Conditional RNA interference Altered Nuclear Transfer and Genome-wide DNA Methylation Analysis

by Alexander Meissner

Conditional RNA Interference, Altered Nuclear Transfer and Genome-wide DNA Methylation Analysis

Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat)

an der Naturwissentschaftlich-Technischen Fakultät

der Universität des Saarlandes

angefertigt am

Whitehead Institute for Biomedical Research

Cambridge, USA

eingereicht von

Dipl. Ing. Alexander Meissner

Februar 2006

Chapter 1- Introduction

Epigenetic reprogramming of the genome

A major goal of current research is focused on understanding the mechanisms that govern nuclear reprogramming, which is defined as the changes in gene expression patterns that expand the developmental potential of a fully differentiated cell to a totipotent state. Nuclear de-differentiation through transplantation of the nucleus into an enucleated oocyte is one experimental approach to reprogram somatic cells. Somatic cell nuclear transfer (SCNT) is ultimately aimed at generating uncommitted stem and progenitor cells that may be useful for cell replacement therapies. The success of reprogramming fully differentiated cells using SCNT has demonstrated that no genetic information is lost during development, with the exception of antigen receptor genes in lymphocytes, and that nuclear totipotency is retained for all cell types thus far studied. The low but reproducible success of SCNT in reprogramming a range of differentiated cells back to totipotency suggested that epigenetic mechanisms of gene regulation and differentiation are responsible for keeping cells in their state of differentiation. Epigenetic refers to mitotically stable modifications of DNA or chromatin that do not alter the primary nucleotide sequence. Epigenetic reprogramming is intended to reset these modifications from a fully differentiated to a less differentiated state, ideally to the totipotent embryonic state that allows differentiation into all lineages.

The goal of the studies described here was to establish a set of methods aimed at ultimately enhancing the efficiency of epigenetic reprogramming. To this end, three experimental avenues were accomplished. The initial study developed an approach to allow conditional regulation of epigenetic modifiers using RNA interference. DNA methylation is probably the best studied epigenetic modification known to regulate the expression of key embryonic “pluripotency genes” (Boiani et al., 2002; Bortvin et al., 2003) such as Oct-4. Consistent with the inverse correlation between DNA methylation and gene expression, DNA hypomethylation of the genome significantly increases the reprogramming efficiency after nuclear transfer (Blelloch et al., submitted). However as discussed below, global hypomethylation also increases the risk of genomic instability and tumor formation. To avoid adverse effects of DNA hypomethylation, we have

developed a Cre-lox based system for conditional gene suppression by RNA interference. It allows in a simple manner to study the effects of transiently down regulating an essential gene, such as Dnmt1, in order to increase epigenetic reprogramming. By subsequently reversing the knockdown and thereby restoring endogenous gene expression many deleterious effects of longer term suppression can be avoided.

The second application of our conditional gene knockdown approach using RNA interference was aimed at testing the notion that development of an embryo derived by SCNT might be restricted by temporarily inactivating a gene essential for development, yet the same embryo might be fully competent for extracting embryonic stem cell lines useful for therapeutic purposes. Due to incomplete epigenetic reprogramming many cloned embryos fail to express (reactivate) one or more of a set of "pluripotency genes", with Oct-4 being one of the best studied members of this class (Boiani et al., 2002; Bortvin et al., 2003). However, many of the abnormalities observed in cloned animals involve also deregulation of gene expression in the placenta (Humpherys et al., 2002). It appears that many placenta-specific genes are also not reactivated after nuclear transfer (Hall et al., 2005). Cdx2 is an essential transcription factor for trophectoderm differentiation and might be involved in some of the cloning phenotypes. Owing to its crucial role it was also an ideal candidate to test a concept called altered nuclear transfer (ANT) that has been proposed as a modification of the current NT technology. We have demonstrated the feasibility of the ANT technique, which now provides a scientific basis for the discussion surrounding alternative ways of deriving stem cells. Our findings confirmed previous results gained through deletion of Cdx2 by gene targeting (Strumpf et al., 2005). The phenotype of our knockdown was indistinguishable from the published knockout phenotype. Importantly, the generation of conditional Cdx2 shRNA expressing ES cells takes only a few weeks compared to at least two months for a knockout of both alleles by gene targeting (see discussion). Moreover, loss of Cdx2 is early embryonic lethal (blastocysts fail to implant) and no conditional knockout has been reported to date. Conditional Cdx2 ES cells will therefore be a useful tool to further investigate the role of Cdx2 in development.

Finally, another approach in studying epigenetic reprogramming is to determine the epigenetic differences between different cellular states, and to further elucidate what

defines the nature of "stemness" within a stem cell. DNA methylation is probably the best studied epigenetic modification that determines patterns of gene expression within a cell. In order to better define the epigenome of different cell types, we have developed an approach for large-scale high resolution DNA methylation analysis. The results presented in this work provide new tools and insights for the study of epigenetic reprogramming. Before describing the main findings of each of these three projects, I will review in more detail the background concepts for the major topics covered by my research.

RNA interference

Overview

In the past decade, small RNAs have emerged as central regulators of gene expression from worms to humans.

Two studies published in 1993 were the first to show that small non-coding RNAs were involved in gene regulation (Lee et al., 1993; Wightman et al., 1993). Lee and colleagues discovered that *lin-4*, a gene involved in the timing of *C. elegans* larval development, did not encode a protein, but rather a 22 nucleotide-long RNA that was predicted to arise from a longer precursor. This group and Wightman *et al*. then discovered that the *lin-4* RNA molecule was complimentary to multiple sites in the 3' UTR of another gene, *lin-14*. Both groups suggested that the binding of the small *lin-4* RNA to the *lin-14* UTR might repress the translation of *lin-14*, which is required for the transition from cell divisions of the first larval stage to those of the second (reviewed in (Bartel, 2004)). It took several years before additional small regulatory RNAs were discovered (Bartel, 2004). Because the first small RNAs to be identified were involved in timing developmental transitions, they were referred to as small temporal RNAs (stRNAs). Later cloning efforts revealed that this new class of RNAs was much broader than originally thought, with many of them expressed in a tissue-specific rather than timing-dependent manner. As a result, stRNAs were renamed microRNAs (miRNAs) (Bartel, 2004).

A few years later, studies in C. *elegans* showed that double-stranded (ds) RNA was a substantially more potent and long-lasting inhibitor of gene expression than single-stranded (ss), or antisense, RNA (Fire et al., 1998). Thus, although the endogenous small RNAs (miRNAs) had already been discovered, their true gene-inhibiting potential was

not appreciated until the study by Fire and colleagues emerged and, ultimately, paved the way for the RNA interference (RNAi) revolution (Fire et al., 1998).

RNAi is now appreciated as one of the most powerful ways to silence gene expression, and this technique has rapidly transformed gene function studies across phyla. RNAi operates through an evolutionarily conserved pathway that is initiated by dsRNA (reviewed in (Dykxhoorn et al., 2003; McManus and Sharp, 2002)). In eukaryotes such as plants and worms, long dsRNA (e.g. 1000 bp) molecules that are introduced into cells are processed into ~21 nt siRNAs by the dsRNA endoribonuclease Dicer (Bernstein et al., 2001). These siRNAs can then associate with a complex known as the RNAi-induced silencing complex (RISC) and direct the destruction of mRNA complementary to one strand of the siRNA. Recent studies have shown that the choice of RNA strand that is incorporated into the RISC is non-random and thus has important consequences for the design of siRNAs (Schwarz et al., 2003). siRNA design will be discussed in detail below.

Although the Dicer pathway is highly conserved, the introduction of long dsRNA (>30 bp) into mammalian cells results in the activation of antiviral signaling pathways, leading to nonspecific inhibition of translation and cytotoxic responses (Stark et al., 1998). One way to circumvent this problem is through the use of synthetic siRNAs that transiently down-modulate target genes without triggering cell death (Elbashir et al., 2001a). The subsequent discovery that plasmid-encoded interfering RNAs could substitute for synthetic siRNAs permitted the stable silencing of gene expression (reviewed in (Mittal, 2004)). In these systems, an RNA polymerase III promoter is used to transcribe a short, inverted stretch of DNA, resulting in the production of a short hairpin RNA (shRNA) that is then processed by Dicer to generate an siRNA. These vectors have been widely used to inhibit gene expression in mammalian cell systems. More recently, several groups have reported the use of similar expression constructs for the generation of RNAi-expressing transgenic mice (Carmell et al., 2003; Kunath et al., 2003; Rubinson et al., 2003), which, in some cases, recapitulate the phenotype of mice genetically deficient for the gene in question (Kunath et al., 2003; Rubinson et al., 2003).

RNAi: siRNAs, shRNAs and miRNAs

As shown in Figure 1, the different groups of small RNAs enter presumably the same RNAi pathway. Long dsRNA, shRNAs and miRNA precursor are all processed by Dicer into small RNA species. Alternatively, synthetically derived 21-nt long dsRNAs can feed directly into the endogenous RNAi processing pathway. Long dsRNAs, although effective in *C. elegans*, cannot be used in mammalian cells (see above) and will therefore not be discussed in more detail.

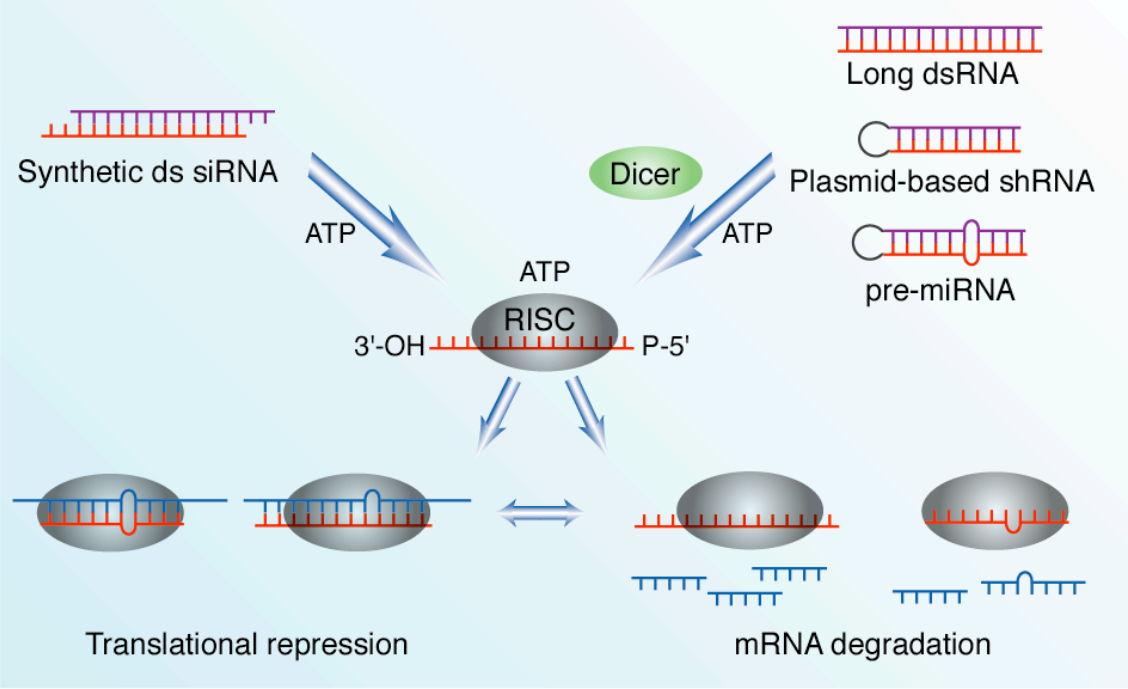

Figure 1: RNA interference. The RNAse III endonuclease, Dicer, generates small RNAs from long dsRNA, shRNAs and pre-miRNAs that become incorporated into the RISC. Depending on the target identity this can lead to translational repression or mRNA degradation.

MicroRNAs are endogenous, small non-coding RNAs that are involved in gene regulation (Bartel, 2004). Most studies suggest that endogenous microRNAs are transcribed by RNA polymerase II as longer primary nuclear transcripts, and are then processed into smaller, ~70-nt long pre-miRNAs by the RNAse III endonuclease, Drosha. These pre-miRNAs are then exported to the cytoplasm. In contrast, shRNAs are typically transcribed by RNA polymerase III using the U6 or H1 promoter. Functional siRNAs, shRNAs and miRNAs all share a preference for A/U basepairs at the 5' terminus of the antisense (AS) strand of the target gene, indicating a strand bias in stability and also the existence of a joint, indistinguishable pathway as illustrated in Figure 1.

In general, it is assumed that miRNAs pair imperfectly to the target sequences and that they work by repressing rather than degrading the target gene. However, recent work in *C. elegans* has demonstrated that this might not be true for all miRNAs (Bagga et al., 2005). In that study, the authors show that *let-7* degrades its target gene *lin-41*, despite only partial base-pairing. The targets of the *lin-4* miRNA, *lin-14* and *lin-28*, are degraded in a similar fashion (Bagga et al., 2005). Doench and colleagues have shown that, like miRNAs, siRNAs can repress the translation of target mRNAs with only partial complementary binding sites (Doench et al., 2003).

Following cleavage and nucleo-cytoplasmic export, the miRNA pathway appears to be indistinguishable from the siRNA pathway. Long dsRNAs, plasmid-based shRNAs, synthetic siRNAs and miRNAs are likely processed through the same endogenous, evolutionary conserved pathway. Although miRNAs and siRNA are processed similarly and are composition-wise indistinguishable, their origin and evolution are different (reviewed in (Bartel, 2004)). Post-transcriptional repression, cleavage or translational repression of mRNA appears to depend on identity of the target sequence rather than the origin of the small RNA (miRNA or siRNA). Notably, when a miRNA cleaves its target, the cleavage occurs at the same position as that cleaved by a siRNA, i.e., between the nucleotides complementary to residues 10 and 11 of the si/miRNA (Bartel, 2004; Elbashir et al., 2001b).

In addition to regulating gene expression at the posttranscriptional level, small RNAs can also suppress gene expression at the genomic level. RNA-directed DNA methylation is well established in plants (Matzke and Birchler, 2005), and has recently been shown to occur in human cells as well (Kawasaki and Taira, 2004). However, the latter report still awaits confirmation in independent studies.

The human genome is estimated to harbor approximately 200-255 miRNA genes. *C. elegans* and Drosophila have approximately 103-120 and 96-124 miRNA genes, respectively (Bartel, 2004). Most of the miRNAs that have been cloned are conserved in closely related species, such as mouse and human. Even more striking, more than a third of *C. elegans* miRNAs have a homolog in humans (Bartel, 2004). Recently identified miRNAs include those that are involved in cell proliferation, cell death, fat metabolism (flies), neuronal patterning (nematodes), differentiation of hematopoietic cells (mice), and

control of leaf and flower development (plants) (reviewed in (Bartel, 2004)). More than 200 microRNAs have been described in humans, although the function of a majority of them is unknown. The 5' region of microRNAs -- in particular the seven nucleotides at position 2-8 -- are essential for siRNA/miRNA function. These "core elements" or "seed" sequences determine the specificity of the siRNA/microRNA. Not surprisingly, these sequences are highly conserved and have thus been used to predict target mRNAs using comparative genetics and computational analysis (Lewis et al., 2003).

Two recent reports implicated miRNAs in tumorigenesis. The first found that a frequent chromosomal amplification (13q31-q32) contains a miRNA cluster that might be involved in the underlying malignancy (Ota et al., 2004). He *et al.* over expressed this miRNA cluster (miR17-92) in a mouse B-cell lymphoma model and showed that it accelerated tumor development when combined with c-myc, suggesting that miRNAs can act as oncogenes (He et al., 2005).

Design of siRNAs and shRNAs

When long dsRNA is used for RNAi, target selection is not required, as an array of siRNAs is produced from the longer dsRNA. One or more of these siRNAs will have the right sequence and effectively knock down the mRNA of interest. However, in mammalian cells, with the exception of early embryos, dsRNA cannot be used, necessitating the design of a single synthetic siRNA or shRNA that can efficiently suppress the target mRNA.

Elbashir *et al.* were the first to demonstrate that synthetic 21-nt duplex RNAs can specifically silence target mRNAs without triggering an interferon response in mammalian cells (Elbashir et al., 2001a). When siRNAs were generated from long dsRNAs by Dicer, a large number of distinct 21-nt long RNAs are produced. Although this approach is effective for gene silencing, it is impossible to determine which of the 21mers is producing the knockdown. The first insights regarding the mechanism of RNAi-mediated cleavage were provided by the same group (Elbashir et al., 2001b). By cloning the Dicer cleavage products of long dsRNAs, they found that nearly half were 21-nt and about a third 22-nt in length (Elbashir et al., 2001b). Based on this initial finding, a number of general design rules have been suggested (Elbashir et al., 2002). The principle design consisted of a 19-nt core sequence with a GC content of 30-70% --

which appeared to lend internal stability -- with 2-nt overhangs on either end: AA at the 5' and TT at the 3' terminus. Even with these rules, the design of siRNAs was still largely based on trial and error. This changed when Schwarz *et al.* showed that the absolute and relative stability at the 5' ends of the siRNA duplex determined which strand would be incorporated into the RISC complex. (Schwarz et al., 2003). This asymmetric incorporation into RISC has important consequences for the design, as the antisense strand of the siRNA can direct cleavage of only the sense strand mRNA target.

The next step towards improving the rational design of siRNAs came from a systematic analysis of the efficiency of 180 siRNAs that targeted two genes at every other base (Reynolds et al., 2004). An analysis of the functional and non-functional siRNAs suggested a new set of design rules, which included low G/C content, a bias towards low internal stability at the sense strand 3'-terminus, a lack of inverted repeats, and sense strand base preferences at certain positions. Each parameter was assigned a numeric score (-1, 0, or +1) with a maximum total score of 10. A score of 6 or higher translated to a five-fold higher probability of achieving 80% inhibition when compared to a random sequence (Boese et al., 2005; Reynolds et al., 2004). In general, applying these rules has significantly improved the average efficiency of target degradation. It should be noted that several of the shRNAs used in this work, such as Cdx2, have been designed according to these rules (see Methods and Materials). Notably, most effective shRNAs that were selected by trial and error followed these rules and had scores of 6 or higher based on the criteria listed. Importantly, although these rules were established using siRNAs, they appear to work equally well for the design of shRNAs.

In our hands, the rate of generating potent knockdowns using these rules ranges from 25-50% of the shRNAs designed. Despite this relatively high success rate, it is advisable to test the knockdown efficiency of shRNAs in reporter assays before proceeding to the actual experiment. There are many different ways to test for functional knockdown (reviewed in (Sandy et al., 2005)); one simple assay used for the selection of Cdx2 shRNAs is described in Figure 12 (Chapter 3).

Finally, it is worth mentioning that the secondary structure of the target mRNA does not appear to play a significant role in determining the efficiency. This conclusion was based on the sorting of siRNA functionality classes to a predicted mRNA target site

accessibility value (Boese et al., 2005). This was further supported by the observation that shifting a siRNA by only two base pairs along the mRNA can dramatically alter the silencing efficiency (Reynolds et al., 2004).

RNAi and its applications

Owing to the ease of synthesizing siRNAs and creating shRNA vectors, RNAi is an attractive tool for large-scale library screens that ultimately allows one to target every gene in the organism of interest (Hannon and Rossi, 2004). Particularly exciting are recent array-formatted, retrovirus-based, barcoded shRNA expression libraries that target human and murine genes (Berns et al., 2004; Paddison et al., 2004). To further enhance the potency of the knockdowns, the next generation of shRNA libraries has applied improved shRNA design and then embedded the shRNA sequence in microRNA sequences to mimic endogenous small RNAs (Silva et al., 2005).

The potency and specificity of RNAi also make this technique attractive for therapeutic purposes. Potential applications include the down-regulation of oncogenes or growth factors for the treatment of cancer (Hannon and Rossi, 2004). Due to the importance of particular base-pairings for effective interference, one could envision designing shRNAs that specifically target a mutant, disease-causing allele. The use of single nucleotide polymorphisms (SNPs) to achieve such specificity has been demonstrated in preliminary studies (Miller et al., 2003). ShRNAs have also been tested for the treatment of viral infections, in particular HIV. The main problem with this strategy is that the high mutation rate of HIV makes it difficult to target the viral RNA. However, targeting cellular HIV co-receptors, such as the CCR5 receptor, have shown promising results in human lymphocytes (reviewed in (Hannon and Rossi, 2004)). Another attractive target for RNAi treatment is viral hepatitis, as there is no vaccine or treatment for hepatitis C virus (HCV), and hepatitis B virus (HBV) can be prevented only by vaccination . McCaffrey and colleagues demonstrated that they could achieve a significant (99%) knockdown of the HBV core-antigen in liver hepatocytes by shRNA (McCaffrey et al., 2003). In a different study Song *et al.* used hydrodynamic injection of siRNAs into mice to silence the Fas-receptor. This led to inhibition of Fas-mediated apoptosis in the liver and the prevention of fulminant hepatitis, which would result in the mouse's death (Song et al., 2003).

These reports, among many others, highlight the potential for the use of RNAi in disease therapy. However, many issues still need to be resolved, including the delivery and specificity of siRNAs. In particular, the potential for off-target effects requires further investigation (Jackson et al., 2003). In general, to reduce the chance of off-target effects BLASTn searches should be performed for all si/shRNAs. When performing these searches, the word size, which determines the search window for identity, should be set to the lowest possible value (7 for BLASTn) and the significance value (expect value; E value) should be set relatively high (e.g.,>1000) to better suit short sequences.

Nuclear transfer

The early days of nuclear transplantation

It is fascinating to consider that a single cell, the zygote, gives rise to an organism with hundreds of different cell types represented by trillions of individual cells. It is even more striking that, with few exceptions, the genetic content of each differentiated cell remains identical to that of the original zygote. These cells therefore retain all the genetic information necessary to generate an entire organism. While this idea of "nuclear totipotency" or "nuclear equivalence" is now well established, it was not clear a few decades ago whether or not cells lose genetic information during the process of differentiation (reviewed in (Hochedlinger and Jaenisch, 2002b)).

