Uspostava rekombinantnih mišjih embrionalnih matičnih stanica u svrhu istraživanja fleksibilnosti DNA

Mikičić, Ivan

University of Zagreb

Faculty of Science

Division of Biology

Ivan Mikičić

Establishment of recombinant mouse embryonic stem cells for use in DNA flexibility assays

Graduation thesis

Zagreb, 2018

This Graduation thesis was conducted in the Laboratory for Epigenetic Regulation and Genome Engineering of the Biotechnological Center at Technische Universität Dresden, under the leadership of Prof. Dr. Francis Stewart. The Graduation thesis is submitted to evaluation to the Department of Biology, Faculty of Science, University of Zagreb in order to acquire the Master of Molecular Biology title.

ACKNOLEDGMENTS

First and foremost, I would like to thank my supervisor Francis Stewart for making the research at BIOTEC, TU Dresden possible and to my immediate supervisor, Sarah Tsurkan for all the experimental help and dedication she invested into my project. Thank you to Konstantinos Anastassiadis for providing me with the cell lines and helping me with Southern blots and to Jun Fu for preparing the plasmid constructs. I am also grateful to other members of the Stewart lab, namely to Andrea Kranz, Frank Groß, Neha Goveas, Katrin Neumann, Deepthi Ashokkumar, Sivaraman Subramaniam and Qinyu Zhang for useful advice as well as Susan Schneider, Isabell Kolbe and Stefanie Weidlich for the technical support and Juliane Wunderlich for the administrative help. Special thanks to Zhen Li for his dedication to re-validation of my gene targeting experiments.

A huge thank you to my co-supervisor Ivana Ivančić Baće for investing a lot of time into helping me with the thesis writing.

Lastly, I would like to thank my friends and family for the care, support and all the logistical help during my time abroad.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek Diplomski rad

USPOSTAVA REKOMBINANTNIH MIŠJIH EMBRIONALNIH MATIČNIH STANICA U SVRHU ISTRAŽIVANJA FLEKSIBILNOSTI DNA

Ivan Mikičić

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Kvantitativnim mjerenjem fleksibilnosti DNA u stanicama HEK293 metodama mjesno-specifične rekombinacije i Southern blota uočeno je da je fleksibilnost DNA *in vivo* dvostruko veća nego *in vitro*, što je pripisano namatanju DNA oko nukleosoma te kromatinskoj remodelaciji. Utišavanje razvojno-specifičnih gena lokalnim remodeliranjem kromatina također bi moglo uzrokovati promjenu fleksibilnosti DNA. Cilj ovog diplomskog rada bio je pripremiti mišje embrionalne matične stanice (mEMS) s ugrađenim različitim genskim kazetama za mjerenje i usporedbu fleksibilnosti DNA s mišjim neuralnim matičnim stanicama (mNMS). Kloniranjem gena proizvedena je serija od 10 plazmida koji sadrže različite genske kazete: mjesta *loxP* različitih udaljenosti (od 70 bp do 2359 bp), 5' i 3' homologiju s kromosomskim genom *mNanog* te gen za otpornost na antibiotik blasticidin. Kloniranjem gena razvijene su probe za provjeru ispravnosti kazeta metodom Southern blot, te su razvijene početnice za provjeru metodom lančane reakcije polimerazom (PCR). Uvođenjem dvolančanog loma potpomognutim metodom CRISPR-Cas9 uspješno su ugrađene kazete svih 10 plasmida u gen *mNanog*, što je potvrđeno metodom Southern blot.

(59 stranica, 15 slika, 21 tablica, 122 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u Središnjoj biološkoj knjižnici

Ključne riječi: fleksibilnost DNA, remodeliranje kromatina, mišje embrionalne matične stanice, neuralne matične stanice, Cre-*loxP*, CRISPR-Cas9, Southern blot

Voditelj: dr. sc. Francis Stewart, profesor genomike

Suvoditelj: dr. sc. Ivana Ivančić Baće, izv. prof.

Ocjenitelji: dr. sc. Ivana Ivančić Baće, izv. prof.; dr. sc. Martina Šeruga Musić, izv. prof.; dr. sc. Željka Vidaković Cifrek, izv. prof.

Rad prihvaćen: 1. veljače 2018.

BASIC DOCUMENTATION CARD

University of Zagreb

Faculty of Science

Division of Biology

Graduation thesis

ESTABLISHMENT OF RECOMBINANT MOUSE EMBRYONIC STEM CELLS FOR USE IN DNA FLEXIBILITY ASSAYS

Ivan Mikičić

Rooseveltov trg 6, 10000 Zagreb, Croatia

Using quantitative measurement of DNA flexibility in HEK293 cells by methods of site-specific recombination and Southern blot, it has been observed that *in vivo* DNA flexibility is double that of *in vitro* DNA, which has been ascribed to nucleosomal DNA wrapping and chromatin remodeling. Silencing of development-specific genes by local chromatin remodeling could also cause a change in DNA flexibility. The aim of this Master's thesis was to prepare the mouse embryonic stem cells (mESCs) with different integrated gene cassettes for measurement and comparison of DNA flexibility with mouse neuronal stem cells (mNSCs). Gene cloning was used to prepare a series of ten plasmids containing the different gene cassettes: *loxP* sites of different distances (from 70 bp to 2359 bp), 5' and 3' homology to chromosomal *mNanog* gene and blasticidin resistance gene. Probes for cassette verification by Southern blotting were produced by gene cloning and primers for verification by polymerase chain reaction (PCR) were designed. Introducing a double-stranded break assisted by CRISPR-Cas9 method, all 10 targeting constructs were successfully inserted into the *mNanog* gene, which was verified by Southern blotting.

(59 pages, 15 figures, 21 tables, 122 references, original in: English)

Thesis deposited in the Central Biological Library

Key words: DNA flexibility, chromatin remodeling, mouse embryonic stem cells, neural stem cells, Cre-*loxP*, CRISPR-Cas9, Southern blot

Supervisor: Francis Stewart, Professor of Genomics, Dr. Sc.

Co-supervisor: Ivana Ivančić Baće, Associate Professor, Dr. Sc.

Reviewers: Ivana Ivančić Baće, Associate Professor, Dr. Sc.; Martina Šeruga Musić, Associate Professor, Dr. Sc.; Željka Vidaković Cifrek, Associate Professor, Dr. Sc.

Thesis accepted: February 1, 2018

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Cre/*loxP*-mediated site-specific recombination (SSR)

Site-specific recombination (SSR) is a type of nonhomologous recombination that includes cleavage and rejoining of DNA strands. The recombination takes place only between specific sequences on DNA and is assisted by site-specific recombinases (Akopian and Stark, 2005; Snyder and Champness, 2007). These recombinases recognize specific DNA sites, cleave DNA backbone, exchange DNA material placed in between the sites and ligate the end products (Kolb, 2002).

SSR serves a wide range of purposes such as resolution of bacterial chromosome and plasmid dimers, integration of phages into bacterial chromosomes, resolution of cointegrates formed during replicative transposition and assembly of mammalian antibody genes during B-cell maturation (Snyder and Champness, 2007; Kolb, 2002). The diverse potential of biological functions carried out by site-specific recombinases has been utilized in various genome engineering applications for decades. These include conditional mutagenesis (Gu et al., 1994; Kuhn et al., 1995; Logie and Stewart, 1995), removal of selectable genes (Gu et al., 1993), gene expression switches (O'Gorman et al., 1991; Lakso et al., 1992; Angrand et al., 1998), recombinase-mediated cassette exchange (Schlake and Bode, 1994; Baer and Bode, 2001; Wallace et al., 2007), lineage tracing (Awatramani et al., 2003) and chromosomal translocations (Smith et al., 1995; Herault et al., 1998; Buchholz et al., 2000; Collins et al., 2000; Liu et al., 2002).

Based on the catalytic mechanism, site-specific recombinases can be divided into two classes: tyrosine (EC 2.7.7.) and serine recombinases (EC 3.1.22.) (Grindley et al., 2006). In genome engineering, the former class is used much more often. Some prominent members of the tyrosine recombinase family include Cre recombinase from phage P1, FLP recombinase from the yeast 2-micron plasmid, λ integrase from phage λ and XerCD from *E. coli* chromosome. The most widely used of the tyrosine recombinases, at least in genome engineering, is the Cre recombinase (Kolb, 2002). Cre's target sites, called *loxP*, are 13 bp inverted repeats (IRs) separated by a nonpalindromic 8 bp central spacer (Anastassiadis et al., 2010), (Figure 1a). One Cre monomer binds to each 13 bp IR and recombination occurs when two Cre dimers bound to two full *loxP* sites contact each other after a random collision, forming a tetramer synaptic complex (Ringrose et al., 1998). Since the 8 bp spacer is nonpalindromic, *loxP* sites are often represented by an arrowhead indicating the spacer direction (Anastassiadis et al., 2010). The spacer is critical for recombination as it is the site of the nucleophilic attack performed by Cre's active-site tyrosine residue, which results in a covalent phosphotyrosyl intermediate. The recombination mechanism of Cre and other tyrosine recombinases includes two rounds of single-strand cleavages, one per round on each DNA helix. These two events are connected via strand exchange leading to the so-called Holliday junction formation and isomerization which

enables resolution of the junction necessary for the second round of cleavage to occur (Grainge and Jayaram, 1999), (Figure 2).

Depending on whether *loxP* sites are located on the same or different DNA molecules and on the orientation of the spacers, there are several possible outcomes of Cre-mediated recombination: excision, integration, inversion and translocation. Since Cre-mediated recombination is inherently reversible and *loxP* sites are more likely to come into close contact if they are located on the same DNA molecule, excision is favored over insertion. Inversions and translocations, generally speaking, occur equally likely in both forward and reverse direction (reviewed in Branda and Dymecki, 2004), (Figure 1b).

There are several reasons for Cre's ubiquitous use in genome engineering. Firstly, it does not require ATP (Kolb, 2002) or any accessory proteins (Warren et al., 2008). Secondly, its small size (38 kDa) and the simplicity of *loxP* sites allow for easy genome manipulation. Thirdly, its single-strand intermediate mechanism avoids creation of side products such as hairpins and three-way junctions (Grindley, 2006) and complications caused by unresolved double-strand breaks. Furthermore, Cre works efficiently in all organisms tested so far, from prokaryotes to mammals (Anastassiadis et al., 2010). Finally, kinetic comparisons with another commonly used tyrosine recombinase, FLP, have shown that Cre has unusually high affinity towards its target site and shows high efficiency of substrate excision even at lower concentrations (Ringrose et al., 1998; Anastassiadis et al., 2009). However, the Cre-*loxP* system is not without its drawbacks. Inherent reversibility of Cre-mediated recombination, particularly pronounced when Cre is used for insertion rather than excision, and binding to cryptic *loxP* sites may cause side effects such as regeneration of the substrate site or unintended gene modifications, respectively. The reversibility problem has been tackled by development of mutant *loxP* sites, for example *lox66*/*lox71*, that enable unidirectionality, albeit at lower recombination efficiency than the wild-type *loxP* (Oberdoerffer et al., 2003). If it is essential to avoid reversibility, large serine recombinases can be used instead of Cre (Anastassiadis et al., 2009).

Diverse applications require diverse recombinases. Decades of SSR application development have yielded a wide range of Cre and *loxP* variants, but also identification of other recombinases and their recognition sites. Examples of *loxP* mutants include the aforementioned *lox66*/*lox71* pair as well as *lox2272* and *lox5171*, which are useful because they interact with each other but not with *loxP*. This means they can be used when several *lox* sites are being introduced in tandem and cross-reactions are undesirable (Araki et al., 1997; Lee and Salto, 1998). Other SSR proteins and their recognition sites have been identified and developed, such as Dre/Rox (Sauer and McDermott, 2004; Anastassiadis et al., 2009) and Vika/Vox

(Karimova et al., 2013). FLP recombinase has been temperature-optimized (FLPe) (Buchholz et al., 1998) and codon-optimized for use in mice (FLPo) (Raymond and Soriano, 2007).

One of the most important improvements of Cre/*loxP* system came via Cre inducibility, which enabled spatial and temporal control of Cre activity (Feil et al., 2009). Cre was fused to the mutated hormone-binding domain of the estrogen receptor, creating a fusion protein called CreER (Feil et al., 1996; Feil et al., 1997; Metzger et al., 1995; Zhang et al., 1996). The most successful version of this fusion protein is called CreERT2 (Feil et al., 2009). The estrogen receptor is retained in the cytoplasm until the administration of estrogen changes its conformation and shuttles it into the nucleus. 4-OHT (4-hydroxytamoxifen) is usually used in Cre inducibility experiments instead of estrogen due to its more than 100-fold increased affinity for the estrogen receptor proteins (Kuhl, 2005).

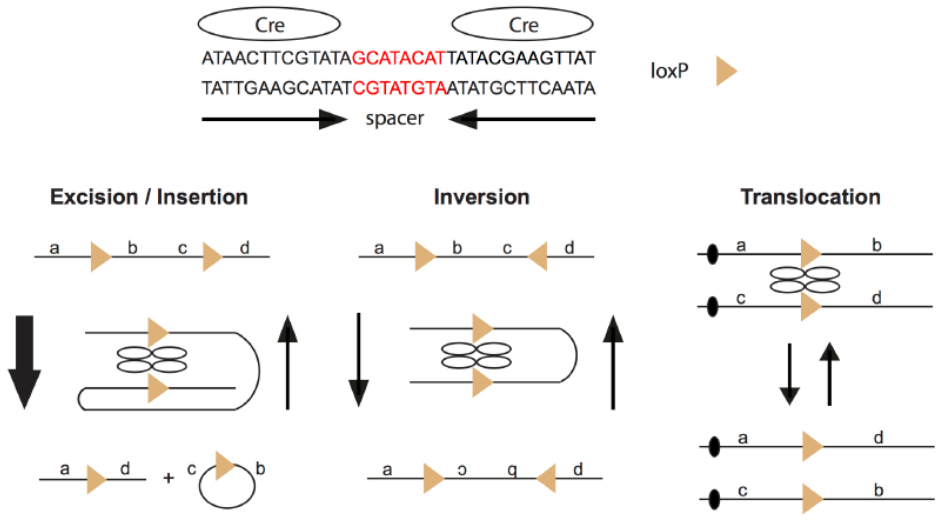


Figure 1. A) The architecture of the *loxP* site, with the Cre-binding sites indicated in black and the spacer indicated in red. **B)** Possible outcomes of Cre/*loxP*-mediated recombination. Beige arrowheads represent spacer orientation and the thickness of the black arrows represents the directionality of recombination events. Adapted from Baker, 2015.

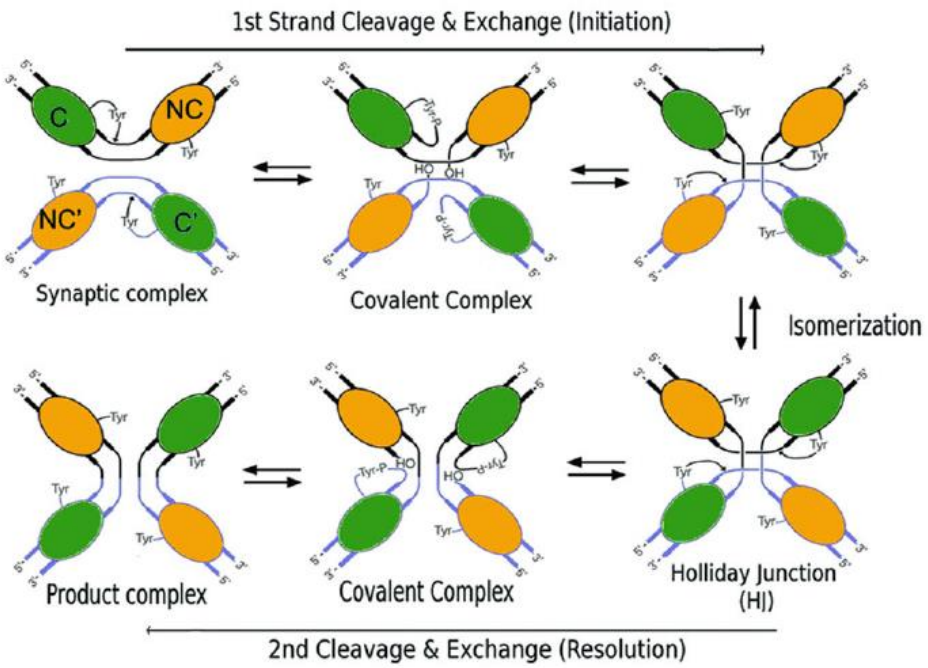


Figure 2. Mechanism of tyrosine recombinase-mediated SSR.
C and C' represent cleaving subunits; NC and NC' represent non-cleaving subunits. The details of the mechanism are explained in the text. Adapted from Abi-Ghanem et al., 2015.

1.2. Mouse embryonic stem cells (mESCs)

Owing to its great genetic similarity with humans, its small size and the speed of reproduction, mouse has become the favorite model organism in biomedicine. Advancements in the ability to manipulate mouse embryos and culture its embryonic stem cells concomitant with the ever more precise genetic engineering tools have fostered great efforts in investigation of molecular genetics of mammalian early development. Mouse embryos start developing through rounds of cell divisions producing first a ball-like 8-cell structure called a morula and subsequently a hollow sphere called a blastocyst. The outer layer of the blastocyst, termed trophoectoderm, contributes to formation of extraembryonic tissues like amniotic sac and the placenta, so called because they form no part of adult mammals. The inner cell mass of the blastocyst, on the other hand, gives rise to embryo proper. The cells of the inner cell mass are initially totipotent, ie. they can give rise to any tissue of the adult organism. A particularly interesting property of the inner cell mass is that it can be isolated and proliferated indefinitely in culture without losing its totipotency, producing embryonic stem cells (ESCs). The huge interest in advancement of ESC manipulation and research on their biological peculiarities is a result of their promising applications in tissue and organ repair as well as the ease and precision with which they can be genetically manipulated and ultimately produce a modified gene

of interest in a whole mouse due to their totipotency. The ease of genetic manipulation refers to their high transfection efficiency and their ability to incorporate foreign DNA into desirable genomic loci at reasonable frequencies as well as the variety of genetic engineering tools (eg. clustered regularly interspaced short palindromic repeats – CRISPR-associated protein 9 method, abbreviated as CRISPR-Cas9, SSR etc.) which are routinely applied on them.

Mouse embryonic stem cells (mESCs) are derived from the inner cell mass of blastocysts at early stage of embryogenesis, from 3.5 days old pre-implantation embryos (Marin, 1981). Like other stem cells, mESCs can be cultured *in vitro* longer than most cell types due to their immortality and ability to retain their developmental potential even after prolonged clonal expansion. mESC culture requires addition of pluripotency-maintaining supplements, such as the leukemia inhibitory factor (LIF; Furue et al, 2005) or inhibitor combinations (glycogen synthase kinase-3 inhibitor CT99021 and ERK signal cascade inhibitor PD184352; Ying et al., 2008). Depending on the desired robustness of pluripotency maintenance and the type of mESCs used, a feeder layer of inactivated mouse embryonic fibroblasts is sometimes used (reviewed in Llames et al., 2015). Several mESC cell lines are in common use. They are derived either from the 129 mouse strain (Downing and Battey, 2004; Simpson et al, 1997) or from C57BL/6J and C57BL/6N mouse strains (Seong et al., 2004; Keskintepe et al., 2007; Hansen et al., 2008). 129 derivatives E14, R1, D3 and J1 are used because of their robustness in cell culture and higher germline transmission rates. On the other hand, C57BL/6 strain is most commonly used as a reference for immunological, neurobiological, behavioral and physiological studies in mice (Bouabe and Okkenhaug, 2013).

