H., Koike-Yusa, H. and Yusa, K. (2018) **Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution** Cell Rep. 24(2): 489–502.

Li, M. A., Pettitt, S. J., Yusa, K., Bradley, A. (2010) **Genome-Wide Forward Genetic Screens in Mouse ES Cells.** Methods in Enzymology (477): 217-242.

Mandegar, M. A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan A.H., Miyaoka, Y., Holmes, K., Spencer C.I. (2016) **CRISPR interference efficiently induces specific and reversible gene silencing in human iPSCs.** Cell Stem Cell 18, 541–553.

Pettitt, S.J., Rehman, F.L., Bajrami, I., Brough, R., Wallberg, F., et al. (2013) **A Genetic Screen Using the PiggyBac Transposon in Haploid Cells Identifies Parp1 as a Mediator of Olaparib Toxicity.** PLoS ONE 8(4): e61520.

Qi, L. S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A. (2013) **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.** Cell 152, 1173–1183.

Shalem, O. Sanjana, N.E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson T., Root D.E., Doench, J.G. (2014) **Genome-scale CRISPR-Cas9 knockout screening in human cells.** Science 343, 84–87.

Yeo, N.C., Chavez, A., Lance-Byrne, A., Chan Y., Menn D., Milanova, C.C., Kuo, X., Guo, S., Sharma, A. (2018) **An enhanced CRISPR repressor for targeted mammalian gene regulation.** Nat. Methods 15, 611-616.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014) **Genetic screens in human cells using the CRISPR-Cas9 system.** Science 343:80-84.

Wurst, W., von Melchner, H. and Noppinger, P.R. (2007) **Splinkerette PCR for more efficient characterization of gene trap events.** Nat. Genetics, 39, 933-934.