In 1952, Briggs and King attempted to address this question by pioneering a novel technique termed nuclear transplantation (Briggs and King, 1952) (Fig. 2). Their approach was to transplant the nucleus of a differentiated cell into an enucleated oocyte as a means of assessing the nuclear potency of the differentiated cell. Using nuclei from frog blastomeres, Briggs and King demonstrated that the nuclei could be reprogrammed to a zygotic state and could generate early cleavage embryos when transplanted into enucleated oocytes (Briggs and King, 1952). Later they used frog cells from several stages of embryonic development as donors to determine whether differentiated cells retain the same developmental potential as the zygote (Briggs and King, 1957). Interestingly, they observed a gradual decline in cloning efficiency with increased donor cell differentiation. In later studies, analysis of cell morphology and marker gene expression were used to more accurately define the origin and differentiation state of the

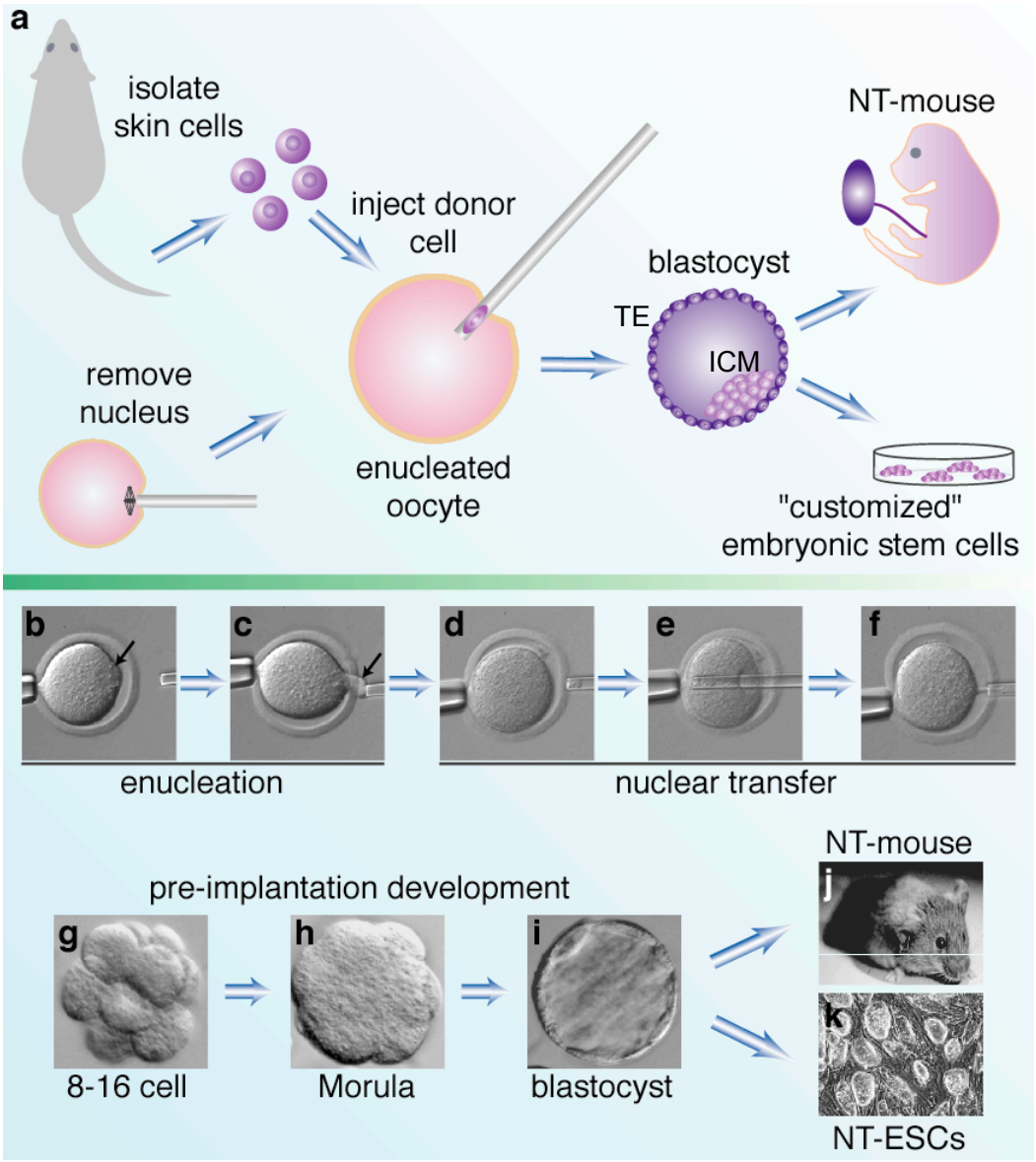

Figure 2: Murine nuclear transfer by microinjection. **a)** Schematic drawing of the NT procedure. The inner cell mass (ICM) will give rise to embryonic stem cells, when explanting the blastocyst onto irradiated feeder cells. Alternatively, when transferred to a synchronized pseudopregnant female, the blastocyst can generate a mouse. The outer cells of the blastocyst, the trophectoderm, will give rise to the extraembryonic tissues (placenta) and the ICM cells will generate the embryo. **b-k)** The same steps as in a) shown by light microscopy.
(ICM: inner cell mass; TE: trophectoderm)

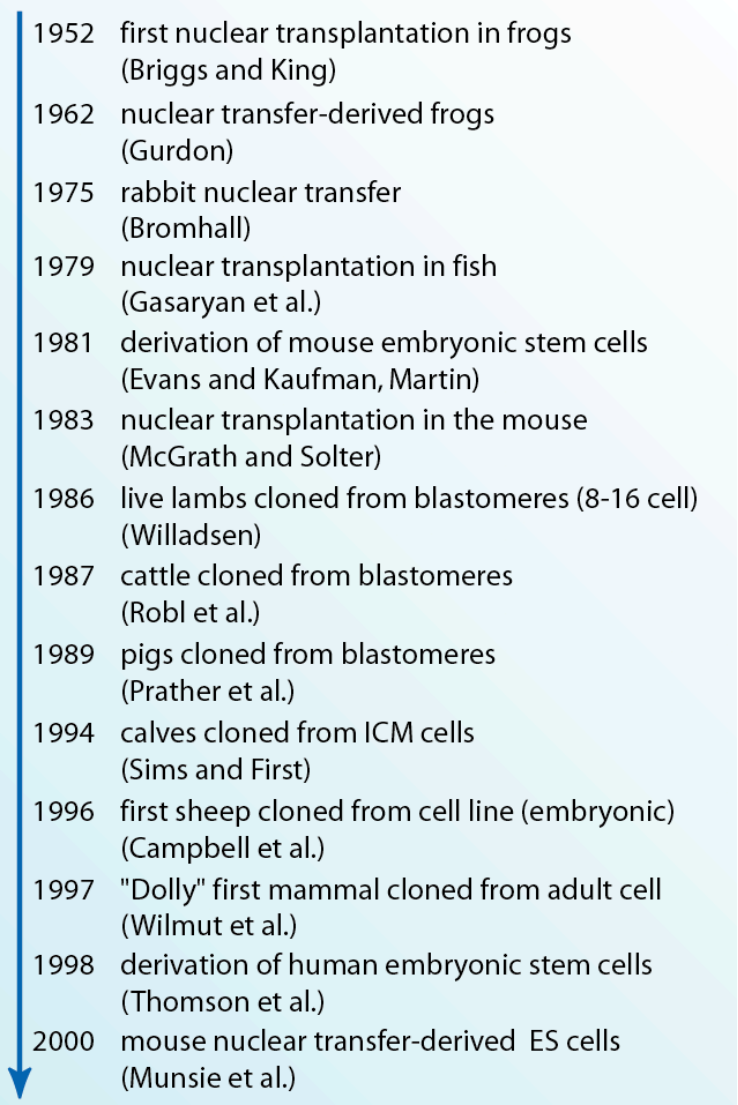

Figure 3: Timeline of key events in the application of nuclear transfer and stem cell research

donor cell -- a critical issue in resolving the question of nuclear potency (Hochedlinger and Jaenisch, 2002b). Although Gurdon succeeded in generating adult frogs from embryonic cells in 1962, adult donor cells have not produced any live frog clones, suggesting that adult cells, in contrast to embryonic cells, lack nuclear totipotency (Hochedlinger and Jaenisch, 2002b).

Despite the early successes in amphibian cloning, the development and application of nuclear transplantation in mammals did not make much headway (Fig. 3). The first attempt to clone mammals was reported in 1975 (Bromhall, 1975). Bromhall used both microinjection and virus-induced fusion of morula cells with unfertilized rabbit

eggs. In both cases, only early cleavage could be observed. Although Bromhall did not enucleate the eggs prior to microinjection, he reported some cases of stochastic self-enucleation. This suggested that the donor nucleus could support early development. In 1983, McGrath and Solter successfully transplanted zygotic nuclei into previously enucleated zygotes using cell fusion. This demonstrated the feasibility of the transplantation technique in mice (McGrath and Solter, 1983). The reconstructed embryos developed at a frequency comparable to wildtype embryos. However, later attempts to use blastomeres as donors failed, leading McGrath and Solter to speculate that cloning in mammals would not work (McGrath and Solter, 1984).

year	species	donor age	reference
1995	sheep	embryo	(Campell et al. 1996)
1996	sheep	foetal	(Wilmut et al. 1997)
	sheep (Dolly)	adult	(Wilmut et al. 1997)
1998	cattle	foetal	(Cibelli et al. 1998)
	cattle	adult	(Kato et al. 1998)
	mice	adult	(Wakayama et al. 1998)
1999	mice	embryo	(Wakayama et al. 1999)
	goat	foetal	(Baguisi et al. 1999)
2000	pig	adult	(Polejaeva et al. 2000)
	gaur	adult	(Lanza et al. 2000)
2001	mouflon	adult	(Loi et al. 2001)
2002	cat	adult	(Shin et al. 2002)
	rabbit	adult	(Chesne et al. 2002)
	zebrafish	embryo	(Lee et al. 2002)
2003	banteng	adult	(ACT, unpublished*)
	rat	foetal	(Zhou et al. 2003)
	mule	foetal	(Woods et al. 2003)
	horse	adult	(Galli et al. 2003)
	deer	adult	(A&M, unpublished**)
2005	dog	adult	(Lee et al. 2005)

Table 1: Animals cloned from embryonic, foetal and adult donor cell by nuclear transfer in the past decade. *ACT: Advanced Cell Technology; Lanza *et al*. ** Texas A&M University

Breakthroughs in mammalian cloning in the years following these early experiments included the generation of live lambs, cattle and pigs. All cases were based on nuclear transplantation using nuclei from blastomeres as donors (reviewed in (Campbell et al., 2005)). Although these reports were milestones in mammalian cloning, they still did not address the question whether adult cells are able to give rise to an entire organism (and therefore retain nuclear totipotency). Ten years after the successful generation of lambs by NT of blastomere nuclei, Campbell and colleagues reported another success: the first sheep to be derived by NT using donor nuclei from an embryonic cell line (Campbell et al., 1996). In 1997, the same group reported a further breakthrough -- the cloning of "Dolly", the first mammal cloned from an adult cell (Wilmut et al., 1997). During the last eight years, considerable progress has been made in the technology of cloning and in the understanding of its species-specific requirements. Researchers have successfully generated cloned offspring from 15 additional species, using both embryonic and adult donor cells (Table 1) (reviewed in (Campbell et al., 2005);(Lee et al., 2005)).

Cloning and differentiation

In their early studies, Briggs and King had observed that differentiating frog cells progressively lose their nuclear potency (Briggs and King, 1957). To date, no frogs have been cloned from adult cells. This raises the question whether mammalian clones, including Dolly, are really derived from differentiated cells or from rare adult stem cells. Dolly's donor cells, as well as the somatic cells used to generate clones in subsequent years, originated from adult animals. There was, however, no conclusive evidence that the donor cells had been fully differentiated (Hochedlinger and Jaenisch, 2002b).

The unequivocal demonstration that terminally differentiated cells can give rise to an NT-derived organism was established using lymphocytes as donor cells (Hochedlinger and Jaenisch, 2002a). In lymphocytes, immunoglobulin (Ig) and T-cell receptor (TCR) gene rearrangements serve as endogenous markers for both the origin and the differentiation state of the donor cell. The NT-derived monoclonal mice displayed unique Ig or TCR rearrangements in every tissue demonstrating that the donor cells were indeed fully differentiated (Hochedlinger and Jaenisch, 2002a). Eggan and colleagues later cloned mice using nuclei form terminally differentiated olfactory neurons. They

demonstrated that even nuclei from post-mitotic donor cells can re-enter the cell cycle and generate viable clones (Eggan et al., 2004). However, due to the low NT efficiency of the donor cells, these studies first derived NT-ES cells. Hochedlinger and Jaenisch then generated cloned mice by implementing a second step in which tetraploid blastocysts were injected with NT-ES cells. This ensured that the resulting offspring was entirely derived from the ES cells. In this procedure, all extra-embryonic tissues were derived from the tetraploid cells. To demonstrate that terminally differentiated cells can generate each and every kind of embryonic and extra-embryonic lineages, Eggan and colleagues subjected the olfactory neuron-derived NT-ES cells to a second round of nuclear transfer. More recently, it was shown that terminally differentiated NKT cells can be directly cloned (Inoue et al., 2005). NKT cells, like B and T cells, have genetic rearrangements that allow retrospective identification of the differentiation state of the cells. It is worth mentioning that, even though T cells and NKT cells are part of the same cell lineage, their respective NT efficiency is significantly different (Hochedlinger and Jaenisch, 2002a; Inoue et al., 2005).

Why is cloning so inefficient?

Based on the progress in cloning over the past decade, it has to be assumed that, given the appropriate conditions, almost any donor cell type can be used for NT. Furthermore, it should be possible to generate offspring using NT from most, if not all, species. However, the overall cloning efficiency remains extremely low: only an estimated 0-5% of reconstructed embryos develop to term (Campbell et al., 2005). The first mouse clones were generated more than 7 years ago (Wakayama et al., 1998) and many different laboratories now conduct murine nuclear transfer experiments. The efficiency of mammalian cloning has not significantly improved, despite great efforts to this effect. In addition to the inefficiency of the cloning procedure, there is the possibility that normal clones do not exist. It is conceivable that the few clones that survive are simply less abnormal than those that do not survive (Jaenisch, 2004). There are several possible explanations for the inability of the donor nucleus to generate functional clones. Technical issues, such as the handling, isolation and type of the donor cell, as well as the subsequent activation and culture conditions, probably impact cloning efficiency. There are also important biological barriers, such as non-equivalence of the maternal and

paternal genomes. These barriers will be more difficult to resolve than any technical limitations (Jaenisch, 2004).

As a result of natural fertilization, the genetic content of a zygote is inherited by all somatic cells of the developing organism. However, only a subset of genes is active in a given cell type. For normal development to proceed, it is crucial that in a particular cell, all essential genes are switched on, while those that are not required are switched off. Many of the genes crucial to early development, such as Oct-4 have already been identified. Recent studies show that somatic cell-derived NT blastocysts often fail to reactivate many of these crucial genes (Boiani et al., 2002; Bortvin et al., 2003).

Gene expression is regulated by epigenetic mechanisms such as DNA methylation and chromatin modifications that impose stable but reversible marks on the genome. Such stable alterations resulting in differential gene expression are often referred to as "epigenetic" (Wang et al., 2004). A key factor in cloning-associated abnormalities probably involves inadequate epigenetic reprogramming of the donor genome. DNA methylation, one of the best-studied epigenetic modifications, is known to be aberrant in many clones (see below). Recent work by Blelloch *et al*. supports the central role of DNA methylation: global hypomethylation of the donor cell before NT was shown to significantly increase the efficiency of ES cell derivation (Blelloch et al., submitted).

Another study showed significant improvements in cloning efficiency and NT-ES cell development after trichostatin-A (TSA) treatment of the reconstructed oocyte. TSA is a compound that increases histone acetylation and DNA demethylation (Kishigami et al., 2006). A better understanding of the reprogramming factors and events that occur in the oocyte should help further improve efficiency of nuclear transfer.

Biological NT applications

Nuclear transplantation was originally developed to study the question of nuclear equivalence (Briggs and King, 1952). The technique allows the *in-vivo* amplification of a single cell into an entire organism. It thus constitutes a powerful tool for dissecting genetic and epigenetic regulatory mechanisms within the genome.

The cloning of mice from lymphocytes, for example, made it possible to study the process of allelic regulation (Gerdes and Wabl, 2004) and secondary Ig rearrangements (Koralov et al., 2005) in monoclonal mice. Similarly, nuclear transfer has been used to

determine whether the variety of olfactory receptors is created by gene rearrangements in a fashion similar to the generation of antibody diversity in the immune system. This question was originally contemplated because of parallels between the immune system and the brain, such as a requirement for the non-homologous end-joining (NHEJ) machinery and Rag1 expression in both lymphocytes and neurons, as well as the presence of a vast number of different receptors in the olfactory epithelium. To address this question, Eggan and colleagues used olfactory sensory neurons as NT donors to produce cloned mice (Eggan et al., 2004). The resulting data suggested that genetic rearrangements do not occur in olfactory neurons.

One intriguing question recently addressed by NT was the contribution of epigenetic changes in the development of cancer and whether such changes could be reprogrammed to alter the cancer phenotype. A study by Hochedlinger *et al*., showed that several tumour cell lines could support the pre-implantation development, but not the derivation of ES cells. One melanoma cell line generated pluripotent ES cells and contributed to chimeras when injected into diploid blastocysts. It is interesting that the tumour phenotype was similar to that of the original cell line, but showed higher penetrance and shorter latency. This demonstrated that the secondary changes in the melanoma cell line are compatible with a wide spectrum of development, but that they also predispose the cells to transformation (Hochedlinger et al., 2004).

In a second study, NT of embryonic carcinoma (EC) cells was used to compare the role of epigenetic and genetic alterations in tumorigenicity and in the developmental potential of EC cells respectively. Although EC cell lines generated NT-ES lines with great efficiency, the tumorigenic and developmental potential in the latter remained the same as in the donor EC cells. This suggested that irreversible genetic changes, rather than epigenetic modifications, determined the phenotype of EC cells (Blelloch et al., 2004).

Commercial NT applications

As well as providing a tool for studying developmental biology, nuclear transfer has great potential for agricultural technology, the propagation of endangered species, and for therapeutic applications.

Transplantation of a nuclear transfer-derived blastocyst to the uterus of a recipient female will, in some cases, generate live offspring. Nuclear transfer in human is a highly controversial theme. Apart from ethical objections, animal experiments have already shown that such endeavours are dangerous. These findings have resulted in the broad consensus among scientists that cloning humans for reproductive purposes would be irresponsible (Jaenisch and Wilmut, 2001).

Many important farm animals such as cattle, pigs and sheep have been successfully cloned using this technique (Campbell et al., 2005). Nuclear transfer also allows the simple propagation of „elite" strains of animals. Among other reasons, the lack of ES or ES-like cells that can contribute to the germline in sheep and in other farm animals made the generation of sheep by NT from a cell line so important (Campbell et al., 1996). Cell lines now allowed similar manipulations of farm animal genomes that are readily available in the mouse. This approach was successfully used to produce transgenic sheep, pigs and cows in the years to follow (Campbell et al., 2005).

It should be emphasized that abnormalities, though present in the cloned animal, are epigenetic and therefore not problematic for the propagation of farm animals. Current evidence indicates that all epigenetic abnormalities can be corrected in the germline and the offspring can therefore be completely normal and healthy (Jaenisch, 2004).

Finally, it is worthy of note that endangered species such as a Gaur have been cloned using domestic animals as oocyte donors (Lanza et al., 2000), demonstrating another useful application of NT.

Therapeutic NT applications

When a NT blastocyst is explanted it can, under certain conditions, generate customized pluripotent NT-ES-cell lines which have great potential for regenerative medicine. Stem cells are generally defined as cells with the ability to self renewal and to generate more restricted, differentiated cells. The two major classes of stem cells are adult and embryonic stem cells. In the adult, many tissues, including the intestine and the haematopoetic system, have a significant turnover. This remarkable ability to regenerate is provided by rare, quiescent stem cells that can give rise to the manifold cell types of their respective tissue. They are generally thought to be restricted to that lineage (Wagers and Weissman, 2004). There have also been a few reports suggesting that adult stem cells

have greater developmental potential, and that they can transdifferentiate into other cell lineages. However, the ability of these cells to generate cells of other lineages remains very controversial (Wagers and Weissman, 2004).

One major limitation of adult stem cells is that, once they have been removed from their biological niches, it is difficult to expand and maintain them in a multipotent state. By contrast, embryonic stem cells grow in culture for unlimited passages, while maintaining the potential to generate every single cell type in the body. Indeed, a single embryonic stem cell can generate an entire mouse (Wang and Jaenisch, 2004).

The first murine ES cells were derived independently by two groups in 1981 (Evans and Kaufman, 1981; Martin, 1981). Shortly after these initial reports, murine ES cells were shown to contribute to all tissues. Above all, they can contribute to the germline and therefore to the offspring of chimeras (Bradley et al., 1984). The first human ES cell line was reported by Thomson and colleagues almost two decades after the derivation of murine ES cells (Thomson et al., 1998). Two years later, Munsie *et al.* reported the derivation of NT-ES cells from somatic cells in the mouse (Munsie et al., 2000).

The first breakthrough in using nuclear transfer for therapy was demonstrated by Rideout and colleagues (Rideout et al., 2002), who derived NT-ES cells from an immunodeficient mouse. The immunodeficiency defect was corrected by homologous recombination. The repaired NT-ES cells were then differentiated *in vitro* into haematopoetic precursors and then transplanted back to the donor mouse. This procedure indeed restored normal lymphocyte populations in the mouse, constituting the proof of principle that NT can be used for therapeutic applications. This was further supported by another study that used NT-ES cell-derived dopaminergic neurons to ameliorate disease in a mouse model of Parkinson disease (Barberi et al., 2003). These studies highlight NT's great potential for stem cell therapies. However, many technical and safety issues must still be addressed before these findings can be translated into human therapy.

Alternative approaches for reprogramming somatic cells

While research in mammalian nuclear transfer is making progress, alternative ways of de-differentiating somatic cells might also prove useful. NT represents the most extreme form of nuclear reprogramming. Alternative approaches, such as cell fusion,

allow the de-differentiation of somatic cells, and thus the generation of pluripotent hybrid cells. When a less differentiated cell is fused with a differentiated cell, the former typically dictates the fate of the hybrid. In mouse and humans, fusion of ES cells with somatic cells results in the formation of pluripotent hybrids (Cowan et al., 2005; Do and Scholer, 2004; Tada et al., 2001). Do and Scholer took the initial fusion studies performed by Tada *et al.* one step further by investigating what part of the ES cell (cytoplast or karyoplast) harboured the reprogramming activity. Their results suggested that the reprogramming activity was present only in the karyoplast (Do and Scholer, 2004).

Although these fusion experiments demonstrate the potential of this approach, they have several drawbacks. The major problem in clinical applications is that the hybrids resulting from cell fusion are tetraploid. Ideally, the embryonic and somatic nucleus should remain separate for the time required to accomplish functional reprogramming. This, however, can only be achieved when efficient reprogramming occurs without DNA replication and cell division. Do and Scholer fused ES cells to neurosphere cells treated with mitomycin C to inhibit DNA replication and cell division. In these experiments, reactivation of the Oct4-GFP reporter was observed only when the neurospheres were pretreated with a demethylating agent. This suggests that reprogramming might depend on DNA replication (Do and Scholer, 2004).

In addition, it is unclear whether the data presented in these studies show complete functional reprogramming of the somatic genome or only partial reprogramming compensated by the pluripotent ES genome. Although numerous reprogramming events, such as reactivation of somatic Oct-4 and other genes, have been demonstrated (Cowan et al., 2005; Tada et al., 2001), this does not necessarily mean that all genes required for a pluripotent state have been reactivated. However, further investigation will provide new insight into the reprogramming events after cell fusion.

Cell free nuclear extracts provide another interesting approach for reprogramming. However, early experiments with human cell lines to determine whether the somatic genome is reprogrammed have not been convincing. It is possible that transient protein uptake from the nuclear extracts takes place rather than a reactivation of endogenous genes and functional reprogramming (Hakelien et al., 2004; Hakelien et al.,

2002; Taranger et al., 2005). Further studies, using more stringent measures of reprogramming, will be required to assess the true reprogramming potential of cell extracts.

Alternative approaches for generating stem cells

Embryonic stem cells are routinely derived from a very early stage embryo termed the blastocyst (Figs. 2 and 4). Ever since Thomson and colleagues reported the isolation of the first human embryonic stem cell (hESC) lines (Thomson et al., 1998), the derivation of new stem cell lines has remained highly controversial. To derive

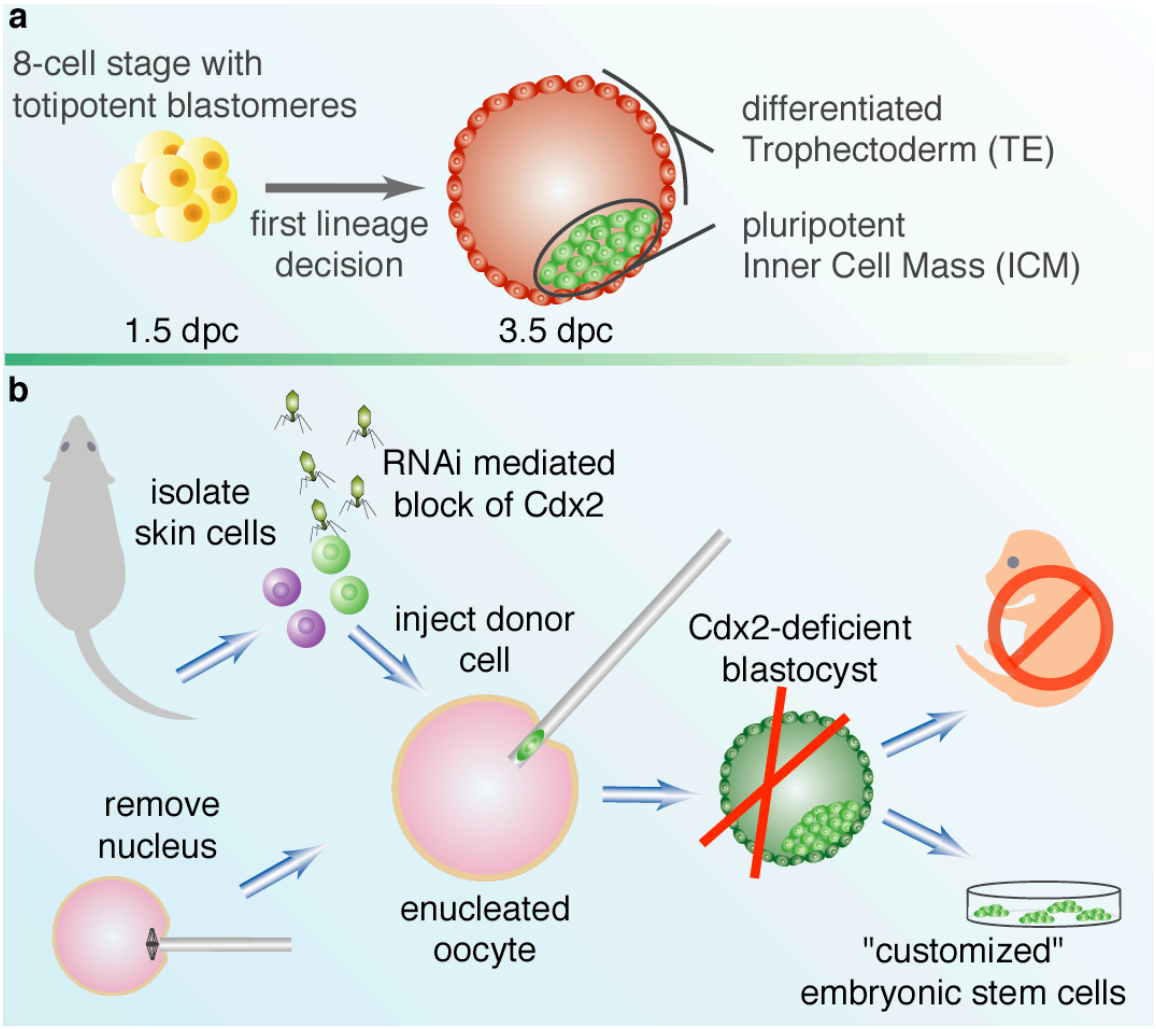

Figure 4: Altered nuclear transfer (ANT). The ANT procedure is very similar to the regular NT described in Figure 2. **a)** The first lineage differentiation takes place after the 8-cell stage. **b)** The only modification required for ANT is the introduction of a conditional RNAi system (contains a GFP cassette; described in Chapter 3) via lentiviral transduction. Due to the presence of the Cdx2 shRNA, no Cdx2 protein can be made and hence the TE cannot form properly. The blastocyst becomes a disorganized clump of cells and cannot implant, yet it still allows derivation of pluripotent of ES cells.

customized ES cells, NT takes the destruction of an intact human blastocyst into account -- one that could potentially generate a human being. This problem is at the root of the whole ethical dilemma.