1.3. Gene targeting in mESCs

DNA sequences can be introduced into genomes either randomly (eg. by using viruses that stably integrate into a genome) (Jaenisch, 1976) or precisely (eg. by providing a template for homologous recombination (HR)). Gene targeting is a technique used for inserting DNA sequences into precise loci in the genome, usually gene bodies, using HR (Doetschman T et al., 1987; Thomas and Capecchi, 1987; Thompson et al., 1989). Gene targeting can be used to delete or insert mutations into genes to determine their function by the loss-of-function approach; introduce human genes into murine genomes to establish models to study human diseases and human-specific genes; introduce pathogen genes or genomes to examine their contribution to pathogenesis; introduce reporter genes to monitor expression of genes of interest (reviewed in Bouabe and Okkenhaug, 2013).

Gene targeting is achieved by introduction of a targeting construct (TC) into mESCs by transfecting the DNA into the cells (Hall at al., 2010). A TC is usually composed of a 5'-homology arm (5'-HA), TC body (containing a selectable marker and additional, application-dependent sequences) and a 3'-homology arm

(3'-HA) (Bouabe and Okkenhaug, 2013). HAs are necessary to facilitate HR and a common rule of thumb states that the longer the HAs the better the targeting efficiency is (Hasty et al., 1991; Deng and Capecchi, 1992). They are usually, but not exclusively, derived from genomic DNA prepared from the same strain that is used for targeting to avoid mismatches (te Riele et al., 1992). Since TCs are often cloned into plasmids, they are often linearized before use to improve targeting efficiency (Thomas et al., 1986). Traditional HR-directed targeting works at low frequencies, so that only 10^{-2} to 10^{-3} of integrations are HR events (Melton, 2002), which makes screening for correct insertions labor-intensive. Double-strand breaks (DSBs) within the locus to be targeted enrich the HR events 100-fold (Jasin and Rothstein, 2013). Recent advancements in gene targeting take advantage of the fact that DSB induction at precise loci has become a routine with development of the zinc finger proteins (ZNFs; Hockemeyer et al., 2009), transcription activator-like effector nucleases (TALENs; Hockemeyer et al., 2011) and in particular the CRISPR-Cas9 system (Wang et al., 2013).

TC bodies are where the diversity of gene targeting applications is realized. The applications include gene knock outs and knock ins, point mutations, protein tagging, various deletions and many more. As previously mentioned, they most commonly consist of a selectable marker, ie. antibiotic resistance gene. This gene can be flanked by *loxP* sites (floxed) to allow for excision when selection is no longer necessary (reviewed in Bouabe and Okkenhaug, 2013). A modification of this approach is called gene trapping. A typical 5' trapping vector consists of a 5' splice acceptor (5' SA) and a promotorless antibiotic resistance gene that can be fused with another gene (Skarnes et al., 2011), eg. a reporter gene or a mutated version of the targeted gene, through internal ribosome entry site (IRES) or viral 2A peptide (Donelly et al., 2001; Szymczak et al., 2004) to facilitate expression of the two proteins as separate entities. Upon insertion of the gene trap into the intron of the targeted gene and expression of the mRNA, the resistance coding sequence (CDS) and the downstream fused CDSs are connected with the upstream exon via splicing mediated through the 5' SA. This essentially inactivates the targeted protein because all the exons located downstream of the inserted gene trap are no longer in frame (for an example of the described TC features, refer to Figure 3). Promotorless resistance genes are used to enrich the targeted cell populations with the gene traps that are correctly inserted into the intron of the targeted gene (Friedel et al., 2005) because the resistance protein can in such cases be expressed only if it is fused with the upstream exon. Apart from positive selection, negative selection is occasionally used to discriminate against non-homologous events. The most common negative selection genes used in gene targeting are thymidine kinase (TK) and diphteria toxin fragment A (DT-A), placed outside of HAs (Bouabe and Okkenhaug, 2013). During HR, sequences outside of HAs are usually lost but are retained upon random integrations. TK and DT-A proteins catalyze the conversion of

gancyclovir or ADP-ribose, respectively, to products that are toxic to the cell (Yagi et al., 1990; Mortensen, 2006).

The ubiquitous expression of transgenes and reporter genes is often desirable. In such cases, the TCs can be targeted into the ROSA26 mouse locus (Zambrowicz et al., 1997). ROSA (reverse orientation splice acceptor β-gal) promoter trap was first used in mESCs for a screen to identify developmental genes in mice (Friedrich and Soriano, 1991). It was subsequently discovered that one of the targeted loci, named ROSA26, exhibited ubiquitous expression in all hematopoetic cells as well as in all tissues of the embryo (Zambrowicz et al., 1997). One candidate for insertion into the ROSA26 locus is CreERT2 (Vooijs et al., 2001). Inducible Cre is especially useful when it comes to conditional mutagenesis, the goal of which is spatial and temporal control over SSR events. CreERT2 can be used both to delete a gene or to activate its expression (Lakso et al, 1992; Torres and Kuhn, 2003). Temporal control over the activation of gene expression is of high importance when it is desirable to avoid harmul effects of the transgene during the mouse embryogenesis (Bouabe and Okkenhaug, 2013) or limit potential toxicity as a result of prolonged high levels of Cre activity (Feil et al., 2009).

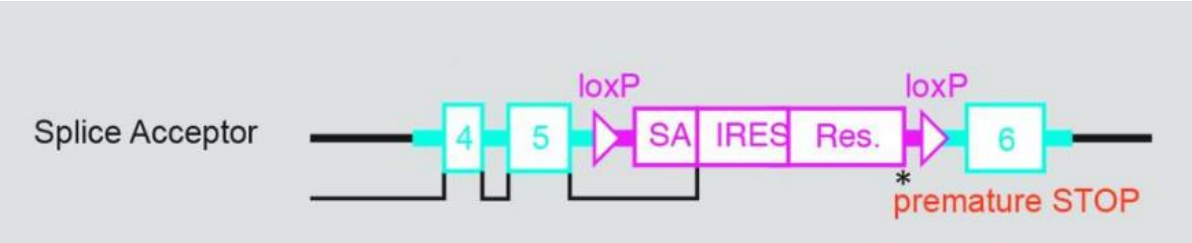


Figure 3. Architecture of a typical gene targeting construct.
Homology arms are depicted left and right of the *loxP* sites, with exons indicated as turquoise-white rectangles. SA – splice acceptor; IRES – internal ribosome entry site; Res. – resistance gene. Adapted from Kaulich and Dowdy, 2015.

1.4. CRISPR/Cas9-mediated genome engineering

Clustered regularly interspaced short palindromic repeats (CRISPR) – CRISPR-associated protein 9 (Cas9) engineering is a genome editing technique that relies on naturally occuring bacterial immunity to viral or plasmid DNA sequences by incorporating stretches of these sequences (termed spacers) into CRISPRs of the bacterial host genome (Barrangou et al., 2007). The spacers act as templates to produce the so called pre-crRNA (Brouns et al., 2008), which upon binding to tracrRNA (Chylinski et al., 2013) matures into crRNA helped by RNase III and guides Cas9 towards the complementary site on the DNA where Cas9 induces a DSB (Gasiunas et al., 2012; Jinek et al., 2012).

The ability to precisely introduce a DSB makes the CRISPR-Cas9 system an ideal tool for genome engineering. Mammalian cells respond to introduced DSBs by activating an error-prone repair process called non-homologous end joining (NHEJ) or, in case a homologous template is provided, homologous recombination (Ma et al., 2014). Instead of using two RNAs (cr- and tracrRNA) to target the desirable mammalian sequence, they are often combined into one single guide RNA (sgRNA; Jinek et al., 2012). In case of the most commonly used version of the CRISPR-Cas system, the one enabled by Cas9 protein from *Streptococcus pyogenes*, the genomic sequence of interest must contain an NGG motif (PAM sequence) 3 bp downstream of the intended DSB site (Qi et al., 2013). Since sgRNA needs to recognize around 20 bp to introduce a cut, any N_{20}-NGG sequence can be used as a target (Ma et al., 2014).

As expected, Cas9 can also cause some off-target DSBs due to complementarity of the sgRNA to more than one sequence in the genome. This problem has been tackled by characterizing the off-target effects (Fu et al., 2013; Kuscu et al., 2014; Hsu et al., 2013), improving the sgRNA design (Doench et al., 2014; Fu et al., 2014; Kim et al., 2015), suggesting the use of longer, more stringent PAM sequences (Müller et al., 2016) and developing Cas9 nickases (Mali et al., 2013; Ran et al., 2013; Shen et al., 2014), which require two sgRNAs and introduce two closely located single-stranded breaks, thus lowering the frequency of unwanted DSBs by doubling the recognition sequence length needed to introduce a DSB.

CRISPR/Cas9 has so far been used for a variety of genome engineering purposes, including gene inactivation (Cong et al., 2013) or activation (Konermann et al., 2015), targeted DNA methylation (Vojta et al., 2016), gene visualization (Chen et al., 2013), genome-wide knockout studies (Wang et al., 2015) and knockin mice generation (Platt et al., 2014; Dow et al., 2015), to name a few. One of the very useful applications of the CRISPR-Cas9 system for the biomedical community, and the one that interests us the most, is the enhancement of gene targeting in mESCs (Zhang et al., 2014). It has recently been reported (Baker et al., 2017) that complex targeting exercises like gene humanization or creation of conditional alleles become a feasible task and require less screening effort by combining the repair of recombinogenic CRISPR-Cas9-induced DSBs with homology templates containing HAs of 1 kb or greater.

1.5. Chromatin structure

Eukaryotic DNA is extremely long – each human cell contains about two meters of DNA. It is intuitively clear that DNA needs to be condensed on multiple levels to fit inside the μm^3-sized eukaryotic nuclei. Indeed, the 10 000-fold DNA condensation is facilitated by packaging proteins which, together with the double-stranded DNA helix, form the nucleoprotein structure called the chromatin. Chromatin is comprised of repeating units called nucleosomes, which are formed by winding of DNA around the proteins called histones. Each nucleosome is an octamer consisting of two H2A-H2B histone dimers and a H3-H4 histone

tetramer, wrapping a stretch of 146 bp of DNA around itself. Further compation is accomplished by the use of H1 linker histone, which packages the nucleosomes into spiraled coils called 30 nm fibers. Nucleosomes can change their positions relative to the DNA seqence they are bound to and thereby substantially contribute to chromatin dynamics and degree of DNA packaging.

Except from being associated with packaging proteins, DNA also extensively contacts proteins and RNA molecules required for gene expression, DNA replication and DNA repair (Alberts et al., 2008). The accessibility of DNA depends on the degree of its histone-mediated condensation. To facilitate binding of proteins needed to execute gene expression, DNA replication and repair, DNA needs to be decondensed by loosening its interation with histones. When histones were first discovered, it was thought that their only role is to compact the genome (Ramakrishnan, 1997). In the meantime it became clear that histones are active players in other processes involving the DNA, and that the regulation of processes such as DNA compaction-dependent transcription involves extensive and precise post-translational modifications (PTMs) of histones, such as acetylation, methylation, phosphorylation and ubiquitination, to name a few (reviewed in Bannister and Kouzarides, 2011). These modifications mostly take place on N-terminal tails of histones, which protrude from the nucleosome structure, making them easily accessible targets of PTMs.

The study of histone PTMs and their functional role, currently one of the most attractive scientific areas, has proposed a well-established hypothesis of the histone code. That is, specific patterns of histone PTMs serve as molecular codes recognized and used by non-histone proteins to regulate a variety of chromatin functions (Tollefsbol, 2012). For example, acetylation of H3 lysine at position 27 (H3K27ac) leads to local chromatin decondensation and transcriptional activation (Tie et al., 2009). Acetylation neutralizes positively charged lysine residues, thus weakening their interation with the negatively charged DNA. Aside from changing the affinity of histone-DNA interactions, PTMs often serve as docking sites for other proteins, such as histone-modifying enzymes, chromatin remodelers and transcription factors.

Chromatin structure is determined not only by histones and their modifying proteins, but also by non-histone architectural proteins. Structural maintenance of chromosome (SMC) proteins can be divided into cohesin and condensin classes. As their names suggest, cohesins regulate the sister chromatid cohesion, while condensins contribute to chromosome condensation. These two protein classes are therefore essential for faithful chromosome segregation (Ball AR Jr and Yokomori, 2001). Topoisomerase II exhibits a decatanating activity, promoting chromosome disentanglement, and cooperates with SMCs as part of the protein network that stabilizes long range-contacts between chromosomal segments (Maeshima and Laemmli, 2003). HP1 proteins demarcate heterochromatin regions. They play an essential role in establishing constitutive heterochromatin at pericentromeric and telomeric sites (Krouwels et al., 2005).

CCCTC-binding factor (CTCF) facilitates the creation of highly conserved genomic functional units called topologically associated domains (TADs) by acting as a domain insulator (Vietri Rudan M et al., 2015). Worth mentioning are also the HMGB (High mobility group B) proteins, which exhibit multiple nucleus- and cytoplasm-associated functions, among which is the induction of local DNA conformational changes, i.e. the increase in DNA flexibility, through their sequence non-specific DNA binding activity (Zhang et al., 2009). All of the aforementioned proteins are found in eukaryotes, with prokaryotes having a different set of proteins responsible for DNA compaction, but whose identities and roles reach outside of the scope of this thesis.

1.6. DNA flexibility and looping

Double-stranded DNA is the primary genetic storage material of all prokaryotic and eukaryotic organisms and some viruses. Being one of the longest natural polymers, it is subject to physical disortions such as bending, twisting, looping (Peters and Maher, 2010) and supercoiling. DNA bending represents sequence-dependent DNA deformation driven by an external force, while curvature refers to the sequence-dependent DNA shape in a hypothetical thermal bath without an external force (Peters and Maher, 2010). DNA supercoiling represents the over- or underwinding of a DNA strand around its axis (Irobalieva et al., 2015). All of these processes are not only a consequence of DNA's physical properties but also play an important role in DNA storage and function. Some of the most prominent processes DNA is involved in include replication, transcription and recombination (Bellomy and Record, 1990; Schleif, 1992; Dillon et al., 1997). All of these depend on DNA looping. Sites of various distances dispersed along the DNA must be able to interact with each other to facilitate such looping, as is the case with transcription factors bound to promoters contacting distant enhancer elements. Apart from being extremely long, DNA is also surprisingly stiff in comparison to other natural polymers. It is precisely this DNA stiffness that makes the looping and packaging into μm^3-range prokaryotic cells or eukaryotic nuclei difficult. Looping efficiency depends on the flexibility of DNA or chromatin and the distance between two sites involved in looping (Ringrose et al., 1999).

Answering the questions surrounding DNA flexibility and looping are of great importance for understanding the above mentioned fundamental biological processes and development of applications such as DNA engineering in nanotechnology (Rothemund, 2006; Seeman, 2005) and DNA delivery in gene therapy (Peters and Maher, 2010). To quantitatively assess the inherent stiffness of a polymer like DNA and the energy cost of deforming it, a measure called persistance length (P, in nm) is used (Manning, 1988). Another useful measure expresses the probability that the two sites on the same DNA molecule will meet by looping as the local molar concentration of one site in respect to the other (j_M, in M) (Kratky and Porod,

1949; Jacobson and Stockmayer, 1950; Hagerman, 1988; Rippe et al., 1995). A single coherent equation for j_M that covers both short and long DNA distances has been developed by Ringrose et al. (1999).

In their work, Ringrose et al. have taken advantage of SSR as a method of choice for DNA flexibility measurement. The frequency at which FLP- or Cre-mediated excition of the DNA substrate occurs depends on the probability that *FRT* or *loxP* sites will contact each other, which is in turn determined by their distance, orientation, accessibility and DNA flexibility. Changes in excision frequency are therefore directly correlated with DNA flexibility and permit its measurement when the sequence distance is experimentally controlled and factors like sequence accessibility and orientation are considered.

The flexibility of DNA *in vitro* has been well studied, yielding a persistence length of 50 nm or around 150 bp (Hagerman, 1988) and an optimum looping distance of 500 bp. DNA in mammalian chromatin is able to form smaller loops and has an optimum looping distance of 200 bp (Figure 4). Furthermore, the apparent persistence length is reduced to 27 nm *in vivo* (Ringrose et al., 1999). This shift in apparent DNA flexibility was attributed to chromatin properties. The propensity towards protein-mediated DNA loop formation is greatly affected by two factors: DNA compaction and binding site occlusion (Hildebrandt and Cozzarelli, 1995). The compaction involves wrapping of DNA around the nucleosomes and higher order packaging and should contribute to increasing apparent DNA flexibility. The presence of nucleosomes and non-histone proteins can, on the other hand, occlude binding of proteins that bring the DNA sites into interaction. The relative contribution of these factors to DNA flexibility changes is based on the length of the DNA distance between the loop ends (Ringrose et al., 1999).

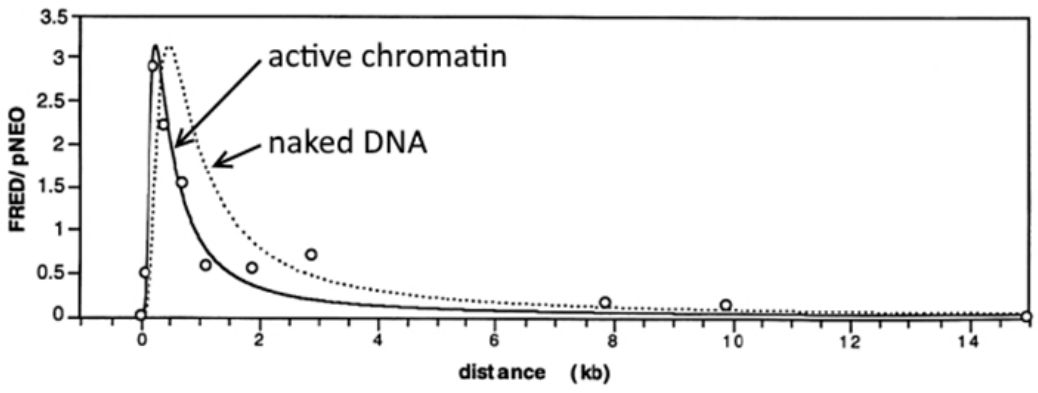


Figure 4. Dependence of DNA flexibility measured by SSR efficiency on the distance between the excision sites on a linear piece od DNA.
The y-axis represents relative recombination frequency. The figure was addapted from Ringrose et al., 1999.

Aside from the factors highlighted by Ringrose et al., there have been many other proposed factors that influence apparent local DNA flexibility and propensity of DNA to loop *in vitro* and *in vivo*, such as the concentrations of multivalent and divalent ions, conformational form of DNA (A-, B- or Z-DNA), histone-mediated DNA compaction, the presence of the linker histone or non-histone architectural proteins (such as the ones reviewed in 1.5., in particular the HMGB proteins) and linker DNA length connecting the neighboring nucleosomes (Perisic and Schlick, 2017.; Perez et al., 2004; Perez et al., 2008; Zhang et al., 2009; Ross et al., 2001). Of particular interest to us are the DNA flexibility changes brought on by local changes in DNA compaction, for instance the ones that occur during transitions from a transcriptionally active to a repressed state. It is believed that DNA compaction is accompanied by alterations in physical properties of chromatin, including its propensity to loop and engage in long-term interactions, which is crucial for regulatory interactions between enhancers and promoters. On top of flexibility changes caused purely by DNA compaction, pluripotent stem cells exhibit a peculiar nucleosomal phenomenon termed hyperactive or breathing chromatin (Zwaka, 2006), that could further influence the special status of DNA flexibility within these cells. To be able to investigate the effects of DNA compaction on *in vivo* DNA flexibility, we will use the *mNanog* gene as the environment for performing Cre-*loxP* based excision experiments. *mNanog* has been chosen because it is transcriptionally active in mESCs, contributing to the major characteristic of embryonic stem cells, ie. their pluripotency (Chambers et al., 2007). However, it is rapidly shut off upon differentiation of ESCs due to repressive chromatin remodelling (Wu et al., 2014) and *Nanog* promoter methylation by methylases Dnmt3a and Dnmt3b (Li et al., 2007), resulting in stable silencing due to heterochromatinization. Indeed, *Nanog* mRNA levels are decreased 10^5-fold in neural stem cells (NSCs; Lee et al., 2010). Differentiation of ESCs into NSCs is routinely performed in the host lab using the method of (Ying et al., 2003). *mNanog* is therefore a particularly suitable gene for bringing heterochromatinization into connection with altered DNA flexibility.