We recently reported an alternative approach to generate NT-ESCs (Meissner and Jaenisch, 2006). This method, called altered nuclear transfer, was proposed by W. Hurlbut, a member of the president's council for bioethics (Hurlbut, 2005). We designed a simple experimental strategy to test the feasibility of this concept in the mouse (see Chapter 3). During the first stages of preimplantation development, all blastomeres remain totipotent. However, after the 8-cell stage, the first lineage differentiation decision takes place. This decision generates the first two cell lineages: the trophectoderm (TE) and the inner cell mass (ICM) (Fig. 4a). The latter gives rise to the embryo as well as to embryonic stem cells when explanted. The TE generates the foetal-maternal interface (placenta) and trophectoderm stem cells (TS) cells when explanted (Tanaka et al., 1998).

The concept of altered nuclear transfer involves disabling a gene that is essential for TE function prior to NT, thus eliminating the developmental potential of the reconstructed embryo (Fig. 4b). In fact, the lack of the TE lineage generates a "biological artifact" (Hurlbut, 2005), since it creates a disorganized clump of cells. Detailed studies from the laboratories of Rossant and Beck provided excellent insight into the role of Cdx2, the earliest known TE-specific transcription factor. Using a recently developed conditional RNAi system (see Chapter 3), we demonstrated that loss of Cdx2 leads to the development of abnormal blastocysts that cannot implant. However, the absence of Cdx2 did not affect the ability to derive pluripotent ES cells, and removal of the Cdx2 inhibitor from the ES cells restored their potency to that of any normal ES cell line.

It has already been shown that human and mouse ES cells can be derived from morula stage embryos (Strelchenko et al., 2004; Tesar, 2005). A new study now reports that single 8-cell stage blastomeres can generate pluripotent ES cell lines (Chung et al., 2006). Although this study is interesting, the reported efficiency was several times lower than the efficiency of ES cell generation using later stage embryos. The authors envisioned this technique as a possible way of solving the ethical dilemma of sacrificing a viable embryo to generate ES cells. They demonstrate that removing a single blastomere from a murine eight-cell embryo does not affect the developmental potential

of the manipulated embryo. The single blastomere could thus be used to generate ES cells and the seven-cell embryo could be transferred to generate a foetus.

Obviously, this procedure, if feasible in humans, would apply only to couples undergoing *in-vitro* fertilization (IVF). In addition, the low efficiency of the procedure means that it is fairly unlikely that every blastomere yields an ES cell line, raising in turn the question of what to do with the embryos that did not yield a matching ES cell line. Moreover, in other species it has been shown that single blastomeres are still totipotent. Thus it could still be argued that destroying a blastomere is equivalent to destroying an embryo. Indeed, people have used similar rationale to argue against the altered nuclear transfer approach (Solter, 2005). While it is clearly impossible to please everybody, it must also be admitted that both approaches provide interesting scientific data as well as a scientific basis for alternative ways of generating ES cells using more ethical approaches.

DNA methylation and epigenetic reprogramming

Overview

DNA methylation is a key epigenetic modification that provides heritable information not encoded in the nucleotide sequence. 5-methylcytosine is the only known covalent modification of DNA in vertebrates (Jeltsch, 2002). Mammalian DNA methylation serves a wide-range of functions including regulation of gene expression, genomic imprinting, and X-chromosome inactivation. It contributes to genomic stability and serves as a defense mechanism against transposable elements (Bestor, 2000b; Jaenisch, 1997; Jaenisch and Bird, 2003; Robertson and Wolffe, 2000). In addition, its role in disease states such as cancer becomes increasingly evident (Feinberg, 2004; Gaudet et al., 2003; Jones and Baylin, 2002; Laird, 2003; Robertson, 2002).

Three catalytically active DNA methyltransferases (Dnmts) have been described that are responsible for establishing and maintaining methylation patterns in mammals (Bestor, 2000b; Li et al., 1992; Okano et al., 1999; Okano et al., 1998a). Dnmt1 has been largely viewed as the maintenance enzyme, owing to its preference for hemimethylated DNA (Robertson and Wolffe, 2000). Dnmt3a and Dnmt3b have no preference and are required for *de novo* methylation activity (Okano and Li, 2002). During murine preimplantation development methylation levels decrease with some notable exceptions including imprinted genes and IAP elements (Jaenisch, 1997; Lane et al., 2003). Around

the time of implantation normal methylation levels are restored by the *de novo* methyltransferases and later maintained by Dnmt1.

Targeted gene disruption for each of the catalytically active Dnmts (1, 3a and 3b) results in a lethal phenotype demonstrating the essential role of DNA methylation in development (Li et al., 1992; Okano et al., 1999). Interestingly, undifferentiated ES cells deficient for Dnmt1, Dnmt3a, Dnmt3b or Dnmt3a/3b do not display any obvious abnormalities (Lei et al., 1996; Okano et al., 1999). Normally in wild-type ES cells most CpG dinucleotides are methylated with the exception of many CpG-islands.

In addition, ES cells and early embryos, but not somatic cells, seem to contain significant amounts of non-CpG methylation (mostly CpA) (Haines et al., 2001; Ramsahoye et al., 2000). Currently the functional role of this non-CpG methylation is not clear.

In general, methylation is found in CpG-poor regions, while CpG-rich areas (CpG islands) seem to be protected from this modification and are generally associated with active genes (Cross and Bird, 1995). This is consistent with the fact that methylated CpG islands are found on the inactive X chromosome and on the silenced allele of imprinted genes (Neumann and Barlow, 1996; Razin and Cedar, 1994; Riggs and Pfeifer, 1992). The methyl group is positioned in the major groove of the DNA where it can easily be detected by proteins interacting with the DNA (Jeltsch, 2002). The effects of DNA methylation on chromatin structure and gene expression are likely mediated by a family of proteins that share a highly conserved methyl CpG-binding domain (MBD) (Wade, 2001). Two of these, MeCP2 and MBD1, have been suggested to be involved in transcriptional repression (Fujita et al., 2000; Fujita et al., 1999; Nan et al., 1997) based on biochemical observations that they form complexes with histone deacetylases and other proteins important for chromatin structure (Jones et al., 1998; Nan et al., 1998; Wade et al., 1999; Zhang et al., 1999).

DNA methylation during development

DNA methylation patterns are extremely dynamic in early mammalian development. Within 1-2 cell divisions after fertilization a wave of global demethylation takes place. It has been suggested that the paternal genome is actively demethylated during the period of protamine-histone exchange and the maternal genome subsequently

becomes demethylated, presumably through a passive DNA replication mechanism (Reik et al., 2001; Santos et al., 2002). By the morula stage, methylation is found only in some repetitive elements and imprinted genes (Reik et al., 2001; Sanford et al., 1987; Walsh et al., 1998). After implantation, genome-wide methylation levels increase dramatically, establishing a differential pattern between the cells of the ICM and those of the trophectoderm (Santos et al., 2002), and ultimately resulting in the formation of methylation patterns found in the adult (Turker, 1999). Primordial germ cells (PGC) also undergo global demethylation. Importantly, in contrast to demethylation during preimplantation, all parental specific epigenetic marks are erased in the PGC by embryonic day 13-14. As a result, PGC and diploid germ cells are the only cell type where the paternal and maternal genomes are equivalent. Upon initiation of gametogenesis, PGC remethylation begins and the parental-specific methylation patterns that will code for monoallelic expression of imprinted genes are established (Lucifero et al., 2002; Reik et al., 2001).

Maintenance and establishment of DNA methylation is accomplished by at least three independent catalytically active DNA methyltransferases: Dnmt1, Dnmt3a and Dnmt3b (Bestor, 2000a; Robertson and Wolffe, 2000). There are two isoforms of Dnmt1, an oocyte specific isoform (Dnmt1o) and a somatic isoform. Somatic Dnmt1 is often referred to as the ‘maintenance’ methyltransferase because it is believed to be the enzyme that is responsible for copying methylation patterns after DNA replication. The oocyte specific isoform of Dnmt1 is believed to be responsible for maintaining but not establishing maternal imprints. The Dnmt3 family (Dnmt3a, 3b, 3l and a number of isoforms) is required for the de novo methylation that occurs after implantation, for the de novo methylation of newly integrated retroviral sequences in mouse ES cells (Bestor, 2000a; Okano et al., 1999), and for the establishment of imprints (Dnmt3l) (Li, 2002). It was recently shown that Dnmt3a has a strong preference for unmethylated DNA (Yokochi and Robertson, 2002).

The essential role of DNA methylation in mammalian development is highlighted by the fact that mutant mice lacking each of the enzymes (generated by gene targeting) are not viable and die either during early embryonic development (*Dnmt1* and *Dnmt3b*) or shortly after birth (*Dnmt3a*) (Li et al., 1992; Okano et al., 1999; Robertson and Wolffe,

2000). The knockout of *Dnmt3l* leads to male infertility and the failure to establish imprints in female eggs (Bourc'his et al., 2001b; Li, 2002). Disruption of *Dnmt2* did not reveal any obvious effects on genomic DNA methylation (Okano et al., 1998b).

DNA methylation and SCNT

Considering the fundamental role of DNA methylation in development, it seems likely that any NT embryo will need to recapitulate a functional pattern of epigenetic modifications in order to proceed through normal embryogenesis. Several groups have investigated DNA methylation patterns in NT embryos and reported finding abnormalities in DNA methylation (Bourc'his et al., 2001a; Dean et al., 1998; Dean et al., 2001; Kang et al., 2001a; Kang et al., 2002; Ohgane et al., 2001). In cloned bovine embryos, satellite sequence methylation levels are more similar to the donor cells than to control embryos (Kang et al., 2001a). However, methylation patterns of single copy gene promoters in cloned bovine blastocysts appeared to be normally demethylated (Kang et al., 2001a). In addition, the satellite sequences, but not the single copy genes, showed more methylation in TE than in the ICM of cloned bovine blastocysts (Kang et al., 2002) (reviewed in (Han et al., 2003)). Using antibodies against 5-methyl cytosine two independent studies showed that the cloned bovine embryos did not undergo global demethylation in early embryogenesis and even showed precocious de novo methylation (Dean et al., 2001), with euchromatin being abnormally hypomethylated and centromeric heterochromatin being abnormally hypermethylated (Bourc'his et al., 2001a). These findings suggest that different chromosomal regions might respond differently to demethylation in the egg cytoplasm. Interestingly, when the same satellite sequences examined in bovine (Kang et al., 2001a) were analyzed in a different species (porcine) methylation levels at the blastocyst stage of cloned embryos were more comparable to those of fertilized control embryos (Kang et al., 2001b), suggesting also species-specific differences. A recent study that analyzed a number of imprinted genes in cloned murine blastocysts showed that most of the examined genes displayed aberrant methylation and expression patterns (Mann et al., 2003).

It would be interesting to know if aberrant methylation patterns during preimplantation development contribute to the low efficiency of generating clones and to what extent the clones can tolerate such variation. Unfortunately, analyzing the

methylation status in preimplantation embryos provides only indirect correlations preventing satisfactory resolution of this question.

In order to establish a potential correlation between global DNA methylation levels and the developmental potential of cloned embryos, Cezar *et al.*, compared the genome-wide methylation status among spontaneously aborted cloned fetuses, live cloned fetuses, and adult clones in bovine (Cezar et al., 2003). When genome-wide cytosine methylation levels were measured by reverse-phase HPLC, they found that a significant number of aborted fetuses lacked detectable levels of 5-methylcytosine. In contrast, when seemingly healthy adult, lactating clones were compared to similarly aged lactating cows produced by artificial insemination, comparable DNA methylation levels were observed. The authors suggested that survivability of cloned cattle is related to the global DNA methylation status. All evidence suggests that a correct global methylation status is required for development. However, subtle changes might be compatible with normal development and result only in minor or no phenotypes. For example, by applying restriction landmark genome scanning (RLGS) in two seemingly healthy cloned mice, it was shown that methylation patterns at several sites in each clone differed from those in the controls (Ohgane et al., 2001).

The reason for the frequent abnormal DNA methylation patterns in cloned embryos is still unclear. It is likely that, because of the epigenetic difference between the somatic donor cell and the gametes, the somatic nucleus responds differently to the egg cytoplasm, affecting subsequent events during embryogenesis. For example, the highly coordinated demethylation process in the pronuclei of the maternal and paternal genome upon fertilization might not happen appropriately in the somatic donor genome following NT. It is not clear whether all of the somatic epigenetic marks imposed by DNA methylation during differentiation can be removed from the donor nucleus. Any failure to demethylate the DNA sequences that are normally demethylated during early cleavage stages of development might be stably passed on to progeny cells. Another possible explanation for the aberrant methylation patterns in clones may result from the ectopic expression of the somatic form of Dnmt1 in the egg and cleavage stage cloned embryos. In the mouse oocyte and preimplantation embryo, the oocyte specific form (Dnmt1o) but not the longer somatic form is expressed. It has been shown that a translocation of

Dnmt1o between nucleus and cytoplasm is tightly regulated during murine preimplantation development (Howell et al., 2001; Ratnam et al., 2002). In contrast, cloned preimplantation mouse embryos were reported to aberrantly express the somatic form of the Dnmt1 gene, and the translocation of Dnmt1o was absent (Chung et al., 2003). As mentioned above, DNA methyltransferases (Dnmt1, 3a, 3b 3l) play important roles in setting up and maintaining DNA methylation patterns. It is reasonable to speculate that dysregulation of any of these enzymes in clones may alter DNA methylation patterns.

These abnormal DNA methylation patterns could result in embryo lethality or phenotypic abnormality. At present, little is known about the developmental role of dynamic changes in DNA methylation during preimplantation, although very recently, the importance of early embryonic methylation patterns in setting up the structural profile of the genome was shown (Hashimshony et al., 2003). This suggests that the failure to establish correct methylation patterns early in development might have far reaching effects on the chromatin structure. Interestingly, mouse embryos deficient for Dnmt1 and Dnmt3b die around E9.5 (Li et al., 1992; Okano et al., 1999), but *Lsh* mutant mice die only after birth despite showing a substantial loss of methylation throughout the genome (Dennis et al., 2001).

Analyzing DNA methylation

The intense interest in the biological functions of DNA methylation and its role in diseases have led to numerous techniques to detect and compare DNA methylation (reviewed in (Laird, 2003; Murrell et al., 2005)). Global methods such as nearest neighbor analysis (NNA) and high-performance liquid chromatography are valuable to quantify the total 5-methylcytosine content of a DNA sample, but information on the position in the genome cannot be gained (Ramsahoye, 2002a; Ramsahoye, 2002b).

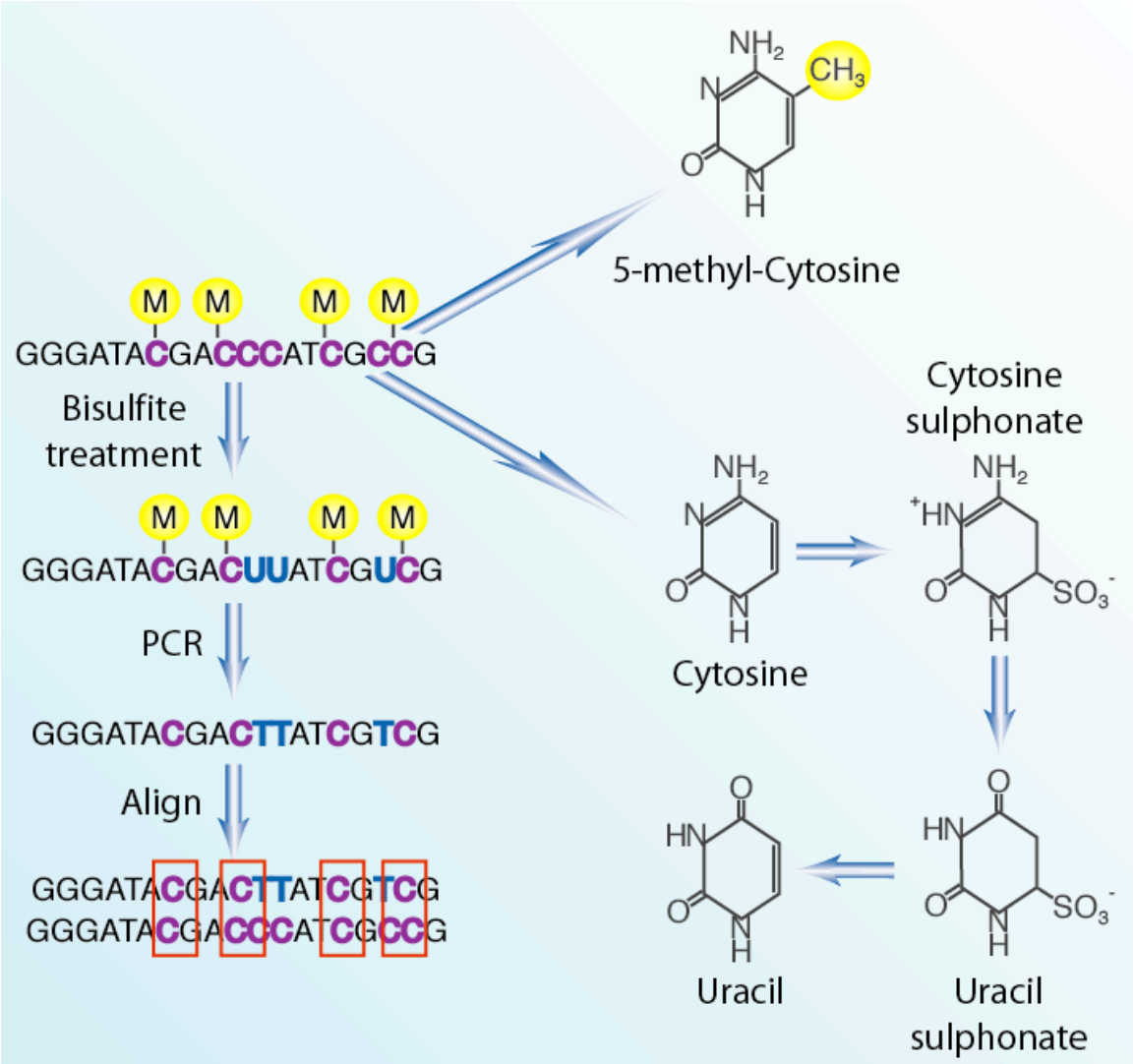

Figure 5: Differential base modification by sodium bisulfite. A hypothetical DNA strand containing cytosine and 5-methyl-cytosine residues. Single stranded DNA is generated using NaOH. After adding sodium bisulfite all cytosine residues are converted to uracil by the mechanism shown. The 5-methyl-group protects against the sulphonation and therefore every methylated cytosine remains unchanged. After complete reaction the DNA strands are not complimentary anymore. Hence, each strand is amplified by PCR using different primer sets that are specific for the converted sequences. The PCR products are cloned into a plasmid vector and sequenced. The sequences can be aligned to the original sequence: C-T mismatches indicate unmethylated cytosines, whereas C-C matches indicate a methylated cytosine

Digestion with methylation-sensitive (or methylation-dependent) restriction enzymes (MSREs) has been used to selectively enrich the methylated and unmethylated DNA fractions, respectively (Bedell et al., 2005; Lippman et al., 2004; Lippman et al., 2005; Strichman-Almashanu et al., 2002; Yamada et al., 2004). Similarly, methylation-dependent restriction in a cloning host has been employed as a filter against methylation-

rich sequences in clone libraries (Rabinowicz et al., 1999). Another, more recent genome-wide approach used immunopreciptation with a methyl cytosine antibody rather than restriction digestion to enrich for the methylated fraction (Weber et al., 2005). The enriched genome fractions are analyzed by sequencing or by array-hybridization (Lippman et al., 2004; Lippman et al., 2005). MSRE-based methods are somewhat indirect in that they discriminate for or against methylation at the recognition site of the particular enzyme used and cannot directly reveal the methylation status of cytosines or CpG dinucleotides outside the restriction site.

In contrast, methylation-sensitive chemical reactions have no specific recognition sequence. Sodium bisulfite efficiently deaminates unmethylated cytosine to uracil without affecting 5-methyl cytosine (Fig. 5). In recent years, PCR amplification and sequencing of bisulfite-converted genomic DNA has emerged as the gold standard for analyzing and comparing methylation patterns at specific loci (Frommer et al., 1992).

Despite these technological advances, in the absence of systematic sequence-based methylation analyses, the genomic methylation landscape in mammals is still largely unexplored. Therefore, the potential diagnostic value of specific methylation differences remains largely untapped.

The human epigenome project (HEP) is aimed at generating a high-resolution DNA methylation map of the human genome (Novik et al., 2002; Rakyan et al., 2004). To achieve this goal the bisulfite sequencing technique has been scaled-up in a targeted fashion using locus-specific PCR primers. We describe a new random approach for large-scale high-resolution DNA methylation analysis termed reduced representation bisulfite sequencing (RRBS) in the last part of Chapter 3.

DNA methylation and disease

Aberrant methylation patterns as a result of incomplete epigenetic reprogramming have been implicated in the abnormal development after NT (see above). ES cells are epigenetically unstable with many imprinted genes affected (Humpherys et al., 2001). Recent work has demonstrated a link between loss of imprinting (LOI), in particular Igf2, and tumorigenesis (Sakatani et al., 2005).

Most human cancers display genome-wide hypomethylation and concomitant promoterspecific tumor suppressor gene hypermethylation (Feinberg, 2004; Gaudet et al.,

2003; Jones and Baylin, 2002; Laird, 2003; Robertson, 2002). Generation of mice carrying a hypomorphic allele for Dnmt1 demonstrated that global hypomethylation resulted in aggressive T cell lymphomas that displayed a high frequency of chromosome 15 trisomy. These results indicated that DNA hypomethylation plays a causal role in tumor formation, possibly by promoting chromosomal instability (Gaudet et al., 2003).

Using conditional inactivation of de novo methyltransferase Dnmt3b in *Apc*$^{min/+}$ mice, our lab has demonstrated that loss of Dnmt3b has no impact on microadenoma formation, which is considered the earliest stage of intestinal tumor formation. However we observed a significant decrease in the formation of macroscopic colonic adenomas. Interestingly many large adenomas showed regions with Dnmt3b inactivation, indicating that Dnmt3b is required for initiation of macroscopic adenomas but is not required for their maintenance. Also Dnmt3b inactivation in normal colonic epithelium had no adverse effect on tissue homeostasis. These results support a role for Dnmt3b in the early stages of macroscopic colonic tumor initiation and suggest that it is not required for later tumor maintenance (Lin et al., 2006).

Consistent with the notion that both promoter hypermethylation and genome-wide hypomethylation are functionally important in tumorigenesis, genetic and or pharmacologic reduction of DNA methylation levels results in suppression or promotion of tumor incidence, respectively, depending on the tumor cell type. For instance, DNA hypomethylation promotes tumors that rely predominantly on loss of heterozygosity (LOH) or chromosomal instability mechanisms, whereas loss of DNA methylation suppresses tumors that rely on epigenetic silencing. Mutational and epigenetic silencing events in Wnt pathway genes have been identified in human colon tumors. *Apc*$^{Min/+}$ mice were recently used to investigate the effect of hypomethylation on intestinal and liver tumor formation. Intestinal carcinogenesis in *Apc*$^{Min/+}$ mice occurs in two stages, with the formation of microadenomas leading to the development of macroscopic polyps. Using *Dnmt1* hypomorphic alleles to reduce genomic methylation, showed an elevated incidence of microadenomas that were associated with LOH at *Apc*. In contrast, the incidence and growth of macroscopic intestinal tumors in the same animals was strongly suppressed. In contrast to the overall inhibition of intestinal tumorigenesis in hypomethylated *Apc*$^{Min/+}$ mice, hypomethylation caused development of multifocal liver

tumors accompanied by *Apc* LOH. These findings support the notion of a dual role for DNA hypomethylation in suppressing later stages of intestinal tumorigenesis, but promoting early lesions in the colon and liver through an LOH mechanism (Yamada et al., 2005).

Parts of this chapter have been published:

Meissner, A. and Jaenisch, R. From frog cloning to the generation of customized ES cells. BIF Futura (2006). In press

Wang, Z.*, **Meissner, A.***, and Jaenisch, R. *Nuclear Cloning and Epigenetic Reprogramming*. Handbook of Stem Cells. Academic Press (2004)

*equal contributing authors

Chapter 2- Materials and Methods

Generation of plasmids

To generate pSico the lox-CMV-GFP-lox cassette was removed from lentilox 3.7 (pLL3.7) (Rubinson et al., 2003) by digesting with *Bfu*AI and *Pci*I, followed by filling-in and religation. The first TATAlox followed by the terminator and by an *Eco*RI was inserted in the resulting plasmid by PCR-mediated mutagenesis using the following oligos: pSico6Eco GAATTCAACGCGCGGTGACCCTCGAGG; and pSico6 ASAAAAAACCAAGGCTT-ATAACTTCGTATAATTTATACTATACGAAGTTATAATTACTTTACAGTTACCC.

To insert the second TATAlox preceded by a *Not*I site the resulting plasmid was digested with *EcoR*I and *Xho*I and ligated to the following annealed oligos: TATALOX F: AATTCGAGAGGCGGCCGCATAACTTCGTATAGTATAAATTATACGAAGTT-ATAAGCCTTGTTAACGCGCGGTGACCC; and TATALOX R: TCGAGGG-TCACCGCGCGTTAACAAGGCTTATAACTTCGTATAATTTATACTATACGAAGT TATGCGGCCGCCTCTCG.

The resulting construct was finally digested with *Eco*RI and *Not*I and ligated to an *Eco*RI-CMV-GFP-*Not*I cassette to generate pSico. A similar strategy was employed to generate the various "test" constructs shown in Figure 7. Primer sequence and details are available upon request.

To generate pSico Reverse (pSicoR) the 5' loxP site present in pLL3.7 was removed by digesting with *Xho*I and *Not*I and replaced with a diagnostic *Bam*HI site using the following annealed oligos: Lox replace for TCGAGTACTAGG-ATCCATTAGGC and Lox replace rev GGCCGCCTAATGGATCCTAGTAC.

A new lox site was inserted 18 nt upstream of the proximal sequence element (PSE) in the U6 promotor by PCR-mediated mutagenesis.

Oligos coding for the various shRNAs were annealed and cloned into *Hpa*I-*Xho*I digested pLL3.7, pSico and pSicoR. Oligo design was as described (Rubinson et al., 2003). The following target regions were chosen: Nucleophosmin (Npm), GGCTGACAAAGACTATCAC; Luciferase, GAGCTGTTTCTGAGGAGCC; DNA methylatransferase 1 (Dnmt1), GAGTGTGTGAGGGAGAAA; and P53, GTACTCTCCTCCCCTCAAT.

The CD8 oligo sequence was the same described in (Rubinson et al., 2003). All constructs were verified by DNA sequencing. To amplify recombined and unrecombined vector the following oligos were used: Loopout F, CCCGGTTAATTTGCATATAA-TATTTC; and Loopout R: CATGATACAAAGGCATTAAAGCAG.

Antibodies, chemicals, flow cytometry and western blotting

Anti-α-tubulin antibody was from Sigma, the p53 antibody was a kind gift by K. Helin (European Institute of Oncology, Milan), and the anti-Npm was a gift from P. G. Pelicci (European Institute of Oncology, Milan) and E. Colombo (European Institute of Oncology, Milan). All mouse antibodies used were monoclonal. Doxorubicin and doxycycline were obtained from Sigma.

To assess expression of CD4 and CD8 in mice, single-cell suspensions of splenocytes were blocked with anti-CD16 CD32 for 10 min on ice. After blocking, the cells were incubated with phycoerythrin-conjugated anti-CD8, allophycocyanin-conjugated anti-CD4, and PerCPCy5.5-conjugated anti-CD3 for 20 min at 4°C (BD Pharmingen, San Diego). Acquisition of samples was performed on a FACScan flow cytometer, and the data were analyzed with CELLQUEST software (BD Immunocytometry Systems, San Jose, CA). Plots were gated on CD3 cells.

For cell-cycle analysis, 10^6 cells were fixed in 70% ethanol, washed in PBS, and resuspended in 20 μg/ml propidium iodide (Sigma) and 200 μg/ml RNAseA in PBS.

For western blotting cells were lysed in a buffer containing 1% TritonX-100, 10mM TrisCl and 140mM NaCl and a protease inhibitor cocktail (SIGMA). Proteins were resolved by SDS-PAGE, transferred to a filter, blocked overnight in 5% fat-free milk in TBS 0.1% Tween (TBS-T). After 1h incubation with the primary antibody filters were washed in TBS-T, incubated 30 minutes with the appropriate HRP-conjugated secondary antibody, washed 3 times in TBS-T and processed using the ECL plus kit and exposed to film.

Luciferase assay

For reporter assay, 293T cells were cotransfected in 12-well plates by using FuGENE 6 with the appropriate shRNA vectors together with pGL3control and pRLSV40. The total amount of transfected DNA was 500 ng per well. Firefly and Renilla luciferase activity were measured 36 h after transfection by using the dual reporter kit

(Promega) according to the manufacturer's instruction. All experiments were performed in triplicate.

Immunocytochemistry

Cells plated on glass coverslips that had been pre-incubated with 0.1% gelatin in PBS at 37˚C for 30 minutes were fixed in 4% paraformaldehyde (in PBS) for 10 minutes, washed with PBS and permeabilized by incubating in PBS 0.1% Triton X-100 for 10 minutes at room temperature. To prevent non-specific binding of the antibodies, cells were then incubated with PBS in the presence of 5% Bovine Serum Albumin (BSA) for 30 minutes. The coverslips were then gently deposited, face down, on 100 µl of primary antibody diluted in PBS 5% BSA. After one-hour, coverslips were washed three times with PBS (5 minutes per wash). Cells were then incubated 30 minutes at RT with the appropriate secondary antibody Cy3 (Amersham), Alexa 488- or Alexa 350-conjugated (Molecular Probes). Coverslips were mounted in a 90% glycerol solution containing diazabicyclo-(2.2.2)octane antifade (Sigma) and examined by fluorescence microscopy. Images were further processed with the Adobe Photoshop software (Adobe).

Southern blot and methylation analysis

DNA was isolated from the indicated ES cell lines. To assess the levels of DNA methylation, genomic DNA was digested with *Hpa*II, and hybridized to pMR150 as a probe for the minor satellite repeats (Chapman et al., 1984), or with an IAP-probe (Walsh et al., 1998). For the methylation status of imprinted genes, a combined bisulfite restriction analysis (COBRA) assay was performed with the CpGenome DNA modification kit (Chemicon) using PCR primers and conditions described previously (Lucifero et al., 2002). PCR products were gel purified, digested with *Bst*UI or *Hpy*CH4 IV and resolved on a 2% agarose gel. NNA was done as previously described (Ramsahoye, 2002b).

Northern blots

For the small RNA Northern blotting, 15 µg total RNA isolated with TRIzol (Invitrogen) according to the manufacturer's instructions, and was resolved on a 15% denaturing polyacrylamide gel, transferred to a nylon membrane, and was cross linked by using the autocrosslink function of a Stratalinker. The membrane was hybridized

overnight to a ^{32}P 5'-labeled DNA probe corresponding to the 19-nt sense strand of the p53 shRNA (GTACTCTCCTCCCCTCAAT). Hybridization and washes were performed at 42˚C.

For detection of the p53 mRNA, 15 µg of total RNA was resolved on an agarose-formaldehyde gel, transferred to a nylon membrane, and hybridized to a probe corresponding to the entire p53 coding sequence.

Cloning and design of shRNAs

shRNAs were designed using the pSicoOligomaker 1.5 (developed by Andrea Ventura/Jacks lab), which is freely available at http://web.mit.edu/ccr/labs/jacks/protocols/pSico.html. Cloning into pSico and pSicoR was done as described on the website.

Generation of lentivirus, infection and Cre-mediated recombination

Lentiviruses were generated essentially as described (Rubinson et al., 2003). Briefly, 5 µg of lentiviral vector and 2.5 µg of each packaging vector were cotransfected in 293T cells by using the FuGENE 6 reagent (Roche Diagnostics). Supernatants were collected 36–48 h after transfection, filtered through a 0.4- µm filter, and used directly to infect MEFs. Two rounds of infection 8 h apart were usually sufficient to infect 90% of cells. GFP-positive cells were sorted 3–4 days after infection. For ES cell infection, the viral supernatant was centrifuged at 25,000 rpm in a Beckman SW41t rotor for 1.5 h, the viral pellet was resuspended in 200 µl of ES cell medium, and was incubated 6 h at 37°C with 10,000–20,000 cells. After infection, ES cells were plated in 10-cm dishes with feeders and GFP-positive colonies were isolated 4–5 days later. On average, 10–30% of ES colonies were GFP-positive. Recombinant adenoviral stocks were purchased from the Gene Transfer Vector Core facility of University of Iowa College of Medicine (Iowa City, IA). Infections were performed by using 100 plaque-forming units of virus per cell. The number of integrations was determined by Southern blot analysis. Genomic DNA was digested with XbaI (single cut in the viral backbone) and probed with an EGFP-probe.

Cre-mediated recombination was achieved by transiently tranfecting a Cre-recombinase containing plasmid. Briefly, after introducing the Cre plasmid into the ES cells by electroporation, cells were cultured for 24h (not longer, to avoid random

integration of plasmid) in ES medium plus puromycin. GFP-negative subclones were picked, expanded and tested for recombination (*Cdx2*^{1Lox}).

Immunohistochemistry and RT-PCR

Cdx2 staining of blastocysts was done as described in (Strumpf et al., 2005). The protocol is available on the Rossant laboratory website (http://www.mshri.on.ca/rossant/protocols/immunoStain.html). Monoclonal anti-Cdx2 (CDX2-88, BioGenex, CA, USA) was used for all Cdx2 stainings. RT-PCR was done with a One-Step RT-PCR Kit (Qiagen) using the following primer: β-actin 5-ggcccagagcaagagaggtatcc-3 (forward) and 5-acgcacgatttccctctcagc-3 (reverse), *Oct-4* (333bp) ggatggcatactgtggacct (forward) and agatggtggtctggctgaac (reverse), *Cdx2* (225bp) AAACCTGTGCGAGTGGATG (forward) and CTGCGGTTCTGAAACCAAAT (reverse). β-actin RT primer were published in ref. (Strumpf et al., 2005), and *Oct-4* and *Cdx2* primer were design using PRIMER3.

Nuclear transfer, embryo transfer, ES cell derivation and 2N/4N blastocyst injections

Nuclear transfer was done as described (Eggan et al., 2001a; Wakayama et al., 2005). Nuclear transfer embryos were transferred at day 3.5 (morula/blastocyst stage) into the uteri of day 2.5 pseudo-pregnant recipient females. For ES cell derivation, the zona pellucida was removed using acidic tyrode (AT) solution and blastocyst were explanted on irradiated feeders in ES medium plus MEK1 inhibitor (PD98059). Diploid and tetraploid blastocyst injections were done as described in (Wang and Jaenisch, 2004).

ES cell manipulation

ES cells were cultivated on irradiated mouse embryonic fibroblasts (MEFs) in DMEM containing 15% fetal calf serum, leukemia inhibiting factor, penicillin/streptomycin, L-glutamine, and non-essential amino acids. All ES cells were depleted of feeder cells for two passages on 0.2% gelatine before isolating DNA.

RRBS library construction and sequencing

Mouse ES DNA (50-100 μg) was digested to completion by overnight incubation with 1,000 units of *Bgl*II and electrophoresed on a 1.8% agarose gel. Marker lanes were

stained with SYBR Green (Invitrogen). A narrow slice containing the 500-600 bp fraction was excised from the unstained preparative portion of the gel. DNA was recovered by electroelution, phenol extraction and ethanol precipitation as described elsewhere (Garnes, 2002). Typical yields were 300-600 ng of size-selected *Bgl*II fragments as measured by PicoGreen fluorescence (Invitrogen). The size-selected *Bgl*II fragments (1-2 pmol) were ligated to 700 pmol *Bgl*II adapter pre-annealed from oligodeoxynucleotides 5'-AGTTATTCCGGACTGTCGAAGCTGAATGCCATGG and 5'-pGATCCCATGGCATTCAGCTTCGACAGTCCGGAAT in 70 µl containing 2,400 units T4 DNA ligase (New England Biolabs) for 16 h at 14°C. Excess adapter was removed by ultrafiltration (Millipore Montage) followed by preparative electrophoresis in 2% agarose and electroelution, yielding 50-100 ng of adapter-ligated material.

Adapter-ligated, size-selected *Bgl*II fragments (50 ng) were bisulfite-treated using the reagents and protocol of the CpGenome DNA modification kit (Chemicon) with the following modifications: the DNA was alkali-denatured for 20 min. at 55°C; the total reaction volume was increased from 650 µl to 750 µl and contained 0.22 g urea (Paulin et al., 1998) ; the mixture was incubated for 24 h at 55°C. After alkaline desulfonation and final desalting, single-stranded uracil-containing reaction products were eluted in 40 µl of TE buffer and converted to double-stranded DNA by PCR with primers 5'-TTGGATTGTTGAAGTTGAATG and 5'-AAACTATCAAAACTAAATACCATAAAATC designed to amplify molecules carrying bisulfite-modified adapter sequences at both ends. For each bisulfite reaction, eight 50 µl PCRs were performed, each containing 2.5 µl bisulfite-treated DNA, 25 pmol of each PCR primer and 2.5 units PfuTurboCx Hotstart DNA polymerase (Stratagene). Thermocycling included eight cycles of "touchdown" (Don et al., 1991) at annealing temperatures from 55°C to 52°C (two cycles at each temperature) followed by 10 cycles at an annealing temperature of 51°C. Denaturation (94°C), annealing and extension (72°C) times were 10 s, 30 s and 3 min., respectively. PCR products were cleaned-up by ultrafiltration followed by preparative electrophoresis on a 2% agarose gel. Typical yields were between 50 and 100 ng for each library. Gel-purified PCR product (4 ng) were incubated for 5 min with 1 µl pCR BluntII TOPO vector and cloned by electroporation of

Escherichia coli TOP10 (Invitrogen). The cloning efficiency was ~2,000 colonies per ng of PCR product.

Plasmid DNA was isolated by standard protocols, and cloned inserts were sequenced using 2.7 pmol M13 reverse primer and 2 µl BigDye3.1 mix (Applied Biosystems) in 10-µl sequencing reactions (25 cycles). Caused by preferential cloning in one orientation, ~80% of the sequences were the G-poor strand. Most inserts that had been cloned in the other orientation (C-poor strand) sequenced poorly, with peak-heights and sequence quality suddenly dropping after 300-400 bases.

Data analysis

In silico digestion of the mouse genome (NCBI Build 33, May 2004) was performed at *Bgl*II sites, followed by selection of fragments ranging from 440 to 640 bases. Cytosines were converted to thymine, with upper/lower case used to differentiate converted from original thymines. Each strand was converted separately. Sequencing reads were mapped to the genome by using NCBI BLAST (without query filtering) to search the database of size-selected and converted *Bgl*II fragments. The top BLAST hit determined the most probable genome location of each read and also permitted identification of original and converted cytosines over the high-scoring segment pair length. The repeat content of the *in silico* reduced representation and the sets of sequencing reads were compared to that of the whole genome using RepeatMasker (http://www.repeatmasker.org). Locations of all sequence reads relative to selected genomic landmarks were determined by comparing fragment coordinates to those of the RefSeq and Ensembl transcript sets and CpG islands from UCSC.

The expected number of redundant RRBS sequences and the sequence overlap between two DNA samples were calculated by composite Poisson statistics in 5 bp bins across the range of insert sizes. D_i is the number of *Bgl*II fragments in the reference genome that fall into bin *i*. N_{ai} is the number of successful sequencing reads from DNA sample *a* that fall into bin *i*. $\sum(N_{ai}^2/2D_i)$ double-hits in sample *a* are expected by random sampling (Altshuler et al., 2000). The expected number of *Bgl*II fragments sequenced at least once in sample *a* and in sample *b* is $\sum[(1-e^{-(Nai/Di)})x(1-e^{-(Nbi/Di)})xD_i]$.

Chapter 3- Results

Cre-lox regulated conditional RNA interference in cells and mice

Abstract: The use of RNA interference for studying gene function has become an essential part of biology. Owing to the dominant nature of RNAi, a major limitation of this approach is that germ-line transmission can be obtained only for shRNAs targeting genes whose knock-down is compatible with animal viability and fertility. Moreover, even for cell-based applications, constitutive knock-down of gene expression by RNAi can limit the scope of experiments, especially for genes whose inhibition leads to cell lethality. To overcome these limitations, and to extend the applications of RNAi in mammalian systems, we have developed a Cre-lox-based approach for the conditional expression of shRNA. Two different strategies were used to generate mouse embryonic fibroblasts (MEFs), embryonic stem (ES) cells and transgenic mice in which the expression of an shRNA is tightly regulated in a Cre-dependent manner. One vector allows for conditional activation of shRNA expression, whereas the other permits conditional inactivation of expression of the hairpin RNA. The ability to efficiently control shRNA expression using these vectors was shown in cell-based experiments by knocking down p53, nucleophosmin and DNA methyltransferase-1. We also demonstrate the usefulness of this approach to achieve conditional, tissue-specific RNA interference in Cre-expressing transgenic mice. Combined with the growing array of Cre expression strategies, these vectors allow spatial and temporal control of shRNA expression *in vivo* and should facilitate functional genetic analysis in mammals.

This part of Chapter 3 has been published:
Ventura A*, **Meissner A***, Dillon CP, McManus M, Sharp PA, Van Parijs L, Jaenisch R, Jacks T. Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci U S A. 2004 Jul 13; 101 (28):10380-5.
*equal contributing authors

As outlined in Chapter 1 the rational for generating a system for conditional gene suppression by RNAi was to investigate conditions, such as global hypomethylation, that could enhance the reprogramming efficiency. The vectors described here allow in a simple manner to study the effects of transiently down regulating essential genes. By subsequently reversing the knockdown and thereby restoring endogenous gene expression many deleterious effects of longer term suppression can be avoided. In the following part, I will describe how the vectors were generated and how the effectiveness of both vectors was tested for a number genes *in vitro* and *in vivo*.

Generation of pSico and pSicoR

The U6 promoter has been widely used to drive the expression of shRNAs and a U6-based lentiviral vector for the generation of transgenic mice has been recently described (Rubinson et al., 2003). To control shRNA expression in a Cre-dependent manner, we decided to modify the mouse U6 promoter by inserting a Lox-STOP-Lox cassette. Similar to other RNA polymerase III promoters, the U6 promoter is extremely compact, consisting of a tightly spaced TATA box, a proximal sequence element (PSE), and a distal sequence element (DSE) (Fig. 6a). Mutagenesis experiments have demonstrated that while the DSE is partially dispensable for transcriptional activity, the PSE and the TATA box are absolutely required. Moreover, the spacing between the PSE and the TATA box (17 nt) and between the TATA box and the transcription start site (25 nt) is critical, as even small changes have been shown to severely impair promoter activity (Paule and White, 2000). A consequence is that to effectively suppress the activity of the U6 promoter, the Lox-STOP-Lox element must be positioned either between the PSE and the TATA box or between the TATA box and the transcription start site. In addition, to reconstitute a functional promoter, after Cre expression, the normal spacing between PSE, TATA box, and transcription start site must be restored. The latter consideration precludes the utilization of a classic lox-STOP-lox cassette because, after Cre-mediated recombination, the residual loxP site (34 nt) would necessarily increase the PSE-TATA or the TATA-start-site spacing, thus resulting in a non-functional promoter (See Fig. 7).

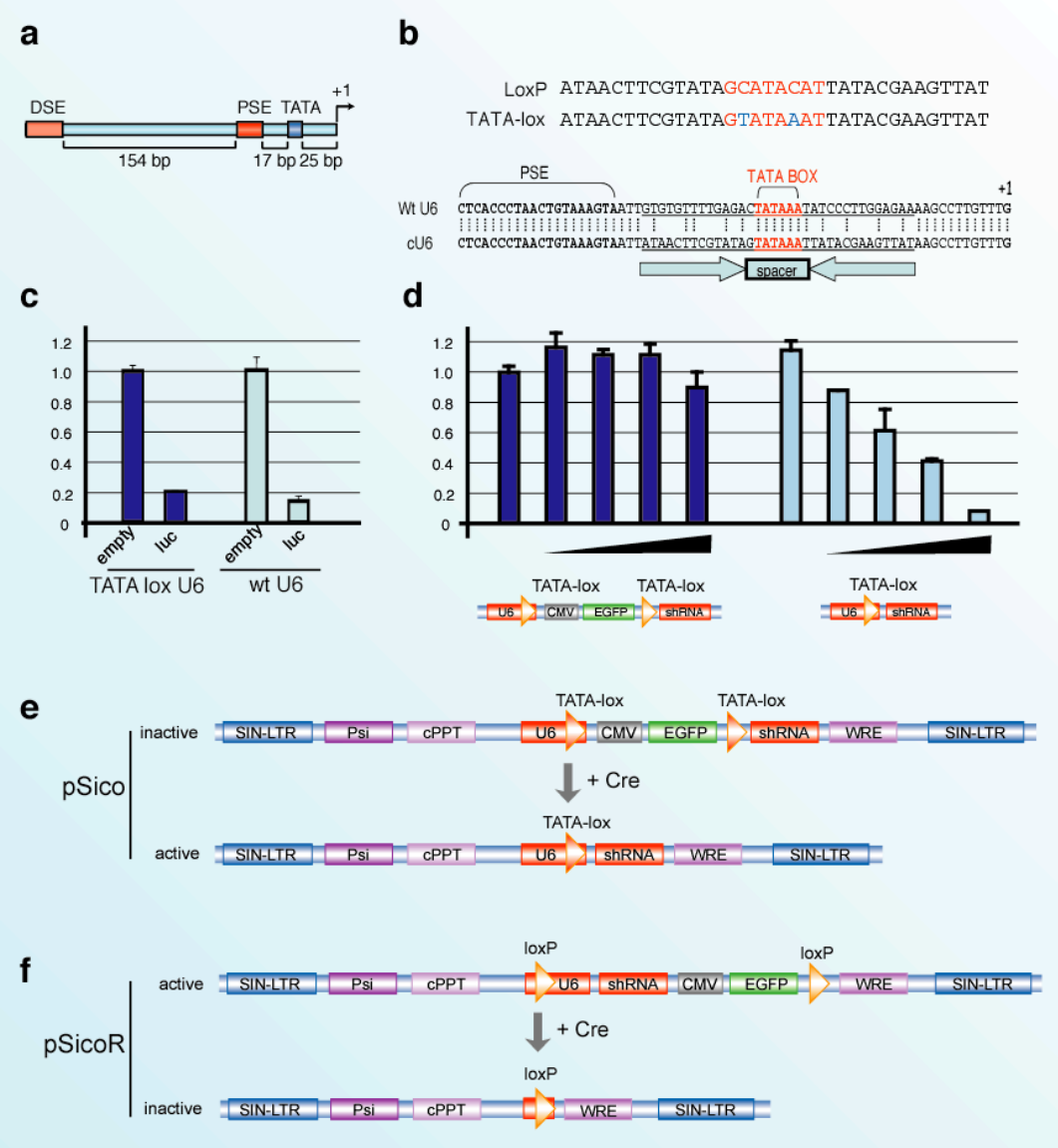

Figure 6: Generation of pSico and pSicoR. **a**) Schematic representation of the mouse U6 promoter. The spacing between the DSE, the PSE, the TATAbox, and the transcription start site (1) is indicated. **b)** Comparison between the sequence of a loxP site and a TATAlox site (*Upper*). Comparison between the sequence of the wild-type mouse U6 promoter and the sequence of the U6 promoter with a TATAlox site replacing the TATA box (*Lower*). **c)** The TATAlox can replace the TATA box in the U6 promoter. Equal amounts of the wild-type U6 promoter and of the TATAlox U6 promoter (empty or driving the expression of shRNA against the firefly luciferase gene) were transfected in 293T cells together with reporter plasmids expressing firefly luciferase and renilla luciferase. Thirty-six hours later, cells were lysed and the ratio between firefly and renilla luciferase activity was measured. **d)** A TATAlox-STOPTATAlox cassette in the U6 promoter efficiently suppresses shRNA expression. Increasing amounts (0–200 ng) of plasmids containing the indicated version of the U6 promoter were transfected in 293T cells together with reporter plasmids, and luciferase activity was measured as in c. **e)** Schematic representation of pSico before and after Cre-mediated recombination. **f)** Schematic representation of pSicoR before and after Cre-mediated recombination. SINLTR, self-inactivating long terminal repeats; Psi, required for viral RNA packaging; cPPT, central polypurine tract; EGFP: enhanced GFP; WRE, woodchuck regulatory element.