As a continuation of the Ringrose et al. experiments, this project will try to improve on the previously used methods by employing the novel advancements in gene targeting (such as CRISPR-Cas9-assisted increase of targeting frequency). This means that our cell lines will have all the TCs necessary for flexibility measurement inserted into the same (*mNanog*) gene, contrary to the previous strategy of random TC insertion into the genome. This will make the project more conclusive, as the possible locus-dependent effects will be avoided. The significance of performing the described measurements and describing the mechanisms guiding the DNA flexibility and looping changes are numerous. We might for example be able to better predict the efficiency with which Cre-*loxP* recombination will occur depending on the gene's degree of compaction, or better estimate how likely the endogenous genetic elements are to contact each other depending on their distance and degree of compation. Furthermore, we will contribute to answering

some of the central questions associated with DNA flexibility, such as the extent to which proteins are needed to make the inherent DNA properties compatible with biological requirements (Peters and Maher, 2010).

2. AIM OF THE THESIS PROJECT

The aim of this lab project was to:

1) Prepare the targeting constructs for *mNanog* targeting in R1 mESCs
2) Test the CreERT2 inducibility in these cells
3) Design and produce 5' and 3' Southern probe for *mNanog* targeting
4) Target *mNanog* with 10 different targeting constructs and screen for correct insertions by Southern blot
5) Develop a strategy for the PCR-based verification of positive colonies

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Instrumentation

Instrument name	*Company*
5415D Benchtop Microcentrifuge	Eppendorf
5415R Refrigerated Benchtop Microcentrifuge	Eppendorf
5804 Cell Culture Centrifuge	Eppendorf
BAS 2040 development cassette with imaging plate	Fuji
Blue light LED transilluminator	Carl Roth
CASY Cell Counter (model TTC)	Roche
Eporator	Eppendorf
FLA-3000 Flourescence and Radioactivity Imager	Fuji
G:Box gel documentation system	Syngene
Gene Pulser Xcell™ Electroporation System	Bio-Rad
Multitron Standard bacterial incubation shaker	Infors HT
NanoDrop ND-1000 spectrophotometer	PeQLab
Thermocycler ep Gradient S	Eppendorf
Thermomixer compact	Eppendorf
UV Stratalinker 1800	Stratagene

3.1.2. Chemicals and reagents

Product name	*Company*
2-mercaptoethanol	Sigma
2-propanol	VWR
4-hydroxytamoxifen (4-OHT)	Sigma-Aldrich
Agarose	Serva
Ammonium Acetate	Sigma-Aldrich
Ampicillin	Sigma-Aldrich
Blasticidin S	Thermo Fisher
Bovine Serum Albumin (BSA)	Sigma-Aldrich
Bromphenol Blue	Merck Millipore
CASYton	Omni Life Science
Chicken Serum	Gibco Invitrogen
CT-99021	ABCR
D-glucose	Merck Millipore
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich
DMEM-High Glucose with GlutaMax	Gibco Invitrogen
ESGRO (LIF)	Millipore
Ethanol	VWR
Ethidium Bromide	Sigma-Aldrich
Ethylenediamine tetraacetic acid disodium salt dihydrate	Merck Millipore
Fetal Calf Serum	GAA
Formaldehyde	Merck Millipore
Gelatin from cold water fish skin	Sigma-Aldrich

Geneticin (G418 sulfate)	Gibco Invitrogen
Glutaraldehyde	Merck Millipore
Hydrochloric acid	VWR
Kanamycin sulfate	Sigma-Aldrich
L-Arabinose	Sigma-Aldrich
L-Glutamine	Gibco Invitrogen
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich
MEM non-essential aminoacids	Gibco Invitrogen
N-Lauroylsarcosine sodium salt	Sigma-Aldrich
PD0325901	ABCR
Penicillin	Gibco Invitrogen
Phenol/chloroform/isoamyl alcohol (25:24:1)	Sigma-Aldrich
Phosphate buffered saline (PBS)	Gibco Invitrogen
Potassium ferricyanide	Sigma-Aldrich
Potassium ferrocyanide	Sigma-Aldrich
Puromycin	Sigma-Aldrich
Radioactive [^{32}P]dCTP – 3000 Ci/mmol	Perkin Elmer
Sodium acetate hydrate	Sigma-Aldrich
Sodium chloride (NaCl)	VWR
Tri-Sodium citrate dihydrate	Merck Millipore
Sodium dodecyl sulfate (SDS)	Carl Roth
Di-Sodium hydrogen phosphate dihydrate	Carl Roth
Sodium hydroxide pellets (NaOH)	Merck Millipore
Sodium pyruvate	Gibco Invitrogen
Spermidine	Sigma-Aldrich
Streptomycin	Gibco Invitrogen
Tetracyclin	Sigma-Aldrich
Tris	Carl Roth
Tris-Acetate-EDTA (TAE)	AppliChem
Tris-Borate-EDTA (TBE)	AppliChem
Trypsin/EDTA	Gibco Invitrogen
X-Gal	Fermentas
Xylene cyanol FF	Sigma-Aldrich

3.1.3. Kits, enzymes, ladders

Product name	*Company*
1 kb DNA ladder	NEB
dNTP mix (10 mM each)	NEB
High Prime DNA labeling kit	Roche
Invisorb Fragment CleanUp	Stratec
Invisorb Spin DNA Extraction Kit	Stratec
Phusion DNA polymerase	NEB
Plasmid Maxi Kit	QIAGEN
Proteinase K	Merck
Restriction nucleases + buffers	NEB
RNase A	QIAGEN
Taq DNA polymerase	NEB
T4 ligase and buffer	NEB

3.1.4. Buffers and media

6x Saline sodium citrate (SSC; Southern blot)

- 300 mL of 20x SSC + 700 mL ddH_2O

Table 1. 20x SSC (Southern blot)

Reagent	**Amount**
ddH_2O	up to 1 l
Na-citrate	87.66 g
NaCl	20 g

Table 2. Southern blot DNA loading dye

Reagent	**Amount**
Bromphenol blue	25 mg
Xylene Cyanol FF	25 mg
Sucrose	4 g
Deionized water	Up to 10 ml

Table 3. Cell lysis buffer (Southern blot)

Reagent	**Amount**
ddH_2O	up to 0.5 l
1 M Tris (pH 7)	5 ml
0.5 M EDTA	10 ml
1 M NaCl	5 ml
N-Lauroylsarcosine	2.5 g
Add fresh before use: **20 mg/ml Proteinase K**	1:100 V/V

Table 4. CG-wash (Southern blot)

Reagent	**Amount**
ddH_2O	908 ml
0.5 M Na_2HPO_4 x $2H_2O$ pH 7.2	40 ml
0.5 M EDTA pH 8	2 ml
20 % SDS	50 ml

Table 5. Denaturation buffer (Southern blot)

Reagent	**Amount**
ddH_2O	up to 1 l
NaCl	87.66 g
NaOH	20 g

Depurination solution (Southern blot)

- 21.5 ml of 37 % HCl and 978.5 ml ddH_2O

Table 6. Fixative solution (X-gal staining)

Reagent	**Amount**
1x PBS	11.78 ml
37 % formaldehyde	675 µl
25 % glutaraldehyde	50 µl

Table 7. Hybridization buffer (Southern blot)

Reagent	**Amount**
ddH_2O	148 ml
0.5 M Na_2HPO_4 x $2H_2O$ pH 7.2	500 ml
0.5 M EDTA pH 8	2 ml
20 % SDS	350 ml

Table 8. mESC medium

Reagent	**Amount**
DMEM-High Glucose + GlutaMax	148 ml
FCS	75 ml
10 mM 2-mercaptoethanol	5 ml
Penicillin-Streptomycin 5000 units/ml	5 ml
200 mM L-Glutamine	5 ml
100x Non-essential aminoacids	5 ml
100 mM Sodium pyruvate	5 ml
self-made recombinant Leukemia inhibitory factor (LIF)	1 ml
10 mM CT-99021	75 µl
10 mM PD-0325901	25 µl
Antibiotic*	**Amount**
50 mg/ml Geneticin G418	2 ml
1 mg/ml Puromycin	500 µl
5 mg/ml Blasticidin S	500 µl

*Up to two mammalian antibiotics can be combined. In case multiple antibiotic resistance needs to be selected for, antibiotic combinations can be changed every four days during medium changes/cell splits.

Table 9. Neutralization buffer (Southern blot)

Reagent	**Amount**
ddH_2O	up to 1 l
Tris	121.14 g
NaCl	87.66 g

Adjust pH to 7 before use.

Table 10. Staining solution (X-gal staining)

Reagent	Amount
1x PBS	9.9 ml
1 M $MgCl_2$	25 µl
50 mM Potassium ferricyanide*	1.25 ml
50 mM Potassium ferrocyanide*	1.25 ml
200 mg/ml X-gal#	62.5 µl

*Can be kept at 4°C for 14 days.
#Heat the solution to 50°C before adding X-gal to avoid crystallization.

Table 11. 0.1 % trypsin for cell culture

Reagent	Amount
PBS	500 ml
2.5 % trypsin	20 ml
Chicken serum	5 ml
EDTA	0.1 g
D-glucose	0.5 g

3.1.5. Primers

Table 12. List of primers

Primer name	Primer sequence 5'→3'	Application
mNanog-Bsd-replacement-F	GCTTACATAAACAGTAATACAAGGGG TGTTATGGCCAAGCCTTTGTCTCAAG	Recombineering
mNanog-Bsd-replacement-R	TGGCAACTAGAAGGCACAGTCGAGGCTG ATCAGCGTTAGCCCTCCCACACATAACC	Recombineering
mNanog 5' F	GTACGAATTCGCGGGTGTCCTTATCACTCT	5' Southern blot probe production
mNanog 5' R	AGACGGATCCGAGCTTCAGACCTTGGCTCC	5' Southern blot probe production
mNanog 3' F	GTACGAATTCTAAGTTAACGTAACTCCTGATGC	3' Southern blot probe production

mNanog 3' R	AGACGGATCCAGGTGGAATTAGTCTGGGCTT	3' Southern blot probe production
850/1566 F	CGGTGGATGTGGAATGTGTG	Gene targeting validation
850/1566 R	CCTGCAGAGTGACAGGACAT	Gene targeting validation
2359 F	TTTTGAAGCGTGCAGAATGCC	Gene targeting validation
2359 R	CTGGGGATAGCTGCAATGGAT	Gene targeting validation

3.1.6. Strains

Table 13. List of *E. coli* and mESC strains

Strain name	**Characteristics**	**Reference**
E. coli		
GB05	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139Δ(ara, leu)7697 galU galK λ⁻ rpsL nupG fhuA::IS2 recET redα, phage T1-resistent*	Fu et al., 2010
mESCs		
E14	E14 ROSA26 CreERT2 ROSA26	Konstantinos Anastassiadis, BIOTEC, Germany
R1 13-18	R1 ROSA26 CreERT2 ROSA26 *loxP*-Neo-*loxP*-*lacZ*	K. Anastassiadis
R1 13-23	R1 ROSA26 CreERT2 ROSA26 *loxP*-Neo-*loxP*-*lacZ*	K. Anastassiadis

3.1.7. Plasmids

Table 14. List of plasmids

Plasmid name	Plasmid description	Application	Reference
pETgA	pSC101-BAD-ETgA-TetR	Recombineering (source of HR-promoting proteins)	Fu et al., 2012
pBsd-replace	pR6K-hyg-NLS-mTIR1-IRES-gb3-Bsd	Recombineering (source of Bsd cassette)	Stewart lab
pNeo-70	pBR322-amp-*mNanog*-Neo-70	Recombineering (source of targeting construct)	Jun Fu, Shandong University, China
pNeo-75	pBR322-amp-*mNanog*-Neo-75		
pNeo-90	pBR322-amp-*mNanog*-Neo-90		
pNeo-120	pBR322-amp-*mNanog*-Neo-120		
pNeo-150	pBR322-amp-*mNanog*-Neo-150		
pNeo-200	pBR322-amp-*mNanog*-Neo-200		
pNeo-400	pBR322-amp-*mNanog*-Neo-400		
pBlue	pBlueScript KS	5'- and 3'-*mNanog* Southern blot probe production	Addgene
pCas9	Caggs-Cas9-NLS	Gene targeting (source of Cas9)	Stewart lab
pgRNA-*mNanog*	pBR322-U6-*mNanog*-tracrRNA-Amp	Gene targeting (source of *mNanog* gRNA)	Stewart lab
pBsd-70	pBR322-amp-*mNanog*-Bsd-70	Gene targeting (targeting contruct)	This work
pBsd-75	pBR322-amp-*mNanog*-Bsd-75		
pBsd-90	pBR322-amp-*mNanog*-Bsd-90		
pBsd-120	pBR322-amp-*mNanog*-Bsd-120		
pBsd-150	pBR322-amp-*mNanog*-Bsd-150		
pBsd-200	pBR322-amp-*mNanog*-Bsd-200		
pBsd-400	pBR322-amp-*mNanog*-Bsd-400		
pBsd-850	pBR322-amp-*mNanog*-Bsd-850		
pBsd-1566	pBR322-amp-*mNanog*-Bsd-1566		
pBsd-2359	pBR322-amp-*mNanog*-Bsd-2359		

3.2. Methods

3.2.1. Genetic engineering methods

Small-scale DNA purification (miniprep)

Plasmid DNA was purified via alkaline lysis miniprep using QIAGEN buffers. A single bacterial colony was picked and incubated overnight at 37°C with 950 rpm shaking in 1.8 ml of LB medium supplemented with appropriate antibiotics. The following day, cultures were centrifuged at 13200 rpm for one minute and the supernatant was decanted. 200 µl of P1 (resuspension) buffer were added to each pellet and shaken at 1300 rpm at room temperature (RT) for 10 minutes to resuspend the cells. 200 µl of P2 (lysis) buffer were added and the samples were shaken three times for five seconds at 1300 rpm. 200 µl of P3 (neutralization) buffer were added and the samples were shaken three times for five seconds at 1300 rpm. The tubes were inverted gently three times. The lysates were then centrifuged at 13200 rpm for 10 minutes. The supernatants were transferred into new 1.5 ml tubes with 500 µl of 100 % isopropanol and vigorously mixed, then centrifuged at 13200 rpm for 20 minutes. The supernatants were decanted and the pellets washed with 500 µl of 70 % ethanol. The DNA pellets were dried at 42°C to remove residual ethanol. They were dissolved in deionized water and stored at 4°C for short-term or at -20°C for long-term storage. The added water volumes depend on the plasmid ori: for pBR322 50 µl, for pBlueScript 60 µl.

Large-scale DNA purification (maxiprep)

The plasmid DNA was purified using the QIAGEN Plasmid Maxi Kit. The bacterial colonies were picked and incubated overnight at 37°C with 200 rpm shaking in 200 ml LB medium supplemented with appropriate antibiotics. The following day, cultures were centrifuged at 4°C at 6000 rpm for 15 minutes and the supernatant was decanted. The pellets were resuspended on ice in 10 ml of P1 (resuspension) buffer by pipetting up and down and vortexing. 10 ml of P2 (lysis) buffer were added, the samples were thoroughly mixed and incubated at RT for five minutes. 10 ml of P3 (neutralization) buffer were added, the samples were throughly mixed and incubated on ice for 20 minutes. During this incubation, QIAGEN-tips 500 were equilibrated with 10 ml QBT (equilibration) buffer using folded qualitative filter papers. The samples were centrifuged at 4°C at 6000 rpm for two minutes and the supernatants were transferred into the QIAGEN-tips 500 and cleared by gravity flow. After the supernatant was cleared, the filter papers were discarded and the QIAGEN-tips washed two times with 30 ml QC (washing) buffer. The QIAGEN-tips were transferred into 50 ml tubes and the DNA was eluted with 15 ml QF (elution) buffer. The DNA was mixed with 10.5 ml of 100 % isopropanol and centrifuged at 4°C at 9000 rpm for one hour. The supernatants were decanted, the pellets washed and resuspended in 2 ml of 70 % ethanol and transferred

into 1.5 ml tubes (2 per one sample). They were centrifuged at RT at 13 200 rpm for 10 minutes. The supernatants were removed with a pipette and the DNA pellets dried at 50°C. They were dissolved in 75 µl of deionized water and stored at 4°C for short-term or at -20°C for long-term storage.

Plasmid restriction digestion

Table 15. Mixture for restriction enzyme-mediated plasmid digestion

Reagent	**Volume per sample (µl)**		
H_2O	Up to 10	Up to 40	Up to 100
DNA	4	10	25
Restriction enzyme	0.2	0.8	2
Buffer	1	4	10
Total volume	**10**	**40**	**100**

Bacterial transformation by electroporation

Bacterial cells were incubated with 1.4 ml of LB at 37°C at 950 rpm overnight. 40 µl of the overnight culture were added to 1.4 ml of fresh LB and incubated at 37°C at 950 rpm for two hours to allow the cells to reach the exponential growth phase. The cultures were centrifuged at 4°C at 9000 rpm for 30 seconds and the supernatants were decanted. 1 ml of cold deionized water was added and the pellets were resuspended by vortexing. Samples were centrifuged at 4°C at 10000 rpm for 30 seconds and all but the 30 µl of supernatant was decanted. The appropriate plasmid was added and the contents were transferred into a cooled, 1 mm electroporation cuvette. The plasmids were electroporated using the Eppendorf Eporator at 1350 V. 1 ml of LB was added and the cells were transferred into a new 1.5 ml tube and recovered at 37°C at 900 rpm for one hour. The bacteria were then streaked out onto 10 ml LB plates supplemented with the appropriate antibiotic(s) to obtain single colonies and incubated at 37°C over night.

Dialysis of DNA

10 ml empty plates were filled with deionized water and a Milipore 0.025 µm dialysis membrane was placed on the water surface. The entire volume of the DNA solution was pipetted onto the middle of the membrane and left to dialyze for 30 minutes. The dialyzed samples were collected into 1.5 ml tubes without disturbing the membrane.

Agarose gel electrophoresis

0.8 % gels were used for electrophoresis of DNA samples. 0.48 g of agarose were added to 60 ml of TBE

buffer and heated until the mixture was homogenously transparent. After short cooling, 2 µl of 10 mg/ml EtBr were added and the mixture was poured into the agarose electrophoresis chamber. The combs were added and the gel was left to solidify. The combs were removed, the chamber was filled with TBE until the gel was completely covered, the samples with 1:6 purple NEB loading dye were loaded and the gel was run at 95 V for 40 minutes or until sufficient band separation was accomplished. 1kb NEB ladder was used for to compare the sized of experimental DNA samples.