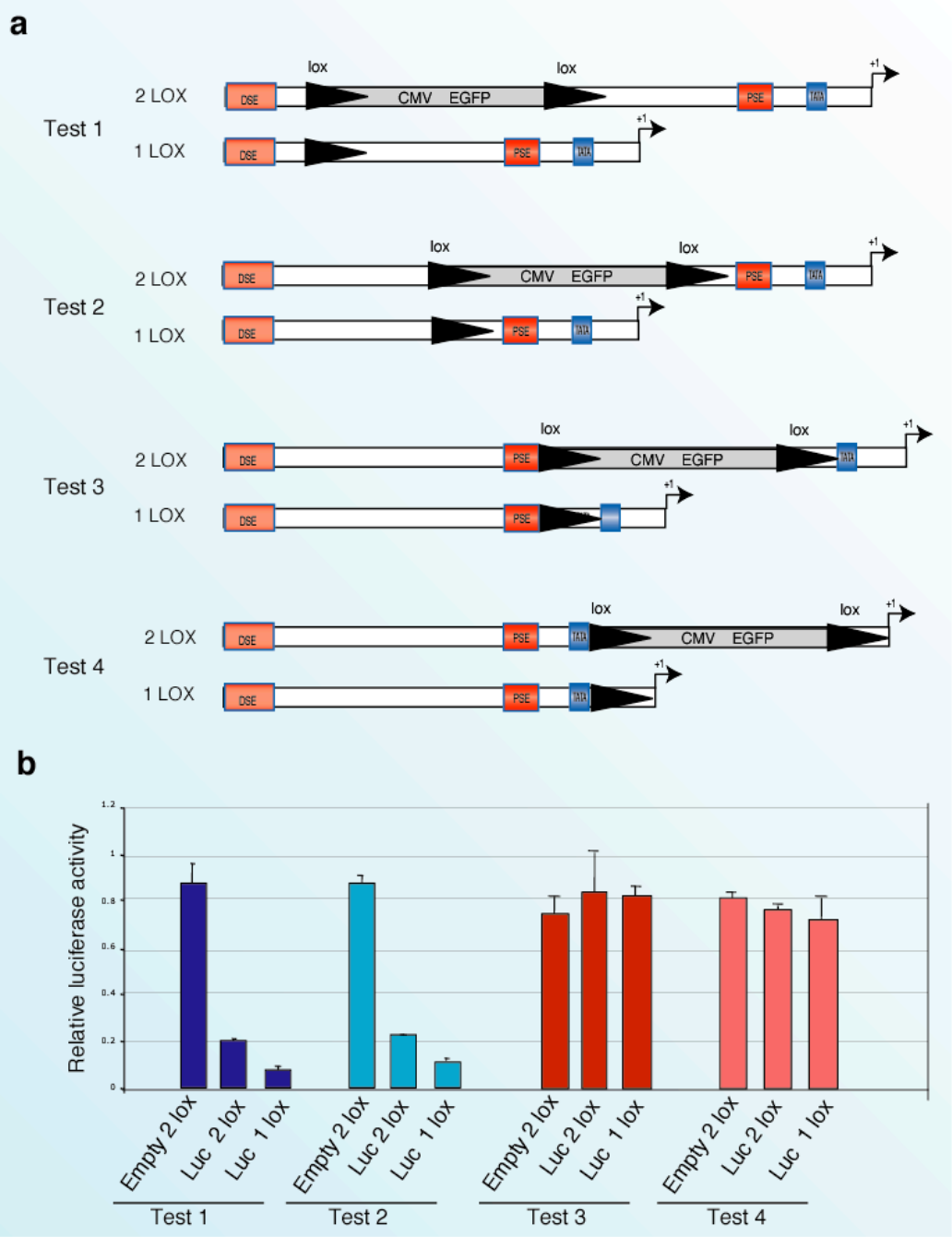

Figure 7: a) Schematic representation of the U6 promoters carrying the lox-CMVGFP-lox tested in panel 6b. The CMV-GFP cassette is not drawn to scale. Test 1 and Test 2 have the lox-STOP-lox cassette between the DSE and the PSE. In Test 3 the cassette is positioned between the PSE and the TATA box and finally in Test 4 it is positioned between the TATAbox and the putative transcription start site. **b)** The indicated U6 constructs were assayed as in Figure 6c for their ability to induce knock-down of the firefly luciferase gene. Note that constructs containing the lox-stop-lox cassette upstream of the PSE are still capable of efficiently repressing luciferase activity (Test 1 and Test 2), while the constructs in which the lox-stop-lox cassette is situated between the PSE and the TATA (Test 3) or between the TATA and the transcription start site (Test 4), are inactive even in the recombined conformation indicating that in both cases the residual lox site negatively affects U6 promoter activity.

To overcome these limitations, we generated a bifunctional lox site (TATAlox), that, in addition to retaining the ability to undergo Cre-mediated recombination, contains a functional TATA box in its spacer region (Fig. 6b-d).

As shown in Figure 6, when the TATAlox replaces the TATA box site in the U6 promoter, the spacing between PSE, TATA and transcriptional start site is not altered (Fig. 6b), and the resulting promoter retains transcriptional activity (Fig. 6c).

To create a conditional U6 promoter, a cytomegalovirus (CMV)-enhanced GFP stop/reporter cassette was inserted between two TATAlox sites so that after Cre-mediated recombination the cassette would be excised, generating a functional U6 promoter with a TATAlox in place of the TATA box (Fig. 6d). A T_6 sequence was positioned immediately upstream of the CMV promoter to serve as a termination signal for RNA polymerase III. The terminator combined with the inserted CMV-GFP cassette completely suppressed the activity of the U6 promoter (Figs. 6d and 8c and 8d). To facilitate the generation of conditional knock-down mice and cell lines, the conditional U6 cassette was inserted into a self-inactivating lentiviral vector derived from pLL3.7 (Rubinson et al., 2003). The resulting plasmid was named pSico (Fig. 6e).

To allow for conditional inactivation of shRNA expression, we generated a second vector named pSicoR (Fig. 6f). In pSicoR, the CMV-GFP reporter cassette is placed downstream of the U6 promoter and does not affect its activity. Two loxP sites in the same orientation are present in this vector; the first positioned immediately upstream of the PSE in the U6 promoter, and the second immediately downstream of the GFP-coding sequence. In contrast to cells infected with pSico, cells infected with pSicoR are expected to constitutively transcribe the desired shRNA until a Cre-mediated recombination event leads to the excision of the CMV-GFP cassette and an essential part of the U6 promoter. Importantly, in both pSico and pSicoR, the CMV-GFP cassette marks infected cells and loss of GFP expression indicates successful Cre-mediated recombination.

Cre-regulated RNAi in cells

The ability of pSico and pSicoR vectors to conditional silence endogenous genes was demonstrated by insertion of a hairpin designed to inhibit expression of the mouse tumor suppressor gene p53. As a control, the same sequence was cloned into the

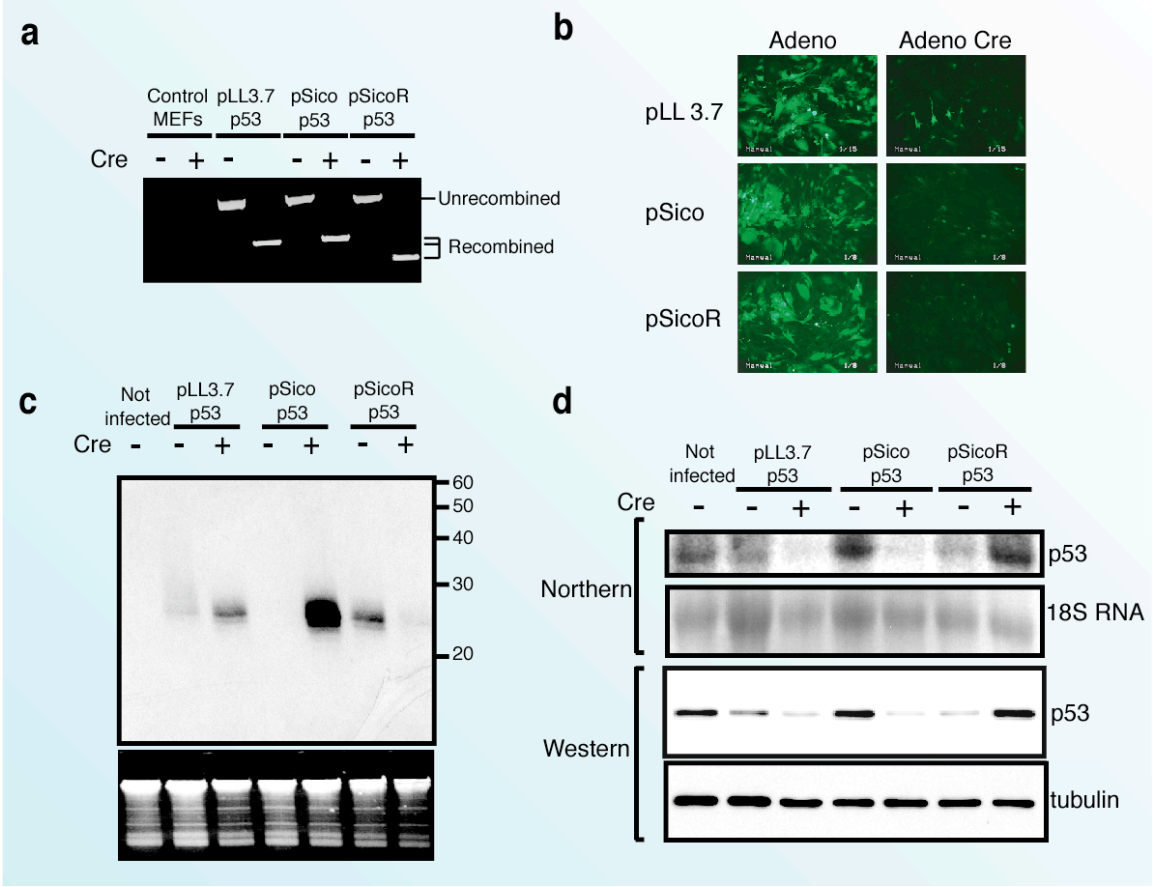

Figure 8: Cre-regulated knockdown of p53. **a)** $p53^{R270H/-}$ MEFs infected with the indicated lentiviruses were sorted for GFP positivity and infected with Ad or Ad-Cre. Four days after infection, genomic DNA was extracted, and a PCR was performed to amplify the recombined and unrecombined viral DNA. **b)** The same cells were analyzed by epifluorescence microscopy to detect GFP. Similar cell density and identical exposure time was used for all images. **c)** Fifteen micrograms of totalRNAextracted from the above indicated MEFs was separated on a 15% denaturing polyacrylamide gel, transferred on a nitrocellulose filter, and hybridized to a radi-labeled 19mer corresponding to the sense strand of the p53 shRNA. Equal RNA loading was assessed by ethidium bromide staining of the upper part of the gel (*Lower*). **d)** Northern (*Upper*) and Western blotting (*Lower*) showing p53 knock-down in the above indicated cells.

constitutive shRNA vector pLL3.7. In pLL3.7 the CMV-GFP cassette is located downstream of the U6 promoter and is flanked by loxP sites such that Cre-mediated recombination is expected to result in loss of GFP expression without affecting shRNA expression (Rubinson et al., 2003). These three constructs were then used to generate lentiviruses and infect MEFs. To simplify the detection of p53, MEFs expressing high basal levels of a transcriptionally inactive point mutant (R270H) p53 allele (K. Olive and T. Jacks, unpublished work) were used in these experiments. High-efficiency transduction by all of these vectors was achieved as indicated by uniform GFP expression in infected cells (Figure 8b and data not shown). As shown in Fig. 8, after superinfection

with a Cre-expressing recombinant adenovirus (Ad-Cre), near complete recombination with concomitant loss of GFP fluorescence was observed for all vectors. One week after Cre expression, high levels of the p53-siRNA were detected in cells infected with pSico-p53 (Fig. 8c), whereas no p53-siRNA was observed in the same cells in the absence of Cre expression, confirming the complete suppression of U6 promoter activity by the TATAlox-STOP-TATAlox cassette. The length of the processed RNA (21-24 nt) was identical in cells infected with pLL3.7-p53, pSico-p53 (after Ad-Cre infection), or pSicoR-p53 (before Ad-Cre infection), indicating that the presence of the TATAlox in pSico does not qualitatively affect shRNA production. Finally, infection with Ad-Cre led to almost complete disappearance of p53-siRNA in pSicoR-p53-infected cells (Fig. 8c).

Consistent with functional p53-siRNA expression by these vectors, Cre-mediated recombination resulted in a dramatic reduction of both p53 mRNA and protein levels in pSico-p53-infected cells (Fig. 8d). Conversely, pSicoR-p53 generated a p53 knock-down that was reversed upon Ad-Cre infection (Fig. 8d). We noticed an unexpected increase in p53-siRNA and p53 knock-down after Cre expression in cells infected with pLL3.7-p53 (Fig. 8c and 8d, lanes 2 and 3). This increase could reflect promoter interference because the CMV and the U6 promoters are in close proximity in pLL3.7 before to Cre-mediated recombination.

As additional proof of concept, we cloned short hairpins directed against the nucleolar protein Npm and Dnmt1 into pSico and pSicoR. Npm is a putative tumor-suppressor gene involved in a number of chromosomal translocations associated with human leukemias and lymphomas, and has been shown to physically and functionally interact with the tumor suppressors p19ARF and p53 (Bertwistle et al., 2004; Colombo et al., 2002). Specific, Cre-dependent knock-down of Npm was observed in both MEFs and ES cell clones infected with pSico-Npm (Figs. 9a and b). The opposite effect, Cre-dependent reexpression of Npm, was observed in pSicoR-Npm infected MEFs (Figures 9a and 9c).

The characterization of ES cells mutant for Dnmt1 has been reported (Li et al., 1992), and demonstrated that Dnmt1 is required for genome-wide maintenance of cytosine methylation. Dnmt1-deficient ES cells are viable and proliferate normally, despite substantial loss of methylation; however, they die upon differentiation. Whereas

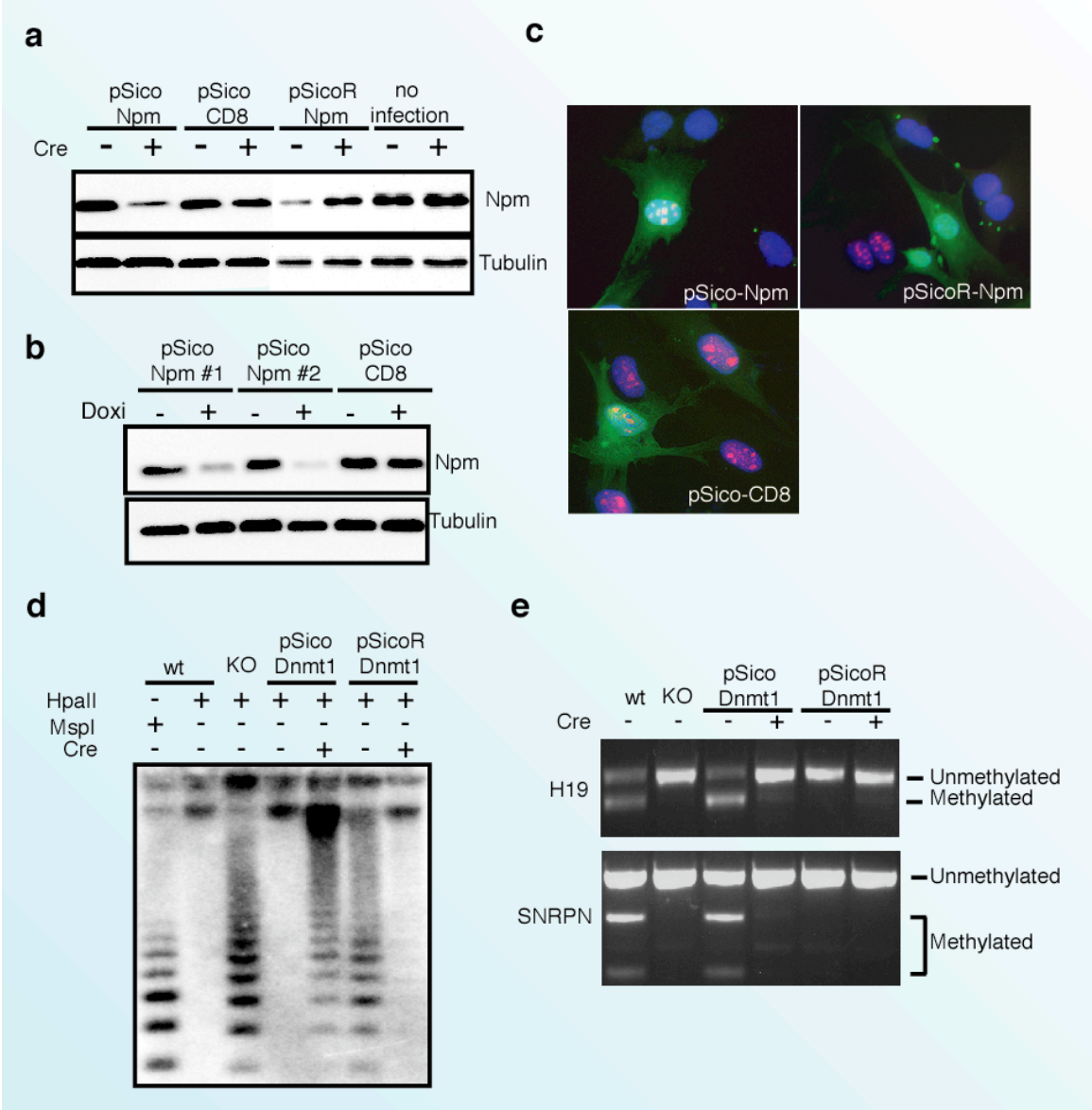

Figure 9: Cre-regulated knockdown of Npm and Dnmt1. **a)** Cre-regulated knock-down of Npm. MEFs were infected with the indicated lentiviruses, and GFP-positive cells were sorted and were superinfected with empty Ad or Ad-Cre. One week later, whole-cell lysates were separated by SDS PAGE, and were subjected to Western blotting against Npm and tubulin. **b)** ES cells carrying a doxycycline-inducible Cre (C. Beard and R.J., unpublished data) were infected with the indicated lentiviruses. GFP-positive clones were isolated, passaged two times, and were either left untreated or were incubated with 2 µg/ml doxycycline for 1 week. Immunoblot analysis was performed as in a. **c)** Immunofluorescence microscopy analysis of MEFs infected with pSico-Npm, pSicoR-Npm or pSico-CD8. After lentiviral infection GFP-positive MEFs were sorted and superinfected with empty Adenovirus or Ad-Cre. One week later cells were co-plated on glass coverslips, fixed and decorated with anti Npm antibody (red). Nuclei were stained with DAPI. **d)** Cre-regulated knock-down of Dnmt1 affects cytosine methylation. Methylation analysis of minor satellite DNA. ES cells carrying a doxycycline-inducible Cre transgene were infected with the indicated lentiviruses. Single GFP-positive clones were isolated, expanded, and passaged five times before being either mock-treated or incubated with 2 µg/ml doxycycline. After five more passages, the genomic DNA was extracted and digested with the indicated enzymes and subjected to Southern blot analysis. **e)** As in d, but the genomic DNA was treated with sodium bisulfite, subjected to PCR to amplify the indicated imprinted regions, and digested with *Bst*UI.

re-expression of the Dnmt1 cDNA in these cells leads to methylation of bulk genomic DNA and nonimprinted genes, the methylation pattern of imprinted loci cannot be restored without germ-line passage (Tucker et al., 1996). We tested whether we could recapitulate the phenotype observed in Dnmt1-deficient ES cells by using pSico-Dnmt1 and pSicoR-Dnmt1. As shown in Figure 9, pSico-Dnmt1-infected ES cells underwent significant loss of CpG methylation of minor satellites (Fig. 9d) and of two imprinted genes tested (Fig. 9e) upon Cre induction. Importantly, the reacquisition of DNA methylation at minor satellites sequences, but not at imprinted loci in pSicoR-Dnmt1 after Cre-mediated recombination, confirms previous results obtained with reexpression of Dnmt1 (Tucker et al., 1996). These results further illustrate the potential for application of the pSicoR vector *in vitro* and *in vivo* to perform "rescue" experiments.

Conditional RNAi in mice

One motivation for incorporating a conditional U6 cassette into a lentiviral vector was to rapidly generate conditional knock-down mice. To demonstrate this application directly, ES cells were infected with pSico-CD8 (Fig. 10a), which was designed to inhibit expression of the T lymphocyte cell surface marker CD8 (Rubinson et al., 2003). Three pSico-CD8 ES clones were used to generate chimeric mice, and transmission of the pSico-CD8 transgene to the progeny was observed for two of them. All transgenic mice were easily identified by macroscopic GFP visualization (Fig. 10b), although we observed some variability in the extent and distribution of GFP expression among littermates. Importantly, all transgenic mice produced normal amounts of $CD4^+$ and $CD8^+$ lymphocytes and were apparently normal and fertile, indicating that the presence of the nonexpressing pSico-CD8 transgene before Cre activation did not affect CD8 expression and was compatible with normal mouse development. To achieve either global or tissue-specific activation of the CD8 shRNA, pSico-CD8 chimeras were crossed to Msx2-Cre or Lck-Cre transgenic mice that express Cre in the oocyte (Gaudet et al., 2004; Sun et al., 2000), or under the control of a T cell-specific promoter (Hennet et al., 1995), respectively. Flourescence-activated cell sorter analysis demonstrated that pSico-CD8;Lck-Cre and pSico-CD8;Msx2-Cre mice had a specific reduction in splenic $CD8^+$, but not $CD4^+$ T lymphocytes as compared to controls (Fig. 10c). As predicted, the pSico-CD8;Msx2-Cre progeny showed complete recombination of

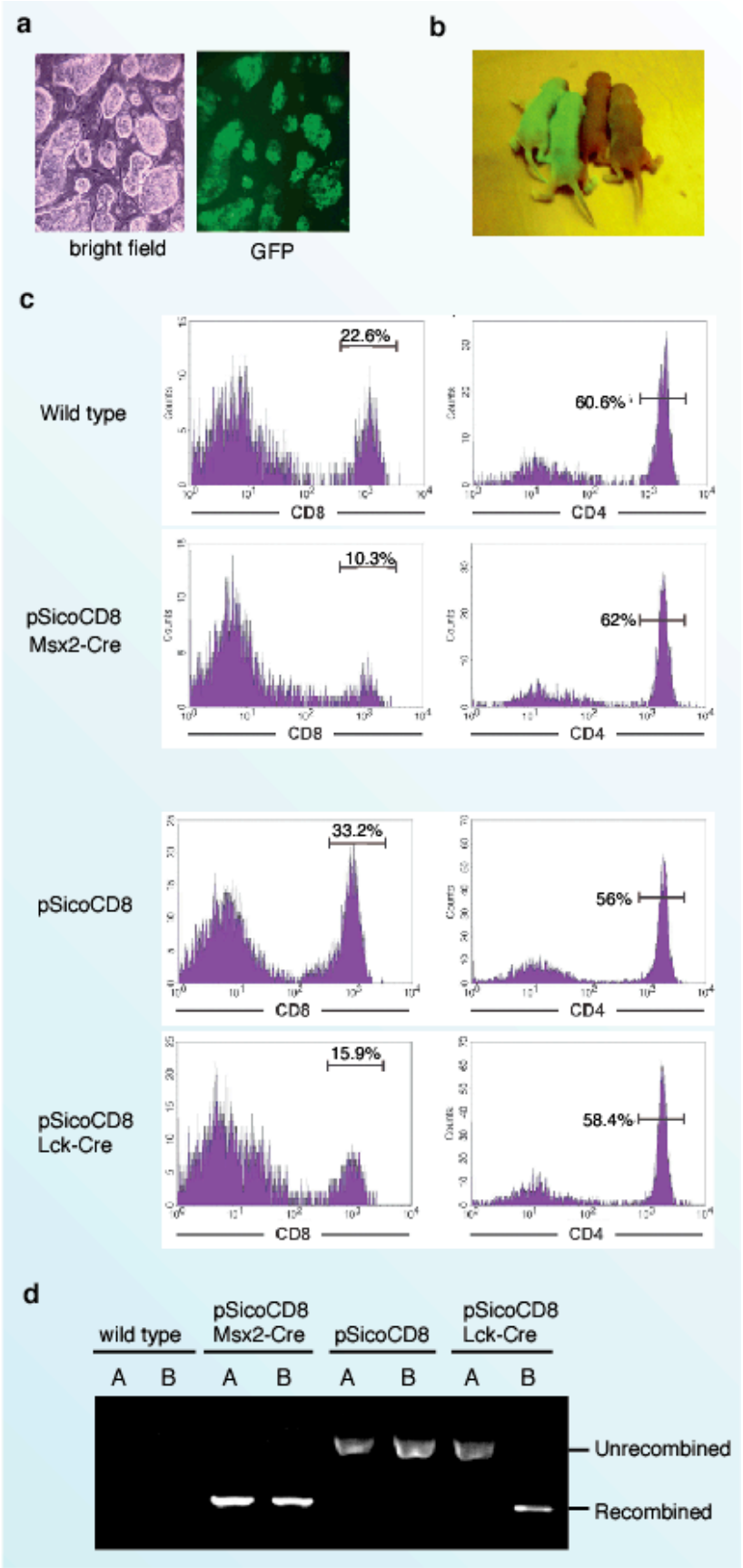

Figure 10: Conditional knockdown of CD8 in transgenic mice. **a)** ES cells infected with pSico-CD8 visualized with an inverted fluorescence microscope. **b)** A litter of newborns derived from a cross between a pSico-CD8 chimera and an Lck-Cre female. Three pups present bright GFP fluorescence, indicating germ-line transmission of the pSico-CD8 transgene. **c)** Knock-down of CD8 in the spleen of Msx2-Cre pSico-CD8 and Lck-Cre pSico-CD8 mice. Chimeras from pSico-CD8-infected ES cells were crossed to Msx2-Cre or Lck-Cre animals. The resulting mice were genotyped for the presence of Cre and pSico. Splenocytes from 1- to 3-week old mice with the indicated genotypes were harvested, stained for CD3, CD4, and CD8 expression, and analyzed by flow cytometry. Only CD3 cells were plotted. One representative example of littermates for each cross is shown. **d)** PCR detection of Cre-mediated recombination of pSico-CD8 in genomic DNA extracted from the tail (*A*) or the thymus (*B*) of mice with the indicated genotypes.

the pSicoCD8 transgene and lacked detectable GFP expression, while in the pSico-CD8;LckCre mice recombination was detected in the thymus but not in other tissues (Fig. 10d and data not shown). Transgenic mice derived from two different ES clones gave similar results.

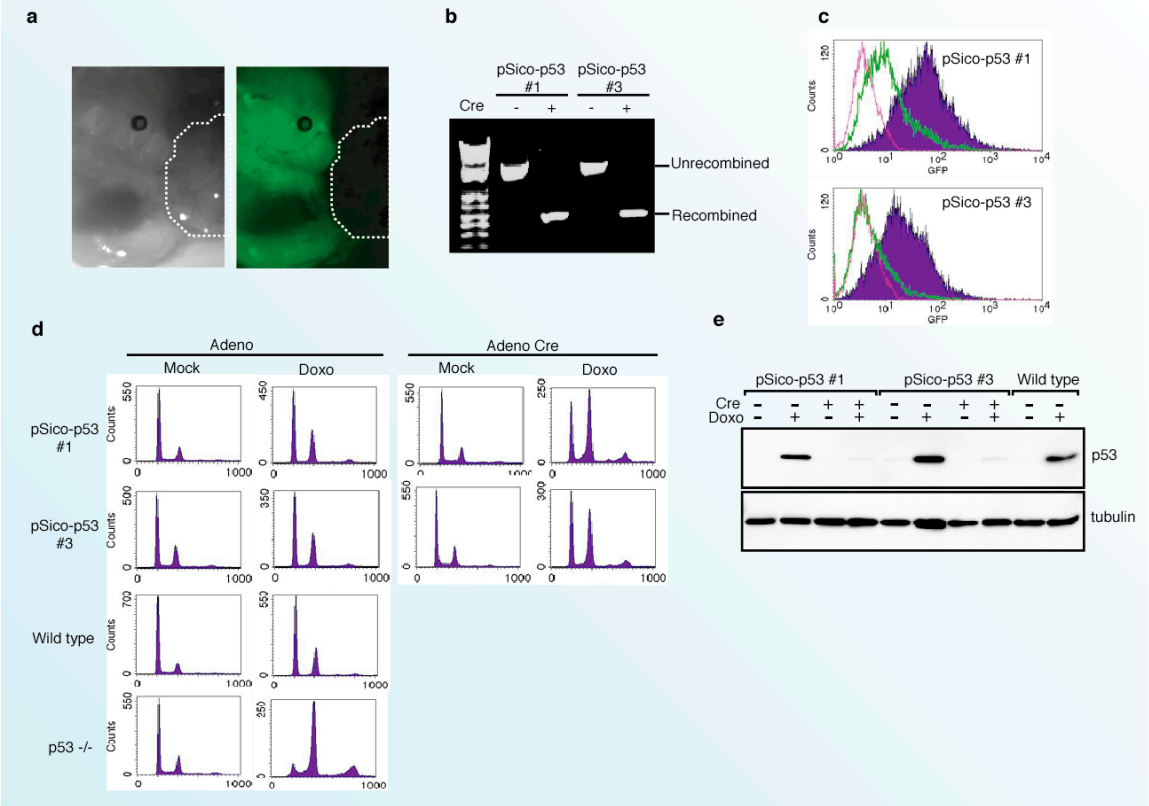

Figure 11: Generation of conditional knockdown embryos by tetraploid complementation. **a)** A postnatal day 14.5 embryo derived by tetraploid complementation using the pSico-p53 #1 ES clone. The area enclosed by the dashed line corresponds to the non-ES cell-derived placenta. **b)** PCR detection of recombination in MEFs derived from the indicated embryos. Genomic DNA was extracted 4 days after Ad or Ad-Cre infection and subjected to PCR. **c)** Histogram overlays showing loss of GFP expression in MEFs derived from pSico-p53#1 (*Upper*) and pSico-p53#3 (*Lower*) embryos 4 days after Ad-Cre (green plot) or Ad empty (purple filled plot) infection. Control, GFP-negative MEFs (red plot) are included as reference. **d)** Cell cycle profile of MEFs derived from embryos with the indicated genotypes infected with Adeno empty or AdenoCre and either mock treated or treated with 1µg/ml doxorubicin for 18 hours. **e)** MEFs derved from the indicated tetraploid complementation pSico-p53 embryos, or from wild-type embryos, were treated with doxorubicin for 18 h and subjected to Western blot against p53 and beta-tubulin.

Tetraploid blastocyst complementation represents a faster alternative to diploid blastocyst injection because it allows the generation of entirely ES-derived mice without passage through chimeras (Eggan et al., 2001a; Tanaka et al., 2001). In principle, this technology applied to pSico-infected ES cells would allow the generation of conditional

knock-down mice in ~5-6 weeks (1 week for cloning the shRNA, 1-2 weeks for ES cells infection and clone selection, and ~2 weeks for tetraploid blastocyst injection and gestation). To test this protocol directly, ES cells were infected with pSico-p53 and two different clones, pSico-p53#1 and pSico-p53#3, were injected into tetraploid blastocysts. As a rapid way to assess the inducibility of the p53 shRNA in ES cell-derived animals, midgestation embryos were recovered from two recipients females. Two apparently normal, GFP positive embryos were recovered; one each from ES clone pSico-p53 #1 and pSico-p53 #3 (Fig. 11a and data not shown). MEFs generated from these embryos were passaged once and infected with Ad or Ad-Cre. As expected, Cre expression induced significant recombination and loss of GFP expression (Figs. 11b and c). Importantly, in Ad-Cre-infected cells, p53 induction and cell-cycle arrest after doxorubicin treatment were significantly inhibited compared to Ad-infected control cells (Figs. 11d and e).

Generation of nuclear transfer-derived pluripotent ES cells from cloned Cdx2*-deficient blastocysts*

Abstract: The derivation of embryonic stem (ES) cells by nuclear transfer holds great promise for research and therapy but involves the destruction of cloned human blastocysts. Proof of principle experiments have shown that "customized" ES cells derived by nuclear transfer (NT-ESCs) can be used to correct immuno-deficiency in mice (Rideout et al., 2002). Altered Nuclear Transfer (ANT) has been proposed as a variation of nuclear transfer because it would create abnormal nuclear transfer blastocysts that are inherently unable to implant into the uterus but would be capable of generating customized ES cells (Hurlbut, 2005). To assess the experimental validity of this concept we have used nuclear transfer to derive mouse blastocysts from donor fibroblasts that carried a short hairpin RNA construct targeting *Cdx2* (pSicoR-Cdx2^{2Lox}). The conditional lentivirus-based vector, pSicoR, was described extensively in the previous part. Cloned blastocysts were morphologically abnormal, lacked functional trophoblast and failed to implant into the uterus. However, they efficiently generated pluripotent embryonic stem cells when explanted into culture.

This part of chapter 3 has been published:
Meissner A, and Jaenisch R. *Generation of nuclear transfer-derived pluripotent ES cells from cloned Cdx2-deficient blastocysts*. Nature. 2006 Jan 12; 439 (7073): 212-5.

The second application of our conditional gene knockdown approach using RNAi was aimed at testing the notion that development of an embryo derived by SCNT might be restricted by temporarily inactivating Cdx2, yet the same embryo might be fully competent for extracting embryonic stem cell lines useful for therapeutic purposes.

Altered nuclear transfer (ANT)

Survival of the normal embryo beyond implantation depends on the formation of the trophectoderm lineage, the extra-embryonic lineage that forms the fetal-maternal interface within the placenta. The second embryonic lineage that forms, the inner cell mass (ICM), gives rise to all subsequent lineages in the embryo proper, and it is the ICM that, upon explanting in culture, gives rise to ES cells. The "altered nuclear transfer" (ANT) concept (Hurlbut, 2005) is based on the premise that the inactivation of a gene crucial for trophectoderm development will eliminate the potential to form the fetal-maternal interface, but will spare the ICM lineage. By genetically altering a somatic donor cell before to nuclear transfer, one could generate cloned blastocysts that have no potential to develop beyond the blastocyst stage because no placenta could be formed. However, such cloned blastocysts could generate NT-ESCs derived from the ICM.

In this study we have performed a proof-of-principle experiment in mice to test the validity of the ANT approach and chose *Cdx2* as a candidate gene, as this gene encodes the earliest-known trophectoderm-specific transcription factor that is activated in the 8-cell embryo and is essential for establishment and function of the trophectoderm lineage (Chawengsaksophak et al., 2004; Strumpf et al., 2005). *Cdx2*-deficient blastocysts fail to maintain a blastocoel, lack epithelial integrity, dysregulate the ICM-specific transcription factors *Oct-4* and *Nanog*, and show increased cell death (Strumpf et al., 2005). Importantly, *Cdx2*-deficient blastocysts are able to form an ICM and generate ES cells when explanted in tissue culture (Chawengsaksophak et al., 2004; Strumpf et al., 2005).

Generation of Cdx2 deficient NT-ES cells

We selected for functional short hairpin (sh)RNAs against *Cdx2* as described in Figures 12 and 13. The experimental scheme, outlined in Figure 14a, involved the introduction of a conditional *Cdx2* shRNA lentiviral vector (Fig. 14b) into primary tail-

tip fibroblasts from neonatal F_1 mice (C57BL/6x129/SvJae). Green fluorescent protein (GFP)-positive

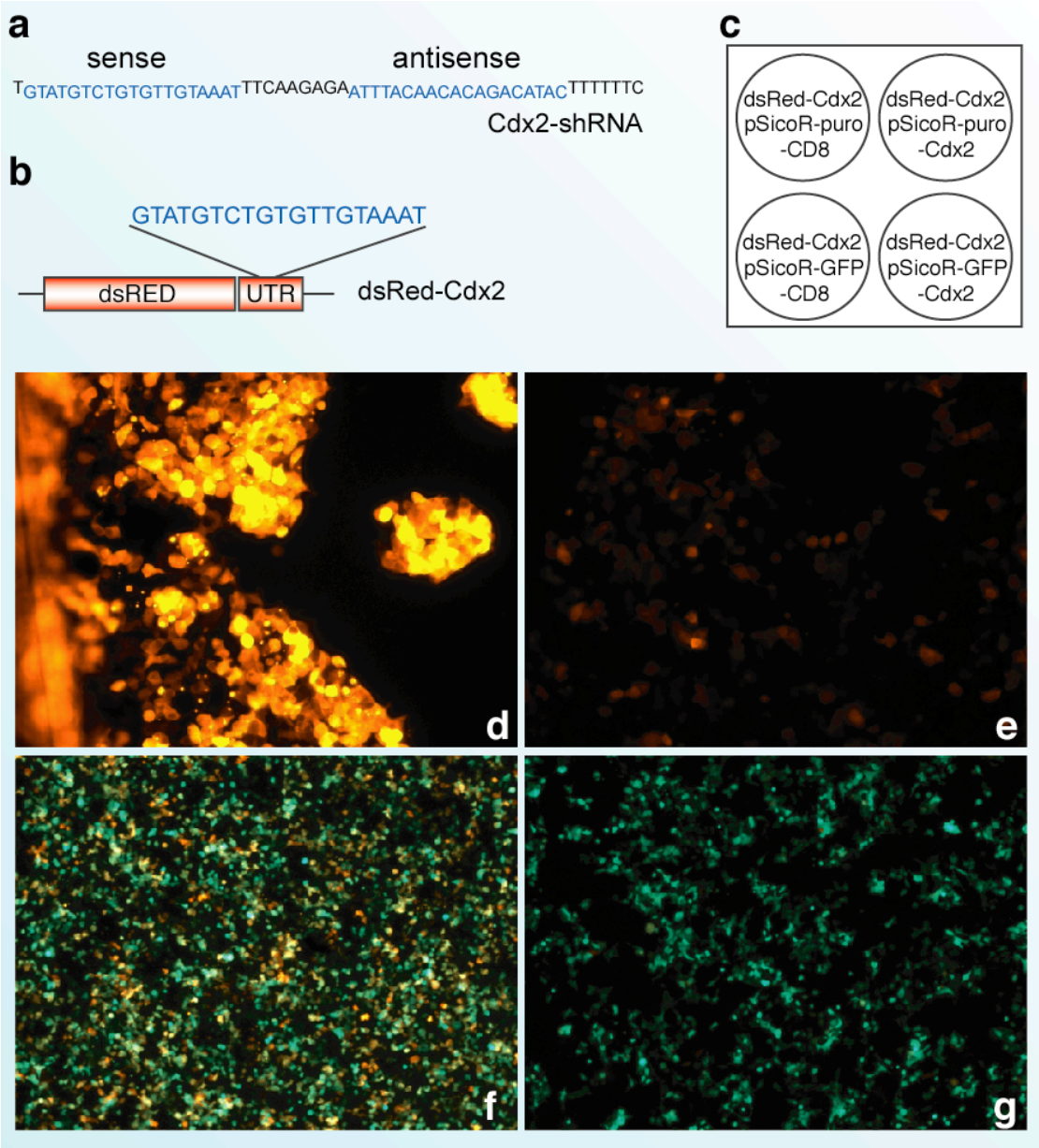

Figure 12: Selection of functional shRNAs. To select functional shRNAs against *Cdx2*, several Cdx2 shRNA target sequences were cloned into the 3'UTR of a DsRed reporter construct. Next transient co-transfections (293 cells) of the various DsRed-Cdx2 plasmids and the respective knockdown constructs were used to determine the potency of each hairpin by fluorescence microscopy and fluorescence activated cell sorting (FACS) analysis **a**) DNA sequence of the Cdx2 shRNA used. **b**) The sense strand sequence was cloned into the 3'UTR of a DsRed reporter gene containing plasmid (Clontech). **c**) DsRed-Cdx2 was mixed with FUGENE (Roche) and split into two. To the first well a pSicoR-CD8 (control, no target sequence present for the CD8 shRNA) was added and to the second well the pSicoR-Cdx2 vector (knockdown, DsRed contains the target sequence). Two different versions of the pSicoR vector were used, one containing a pgk-puro selection gene (**d** and **e**) and the other one a CMV-EGFP cassette (**f** and **g**). **d**-**g**) The knockdown efficiency was determined 24-48h post infection by fluorescence microscopy (images shown were taken with identical exposure time) and fluorescence activated cell sorting (FACS) analysis (data not shown).

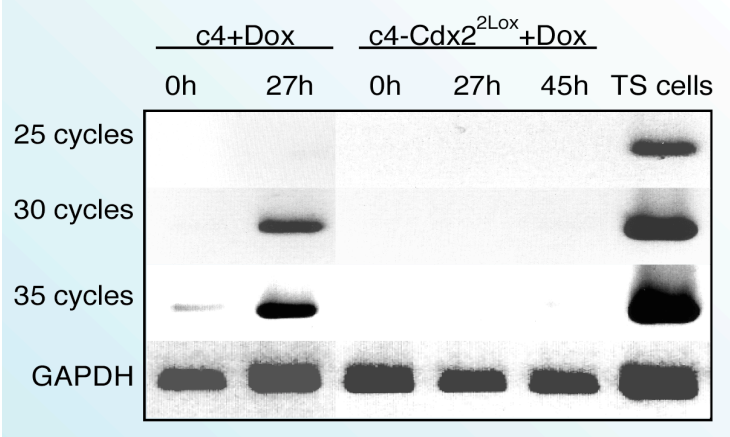

Figure 13: Knockdown of endogenous Cdx2. The knockdown efficiency against endogenous *Cdx2* was demonstrated using ZHBTc4 ES cells that upon downregulation of Oct-4 (doxycycline dependent) differentiate into trophectoderm with concomitant upregulation of Cdx2 expression. Semi-quantitative RT-PCR was used with 25, 30 and 35 amplification cycles for *Cdx2*. Cdx2 was readily detectable after 27h in the uninfected control cells. In the presence of the shRNA no Cdx2 was detected even after prolonged repression of Oct-4, which was confirmed by immunohistochemistry (data not shown). In the control cells, but not in the Cdx2 knockdown cells we detected a faint Cdx2 signal after 35 cycles of amplification. The absence of the signal in the knockdown cells at 0h suggests that it might be real and could originate from a small fraction of the ES cells that were differentiated. (ZHBTc4: Niwa et al., 2000)

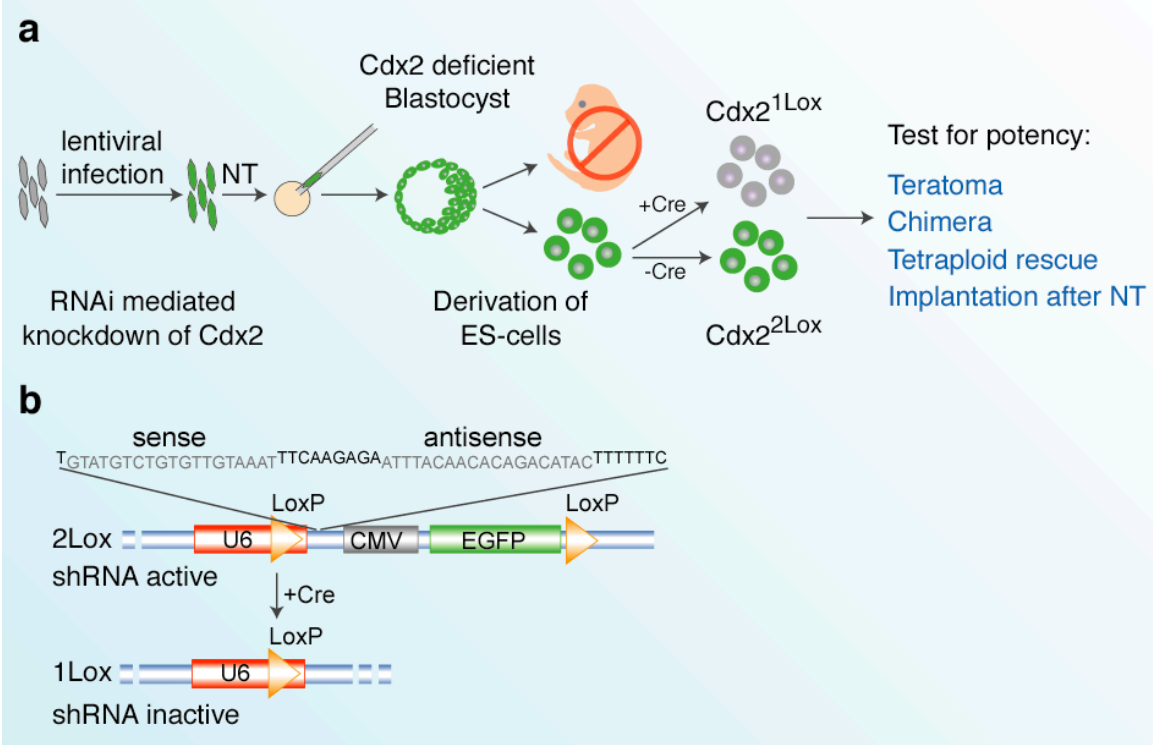

Figure 14: Derivation of NT-ESCs from Cdx2-deficient blastocysts. **a)** Primary tail-tip fibroblasts were infected with a conditional lentiviral RNA interference (RNAi) construct targeting Cdx2 before nuclear transfer (NT). Blastocysts deficient for Cdx2 were morphologically abnormal and unable to implant but gave rise to NT-ESCs. After initial expansion of the Cdx2 knockdown NT-ESCs (Cdx2^{2Lox}) we used transient Cre expression to generate subclones (Cdx2^{1Lox}) with a deleted hairpin. To test the potency of ES lines before and after 'loop-out' we used teratoma formation, diploid and tetraploid blastocyst injections as well as nuclear transfer. **b)** The conditional RNAi system (pSicoR) has been described above. The shRNA, which targets nucleotides 1890–1908 located in the 30 UTR of Cdx2, was cloned into the conditional RNAi vector generating pSicoR-Cdx2^{2Lox}. This vector carries the Cdx2 shRNA construct and an enhanced green fluorescence protein (EGFP) gene flanked by two LoxP sites (2Lox), which allows for Cre-mediated deletion of the shRNA and the EGFP sequences.

Cdx2^{2Lox} tail-tip fibroblasts were selected and used as donors for nuclear transfer. *Cdx2*-deficient blastocysts derived from the manipulated donor cells were tested for their potential to implant into the uterus and to generate pluripotent ES cells.