Polymerase chain reaction (PCR)

Table 16. Mixture and cycling parameters for the Taq-based PCR

Reagent	**Volume per sample (µl)**	**Step**	**Temperature (°C)**	**Time (s)**
dH_20	37.5	Initial denaturation	98	10
10x Taq buffer	5	30 cycles	98	5
dNTPs (10 mM)	2		60	20
DMSO	2		68	30
forward primer (10 µM)	1	Final extension	68	60
reverse primer (10 µM)	1	Hold	4	
5prime Taq	0.5			
DNA	1			
Total volume	50			

Table 17. Mixture and cycling parameters for the Phusion-based PCR

Reagent	**Volume per sample (µl)**	**Step**	**Temperature (°C)**	**Time (s)**
H_20	34.5	Initial denaturation	98	30
5x Phusion buffer	10	30 cycles	98	5
dNTPs (10 mM)	2		55	10
forward primer (10 µM)	1		72	20
reverse primer (10 µM)	1	Final extension	72	60
Phusion Pol	0.5	Hold	4	
DNA (100 ng/µl)	1			
Total volume	50			

DNA purification from the PCR reaction mix

The PCR products were transferred into 1.5 ml tubes and 250 µl of binding buffer were added. The contents were mixed and transferred into a spin filter inside a 2 ml tube. After centrifugation at 11000 rpm for two minutes, the flowthrough was discarded and the spin filters were placed into a new 1.5 ml tube. 20 µl of deionized water were placed onto the center of the filter and incubated at RT for 10 minutes. After centrifugation at 11000 rpm for one minute, the purified DNA was stored at -20°C.

DNA ligation and blue-white screen

Linearized and restricted pBlue and PCR product were incubated at a mass ratio of 1:3 at RT for one hour. The ligation mix contained the following:

Table 18. Mixture for DNA ligation

pBlue	2 µl (25 ng)
PCR product	1 µl (10 ng)
Ligation buffer	2 µl
dH_20	15 µl
T4 ligase	1 µl

After ligation, T4 ligase was heat-inactivated at 65°C for 10 minutes.

LB agar plates were prepared for the blue-white screen by adding a mixture of 40 µl of LB, 10 µl of IPTG and 20 µl of X-gal onto their surface and distributing it evenly. They were kept in the dark until streaking. The ligated DNA was dialysed as previously described. The DNA was transformed into GB05 *E.coli* by electroporation and streaked as previously described. Plates were screened the next day and 12 white colonies per ligation event were picked and miniprepped for purposes of gene cloning validation by plasmid restriction digestion and sequencing.

Purification of DNA fragments from the gel

Linearized plasmid fragments that were subsequently used for [^{32}P]dCTP-labelled Southern probe synthesis were visualized using the blue light LED transilluminator, excised from the gel and transferred into 2 ml tubes. 1 ml of Gel Solubilizer S per tube was added and the samples were incubated at 50°C shaking at 300 rpm for 10 minutes. 500 µl of Binding Enhancer were added, mixed and 750 µl of the mix were loaded into purification columns inside receiver tubes. The samples were centrifuged at 11000 rpm for one minute. The

loading and centrifugation were repeated until the entire mix volume was used. 500 µl of the Wash Buffer were added and centrifuged at 12 000 for 30 seconds. The washing step was repeated. The samples were centrifuged at 13200 rpm for four minutes. 25 µl of deionized water was added onto the filter, incubated at RT for five minutes and centrifuged into a receiver tube at 12 000 rpm for one minute. The samples were stored at 4°C or -20°C.

DNA precipitation

1:10 (V/V) of 3 M sodium acetate and three volumes 100 % ethanol were added to the DNA sample to be precipitated and kept at -80°C for 10 minutes. The samples were centrifuged at 4°C at 13200 rpm for 20 minutes. The supernatant was discarded, the samples were washed with 300 µl of 70 % ethanol and dried at 42°C for 30 minutes. The DNA was resolved in 50 µl of deionized water.

Recombineering

E. coli GB05 cells containing the pETgA plasmid were grown in LB-Tet40 overnight at 30°C and 950 rpm. As a template for the 50 µL PCR reaction, pBsd-replace plasmid was used. The primers used for Bsd cassette amplification are *mNanog*-Bsd-replacement-F and *mNanog*-Bsd-replacement-R. 1 µM of primers and 100 ng of the template were used in the Phusion-mediated PCR reaction. The fresh GB05 pETgA culture was prepared by incubating the overnight culture for two hours in LB-Tet40 at 30°C and 950 rpm. The PCR products were column purified. The expression of recombineering proteins was induced by adding 20 µL of 10 % arabinose for 40 minutes at 37°C and 950 rpm. 70 µg of the PCR product and 150 ng of the NcoI-digested pNeo-(75-400) constructs were used for electroporation into GB05 pETgA cells. The cells were plated onto LS-LB + Bsd plates and subsequently examined for correct recombineering events by DNA isolation, PvuII-mediated digestion and DNA sequencing.

3.2.2. Mammalian cell culture methods

Mouse embryonic stem cell thawing

R1 ROSA26 CreERT2 ROSA26 *loxP*-Neo-*loxP*-*lacZ* mESCs kept in liquid nitrogen were thawed rapidly in a 37°C water bath. The cryo tube contents were added into 4 ml of mESC medium. They were centrifuged at 1000 rpm for five minutes to remove the DMSO. After aspirating the supernatant, the pellet was resuspended in 10 ml of mESC medium, plated onto 10 cm plate and incubated in the 5 % CO_2 incubator.

Mouse embryonic stem cell passaging

The medium was aspirated and the cells were washed with 5 ml of PBS. 1 ml of 0.1 % trypsin was added and incubated for five minutes at 37°C. mESC medium was added to 10 ml of cell suspension. 20 µl of cell suspension were mixed with 8 ml of CASYton buffer and counted using the CASY Cell Counter. The

remainder of the cell suspension was centrifuged at 1000 rpm for five minutes. Cells were resuspended and a defined amount was transferred to new 10 cm dishes with 10 ml mESC medium to achieve the desired cell number, depending on the application. The cells were incubated in the CO_2 incubator.

Mouse embryonic stem cell freezing

The cells were incubated with 1 ml of 0.1 % trypsin for five minutes in the CO_2 incubator and centrifuged at 1000 rpm for five minutes. The supernatant was aspirated and the pellet resuspended in 500 µl of mESC medium. 500 µl of 2x freezing medium (DMEM supplemented with 20 % FCS and 20 % DMSO) were added dropwise to the cell suspension, transferred into cryo tubes and flash frozen at -80°C and moved to liquid nitrogen witin a week.

Transformation of mouse embryonic stem cells by electroporation

The mESCs were washed with 5 ml of PBS, trypsinized and resuspended in mESC medium to inactivate the trypsin. They were counted as described in the section on mESC passaging and centrifuged at 1000 rpm for five minutes. After resuspension in 10 ml of PBS, a desired amount of cells was taken to get $5*10^6$ cells per electroporation. They were centrifuged again at 1000 rpm for five minutes and resuspended in 800 µl of PBS + 10 µg of linear TC + 5 µg of pCas9 plasmid + 15 µg of pgRNA-*mNanog* plasmid. The samples were put into ice-cold 4 mm electroporation cuvettes. Electroporation was conducted using exponential protocol, 250 V, 500 µF and no set capacitance. The cells were recovered at RT for five minutes and resuspended in 4.5 ml of mESC medium. 1 ml of cell suspension (about $1x10^6$ cells) and 9 ml of mESC medium were plated and incubated in the CO_2 incubator. Five plates were used per electroporated sample.

CreERT2 inducibility test

Six samples of R1 ROSA26 CreERT2 ROSA26 *loxP*-Neo-*loxP*-*lacZ* cells were plated in a 6-well plate at a density of $2.5*10^5$ cells/well and selected with Puro. After 24 hours, two of the samples they were treated with mESC + Puro medium suppl. with 10^{-7} M 4-OHT and medium was changed for the other wells without adding the 4-OHT. After 24 hours, the treatment was repeated for the mentioned two samples and two more samples were treated the same way, while the medium was changed without adding 4-OHT for the negative control samples. After 24 hours, all the wells were washed with PBS and the fixative solution was added for two minutes. The wells were washed three times with PBS, the staining solution was added and samples were incubated overnight at 37°C. The following day, cells were washed with PBS and imaged with a bright-field microscope.

3.2.13. Southern blotting

Cells grown in 10 cm plates were incubated with 1ml of Lysis Buffer suppl. with 10 µl of 10 mg/ml Proteinase K at 55°C at 900 rpm overnight. For 24-well plate samples, 200 µl of Lysis Buffer supplemented with proteinase K were used per well.

For 10 cm plate samples, DNA was isolated by adding 500 µl of phenol/chloroform per 500 µl of cell lysate. The mixture was incubated at 4°C at 7 rpm for one hour and then centrifuged at RT at 10000 rpm for 10 minutes. The upper phase was transferred into a new tube, incubated with 500 µl of phenol/chloroform/isoamyl at 4°C at 7 rpm for one hour and then centrifuged at RT at 10000 rpm for 10 minutes. The upper phase was transferred into a tube containing 350 µl of 100 % isopropanol and 20 µl of 3M sodium acetate, mixed and centrifuged at RT at 13200 rpm for 20 minutes. The supernatant was decanted and DNA washed with 500 µl of 70 % ethanol. The pellet was dried at 42°C for one hour. The DNA was dissolved in 200 µl of deionized water and stored at 4°C.

In case of 24-well plate samples, DNA was isolated by adding 400 µl of ice-cold 100 % ethanol suppl. with 20 % of 3M sodium chloride per well and incubating at RT for 30 minutes. The solution was discarded and the wells were washed two times with 300 µl of 70 % ethanol. The plates were dried at 50°C for one hour.

DNA was digested using the following mixture (Table 19):

Table 19. Mixture for mESC gDNA digestion used for Southern blotting

Reagent	**Amount per 1.5 ml tube or per well**
gDNA	20 µg
CutSmart buffer	4 µl
BSA 100x	0.4 µl
Spermidin 100 mM	0.4 µl
Rnase A 10 mg/ml	0.4 µl
Restriction enyzme	3 µl
Deionized water	31.8 µl

The samples were incubated overnight at 37°C. In case of samples from 24-well plates, the entire amount of DNA available in the well after precipitation was used for digestion because less than 20 µg were usually available.
Digested DNA was analysed on a 0.8 % agarose gel prepared using 400 ml of TBE buffer. The 1 kb NEB ladder was diluted 1:50 and 3 µl were mixed with 5 µl of Southern blot DNA loading dye. 5 µl of the dye were added per sample. The gel was run at 80 V for five hours.

The gel was stained for 15 minutes with 5 % EtBr solution and examined for transfer efficiency using the G:box gel documentation system. It was incubated 2 x 15 minutes with 500 µl of denaturation solution, washed with water and incubated 15 minutes with neutralization solution. The transfer system was assembled by laying the plexi-glass cover onto the transfer chamber filled with 400 ml of 20x SSC buffer. The glass was covered with a 2-layer Whatmann paper bridge immersed in the 20x SSC buffer from both ends. The bridge was topped with the gel (flipped upside down), the positively charged nylon transfer membrane, one wet and two dry blotting papers, paper towels and 250 g of weight per gel and left to transfer overnight at RT. The membrane was rinsed with 6x SSC and dried for five minutes at RT. It was UV-crosslinked using the UV Stratalinker 1800 set at auto-crosslink mode (1200 µJ x100) and subsequently baked at 80°C for two hours. It was stored at RT.

The hybridization buffer was pre heated to 55°C. The membrane was placed in the hybridization tube, DNA side facing the inside of the tube. It was rinsed with 30 ml of 6x SSC to make it wet and fix it to the hybridization tube. The 6x SSC was decanted and 8 ml of hybridization buffer were added. The membrane was prehybridized at 63°C at 15 rpm for one hour.

During prehybridization, 100 ng of the probe template DNA were pipetted, filled with deionized water up to 13 µl and incubated at 99°C for 10 minutes. The tube was then immediately cooled down on ice for one minute and short-spin centrifuged. 4 µl of the High Prime mix were added and the mixture was put on ice. 3 µl of CTP [32P-α] were added and incubated at 37°C for 30 minutes.

A G-25 sepharose column was provided and the liquid was removed from the column by centrifugation in a 15 ml falcon with a 1.5 ml collection tube inside at 1000 rpm for two minutes. The collection tube was replaced and the Southern probe was added to the middle of the column and spun down at 1000 rpm for four minutes. The column was discarded and the Southern probe transferred into a fresh 1.5 ml tube. 1:100 dilution of the Southern probe was made to measure the counts per minute (CPM). The rest was incubated at 99°C for 10 minutes, then cooled down on ice and spun down. The hybridization buffer was decanted and 7 ml of the fresh hybridizatin buffer was added to the hybridization tube. The Southern probe was added to the bottom of the tube. The system was incubated at 63°C at 15 rpm overnight.

The hybridization buffer was decanted and the membrane incubated with 20 ml of CG-wash at 61.5°C at 15 rpm for 20 minutes. The washing step was repeated. The CG-wash was decanted and the membrane was wrapped in a plastic wrap and put into the BAS development cassette. The imaging plate was exposed to bright light twice for 30 minutes to remove old readings and put on top the the wrapped membrane, inside the BAS development cassette for 24 hours. The film was developed using the FLA-3000 Imager (16 bits/pixel).

4. RESULTS

4.1. Cell line validation

Two lines of R1 ROSA26 CreERT2 ROSA26 *loxP*-Neo-*loxP*-*lacZ* cells were obtained from Konstantinos Anastassiadis (BIOTEC, Dresden, Germany), namely R1 13-18 and R1 13-23. Special genetic elements relevant for subsequent experiments that are present in those lines are shown in Figure 5. The cells were first weaned off of MEF mitomycin-C inactivated feeder layer and then checked for inducibility of CreERT2 using the X-gal test, as indicated in Figure 6. Upon 4-OHT-mediated CreERT2 induction for 24 or 48 hours, the neo gene flanked by *loxP* sites should be excised. This allows for expression of β-galactosidase, which reacts with X-gal to produce an indigo colored product. Both R1 13-18 and R1 13-23 showed negligible background CreERT2 nuclear activity and strong CreERT2 inducibility upon 4-OHT administration. Because both lines showed the same inducibility, R1 13-18 line was chosen at random for the targeting experiments.

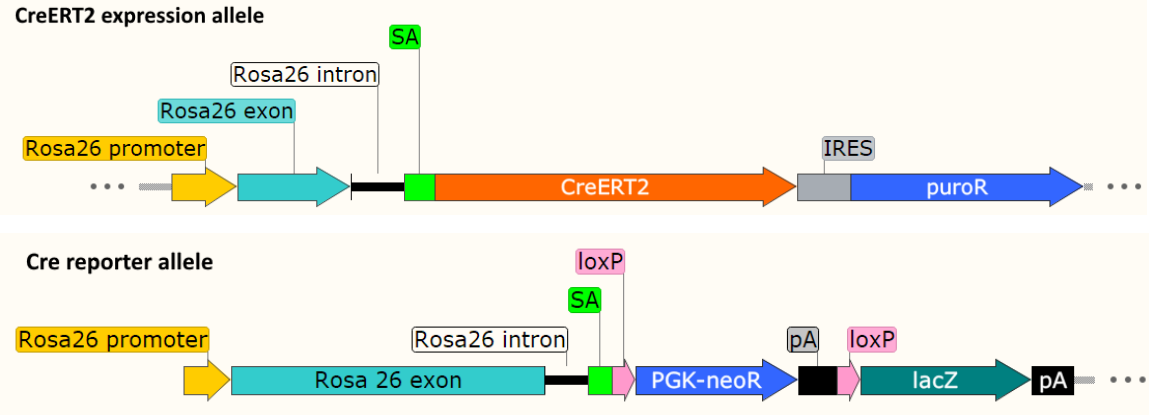


Figure 5. Special genetic elements inserted into each ROSA26 allele of the R1 mESCs.
The upper picture represents the CreERT2 expression cassette. CreERT2 is constitutively expressed from the ROSA26 promoter and its cellular localization is controlled by 4-OHT administration. The PuroR gene was used during the cell line production for positive selection of the CreERT2 cassette insertion and was continuously selected for later on to prevent the loss of the CreERT2 expression. The lower picture represents the *lacZ* Cre reporter cassette. The role of this allele in validation of CreERT2 inducibility is described in the text. This allele will also be used for normalization in DNA flexibility measurements, since the distance of *loxP* sites within this locus is the same in all the targeted lines that we have subsequently produced.

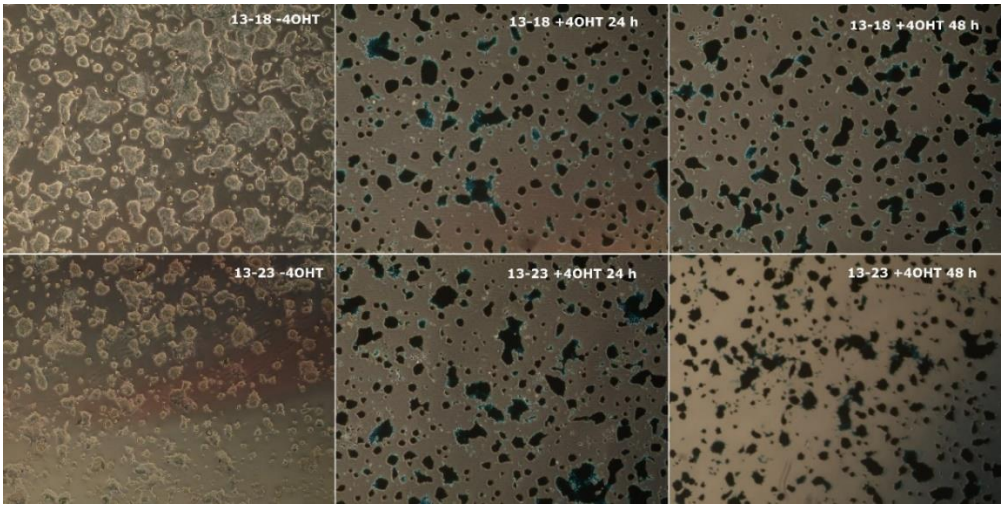


Figure 6. X-gal test of CreERT2 inducibility on R1 13-18 and R1 13-23 mESC.
The leftmost samples represent negative control and were not treated with 4-OHT. The samples in the middle column were treated with 4-OHT for 24 hours and the rightmost samples were treated for 48 hours, as indicated on the picture.

4.2. Targeting construct preparation

To obtain TCs suitable for targeting of R1 13-18 mESCs, the Neomycin resistance gene (Neo) contained in the TCs obtained from Jun Fu (Shandong University, State Key Laboratory of Microbial Technology, China) had to be exchanged for a Blasticidin S resistance gene (Bsd). This was done because Neo is already integrated in the genome of the aforementioned cell line so another resistance gene was necessary to select for targeting events. To achieve this, we used recombineering (homologous recombination-mediated genetic engineering *in E. coli*), a genetic engineering method partly developed (Murphy, 1998; Zhang et al., 1998) and commonly used in our lab. The method takes advantage either of phage lambda HR-mediating proteins Redα and Redβ or prophage Rac proteins RecE and RecT, which carry out the same function. To inhibit *E. coli*'s exonuclease activity, its RecBCD exonuclease and recombination protein complex needs to be inhibited by introduction of the lambda-encoded Gam protein. Most *E. coli* strains used for cloning are RecA-deficient to eliminate unwanted recombination events and thus increase the stability of the cloned DNA. However, it has been noticed (Wang J et al., 2006) that transient expression of RecA during recombineering events enhances the number of successful recombinations, presumably by allowing *E. coli* to better survive the DNA transformation procedures.

We used the temperature-sensitive pETgA plasmid (see Materials) for recombineering purposes in *E. coli* GB05 strain. The ETgA operon (RecE RecT Gam RecA) is controlled by the arabinose-inducible pBAD promoter. The TC plasmids containing Neo cassettes that needed to be exchanged for Bsd are pBR322-amp-*mNanog*-Neo-(70-400), which we abbreviate as pNeo-(70-400). All the other TCs already contained

a Bsd gene (pBsd-850, -1566 and -2359). The numbers in the parentheses reflect the distance of one *loxP* site to another, namely 70, 75, 90, 120, 150, 200, 400, 850, 1566 and 2359. To be more precise, this distance is the size of one *loxP* site plus the spacer between the *loxP* sites that has been left after FLP-*FRT* and Dre-*rox* recombination. TC schemes are depicted in Figure 7. The most complex constructs (ie. pBsd-850, -1566, and -2359) additionally contain sites for FLP-*FRT* and Dre-*rox* recombination, which enables us to increase the effective number of *loxP* distances from 10 to 23. For further clarification, refer to Table 20.

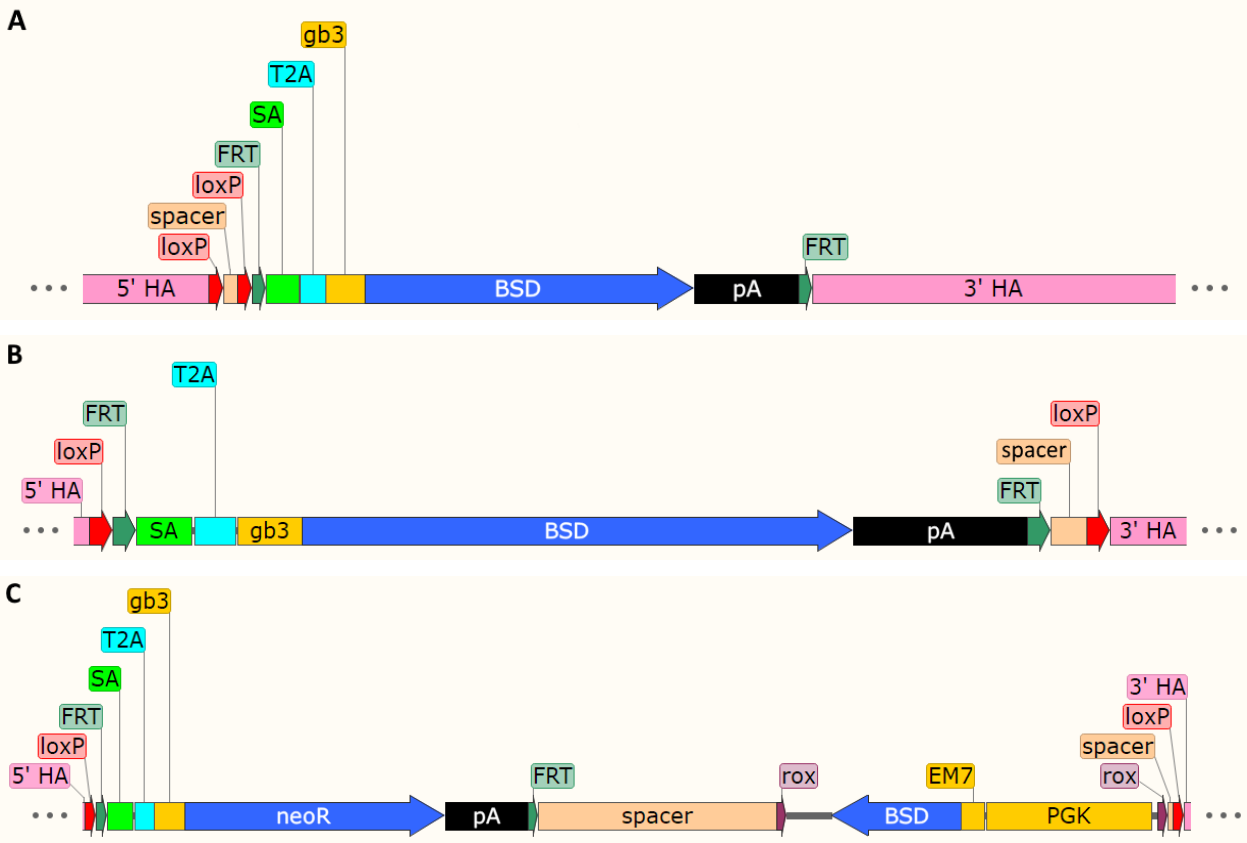


Figure 7. Gene targeting construct schemes.

A) *mNanog*-Bsd-(70-90) scheme. Numbers in parentheses (70, 75, 90) represents the size of one *loxP* site and intervening DNA up to beginning of another *loxP* site. 3'-HA is 4519 bp and 5'-HA is 2140 bp long. They are not fully shown. SA: 5' splice acceptor; T2A: viral T2A peptide; gb3: mammalian globotriaosylceramide promoter; pA: polyA. **B)** *mNanog*-Bsd-(120, 150, 200, 400) scheme. For more information, see 7.A. **C)** *mNanog*-Neo-Bsd-(850, 1566, 2359) scheme. EM7: synthetic T7-promoter-based bacterial promoter for Bsd expression in bacteria; PGK: mouse phosphoglycerate kinase 1 promoter. For more information, see 7.A. Note: vectors for TC cloning additionally contain pBR322 ori and ampicillin resistance gene.

Table 20. Summary of *loxP* distances present in targeting constructs.

For constructs which additionally contain *FRT* and *rox* sites, intermediate *loxP* distances are possible after FLP-*FRT* or Dre-*rox* recombination. The intermediate distances differ depending on whether FLP-*FRT* or Dre-*rox* recombination is conducted first. Distances which are used for plasmid naming are colored in red. Note that the total number of possible *loxP* distances is 23.

Size in bp before any excision	Intermediate size in bp if *Dre* excised first	Intermediate size in bp if *FRT* excised first	Size in bp after *Dre* and *FRT* excision
70	No *Dre* or *FLP* sites present		
75	No *Dre* or *FLP* sites present		
90	No *Dre* or *FLP* sites present		
1045	Only *FLP* sites present		120
1075	Only *FLP* sites present		150
1125	Only *FLP* sites present		200
1325	Only *FLP* sites present		400
3359	2183	2026	850
4075	2899	2742	1566
4868	3692	3535	2359

Linear plus linear recombineering (Fu et al., 2012) was used to replace the Neo gene with the Bsd gene. The recombineering protocol is schematically represented and described in Figure 8**.** After the completion of the protocol, seven Tet-resistant colonies per construct were miniprep-isolated as described in Genetic engineering methods. The DNA was digested with PvuII to verify the accuracy of the recombineering procedure. Figure 9 depicts the clones that showed the desired digestion pattern and that were further verified by DNA sequencing of junctions between the 5'- or 3'-end of the Bsd gene and the surounding TC DNA (Eurofins Genomics Sequencing Service). These clones were subsequently used for targeting of the *mNanog* gene.

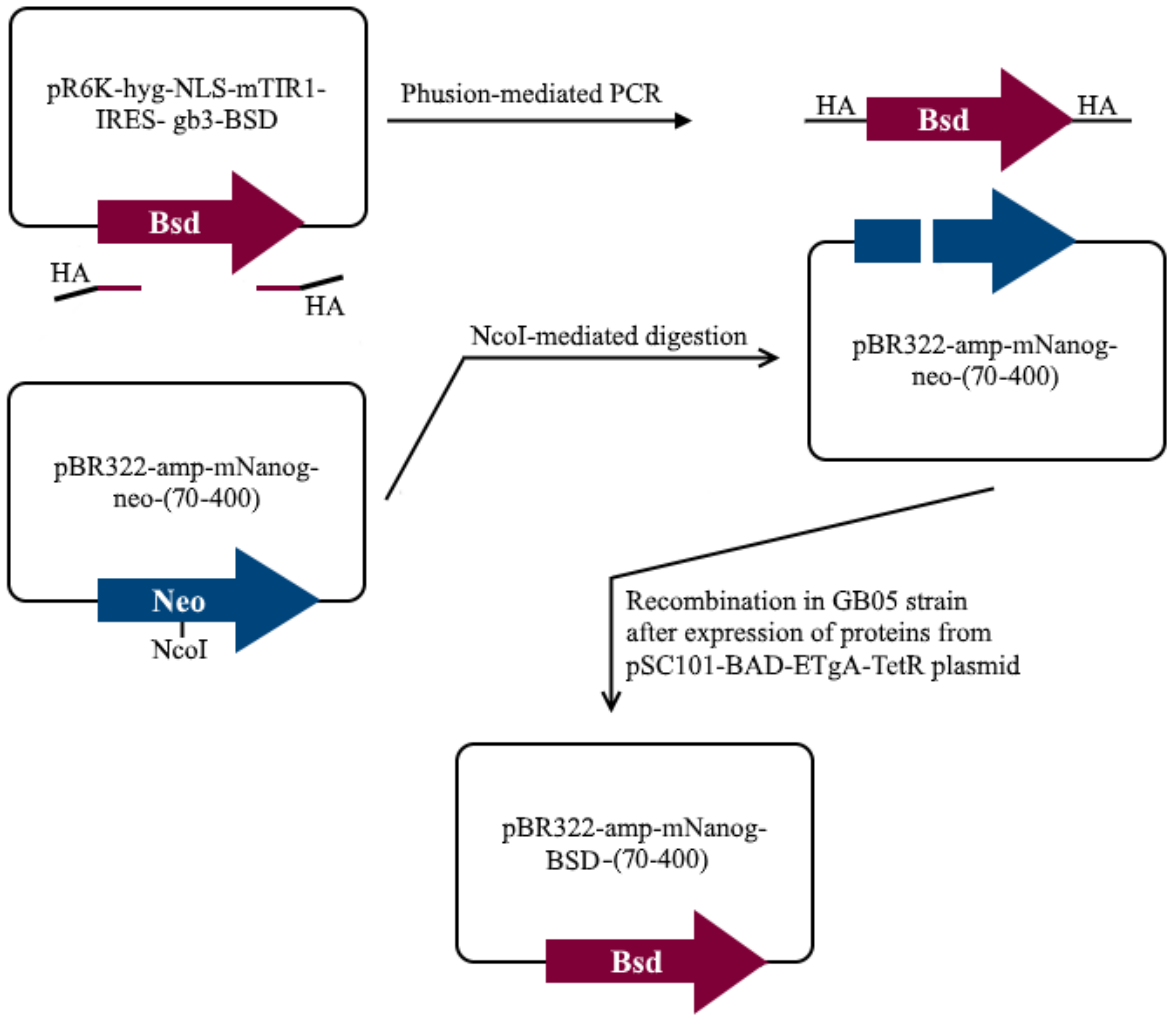


Figure 8. Scheme of the recombineering experiment used to exchange the Neo for the Bsd cassette in 7 out of 10 plasmids used for *mNanog* gene targeting.

Homology arms (HA) present on the primers are homologous to the pNeo-(70-400) sites surrounding the Neo cassette and facilitate homolohous recombination. For detailed description of the protocol, refer to Genetic engineering methods.

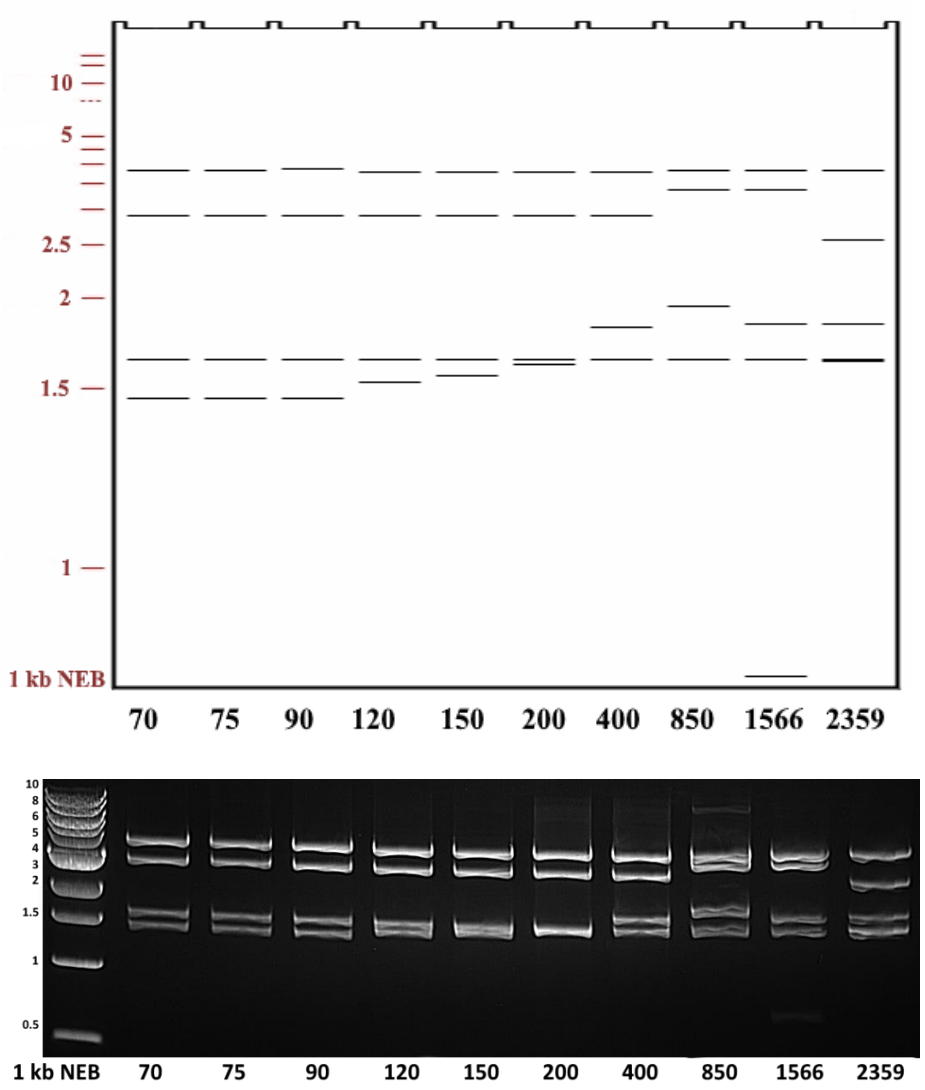


Figure 9. PvuII-digested pBsd-(70-2359) after the recombineering experiment.

The upper picture represents the expected PvuII digestion pattern for all 10 plasmids produced with the Gene Construction Kit (Textco Biosoftware). The lower picture represents the experimentally obtained digestion pattern.

4.3. 5' and 3' Southern probe template production

Screening of the targeted mESCs by Southern blot often requires two probes to ensure that both ends of the TC are faithfully integrated into the mESC genome. To obtain PCR products used for the subsequent 5' and 3' probe template production, the following primers were used: *mNanog* 5' probe F, *mNanog* 5' probe R, *mNanog* 3' probe and *mNanog* 3' probe R. Taq-based PCR was performed using the wild type E14 mESC DNA and the above mentioned primers and the amplification was assessed by agarose electrophoresis. The PCR products were purified from the PCR mix. pBlue plasmids from GB05 cells were miniprepped and digested with BamHI-HF and EcoRI-HF. The same digest was done with the PCR products to produce the

same sticky ends. After agarose electrophoresis and gel extraction, pBlue plasmid and the PCR product were ligated using the T4 ligase, electroporated into GB05 cells and plated onto LB plates with IPTG and X-gal. White colonies were picked, grown overnight, and miniprepped. The cloned plasmids were digested using BamHI-HF and EcoRI-HF to verify the insert size. The majority of the digested cloned plasmids showed the expected pattern (Figure 10). After sequencing, 5' clone #4 and 3' clone #13 were chosen for further experiments.

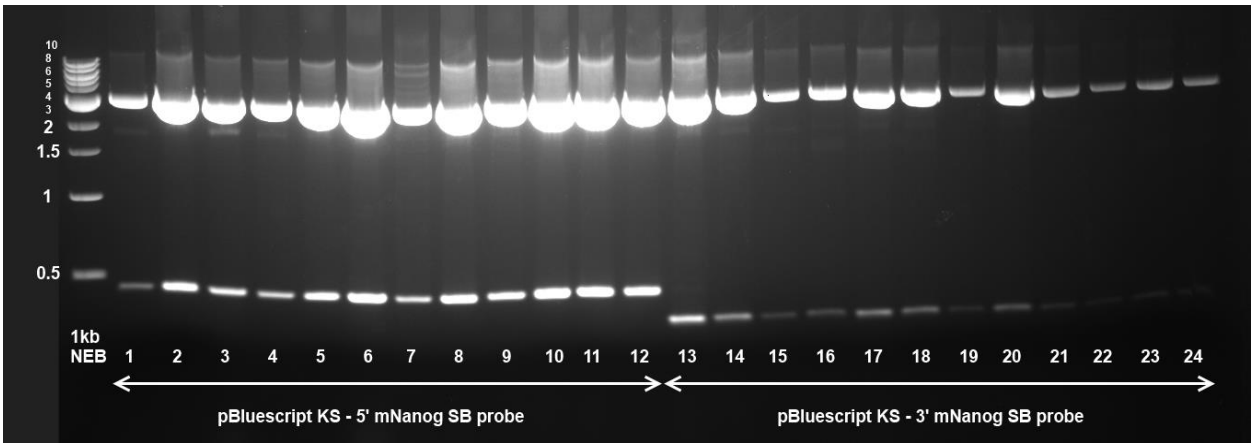


Figure 10. BamHI and EcoRI digest of BluescriptKS + 5'/3' *mNanog* SB probe.
Expected product sizes are 435 bp for 5' probe and 315 bp for 3' probe. The plasmid backbone yields a 3 kb product.

4.4. *mNanog* targeting and screening

mNanog gene targeting experiment involved the electroporation of each of the ten circular maxiprepped plasmids carrying the TC, namely pBsd-(70-2359) together with pCas9 and pgRNA-*mNanog* plasmids. The sequence of crRNA portion of the sgRNA is TAAGAATGTTTCTACTTAGG. The gene targeting mechanism is shown in Figure 11. Figure 12 depicts the smallest TC cassette (Bsd-70) inserted into the *mNanog* background, with the binding site of the 5'-*mNanog* probe used for Southern blot testing indicated within the first intron. The 3'-*mNanog* probe was not used for the testing (see 4.5) due to low specificity for the genomic binding site.

One day after eletroporation, cells were selected with Bsd and G418. The medium was changed every second day, with selection alternating between Bsd + G418, Bsd + Puro and G418 + Puro to avoid loss of intergrated genetic elements. Eight days after plating, the resistant colonies reached the satisfactory size and were picked and plated into 96-well plates, one colony per well. To ensure that positive colonies are encountered, at least 48 colonies were picked per TC. After four days of growing, the colonies were expanded into two 96-well plates and further grown for three days. Cells from one plate were frozen and cells from the second plate were expanded to a 24-well format and grown until confluency was reached.

What followed was a Southern blot protocol with the purpose of testing for correct TC insertion into the *mNanog* gene, described in detail in Methods. The results of the Southern blot-based screen are shown in Figure 13 and the summary of all the isolated positive lines is given in Table 21.