Of a total of 526 reconstructed oocytes, 350 formed pronuclei, of which 61 cleaved and developed into nuclear transfer morula/blastocysts. *Cdx2* knockdown nuclear transfer embryos showed no delay in developing to the early blastocyst stage compared to

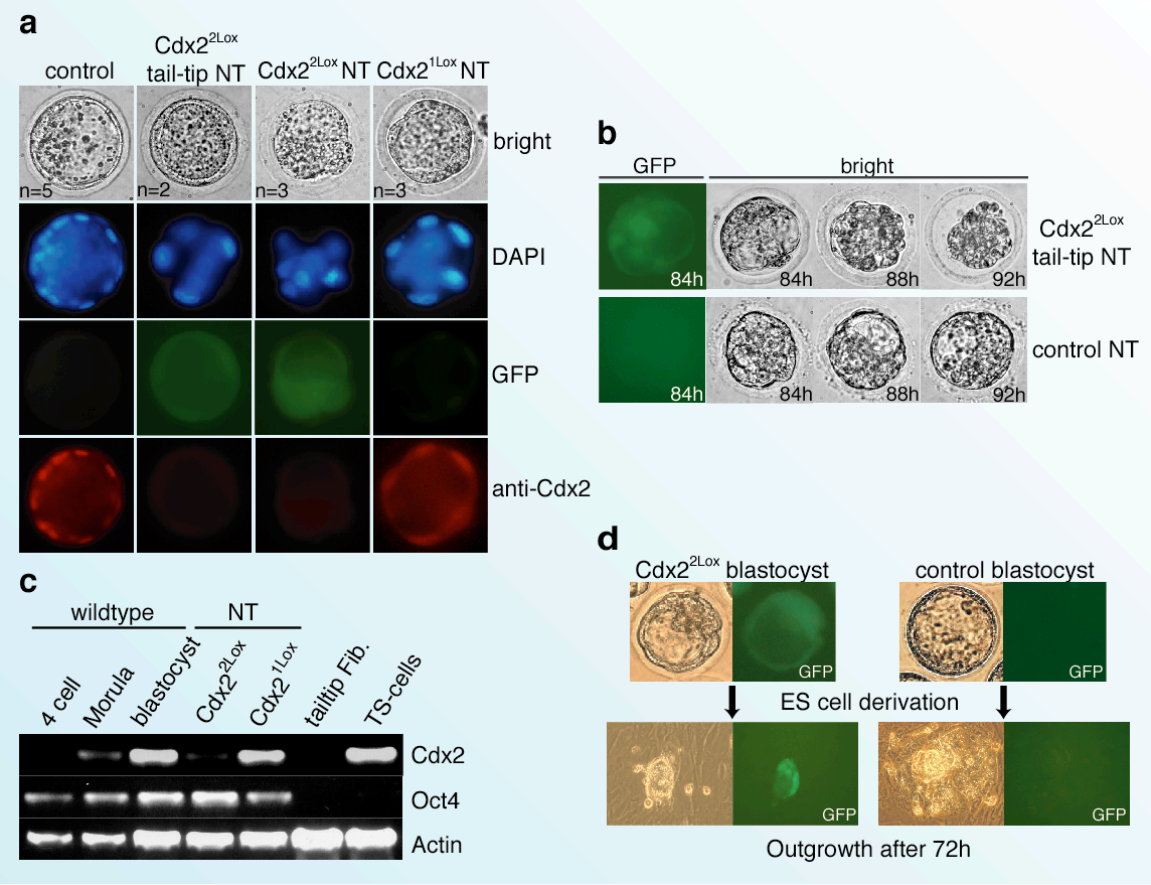

Figure 15: Cdx2-deficient blastocysts and ES cell derivation. **a)** Cdx2 immunostaining of day 3.5–4.5 wild-type and nuclear transfer blastocysts. The following donor cells were used for the nuclear transfer (from second column, left to right): Cdx2^{2Lox} tail-tip, Cdx2^{2Lox} ES cells, and Cdx2^{1Lox} ES cells. **b)** A typical Cdx2^{2Lox} tail-tip nuclear transfer blastocyst is shown 84 h after activation of the reconstructed oocytes. Cdx2-deficient blastocysts initially cavitated but failed to maintain the blastocoel and collapsed. Below, an expanded nuclear transfer blastocyst derived from control cells is shown. **c)** RT–PCR analysis of normal and Cdx2-deficient nuclear transfer preimplantation morula/blastocysts. Four 4-cell embryos were pooled andRNA was extracted for reverse transcription. All other samples were prepared from single morulae or blastocysts. Tail-tip fibroblasts (lane 6) express neither Cdx2 nor Oct-4. Trophectoderm stem (TS) cells (lane 7) express Cdx2, but no Oct-4. A faint Cdx2-specific band, such as that seen in the blastocyst containing the shRNA construct targeting Cdx2 shown in the figure, was detected in less than half of the tested embryos; most gave no signal in this test. **d)** Derivation of ES cells from Cdx2-deficient blastocysts. A Cdx2^{2Lox} tail-tip nuclear transfer-derived blastocyst with its initial outgrowth after 72 h (left) and a wild-type blastocyst (right) with its initial outgrowth are shown.

nuclear transfer embryos expressing a shRNA targeting CD8 ((Ventura et al., 2004); data not shown). Figure 15a shows that GFP-positive *Cdx2*^{2Lox} nuclear transfer blastocysts did not express *Cdx2* as assessed by immunohistochemistry, in contrast to wild-type blastocysts (column 1 and 2, Fig. 15a). Figure 15b shows that, when compared to control nuclear transfer blastocysts, *Cdx2*^{2Lox} nuclear transfer blastocysts were morphologically abnormal and failed to maintain a blastocoel cavity during *in vitro* cultivation, similar to previous results with *Cdx2* knockout blastocysts (Strumpf et al., 2005). Using semi-quantitative polymerase chain reaction with reverse transcription (RT-PCR), we confirmed the deficiency of *Cdx2* expression in *Cdx2*^{2Lox} nuclear transfer blastocysts, whereas control morulae and blastocysts showed robust *Cdx2* expression (Fig. 15c).

Developmental potential of Cdx2^{2Lox} NT-ES cells

To assess whether Cdx2 deficiency interfered with postimplantation development, *Cdx2*^{2Lox} nuclear transfer morulae/blastocysts were transferred into the uteri of pseudo-pregnant females. The uteri were removed at embryonic day E6.5 and examined for sites of implantation. Figure 16a shows no implantations in the uterus from a foster mother transplanted with five *Cdx2*^{2Lox} nuclear transfer blastocysts, in contrast to a uterus transplanted with five nuclear transfer control blastocysts that resulted in successful implantations (Fig. 16b). As summarized in Table 2, none of the *Cdx2*^{2Lox} nuclear transfer blastocysts formed visible implantation sites (0 out of 40), in contrast to control nuclear transfer blastocysts that were derived from fibroblasts carrying the CD8 control shRNA (6 out of 15). In addition, no evidence for delayed implantation was obtained, as we failed to detect implantation sites at E7 or E8 in females transplanted with a total of 18 *Cdx2* knockdown nuclear transfer embryos (data not shown). These results demonstrate that nuclear transfer from donor fibroblasts carrying the pSicoR-Cdx2^{2Lox} virus resulted in morphologically abnormal *Cdx2*-deficient nuclear transfer blastocysts that failed to implant upon transfer into foster mothers.

To investigate whether *Cdx2*-deficient blastocysts can generate ES cells upon explantation in culture, nuclear transfer *Cdx2*^{2Lox} blastocysts were transferred onto feeder cells. While control nuclear transfer blastocysts formed trophoblastic outgrowths characteristic of the trophectoderm lineage, the *Cdx2*^{2Lox} nuclear transfer blastocysts failed to generate any trophoblast cells (Fig. 15d). Consistent with previous observations

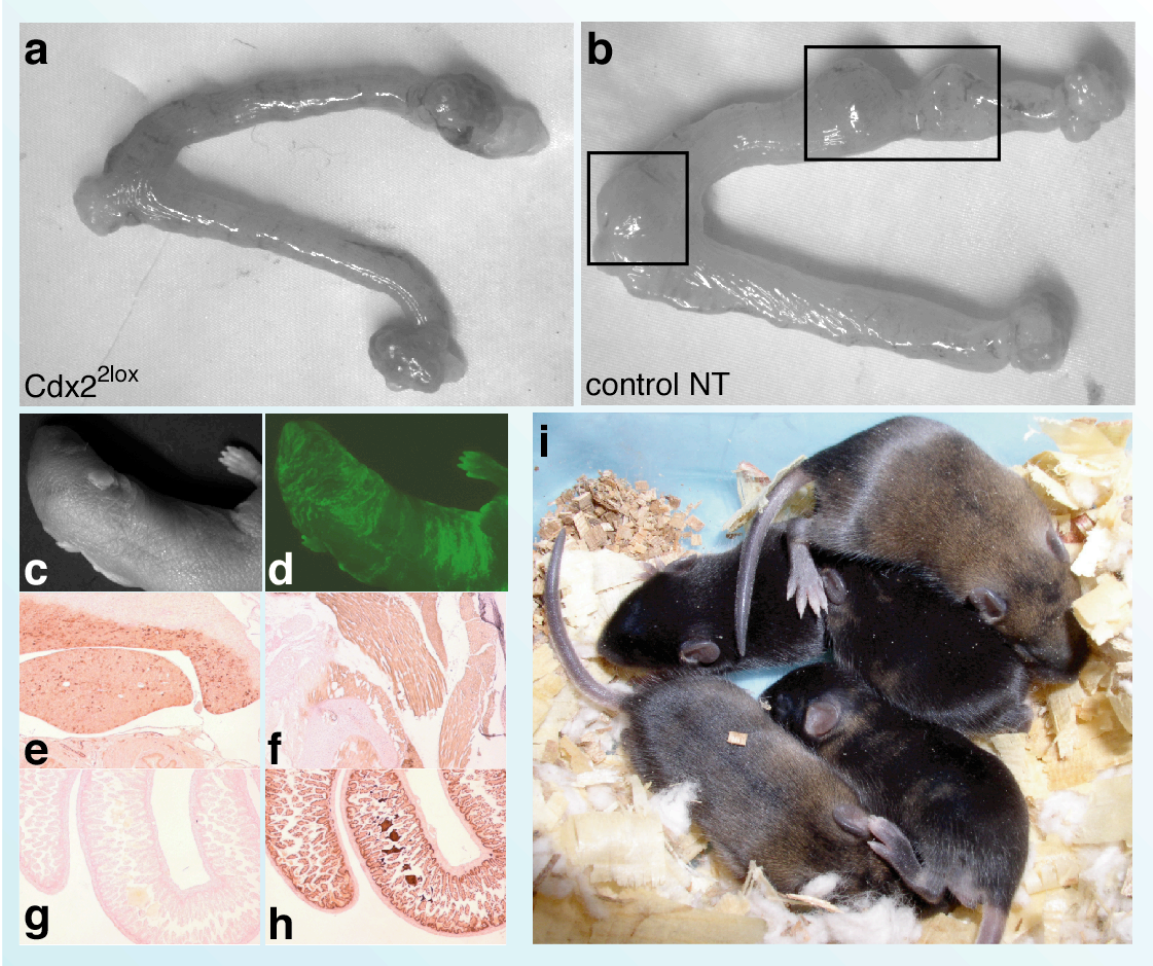

Figure 16: Cdx2-deficient cells maintain developmental potential but are unable to implant after nuclear transfer. **a-b)** In each example shown, five nuclear transfer blastocysts were transferred at day 3.5 into the uterus of a day 2.5 pseudo-pregnant female. **a)** Cdx2-deficient blastocysts fail to implant. A representative uterus isolated at day 6.5 is shown. No deciduae were detectable from transplanted Cdx2-deficient blastocysts. **b)** Control nuclear transfer blastocysts showed normal implantation sites at day 6.5. **c)** Bright-field image of a postnatal Cdx2^{2Lox} ES chimaera. **d)** GFP signal indicates a contribution from Cdx2^{2Lox} ES cells. **e–g)** Histological sections and anti-GFP staining from a newborn Cdx2^{2Lox} chimaera. There was a contribution to the liver (endoderm; **e**) and muscle (mesoderm; **f**) but not to the intestine (**g**). **h)** Anti-Cdx2 staining of the intestine shown in **g**. **i)** Coat colour contribution of Cdx2^{2Lox} ES cells. Recipient blastocysts have a C57BL/6 x DBA/2 F1 background and the Cdx2^{2Lox} ES cells a C57BL/6 x 129SvJae background. The presence of agouti (129/SvJae) fur indicates donor cell contribution. A litter with one wild type (black mouse below the top agouti), two low-contribution (middle) and two high-contribution chimaeras are shown.

Table 2: Survival of clones to blastocyst and post-implantation stage after nuclear transfer from different donor cells.

Donor cells	pSicoR genotype	Number of clones with pseudo-pronuclei	Number of morulae/blastocysts	Number of implants at E6.5 (number of foster mothers)
Fibroblasts	*Cdx2*^{2Lox}	211	40	0 (7)
Fibroblasts	*CD8*^{2Lox}	76	15	6 (3)
ES cells	*Cdx2*^{2Lox}	177	18	0 (4)
ES cells	*Cdx2*^{1Lox}	199	22	11 (5)
ES cells	*CD8*^{2Lox}	103	15	7 (3)

Shown are the number of reconstructed oocytes with pseudo-pronuclei after 5–6 h of activation. Morula/blastocyst transfers were done on day 3.5. pSicoR-CD8^{2Lox} fibroblasts carry a shRNA against CD8 (see above).

(Chawengsaksophak et al., 2004; Strumpf et al., 2005), *Cdx2*-deficient blastocysts generated ICM outgrowths that grew into stable, GFP-positive nuclear transfer *Cdx2*^{2Lox} ES cell lines with an efficiency that was comparable to that of nuclear transfer blastocysts derived from wild-type fibroblasts (14% of explanted blastocysts). As criterion for pluripotency, we tested the ability of the nuclear transfer *Cdx2*^{2Lox} ES cell lines to form chimeras when injected into diploid blastocysts. The GFP-labeled cells contributed extensively to neonatal chimeras (Figs. 16c, d) and formed high-grade postnatal chimeras (Fig. 16i, summarized in Table 3) with high contributions to most tissues (Fig. 16e, f), with the notable exception of the intestine (Fig. 16g), which was entirely composed of *Cdx2*-positive cells derived from the host blastocyst (Fig. 16h). This is in agreement with previous reports, as it has been shown that Cdx2 is required for normal development of the gastro-intestinal tract (Chawengsaksophak et al., 1997). We further explored the developmental potency of the NT-ESCs using tetraploid complementation, which

Table 3: Developmental potential of ES cells deficient (Cdx2^{2Lox}) or proficient (Cdx2^{1Lox}) for Cdx2 expression.

	Injected into 4N blastocysts			Injected into 2N blastocysts	
		Embryos live at			
Donor ES cells	Number injected	E14	Term	Number injected	Number of chimaeras
Cdx2^{2Lox}	112	0	0	65	12/23
Cdx2^{1Lox}	36	ND	6	ND	ND

Thirty-three sites of implantation were detected in the tetraploid *Cdx2*^{2Lox} recipients, but no or only reabsorbed embryos (at 14.5) were present. Twelve of the 23 born diploid pups were chimaeras. ND, not done.

represents the most stringent test for pluripotency, as the resulting "ES mice" are entirely composed of cells derived from the injected ES cells (Eggan et al., 2001b). Consistent with previous results (Chawengsaksophak et al., 2004), transfer of the *Cdx2*^{2Lox} ES cells resulted in no live embryos at E14 (Table 3). These data indicate that nuclear transfer using pSicoR-Cdx2^{2Lox} fibroblasts generates abnormal blastocysts that are inherently unable to implant and grow into a fetus but are able to generate pluripotent ES cells that have a diminished developmental potency as compared to wild-type ES cells.

Restoring Cdx2 function

To assess whether NT-ESCs derived from *Cdx2*-deficient blastocysts could have the same pluripotency as wild-type ES cells, we investigated whether the block to normal developmental potential could be relieved by reversing the effects of the *Cdx2* gene knock-down. Normal *Cdx2* gene function was restored in *Cdx2*^{2Lox} ES cells by transient transfection of a Cre plasmid, resulting in the deletion of the *Cdx2* shRNA and EGFP marker gene (*Cdx2*^{2Lox} to *Cdx2*^{1Lox}; compare Fig. 14b), and rendering the cells Cdx2 competent and GFP negative. Nuclear transfer from *Cdx2*^{1Lox} donor cells generated GFP negative normal appearing nuclear transfer blastocysts that expressed wild type levels of Cdx2, as shown by immunostaining (Fig. 15a, right column) and RT-PCR (Fig. 15c, lane 5). To test whether deletion of the shRNA would restore pluripotency, the *Cdx2*^{1Lox} ES cells were injected into tetraploid blastocysts. As shown in Table 2, *Cdx2*^{1Lox} ES cells efficiently generated ES mice in contrast to the *Cdx2*^{2Lox} ES cells that were unable to give rise to ES mice. These results show that the deletion of the *Cdx2* shRNA sequences creates ES cells that can generate all somatic tissues including normal intestinal cells, which cannot be derived from the *Cdx2*^{2Lox} parental ES cells (compare Fig. 16g, h). Finally, to test whether totipotency of *Cdx2*^{1Lox} ES cell nuclei was recovered, we transplanted *Cdx2*^{1Lox} blastocysts derived by nuclear transfer using *Cdx2*^{1Lox} donor ES cells into pseudo-pregnant foster mothers. As summarized in Table 2, normal-sized implants were detected at E6.5. These results confirm that *Cdx2* deficiency was responsible for the failure of clones to generate functional blastocysts and exclude other genetic alterations acquired during *in vitro* manipulation of the cells in the characteristic block to implantation. Most importantly, our data demonstrate that ES cells competent to generate all lineages can be derived from abnormal nuclear transfer blastocysts.

Reduced Representation Bisulfite Sequencing for comparative high-resolution DNA methylation analysis

Abstract: To improve the study of epigenetic differences between cell types more global approaches for analyzing epigenetic modifications are required. We have developed a large-scale random approach termed reduced representation bisulfite sequencing (RRBS) for analyzing and comparing genomic methylation patterns. *Bgl*II restriction fragments were size-selected to 500-600 bp, equipped with adapters, treated with bisulfite, PCR amplified, cloned and sequenced. We constructed RRBS libraries from murine ES cells and from ES cells lacking DNA methyltransferases Dnmt3a and 3b and with knocked-down (kd) levels of Dnmt1 (*Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$]). Sequencing of 960 RRBS clones from *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells generated 343 kb of non-redundant bisulfite sequence covering 66,212 cytosines in the genome. All but 38 cytosines had been converted to uracil indicating a conversion rate of >99.9%. Of the remaining cytosines 35 were found in CpG and 3 in CpT dinucleotides. Non-CpG methylation was >250-fold reduced compared to wild-type ES cells, consistent with a role for Dnmt3a and/or Dnmt3b in CpA and CpT methylation. Closer inspection revealed neither a consensus sequence around the methylated sites nor evidence for clustering of residual methylation in the genome. Our findings indicate random loss rather than specific maintenance of methylation in *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells. Near-complete bisulfite conversion and largely unbiased representation of RRBS libraries suggest that random shotgun bisulfite sequencing can be scaled to a genome-wide approach.

This part of Chapter 3 has been published:
Meissner A*, Gnirke A*, Ramsahoye B, Bell G, Lander ES, Jaenisch R. *Reduced Representation Bisulfite Sequencing for comparative high-resolution DNA methylation analysis.* Nucleic Acid Research. 2005 Oct 13; 33 (18): 5868-77.
* equal contributing authors

Another approach in studying epigenetic reprogramming is to determine the epigenetic differences between different cellular states, and to further elucidate what defines the nature of "stemness" within a stem cell. DNA methylation is probably the best studied epigenetic modification that determines patterns of gene expression within a cell. In order to better define the epigenome of different cell types, we have developed an approach for large-scale high resolution DNA methylation analysis.

Reduced representation bisulfite sequencing

RRBS is analogous to the reduced representation shotgun sequencing (RRS) used for single nucleotide polymorphism (SNP) discovery (Altshuler et al., 2000). The method is based on size selection of restriction fragments to generate a "reduced representation" of the genome of a strain, tissue or cell type.

For this study, we digested genomic DNA with *Bgl*II and purified fragments between 500 and 600 bp in size on an agarose gel. Based on the available mouse genome sequence, *Bgl*II digestion is expected to generate 21,939 *Bgl*II fragments in this size range comprising ~12 Mb (0.5%) of the genome. Size-selected *Bgl*II fragments were equipped with end adapters, denatured and treated with bisulfite to convert all unmethylated cytosines to uracil. Bisulfite-converted DNA remains single-stranded as the two strands are no longer complementary. Primers specific for the converted adapter sequence and a proofreading thermostable DNA polymerase were used to synthesize the second strand and to PCR amplify the bisulfite-converted material. Blunt-end PCR products were cloned in a plasmid vector and sequenced (Fig. 17).

For analysis of the bisulfite sequences and to identify the corresponding genomic sequence we searched RRBS reads against a reduced representation database of the mouse genome that contained both strands of *Bgl*II fragments that had been size-selected and bisulfite-converted *in silico*. When aligned to the original genome sequence, a 5-methylcytosine is thus displayed as a matching C in the bisulfite sequence, and C to T transitions indicate unmethylated cytosines.

Even though bisulfite sequencing is a widespread technique, some concerns persist. Since bisulfite converts single-stranded but not double-stranded DNA, incomplete denaturation or reannealing leads to incomplete conversion. This complicates the data analysis, as it is not always possible to determine whether an unconverted cytosine

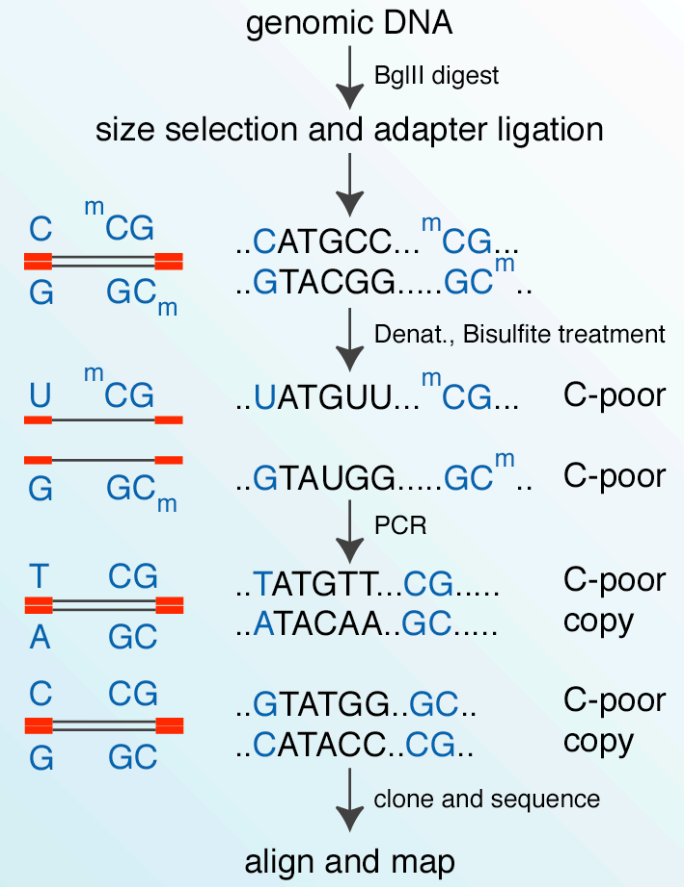

Figure 17: Reduced representation bisulfite sequencing. Genomic DNA is digested to completion using a restriction enzyme (here BglII). After size-selection an adapter is added. The DNA is denatured, and unmethylated cytosines are bisulfite-converted to uracil. The two resulting C-poor strands are no longer complementary to each other. Primers specific for the converted adapter sequence are used to fill-in the second (G-poor) strand and for PCR amplification. PCR products are cloned and sequenced. Sequences generated from RRBS libraries are projected onto the genome by searching against a reduced representation database of BglII fragments that had been size-selected and bisulfite-converted in silico.

represents *bona fide* methylation or an experimental artifact. Another potential problem is depurination, strand breakage and DNA degradation caused by the harsh reaction conditions, which lower the yield of full-length *Bgl*II fragments significantly. It has been estimated that >90% of the input DNA is lost due to DNA degradation during the first hour of a bisulfite reaction (Grunau et al., 2001). However, to maximize the conversion rate, the reaction is usually carried out overnight, necessitating extensive PCR amplification before cloning or sequencing to compensate for the inevitable loss of DNA. Moreover, since most proofreading enzymes stall at uracil residues in the template strand, non-proofreading *Taq* polymerase is usually prescribed for second-strand synthesis and PCR amplification which can lead to PCR-induced sequencing errors.