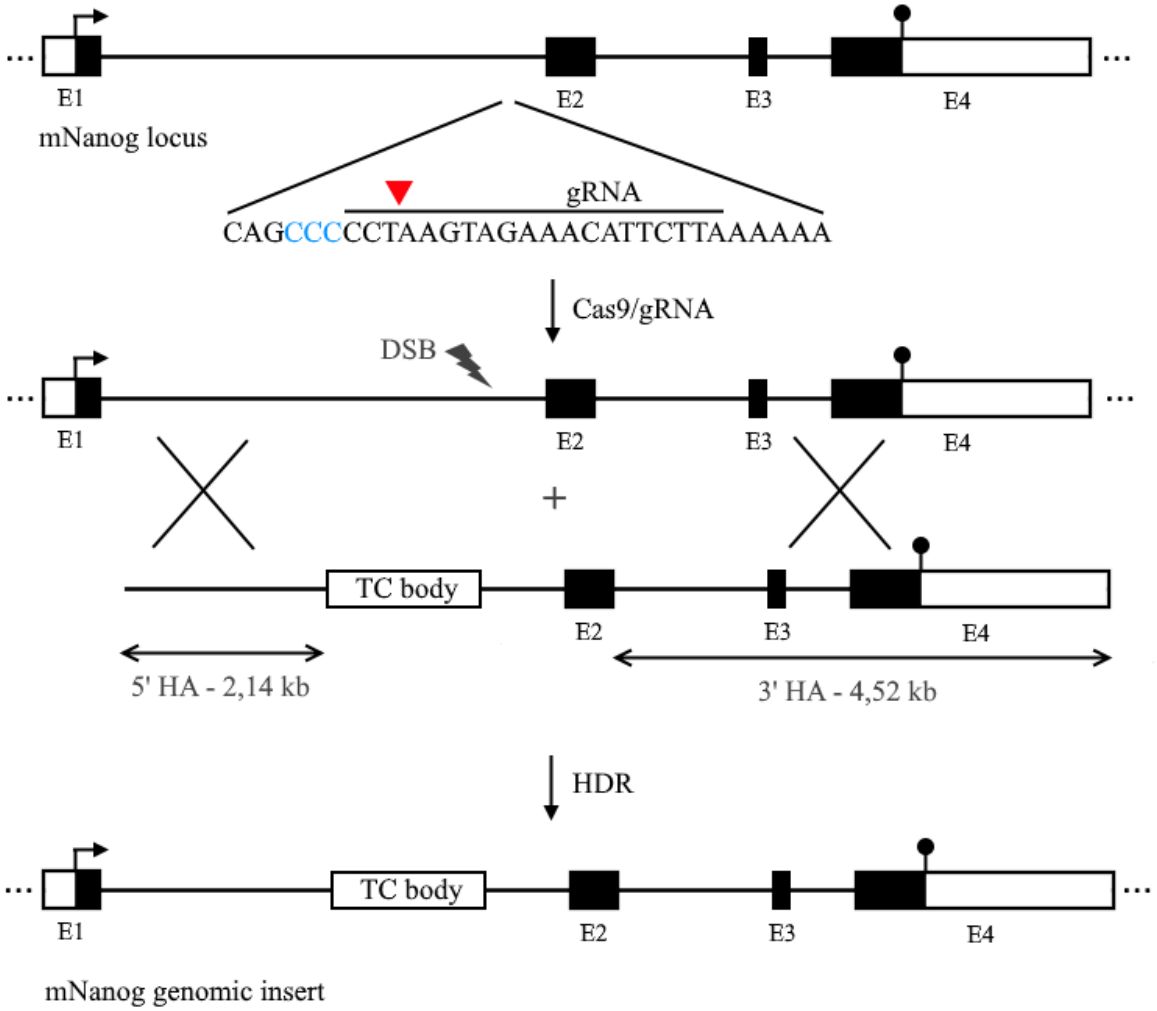


Figure 11. Workflow of the *mNanog* gene targeting experiment.
R1 13-18 mESCs are co-electroporated with one of the 10 prepared TCs together with gRNA- and Cas9-producing plasmids. After the DSB has been induced in the first intron of the *mNanog* gene by gRNA/Cas9 complex, the TC is inserted into the gene by HR. Nucleotides highlighted in blue represent the PAM sequence. The red triangle points to the exact location where the DSB occurs on the DNA strand complementary to the shown gRNA. E1-4 labels are *mNanog* exons and the intervening sequences are introns. The protein coding portion of the exons is depicted in black.

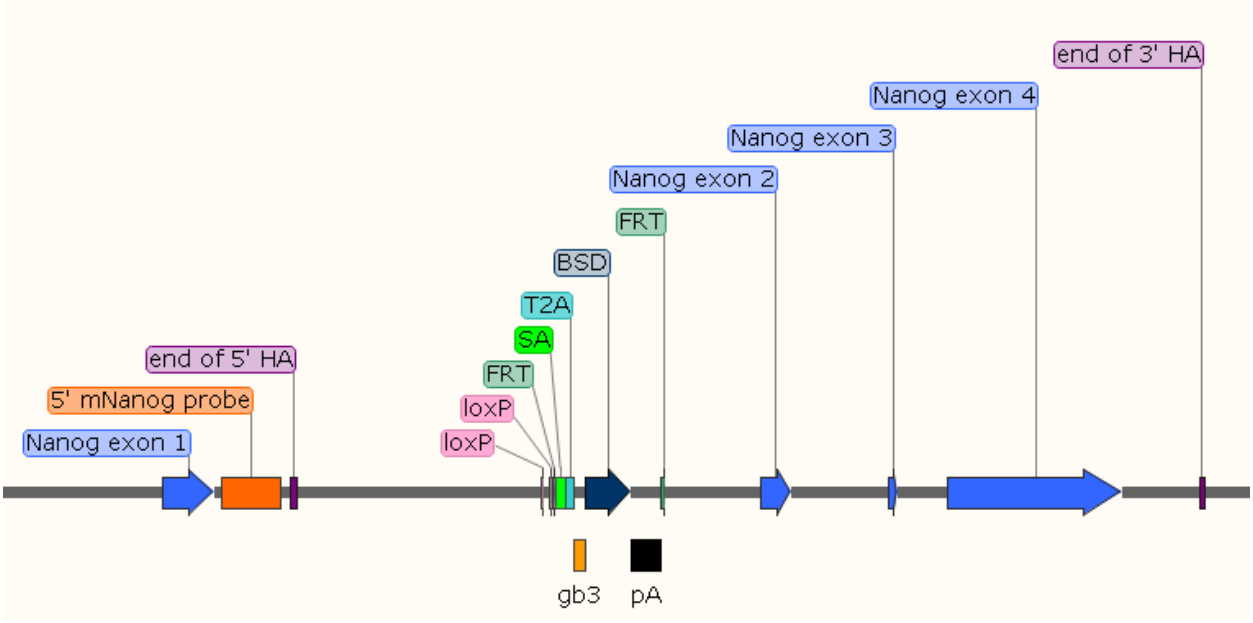
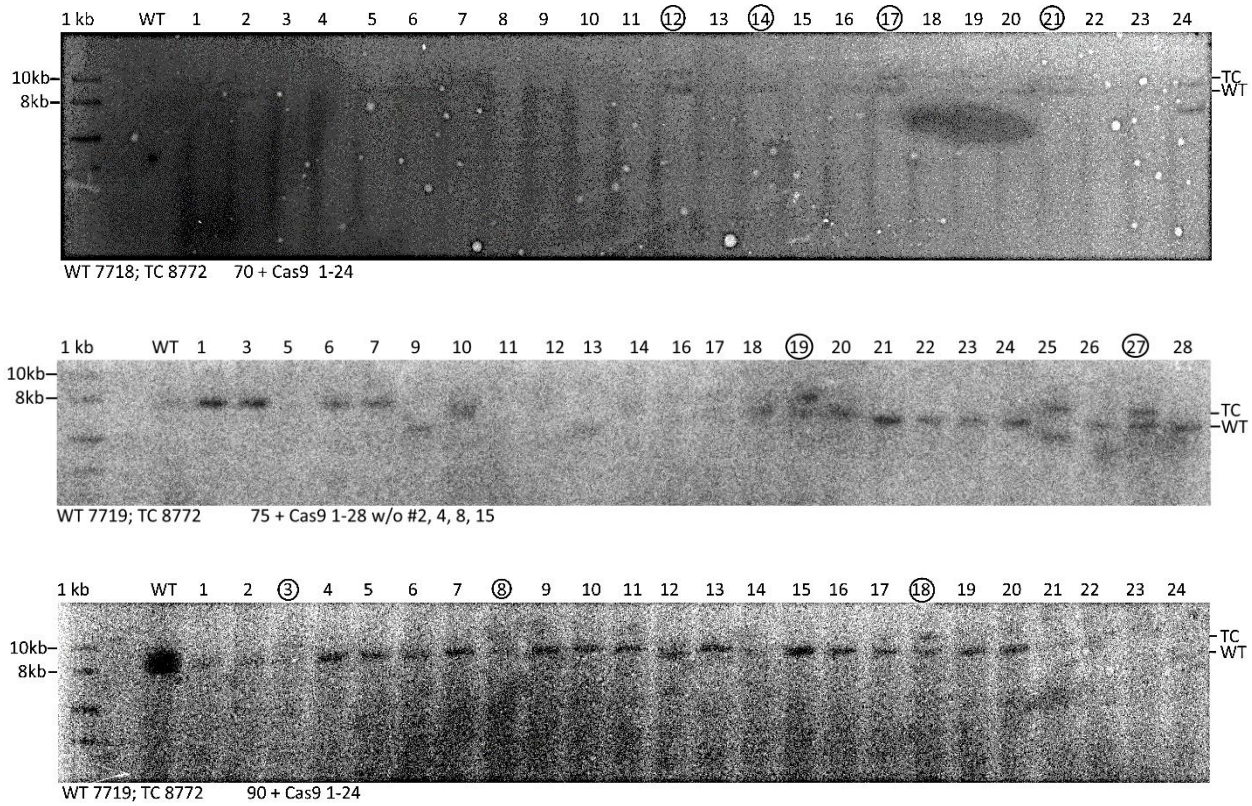
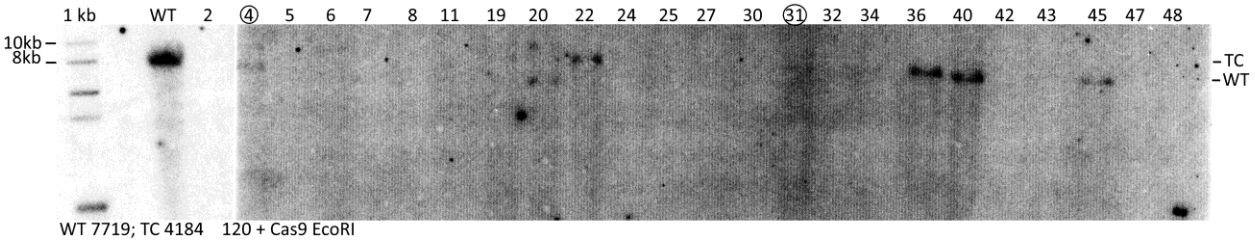
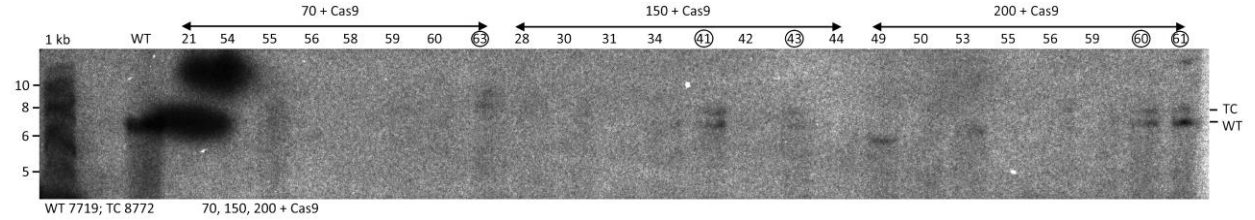


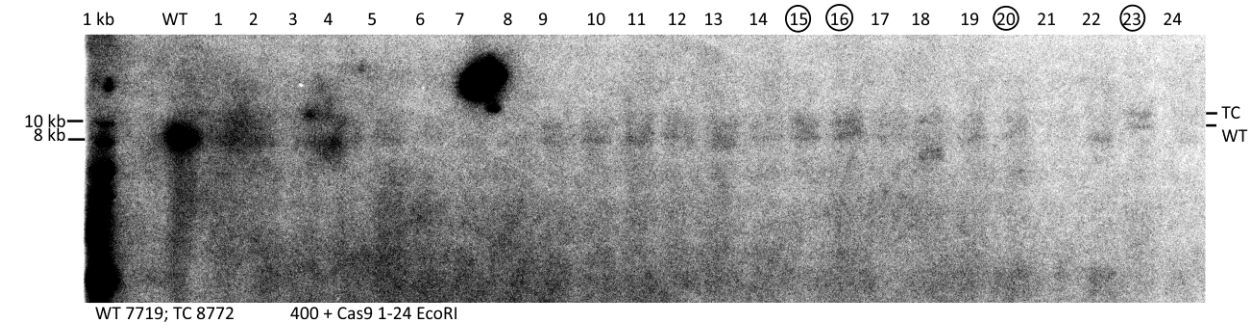
Figure 12. *mNanog*-Bsd-70 cassette inserted into the *mNanog* gene.

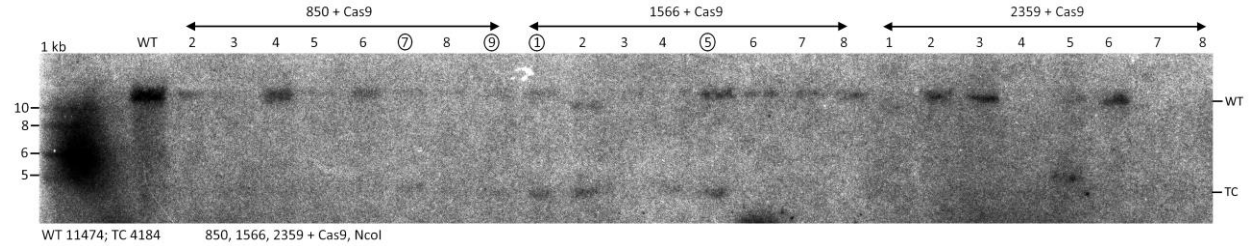
The 5' *mNanog* probe binding site is highlighted in orange. Ends of TC homology arms are highlighted in purple.

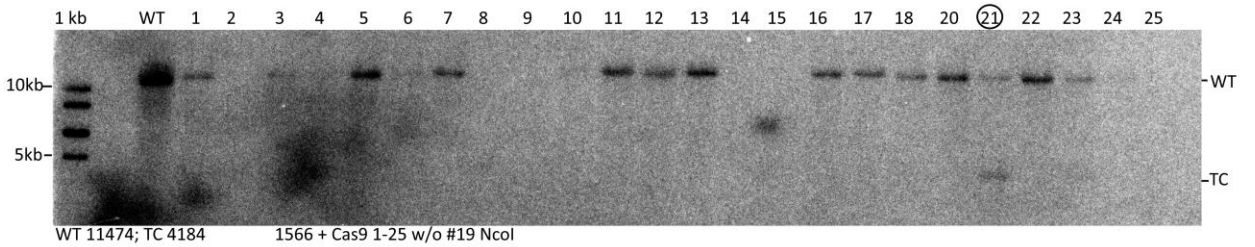



1 kb WT 2 4 5 6 7 8 11 19 20 22 24 25 27 30 31 32 34 36 40 42 43 45 47 48
10kb
8kb
TC
WT
WT 7719; TC 4184 120 + Cas9 EcoRI


70 + Cas9
150 + Cas9
200 + Cas9
1 kb WT 21 54 55 56 58 59 60 63 28 30 31 34 41 42 43 44 49 50 53 55 56 59 60 61
10
8
6
5
TC
WT
WT 7719; TC 8772 70, 150, 200 + Cas9


1 kb WT 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
10 kb
8 kb
TC
WT
WT 7719; TC 8772 400 + Cas9 1-24 EcoRI


850 + Cas9
1566 + Cas9
2359 + Cas9
1 kb WT 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8
10
8
6
5
WT
TC
WT 11474; TC 4184 850, 1566, 2359 + Cas9, NcoI


1 kb WT 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20 21 22 23 24 25
10kb
5kb
WT
TC
WT 11474; TC 4184 1566 + Cas9 1-25 w/o #19 NcoI

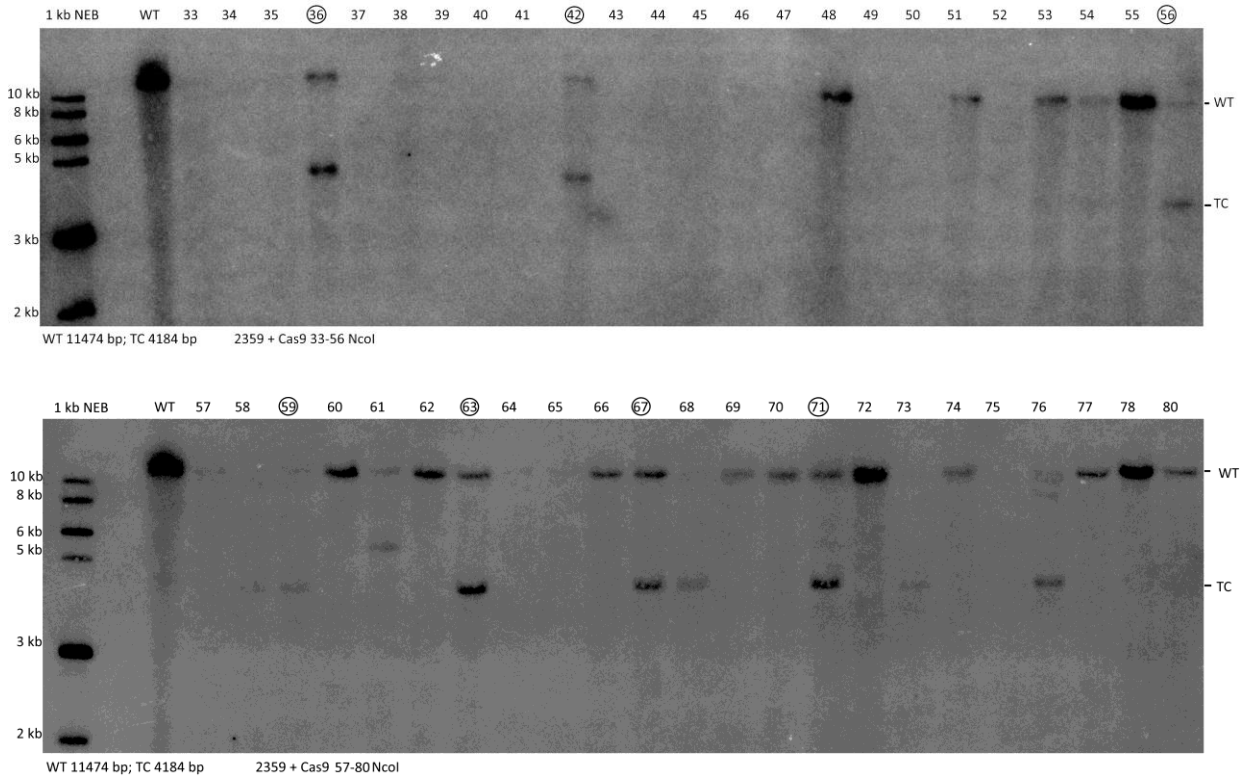


Figure 13. Results of the screening for successful *mNanog* gene targeting events by Southern blot.
Each picture is labeled with the TC number (eg. 70 + Cas9) in the lower left corner. WT and TC expected band sizes are indicated and positive clones are rounded.

Table 21. Synopsis of the positive clones isolated during the Southern blot screening of targeted 13-18 R1 cells.

Targeting construct (*loxP-loxP* distance in bp)	**Positive samples isolated by Southern blot screening**	**Targeting construct (*loxP-loxP* distance in bp)**	**Positive samples isolated by Southern blot screening**
70	14, 17, 63	200	60, 61
75	17, 18, 19, 27	400	15, 16, 20, 23
90	3, 8, 18	850	7, 9, 13, 20
120	4, 20, 31	1566	1, 2, 5, 21, 23, 26, 27
150	11, 41, 43	2359	36, 42, 56, 58, 59, 63, 67, 71, 73

4.5. Further verification of the positive targeting events

We initially wanted to use both 5'- and 3'-*mNanog* probe for targeting verification. However, the 3'-*mNanog* probe showed unspecific binding, manifested as a smear on the gel, to E14 mESC DNA digested with either EcoRI or AfeI (Figure 14). We thus decided to use a different strategy to test the gene targeting cassette-*mNanog* gene 3' junction, ie. PCR-based verification. To this end, we developed the PCR primers appropriate for testing the 3' junctions of the three largest constructs (*mNanog*-Bsd-850, 1566 and 2359). The criteria we used during the design are following: a band of correct size should be produced only if the TC was correctly inserted; a band should not be produced if the sample is WT, if the TC was inserted at the wrong locus or if it was partially inserted, leaving out the 3' end; the primers should be able to later be used for sequencing of the 3' junctions (the product length thus needs to be considerable, ideally above 500 bp).

For 850 and 1566 constructs, we obtained the primers named 850/1566 F and R with the expected product size being 698 bp. For 2359 construct, we obtained the primers named 2359 F and R with the expected product size being 815 bp. Results of the primer testing are shown in the Figure 15.

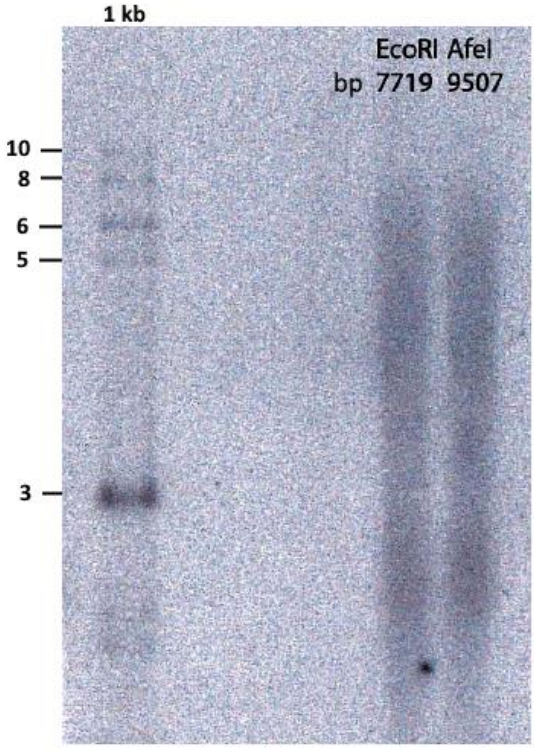


Figure 14. Testing of the 3'-*mNanog* probe on E14 mESC WT DNA digested with EcoRI and AfeI.
The expected product sizes are 7719 bp for EcoRI and 9507 bp for AfeI.

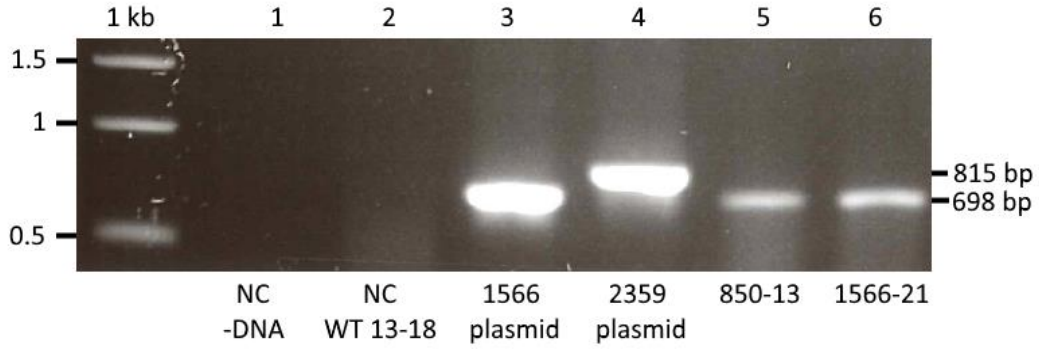


Figure 15. Testing of primers for TC-*mNanog* gene3' junction verification.

Lane 1: negative control (no DNA added, both primer pairs added); lane 2: negative control ("WT" R1 13-18 DNA tested, both primer pairs added); lane 3: positive control (circular pBsd-1566 plasmid tested with 850/1566 primers); lane 4: positive control (circular pBsd-2359 plasmid tested with 2359 primers); lane 5: positive targeting event of pBsd-850 TC into R1 13-18 mESCs tested on clone #13 with 850/1566 primers; lane 6: positive targeting event of pBsd-1566 TC into R1 13-18 mESCs tested on clone #21 with 850/1566 primers). Expected product sizes: 698 bp for lane 3, 5 and 6; 815 bp for lane 4.

5. DISCUSSION

DNA double helix is surprisingly resistant to twisting and bending in comparison to other biopolymers. It is of great interest to elucidate the mechanisms responsible for DNA's remarkable ability to loop and condense and connect these mechanisms with its biological roles. Major advancements have been made in this particular field of biophysics when it comes to DNA flexibility *in vitro*. However, much remains to be discovered regarding the observed changes of DNA flexibility in different *in vivo* contexts. For example, mechanisms involved in possible changes of DNA flexibility *in vivo* upon heterochromatinization have not been comprehensively studied yet. Acquiring knowledge of these concepts will contribute to our understanding of gene expression regulation during development, in physiological contexts as well as desease and it will facilitate improvements in genetic engineering and bionanotechnology. To experimentally test the differences in DNA flexibility upon heterochromatinization we have decided to employ a method of DNA flexibility measurement based on the site-specific recombination and quantitative Southern blot, previously developed by Ringrose et al.,1999. The aim of this thesis was to produce the mouse embryonic stem cell clones containing the site specific recombination sites (*loxP*) of various distances which will be used for purposes of the aforementioned DNA flexibility comparison.

In this project we have successfully tested the mouse embryonic stem cell Cre reporter line for CreERT2-inducibility, prepared the plasmids containing the targeting construct for *mNanog* gene targeting, produced the 5'- and 3'-*mNanog* probe template for Southern blotting, targeted the mentioned cell line with 10 targeting constructs, isolated positive clones for all of them by Southern blotting and developed a PCR-based strategy for further verification of positive clones.

We have managed to develop all 10 desired mESC cell lines achieving frequency of the successfully targeted events of 13.3 %. This frequency is strikingly similar to the one previously reported by our group in Baker et al., 2017, which amounted to 14 %. The authors used the same principle of CRISPR-Cas9-assisted gene targeting, although their TCs significantly differed from ours in size. The aim of their experiment was to humanize the 39.5 kb long mouse *Kmt2d* gene. Their TCs contained a 4 kb 5'-HA (our 5'-HA is 2.14 kb long) and an extremely long 3' HA, the length of which is 42 kb (compared to our 4.52 kb long 3' HA). The length of our largest TC without homology arms is 4.9 kb, which is roughly 10 times less than the size of the mouse *Kmt2d* gene. The conclusion we draw from this is that TCs of vastly different sizes can be targeted with similar frequencies provided that homology arm lengths are adjusted according to the TC body size. Taking this into consideration, it is interesting to notice that we did not observe any reduction in the targeting efficiency when comparing our longest and shortest constructs, although the lengths of homology arms were held constant. On the contrary, our longest constructs Bsd-1566 and -2359

integrated with the highest frequency. This might however not reflect true differences in targeting efficiencies among our constructs of different sizes, but rather improvements in the Southern blot method we implemented during the screening procedure (the longest constructs were screened the last). In light of this, we suspect that the true frequency of our targeting exercise might be higher due to some putative false negatives caused by inconsistent quality of Southern blot results.

There are various ways the Southern blot protocol could be improved. For example, the amount of genomic DNA could be higher. For our purposes this was difficult to achieve because it would include further expansion of targeted clones into 6-well plates, which would be highly time consuming. Moreover, the DNA isolation could be enhanced. We have already investigated this possibility by replacing the invert-and-spill method used while precipitating the DNA with ethanol and sodium chloride with the manual one-by-one aspiration of the mixture, without significant increases in DNA yields. We have also tried replacing sodium chloride with ammonium acetate with similar results. The parameter we concluded was critical during the Southern blot testing is the amount of the probe template used. Not only was it possible to use the amount that is too low, but it also seemed that higher concentrations of the probe template inhibit the radioactive probe synthesis. It is therefore crucial to finetune the amount of the template used. Running a gel overnight at lower voltage (eg. 30 V) might have been helpful in obtaining clearer, more compact bands. The transfer of the DNA to the nylon membrane could be enhanced by taping the ends of the gel with parafilm to avoid the water bridge formation. Furthermore, it might be worth trying to omit washing the membrane in 6x SSC after the transfer, which could help to avoid washing off the DNA before fixing it by UV-irradiation and baking. A standard way to improve the Southern blot efficiency is by decreasing the hybridization or CG-wash temperature, albeit at the risk of introducing higher background. We have also tried modifying these temperatures without much success. Furthermore, we have tested longer development time, for example over the weekend instead of over night, with moderate increases in yields. It is worth noting that our 1 kb ladders were consistently producing bands of reasonable intensity, so the problem does not seem to stem from the transfer or the post-transfer steps but the amount of preparation of the genomic DNA.

Another commonly used method for gene targeting screens is the qPCR. We have avoided employing this method due to difficulties in designing the primers which would unequivocally distinguish between true positive events and all other possible scenarios.

Finishing the screening procedure would require testing the 3' TC-*mNanog* junction of all of the positive clones either by Southern blot, which would include the development of the new, more specific 3'-*mNanog*

probe template, or by PCR and sequencing. Based on the results we presented, it seems that the PCR-based screening is a satisfying strategy for colonies that are previously confirmed as positive by Southern blot.

Figure 15 shows a preliminary test which requires additional controls and samples. For negative controls, the primer pairs need to be tested separately and not added simultaneously to better reflect the conditions under which positive controls and targeted samples were tested. One of the positive controls is missing, ie. the 850 plasmid sample, due to the lack of the miniprep sample at the time of testing. Also, one of the positive 2359 + Cas9 clones should have been included in the test but this was not possible as the test was performed before the first positive 2359 + Cas9 clones were obtained.

Future work on this project will include the verification of positive colonies either by PCR or by using a functioning 3'-*mNanog* probe, 4-OHT-induced *loxP* excision test and preparation of a larger variety of *loxP* distances by Dre/*rox* and FLP/*FRT* recombination on the existing positive clones, optimization and measurement of DNA flexibility in mESCs based on the relative quantitative Southern blot assay and comparison with the previously published results (Ringrose et al., 1999), mESC differentiation into NSCs and their DNA flexibility measurement, and finally, comparison of dependence of DNA flexibility on *mNanog* gene compaction that occurs during differentiation into NSCs.

6. CONCLUSIONS

1) Ten different gene targeting constructs containing the Bsd gene and *loxP* sites of different distances were prepared by recombineering and verified by restriction digestion and sequencing.

2) Templates for 5'- and 3'- Southern blot probes were prepared by gene cloning and verified by PCR. 3'-probe needs to be redesigned due to low specificity for the target genomic site.

3) *mNanog* targeting in the mouse embryonic stem cells was successful and positive colonies were found by Southern blot using the 5'-probe. The efficiency of targeting was 13.3 %. The Southern blot screening procedure needs further optimization so that false negative results can be avoided.

4) A PCR strategy for the validation of the 3' targeting construct-*mNanog* junction was successfully developed and tested and can be used in further experiments as a substitute for the Southern testing with the 3'-*mNanog* probe.

7. LITERATURE

1. Abi-Ghanem J, Samsonov SA, Pisabarro MT (2015): Insights into the Preferential Order of Strand Exchange in the Cre/*loxP* Recombinase System: Impact of the DNA Spacer Flanking Sequence and Flexibility. Journal of Computer-Aided Molecular Design, 29: 271-282.

2. Akopian A, Stark WM (2005): Site-Specific DNA Recombinases as Instruments for Genomic Surgery. Advances in Genetics, 55: 1-23.

3. Anastassiadis K, Fu J, Patsch C, Hu S, Weidlich S, Duerschke K, Buchholz F, Edenhofer F and Stewart AF (2009): Dre Recombinase, Like Cre, Is a Highly Efficient Site-Specific Recombinase in E. coli, Mammalian Cells and Mice. Disease Models & Mechanisms, 9-10: 508-515

4. Anastassiadis K, Glaser S, Kranz A, Bernhardt K, and Stewart AF (2010): A Practical Summary of Site-Specific Recombination, Conditional Mutagenesis, and Tamoxifen Induction of CreERT2. Methods in Enzymology, 477: 109-122.

5. Angrand PO, Woodroofe CP, Buchholz F and Stewart AF (1998): Inducible Expression Based on Regulated Recombination: a Single Vector Strategy for Stable Expression in Cultured Cells. Nucleic Acids Research, 26: 3263-3269.

6. Araki K, Araki M, Yamamura K (1997): Targeted Integration of DNA Using Mutant *lox* Sites in Embryonic Stem Cells. Nucleic Acids Research, 25: 868-872.

7. Awatramani R, Soriano P, Rodriguez C, Mai JJ and Dymecki SM (2003): Cryptic Boundaries in Roof Plate and Choroid Plexus Identified by Intersectional Gene Activation. Nature Genetics, 35: 70-75.

8. Baer A and Bode J (2001): Coping with Kinetic and Thermodynamic Barriers: RMCE, an Efficient Strategy for the Targeted Integration of Transgenes. Current Opinion in Biotechnology, 12: 473-480.

9. Baker O (2015): Adapting Nature's Toolbox for Genome Engineering. PhD thesis

10. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, and Horvath P (2007): CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science, 315: 1709-1712.

11. Bellomy GR, Recort MT (1990): Stable DNA Loops in vivo and in vitro: Roles in Gene Regulation at a Distance and in Biophysical Characterization of DNA. Progress in Nucleic Acid Research and Molecular Biology, 39: 81-128.

12. Bouabe H and Okkenhaug K (2013): Gene Targeting in Mice: a Review. Methods in Molecular Biology, 1064: 315-336.

13. Bradford H, Advait L, and Kulkarni AB (2009): Overview: Generation of Gene Knockout Mice. Current Protocols in Cell Biology, 44: 19.12: 19.12.1–19.12.17.

14. Branda CS, Dymecki SM (2004): Talking about a Revolution: The Impact of Site-Specific Recombinases on Genetic Analyses in Mice. Developmental Cell, 6: 7-28.

15. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV and van der Oost J (2008): Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. Science, 321: 960-964.

16. Buchholz F, Angrand PO, Stewart AF (1998): Improved Properties of FLP Recombinase Evolved by Cycling Mutagenesis. Nature Biotechnology, 16: 657-662.

17. Buchholz F, Refaeli Y, Trumpp A and Bishop JM (2000): Inducible Chromosomal Translocation of AML1 and ETO Genes Through Cre/loxP-Mediated Recombination in the Mouse. EMBO Reports, 1: 133-139.

18. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A (2007): *Nanog* Safeguards Pluripotency and Mediates Germline Development. Nature, 450: 1230-1234.

19. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS and Huang B (2013): Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. Cell, 155: 1479-1491.

20. Chylinski K, Le Rhun A, Charpentier E (2013): The tracrRNA and Cas9 Families of Type II CRISPR-Cas Immunity Systems. RNA Biology, 10: 726-37.

21. Collins EC, Pannell R, Simpson EM, Forster A and Rabbitts TH (2000): Interchromosomal Recombination of Mll and Af9 Genes Mediated by Cre-*loxP* in Mouse Development. EMBO Reports 1: 127-132.

22. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013): Multiplex Genome Engineering Using CRISPR/Cas Systems. Science, 339: 819-823.

23. Deng C, Capecchi MR (1992): Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the Target Locus. Molecular Cell Biology, 12: 3365-3371.

24. Dillon N, Trimborn T, Strouboulis J, Fraser P, Grosveld F (1997): The Effect of Distance on Long-Range Chromatin Interactions. Molecular Cell, 1: 131-399.

25. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ and Root DE (2014): Rational Design of Highly Active sgRNAs for CRISPR-Cas9-Mediated Gene Inactivation. Nature Biotechnology, 32: 1262-1267.

26. Doetschman T, Gregg RG, Maeda N, Hooper ML, Melton DW, Thompson S, Smithies O (1987): Targetted Correction of a Mutant HPRT Gene in Mouse Embryonic Stem Cells. Nature, 330: 576-578.

27. Donnelly ML, Luke G, Mehrotra A, Li X, Hughes LE, Gani D, Ryan MD (2001): Analysis of the Aphthovirus 2A/2B Polyprotein 'Cleavage' Mechanism Indicates not a Proteolytic Reaction, but a Novel Translational Effect: a Putative Ribosomal 'Skip'. The Journal of General Virology, 82: 1013-1025.

28. Downing GJ, Battey JF Jr (2004): Technical Assessment of the First 20 Years of Research Using Mouse Embryonic Stem Cell Lines. Stem Cells, 22: 1168-1180.

29. Feil R, Brocard J, Mascrez B, LeMeur M, Metzger D, Chambon P (1996): Ligand-Activated Site-Specific Recombination in Mice. Proceedings of the National Academy of Sciences of the United States of America, 93: 10887-10890.

30. Feil R, Wagner J, Metzger D, Chambon P (1997): Regulation of Cre Recombinase Activity by Mutated Estrogen Receptor Ligand-Binding Domains. Biochemical and Biophysical Research Communications, 237: 752-7.

31. Feil S, Valtcheva N, and Feil R (2009): Inducible Cre Mice. Methods in Molecular Biology, 530: 343-363.

32. Friedel RH, Plump A, Lu X, Spilker K, Jolicoeur C, Wong K, Venkatesh TR, Yaron A, Hynes M, Chen B, Okada A, McConnell SK, Rayburn H and Tessier-Lavigne M (2005): Gene Targeting Using a Promoterless Gene Trap Vector (‘‘Targeted Trapping’’) is an Efficient Method to Mutate a Large Fraction of Genes. Proceedings of the National Academy of Sciences of the USA, 102: 13188-13193.

33. Friedrich G, Soriano P (1991): Promoter Traps in Embryonic Stem Cells: a Genetic Screen to Identify and Mutate Developmental Genes in Mice. Genes & Development, 5: 1513-1523.

34. Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Müller R, Stewart AF, Zhang Y (2012): Full-Length RecE Enhances Linear-Linear Homologous Recombination and Facilitates Direct Cloning for Bioprospecting. Nature Biotechnology, 30: 440-446.

35. Fu J, Teucher M, Anastassiadis K, Skarnes W, Stewart AF (2010): A Recombineering Pipeline to Make Conditional Targeting Constructs. Methods in Enzymology, 477: 125-144.

36. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK and Sander JD (2013): High-Frequency Off-Target Mutagenesis Induced by CRISPR-Cas Nucleases in Human Cells. Nature Biotechnology, 31: 822-826.

37. Fu Y, Sander JD, Reyon D, Cascio VM and Joung JK (2014): Improving CRISPR-Cas Nuclease Specificity Using Truncated Guide RNAs. Nature Biotechnology, 32: 279-284.

38. Furue M, Okamoto T, Hayashi Y, Okochi H, Fujimoto M, Myoishi Y, Abe T, Ohnuma K, Sato GH, Asashima M, Sato JD (2005): Leukemia Inhibitory Factor as an Anti-Apoptotic Mitogen for Pluripotent Mouse Embryonic Stem Cells in a Serum-Free Medium Without Feeder Cells. In Vitro Cell & Developmental Biology – Animal, 41: 19-28.

39. Grainge I, Jayaram M (1999): The Integrase Family of Recombinases: Organization and Function of the Active Site. Molecular Microbiology, 33: 449-456.

40. Grindley ND, Whiteson KL, and Rice PA (2006): Mechanisms of Site-Specific Recombination. Annual Review of Biochemistry, 75: 567-605.

41. Gu H, Marth JD, Orban PC, Mossmann H and Rajewsky K (1994): Deletion of a DNA Polymerase Beta Gene Segment in T Cells Using Cell Type-Specific Gene Targeting. Science, 265: 103-106.

42. Gu H, Zou YR and Rajewsky K (1993): Independent Control of Immunoglobulin Switch Recombination at Individual Switch Regions Evidenced Through Cre-*loxP* Mediated Gene Targeting. Cell, 73: 1155-1164.

43. Hagerman PJ (1988): Flexibility of DNA. Annual Review of Biophysics and Biophysical Chemistry, 17: 265-86.

44. Hansen GM, Markesich DC, Burnett MB, Zhu Q, Dionne KM, Richter LJ, Finnell RH, Sands AT, Zambrowicz BP and Abuin A (2008): Large-Scale Gene Trapping in C57BL/6N Mouse Embryonic Stem Cells. Genome Research, 18: 1670–1679.

45. Hasty P, Rivera-Perez J, Bradley A (1991): The Length of Homology Required for Gene Targeting in Embryonic Stem Cells. Molecular Cell Biology, 11: 5586-5591.

46. Herault Y, Rassoulzadegan M, Cuzin F and Duboule D (1998): Engineering Chromosomes in Mice through Targeted Meiotic Recombination (TAMERE). Nature Genetetics, 20: 381-384.

47. Hildebrandt ER, Cozzarelli NR (1995): Comparison of Recombination in vitro and in E. coli Cells: Measure of the Effective Concentration of DNA in vivo. Cell, 81: 331-340.

48. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, Meng X, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2009): Efficient Targeting of Expressed and Silent Genes in Human ESCs and iPSCs Using Zinc-Finger Nucleases. Nature Biotechnology, 27: 851-857.

49. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2011): Genetic Engineering of Human Pluripotent Cells Using TALE Nucleases. Nature Biotechnology, 29: 731-734.

50. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013): DNA Targeting Specificity of RNA-Guided Cas9 Nucleases. Nature Biotechnology, 31: 827-832.

51. Irobalieva RN, Fogg JM, Catanese DJ Jr, Sutthibutpong T, Chen M, Barker AK, Ludtke SJ, Harris SA, Schmid MF, Chiu W and Zechiedrich L (2015): Structural Diversity of Supercoiled DNA. Nature Communications, 6: 8440.

52. Jacobson H, Stockmayer WH (1950): Intramolecular reaction in polycondensation. I. Theory of Linear Systems. Journal of Chemical Physics, 18:1600–1606

53. Jaenisch R (1976): Germ Line Integration and Mendelian Transmission of the Exogenous Moloney Leukemia Virus. Proceedings of the National Academy of Sciences of the United States of America. 73: 1260-1264.

54. Jasin M and Rothstein R (2013): Repair of Strand Breaks by Homologous Recombination. Cold Spring Harb Perspect Biol.

55. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012): A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science, 337: 816-21.

56. Keskintepe L, Norris K, Pacholczyk G, Dederscheck SM, Eroglu A (2007): Derivation and Comparison of C57BL/6 Embryonic Stem Cells to a Widely Used 129 Embryonic Stem Cell Line. Transgenic Research, 16: 751-758.

57. Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim JI and Kim JS. (2015): Digenome-Seq: Genome-Wide Profiling of CRISPR-Cas9 Off-Target Effects in Human Cells. Nature Methods, 12: 237-243.

58. Kolb AF (2002): Genome Engineering Using Site-Specific Recombinases. Cloning and Stem Cells, 4: 65-80.

59. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O and Zhang F (2015): Genome-Scale Transcriptional Activation by an Engineered CRISPR-Cas9 Complex. Nature, 517: 583-588.

60. Kratky O and Porod G (1949): Röntgenuntersuchung gelöster Fadenmoleküle. Recueil des Travaux Chimiques des Pays-Bas, 68: 1106-1122.

61. Kuhn R, Schwenk F, Aguet M and Rajewsky K (1995): Inducible Gene Targeting in Mice. Science, 269: 1427-1429.

62. Kuscu C, Arslan S, Singh R, Thorpe J and Adli M (2014): Genome-Wide Analysis Reveals Characteristics of Off-Target Sites Bound by the Cas9 Endonuclease. Nature Biotechnology, 32: 677–683.

63. Lakso M, Sauer B, Mosinger B Jr, Lee EJ, Manning RW, Yu SH, Mulder KL and Westphal H (1992): Targeted Oncogene Activation by Site-Specific Recombination in Transgenic Mice. Proceedings of the National Acadademy of Sciences of the USA, 89: 6232-6236.

64. Lee SH, Jeyapalan JN, Appleby V, Mohamed Noor DA, Sottile V, Scotting PJ (2010): Dynamic methylation and expression of Oct4 in early neural stem cells. Journal of Anatomy, 217: 203-213.

65. Li J-Y, Pu M-T, Hirasawa R, et al (2007): Synergistic Function of DNA Methyltransferases Dnmt3a and Dnmt3b in the Methylation of *Oct4* and *Nanog* . Molecular and Cellular Biology, 27: 8748-8759.

66. Liu P, Jenkins NA and Copeland NG (2002): Efficient Cre-*loxP*-Induced Mitotic Recombination in Mouse Embryonic Stem Cells. Nature Genetics, 30: 66-72.

67. Llames S, García-Pérez E, Meana A, Larcher F and del Río M (2015): Feeder Layer Cell Actions and Applications. Tissue Engineering Part B, Reviews, 4: 345-53.

68. Logie C and Stewart AF (1995): Ligand-Regulated Site-Specific Recombination. Proceedings of the National Acadademy of Sciences of the USA, 92: 5940-5944.

69. Ma Y, Zhang L and Huang X (2014): Genome Modification by CRISPR/Cas9. FEBS Journal, 281: 5186-5193.

70. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L and Church GM (2013): Cas9 Transcriptional Activators for Target Specificity Screening and Paired Nickases for Cooperative Genome Engineering. Nature Biotechnology, 31: 833-838.

71. Martin GR (1981): Isolation of a Pluripotent Cell Line from Early Mouse Embryos Cultured in Medium Conditioned by Teratocarcinoma Stem Cells. Proceedings of the National Academy of Sciences of the USA, 78: 7634-7638.

72. Melton DW (2002): Review: Gene-Targeting Strategies. Methods in Molecular Biology, 180: 151-173.

73. Metzger D, Clifford J, Chiba H, Chambon P (1995): Conditional Site-Specific Recombination in Mammalian Cells Using a Ligand-Dependent Chimeric Cre Recombinase. Proceedings of the National Academy of Sciences of the USA, 92: 6991-6995.

74. Mortensen R (2006): Overview of Gene Targeting by Homologous Recombination. Current Protocols in Molecular Biology, 23: 19-12.

75. Murphy KC (1998): Use of Bacteriophage λ Recombination Functions to Promote Gene Replacement in *Escherichia Coli*. Journal of Bacteriology. 180: 2063-2071.

76. Müller M, Lee CM, Gasiunas G, Davis TH, Cradick TJ, Siksnys V, Bao G, Cathomen T and Mussolino C (2016): Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. Molecular Therapy, 24: 636-644.

77. O'Gorman S, Fox DT and Wahl GM (1991): Recombinase-Mediated Gene Activation and Site-Specific Integration in Mammalian Cells. Science, 251: 1351-1355.

78. Oberdoerffer P, Otipoby KL, Maruyama M and Rajewsky K (2003): Unidirectional Cre-Mediated Genetic Inversion in Mice Using the Mutant *loxP* Pair *lox*66/*lox*71. Nucleic Acids Research, 31(22): e140.

79. Pérez A, Lankas F, Luque FJ, Orozco M (2008): Towards a Molecular Dynamics Consensus View of B-DNA Flexibility. Nucleic Acids Research, 36: 2379-2394.

80. Pérez A, Noy A, Lankas F, Luque FJ, Orozco M (2004): The Relative Flexibility of B-DNA And A-RNA Duplexes: Database Analysis. Nucleic Acids Research, 32: 6144-6151.

81. Perisic O and Schlick T (2017): Dependence of the Linker Histone and Chromatin Condensation on the Nucleosome Environment. The Journal of Physical Chemistry B, 121: 7823-7832.

82. Peters JP and Maher LJ (2010): DNA Curvature and Flexibility *In Vitro* And *In Vivo*. Quarterly Reviews of Biophysics, 43: 23-63.

83. Pluck A and Klasen C (2009): Generation of Chimeras by Microinjection. Methods in Molecular Biology, 561: 199-217.

84. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013): Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell, 152: 1173-1183.

85. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y and Zhang F (2013): Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell, 154: 1380-1389.

86. Raymond CS, Soriano P (2007): High-Efficiency FLP and PhiC31 Site-Specific Recombination in Mammalian Cells. PLoS One, 2: e162.

87. te Riele H, Maandag ER, Berns A (1992): Highly Efficient Gene Targeting in Embryonic Stem Cells through Homologous Recombination with Isogenic DNA Constructs. Proceedings of the National Academy of Sciences of the United States of America, 89: 5128–5132.

88. Ringrose LH, Chabanis S, Angrand PO, Woodroofe C, Stewart AF (1999): Quantitative Comparison of DNA Looping in vitro and in vivo: Chromatin Increases Effective DNA Flexibility at Short Distances. EMBO Journal, 18: 6630-6641.

89. Ringrose LH, Lounas V, Ehrlich L, Buchholz F, Wade R and Stewart AF (1998): Comparative Kinetic Analysis of FLP and Cre Recombinases: Mathematical Models for DNA Binding and Recombination. Journal of Molecular Biology, 284: 363-384.

90. Rippe K, von Hippel PH, Langowski J (1995): Action at a Distance: DNA-Looping and Initiation of Transcription. Trends in Biochemical Sciences, 20: 500-506.

91. Ross ED, Hardwidge PR, Maher LJ 3rd (2001): HMG Proteins and DNA Flexibility in Transcription Activation. Molecular and Cellular Biology, 21: 6598-6605.

92. Sauer B, McDermott J (2004): DNA Recombination with a Heterospecific Cre homolog Identified from Comparison of the Pac-c1 Regions of P1-Related Phages. Nucleic Acids Research, 32: 6086-6095.

93. Schlake T and Bode J (1994): Use of Mutated FLP Tecognition Target (*FRT*) Sites for the Exchange of Expression Cassettes at Defined Chromosomal Loci. Biochemistry, 33: 12746-12751.

94. Schleif R (1992): DNA Looping. Annual Review of Biochemistry, 61:199-223.

95. Seong E, Saunders TL, Stewart CL, Burmeister M (2004): To Knockout in 129 or in C57BL/6: That is the Question. Trends in Genetics, 20: 59-62.

96. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X and Skarnes WC (2014): Efficient Genome Modification by CRISPR-Cas9 Nickase with Minimal Off-Target Effects. Nature Methods, 11: 399-402.

97. Simpson EM, Linder CC, Sargent EE, Davisson MT, Mobraaten LE, Sharp JJ (1997): Genetic Variation Among 129 Substrains and its Importance for Targeted Mutagenesis in Mice. Nature Genetics, 16: 19-27.

98. Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, Mujica AO, Thomas M, Harrow J, Cox T, Jackson D, Severin J, Biggs P, Fu J, Nefedov M, de Jong PJ, Stewart AF, Bradley A (2011): A Conditional Knockout Resource for the Genome-Wide Study of Mouse Gene Function. Nature, 474: 337-342.

99. Smith AJ, De Sousa MA, Kwabi-Addo B, Heppell-Parton A, Impey H and Rabbitts P (1995): A Site-Directed Chromosomal Translocation Induced in Embryonic Stem Cells by Cre-*loxP* Recombination. Nature Genetics, 9: 376-385.

100. Snyder L and Champness W (2007): Molecular Genetics of Bacteria, third edition. ASM Press.

101. Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF, Vignali DA (2004): Correction of Multi-Gene Deficiency in vivo Using a Single 'Self-Cleaving' 2A Peptide-Based Retroviral Vector. Nature Biotechnology, 22: 589-94.

102. Thomas KR, Capecchi MR (1987): Site-Directed Mutagenesis by Gene Targeting in Mouse Embryo-Derived Stem Cells. Cell, 51: 503-512.

103. Thomas KR, Folger KR, Capecchi MR (1986): High Frequency Targeting of Genes to Specific Sites in the Mammalian Genome. Cell, 44: 419–428.

104. Thompson S, Clarke AR, Pow AM, Hooper ML, Melton DW (1989): Germ Line Transmission and Expression of a Corrected HPRT Gene Produced by Gene Targeting in Embryonic Stem Cells. Cell, 56: 313-321.

105. Torres RM and Kuhn R (2003): Laboratory Protocols for Conditional Gene Targeting. Oxford Uninversity Press.

106. Warren D, Laxmikanthan G, Landy A (2008): A Chimeric Cre Recombinase with Regulated Directionality. Proceedings of the National Academy of Sciences of the United States of America, 105: 18278-83

107. Vojta A, Dobrinić P, Tadić V, Bočkor L, Korać P, Julg B, Klasić M and Zoldoš V (2016): Repurposing the CRISPR-Cas9 System for Targeted DNA Methylation. Nucleic Acids Research, 44: 5615–5628.

108. Vooijs M, Jonkers J, Berns A (2001): A Highly Efficient Ligand-Regulated Cre Recombinase Mouse Line Shows that loxP Recombination is Position Dependent. EMBO Reports, 2: 292-297.

109. Wallace HA, Marques-Kranc F, Richardson M, Luna-Crespo F, Sharpe JA, Hughes J, Wood WG, Higgs DR, Smith AJ (2007): Manipulating the Mouse Genome to Engineer Precise Functional Syntenic Replacements with Human Sequence. Cell, 128: 197-209.

110. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013): One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell, 153: 910-918.

111. Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF, Zhang Y (2006): An Improved Recombineering Approach by Adding RecA to Lambda Red Recombination. Molecular Biotechnology, 32:43-53.

112. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM (2015): Identification and Characterization of Essential Genes in the Human Genome. Science, 350: 1096-1101.

113. Wu CY, Feng X, Wei LN (2014): Coordinated Repressive Chromatin-Remodeling of Oct4 And Nanog Genes in RA-Induced Differentiation of Embryonic Stem Cells Involves RIP140. Nucleic Acids Research, 42: 4306-4317.

114. Yagi T, Ikawa Y, Yoshida K, Shigetani Y, Takeda N, Mabuchi I, Yamamoto T, Aizawa (1991): Homologous Recombination at c-fyn Locus of Mouse Embryonic Stem Cells with Use of Diphtheria Toxin A-Fragment Gene in Negative Selection. Proceedings of the National Academy of Sciences of the United States of America, 87: 9918-9922.

115. Ying QL, Stavridis M, Griffiths D, Li M, Smith A (2003): Conversion of Embryonic Stem Cells into Neuroectodermal Precursors in Adherent Monoculture. Nature Biotechnology, 21: 183-186.

116. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008): The Ground State of Embryonic Stem Cell Self-Renewal. Nature, 453: 519-523.

117. Zambrowicz BP, Imamoto A, Fiering S, Herzenberg LA, Kerr WG, Soriano P (1997): Disruption of Overlapping Transcripts in the ROSA Beta Geo 26 Gene Trap Strain Leads to Widespread Expression of Beta-Galactosidase in Mouse Embryos and Hematopoietic Cells. Proceedings of the National Academy of Sciences of the United States of America, 94: 3789–3794.

118. Zhang J, McCauley MJ, Maher LJ, Williams MC, Israeloff NE (2009): Mechanism of DNA Flexibility Enhancement by HMGB Proteins. Nucleic Acids Research, 37: 1107-1114.

119. Zhang Y, Buchholz F, Muyrers JP, Stewart AF (1998): A New Logic for DNA Engineering Using Recombination in Escherichia coli. Nature Genetics 20: 123-128.

120. Zhang Y, Riesterer C, Ayrall AM, Sablitzky F, Littlewood TD, Reth M (1996): Inducible Site-Directed Recombination in Mouse Embryonic Stem Cells. Nucleic Acids Research, 24: 543-8.

121. Zhang Y, Vanoli F, LaRocquea JR, Krawczyka PM, and Jasina M (2014): Biallelic Targeting of Expressed Genes in Mouse Embryonic Stem Cells Using The Cas9 System. Methods, 69: 171-178.

122. Zwaka TP (2006): Breathing Chromatin in Pluripotent Stem Cells. Developmental Cell, 10: 1-2.

8. CURRICULUM VITAE

IVAN MIKIČIĆ

1. EDUCATION

Master's program in Molecular Biology

-University of Zagreb, Faculty of Science | September 2015 – February 2018

- **Erasmus+ exchange year** (study visit and internship)

-Technische Universität Dresden, BIOTEC | October 2016 – October 2017

Bachelor's program in Molecular Biology

-University of Zagreb, Faculty of Science | October 2012 – September 2015

IV. Gymnasium in Zagreb

-Bilingual Croatian-English program | September 2008 – July 2012

2. PERSONAL SKILLS

A) Language skills

- Croatian – native proficiency
- English – proficiency (Cambridge Certificate in Advanced English, 2012, level C2)
- German – advanced (TU Dresden German course, 2017, level C1/C2; Summer German course at the Ludwig-Maximilian University in Munich, 2015, level B2)

B) Laboratory skills

Southern blotting, mammalian cell culture (mouse embryonic stem cells, MEF feeder cells and U2OS cells), PCR, bacterial cell culture, genetic engineering techniques, Western blot, CRISPR-Cas9, Immunoprecipitation, fluorescence and bright-field microscopy

C) Computer skills

- Good command of Microsoft Office tools
- Good command of basic bioinformatics tools (databases, genome browsers, sequence alignment, homology search, phylogenetic trees, molecular graphics)
- Picture Editing Software
- Plasmid Mapping Software (eg. SnapGene)

D) Organizational/communication skills

- Leader of Section for Bioinformatics at Biology Students Association (2016)
- Organization of the Night of Biology at the University of Zagreb (2016, 2014)
- Organization of the Open Day of the Division of Chemistry at the University of Zagreb (2015)

3. ADDITIONAL INFORMATION

A) Theses

- **Master's thesis**: Establishment of recombinant mouse embryonic stem cells for use in DNA flexibility assays | Mentors: Dr. Francis Stewart and Dr. Ivana Ivančić Baće | February 2018
- **Bachelor's thesis**: Experimental Methods in Functional Genomics | Mentor: Kristian Vlahoviček | September 2015

B) Jobs

- Testing of programs for secondary analysis of next generation sequencing reads | Employer: Slobodan Obrenović (freelance programmer) | February – June 2016, Zagreb

C) Summer schools

- International Summer School at the Institute of Molecular Biology, Mainz, Germany | Phosphoproteomics - Cellular response to UV light-induced DNA damage | Mentor: Petra Beli | August – September 2016
- Summer Undergraduate Research Programme at the University of Lausanne, Switzerland | Research topic: Plant developmental genetics | Mentor: Christian Hardtke | July – August 2014

D) Internships

- Translational Medicine lab, Ruđer Bošković Institute, Zagreb | Gateway Cloning for establishment of plasmids for high-throughput protein-protein interaction screens | Mentor: Oliver Vugrek | October 2015 – February 2016
- Microbiology lab, University of Zagreb | Research topic: Bacterial and viral plant pathogen genetics | Mentor: Martina Šeruga Musić | February – June 2015

E) Honors and awards

- Student scholarship awarded by the City of Zagreb | Academic year 2016/2017 and 2015/2016
- Student scholarship awarded by the University of Zagreb | Academic year 2014/2015
- Special award "Mitko Naumovski" for academic excellence by Rotary Club Zagreb Center | 2014
- Scholarship for excellence "Gordan Čačić" by Rotary Club Zagreb Center | 2011 – 2014