These limitations are less worrisome for single-copy loci, but could be significant in a genome-wide setting, where no preselection against fast-reannealing repetitive sequences is made and where amplification bias and skewed sequence representation creates serious sampling problems. Indeed, our preliminary attempts were plagued by DNA degradation, incomplete conversion and poor efficiency of PCR amplification, most

likely caused by the re-annealing of repetitive sequences including the common adapter sequence at the ends of each DNA molecule. Moreover, certain sequences were clearly overrepresented in the resulting libraries indicating amplification bias during the PCR. These initial problems were largely remedied by performing the bisulfite reaction in the presence of urea as suggested by Paulin *et al.* (Paulin et al., 1998) and by fine-tuning experimental parameters such as DNA concentration, time and temperature of the bisulfite reaction, and number of PCR cycles for the double-strand rescue and amplification by a proofreading thermostable DNA polymerase engineered to accept uracil in the template strand (Fogg et al., 2002).

To test if our optimized protocol was sufficient to achieve complete genome-wide bisulfite conversion without compromising library complexity and representation, we wished to construct and sequence RRBS libraries from genomic DNA that was largely free of methylation. To this end we generated ES cells deficient in all three major DNA methyltransferases.

ES cells deficient for Dnmt1, Dnmt3a and Dnmt3b

We combined knockouts for the *de novo* Dnmts (Dnmt3a and Dnmt3b) with RNAi-induced knockdown of Dnmt1 (Fig. 18a) using a lentivirus-based system for stable short hairpin RNA (shRNA) expression (Ventura et al., 2004). The Dnmt1 knockdown resulted in a significant albeit not complete loss of Dnmt1 protein compared to the *Dnmt*[$3a^{-/-}$,$3b^{-/-}$] control cells (Fig. 18b).

To determine whether the decrease in Dnmt1 levels led to efficient demethylation, we analyzed the methylation status of minor satellite repeats and IAP elements in a number of control and knockdown ES cell lines by MSRE analysis. Significant repeat demethylation was observed when Dnmt1 was knocked down, and the methylation levels in the *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] ES cells closely resembled the digest of genomic DNA with *Msp*I which cuts irrespective of the methylation status (Fig. 19a and b). Loss of methylation at these repeat elements appears to be primarily caused by the lack of Dnmt1 and largely independent of the *de novo* Dnmts at these early passages. Using a COBRA assay (Eads and Laird, 2002) we observed loss of imprinting at four imprinted genes following Dnmt1 knockdown as compared with the controls (Fig. 19c). Taken together,

these experiments showed that Dnmt1 knockdown resulted in significant loss of methylation at specific genes and repeat elements.

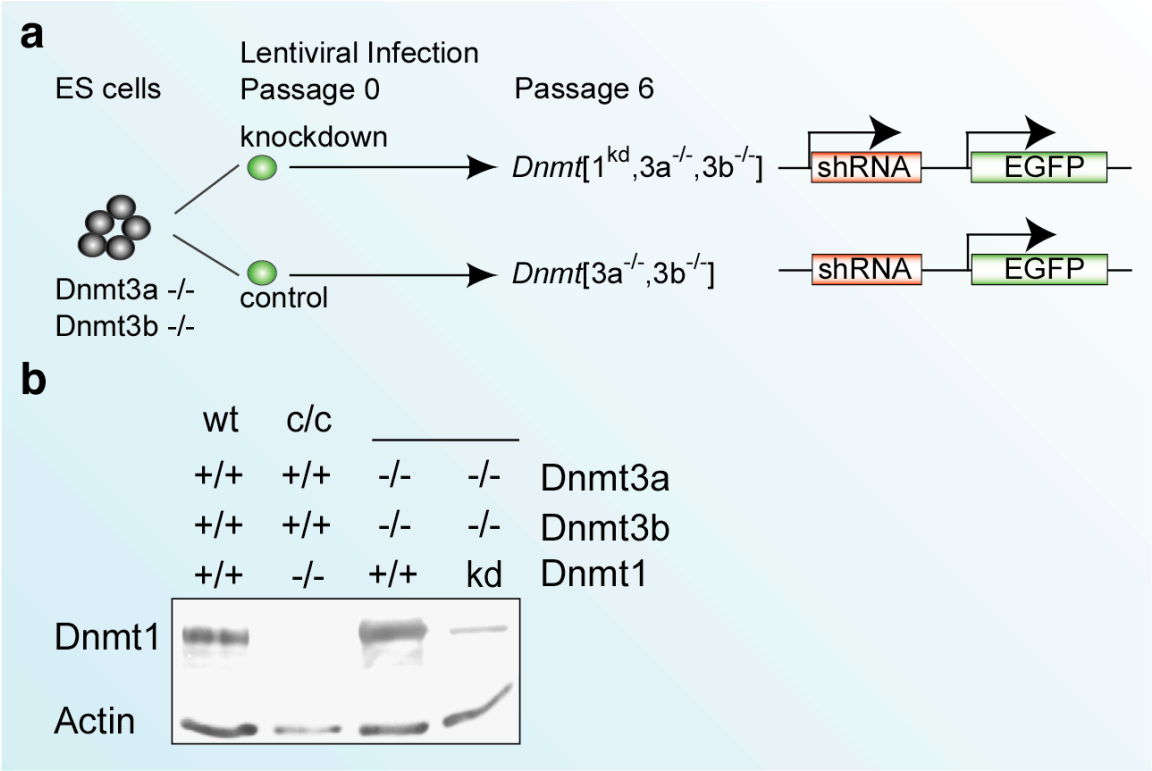

Figure 18: Generation of Dnmt1, Dnmt3a and Dnmt3b deficient ES cells. **a)** Dnmt3a/3b homozygous double knockout ES cells have been described previously (Okano et al., 1999). The knockdown virus is expressing a Dnmt1 shRNA, whereas the control is not. The infection was termed Passage 0. After the infection ES cells were passaged four times on feeders followed by two additional passages under feeder-free conditions (Passage 6). Number of viral integrations were determined by Southern blotting and clones with single integration were selected (data not shown). **b)** Western blot analysis. The status of the different Dnmts is indicated above. The knockdown ES cells showed a significant reduction in Dnmt1 levels compared with their sister clone. c/c is a previously reported Dnmt1 null ES line (Lei et al., 1996).

To better quantify these results, we used NNA, which allowed to determine the global amounts of CpG methylation in wild-type and mutant ES cells. We detected ~2% residual CpG methylation in the *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells compared to 22% in the *Dnmt*1 null ES cells and 75% in wild-type ES cells (Fig. 19d). Dnmt3b heterozygous and homozygous ES cells displayed wild-type methylation levels in the presence of Dnmt1 and showed similar loss of methylation within six passages of Dnmt1 knockdown (Fig. 19d and data not shown) confirming the potency of the shRNA.

To test the RRBS approach and to determine whether specific sequences were retaining methylation we generated and sequenced *Bgl*II RRBS libraries from wild-type and Dnmt-deficient ES cells.

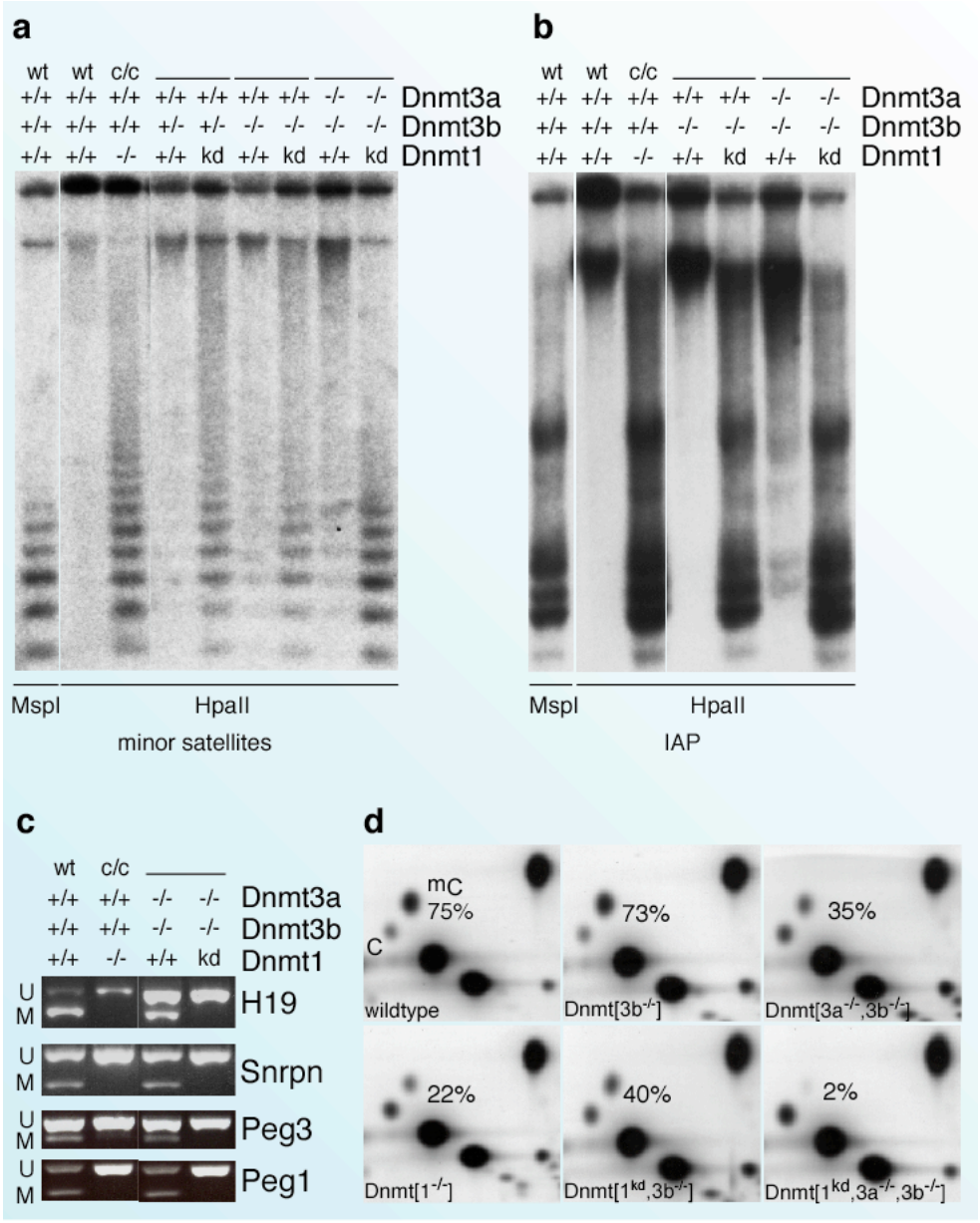

Figure 19: Methylation status of the Dnmt-deficient ES cells. All knockdown and control ES cells were analyzed at Passage 6 after infection. **a)** Minor satellite repeat methylation. HpaII digests of genomic DNA were hybridized to minor satellite probe pMR150. MspI is an isoschizomere of HpaII and cuts irrespective of the methylation status (i.e. appearance of a ladder in HpaII lane indicates loss of methylation). The status of the different Dnmts is shown above the Southern blot. All knockdown and control ES cell lines were generated as described in Figure 18. Each knockdown line contains a single lentiviral integration (data not shown). **b)** IAP methylation. HpaII-digested genomicDNAwas hybridized to an IAP probe. **c)** COBRA analysis for imprinted genes. Genomic DNA was bisulfite treated and after PCR amplification of H19, Snrpn, Peg1 and Peg3 a restriction digest was performed to analyze the methylation status of the differentially methylated regions (U=unmethylated, M=methylated). The second (smaller) fragment of the methylated and digest product is not shown. **d)** Total mCpG quantification by NNA. The spots corresponding to CpG and mCpGare indicated in the upper left panel. The per cent mCpG/(CpG+mCpG) are displayed in each panel (estimated error 5%).

Sequencing of RRBS libraries

In preliminary experiments we noticed that sequencing RRBS clones with reverse primer had a significantly higher success rate and produced longer reads on average than sequencing with forward primer. We therefore sequenced the RRBS clones single-pass using reverse primer. Only clones with high-quality sequence across the entire length of the insert were used for the final methylation analysis. Table 4 summarizes the sequencing statistics from 960 RRBS clones from Dnmt-deficient cells and 192 clones from wild-type ES cells.

Table 4: Sequencing and methylation statistics

ES cell line	Dnmt[1^{kd}, $3a^{-/-}$,$3b^{-/-}$]	wild-type
Colonies picked	960	192
Bisulfite sequencing reads[a]	876	186
Insert in plus orientation[b]	153	50
Plus read complete[c]	38	23
Insert in minus orientation[d]	723	136
Minus read complete[c]	719	134
Complete bisulfite sequencing reads[c]	757	157
Genome hits	676	148
Non-redundant genome hits[e]	609	147
Total bp of non-redundant genome hits	342 556	80 692
Cytosines in aligned genome sequence	66 212	15 296
5-Methylcytosine (mC)	38 (0.06%)	707 (4.6%)
CpG in aligned genome sequence	3458	594
mCpG	35 (1.0%)	533 (90%)
CpA in aligned genome sequence	23 046	5601
mCpA	0 (0%)	135 (2.4%)
CpT in aligned genome sequence	25 505	5924
mCpT	3 (0.01%)	39 (0.7%)
CpC in aligned genome sequence	14 203	3177
mCpC	0 (0%)	0 (0%)

[a]Excludes growth failures, sequencing failures, mixed clones, vector-only clones and a total of nine reads that showed no bisulfite conversion at all.
[b]Sequenced strand is the bisulfite-converted C-poor strand.
[c]High-quality sequence across entire length of BglII fragment.
[d]Sequenced strand is the G-poor complementary strand of the bisulfite-converted strand.
[e]Includes sequences that are duplicated in the genome. BglII fragments that were hit more than once were counted only once.

Although blunt-ended PCR products can insert in either orientation into the cloning vector, only a minority had inserts in the orientation that resulted in the C-poor sequence, *i.e.* the strand that has been modified by bisulfite (153 out of 876 RRBS reads from the *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] library). The vast majority of the clones produced the

complementary G-poor reads. Notably, the sequence quality was also significantly different for the two orientations. Almost all G-poor reads were high-quality across the entire insert whereas peak heights and quality of many C-poor reads dropped after a few hundred bases, leaving relatively few complete C-poor sequences for the methylation analysis. Preferential cloning in one orientation and high drop-out rate for C-poor strands were more pronounced in the *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] library which has an extremely asymmetric base distribution. Of the sequenced inserts 96% from this library consisted solely of three bases, *i.e.*, either A, G and T or A, C and T due to complete absence of methylated cytosine in the corresponding genome loci. Directional cloning and sequencing bias has been observed before with bisulfite-treated DNA (Grunau et al., 2001) and is therefore not a RRBS specific phenomenon.

Of the complete RRBS reads from *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells (89%) found a near-perfect match in the reduced representation reference-sequence database and could be placed with high confidence on the mouse genome. The rate of genome alignments for sequences from wild-type ES was slightly higher (94%). Overall, the success rate of full-length, mapped bisulfite sequence was 72% of all clones picked. A schematic of the distribution of RRBS sequences along the mouse chromosomes is available in Figures 20 and 21. In addition we have developed a genome browser that allows a more comprehensive view of the genomic environment of the RRBS libraries and the data generated (for a sample screenshot see Fig. 22).

Fifty-six loci were hit by more than one RRBS sequence from the *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] library. Ten of these potentially represent sequences that occur more than once in the genome. The remaining 46 appear to be unique loci that have indeed been cloned and sequenced twice. This is more than the 23 double-hits expected by random sampling of an ideal library, possibly indicating a slight cloning or sequencing bias. Consistent with random cloning, the much smaller number of wild-type RRBS sequences produced only one double-hit. Eleven fragments were sequenced in both cell lines, compared to eight sequence overlaps expected given the number and size distribution of successful reads from each library (Fig. 23). The total length of non-redundant and mapped RRBS sequences was 342,556 bp for *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] and 80,692 bp for wild-type ES cells.

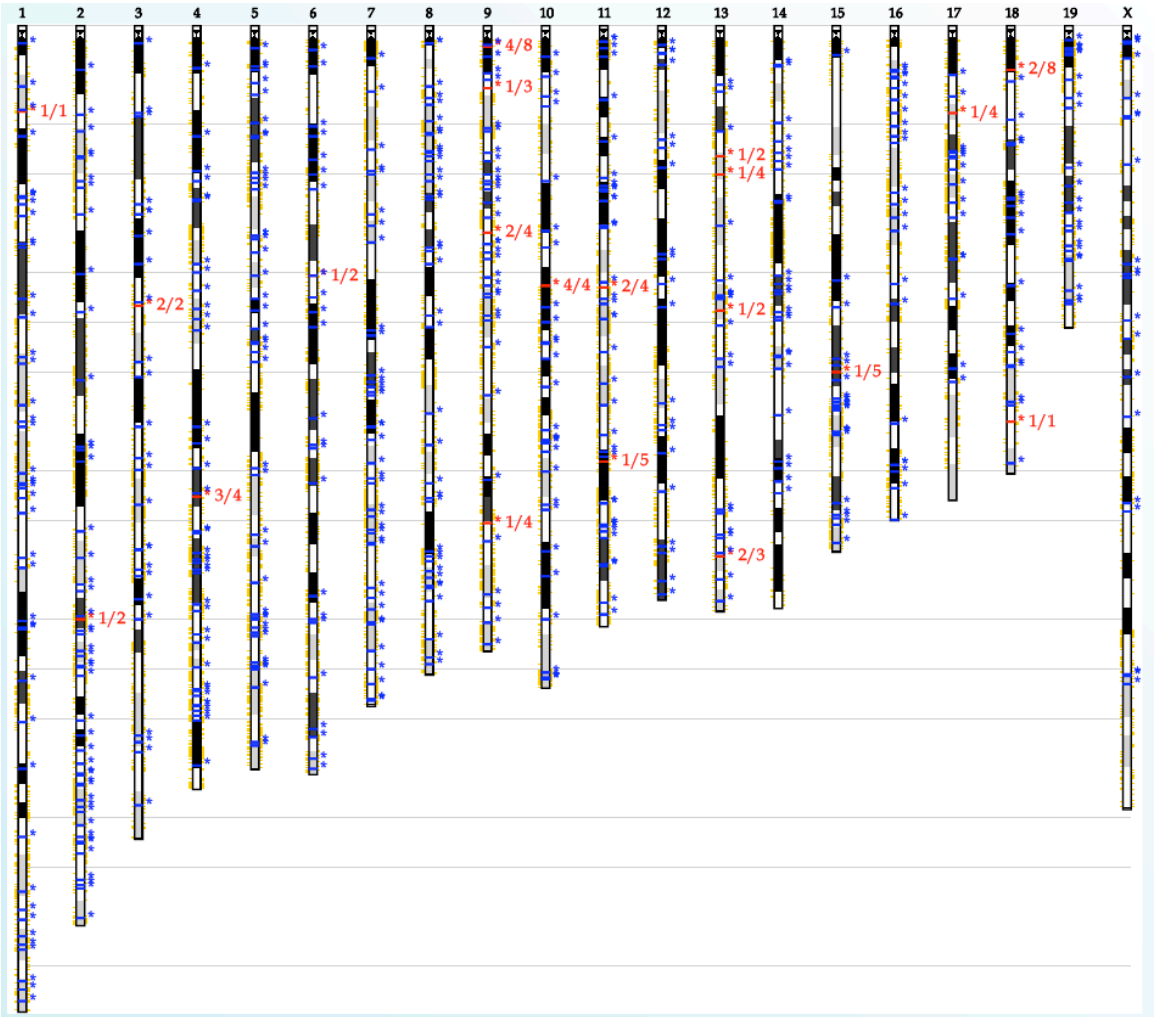

Figure 20: Genomic location and methylation status of 609 non redundant RRBS sequences from *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] ES cells. The blue asterisk indicates reads that were free of methylation. All the reads that contained methylation have the number of mCpG/CpG displayed. Yellow indicates transcript locations. The Y chromosome was also hit by RRBS sequences, but is not included in the mapping tool used.

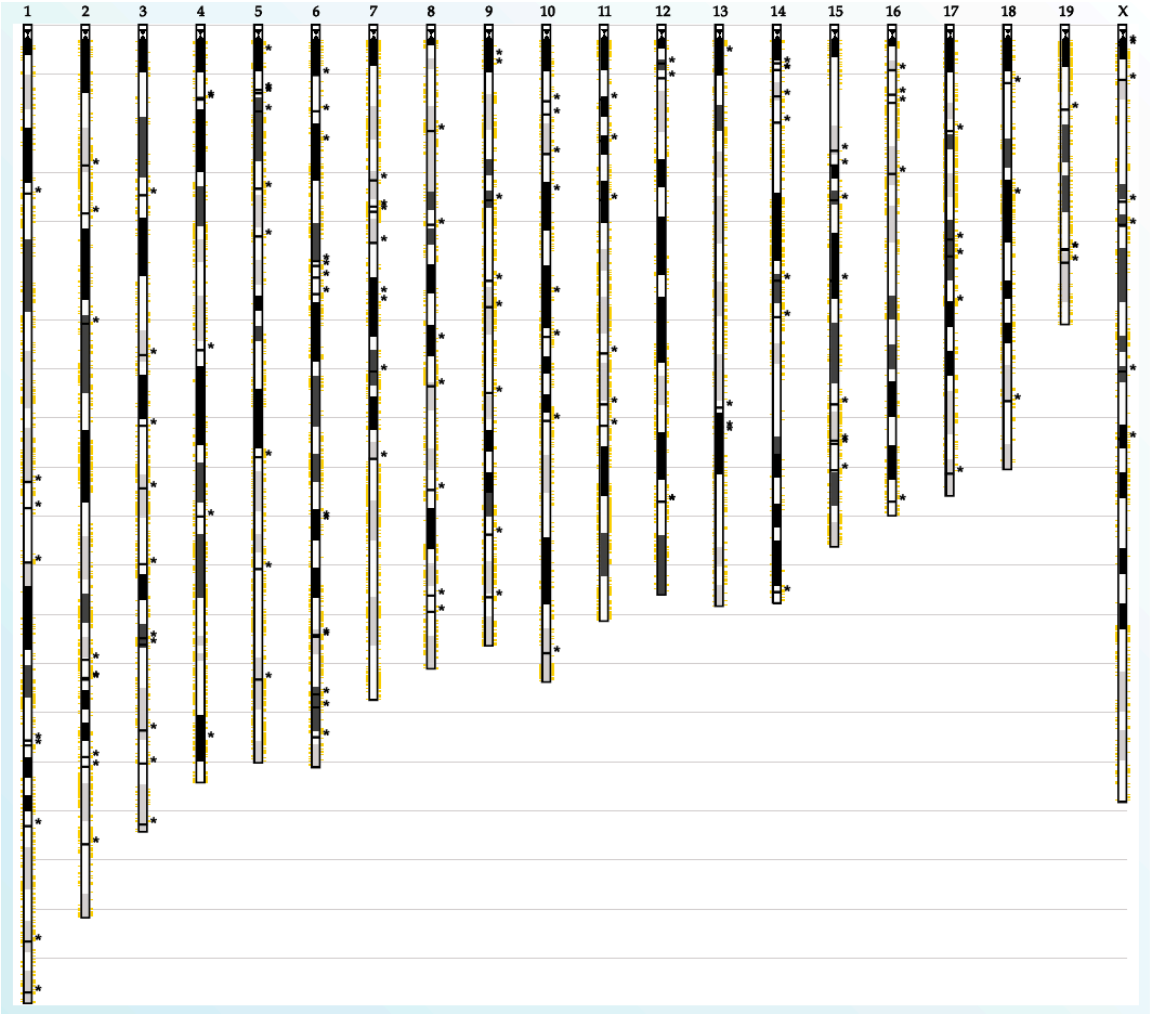

Figure 21: Genomic location of 148 RRBS sequences from wildtype ES cells. Yellow indicates transcript locations. The Y chromosome was also hit by RRBS sequences, but is not included in the mapping tool used.

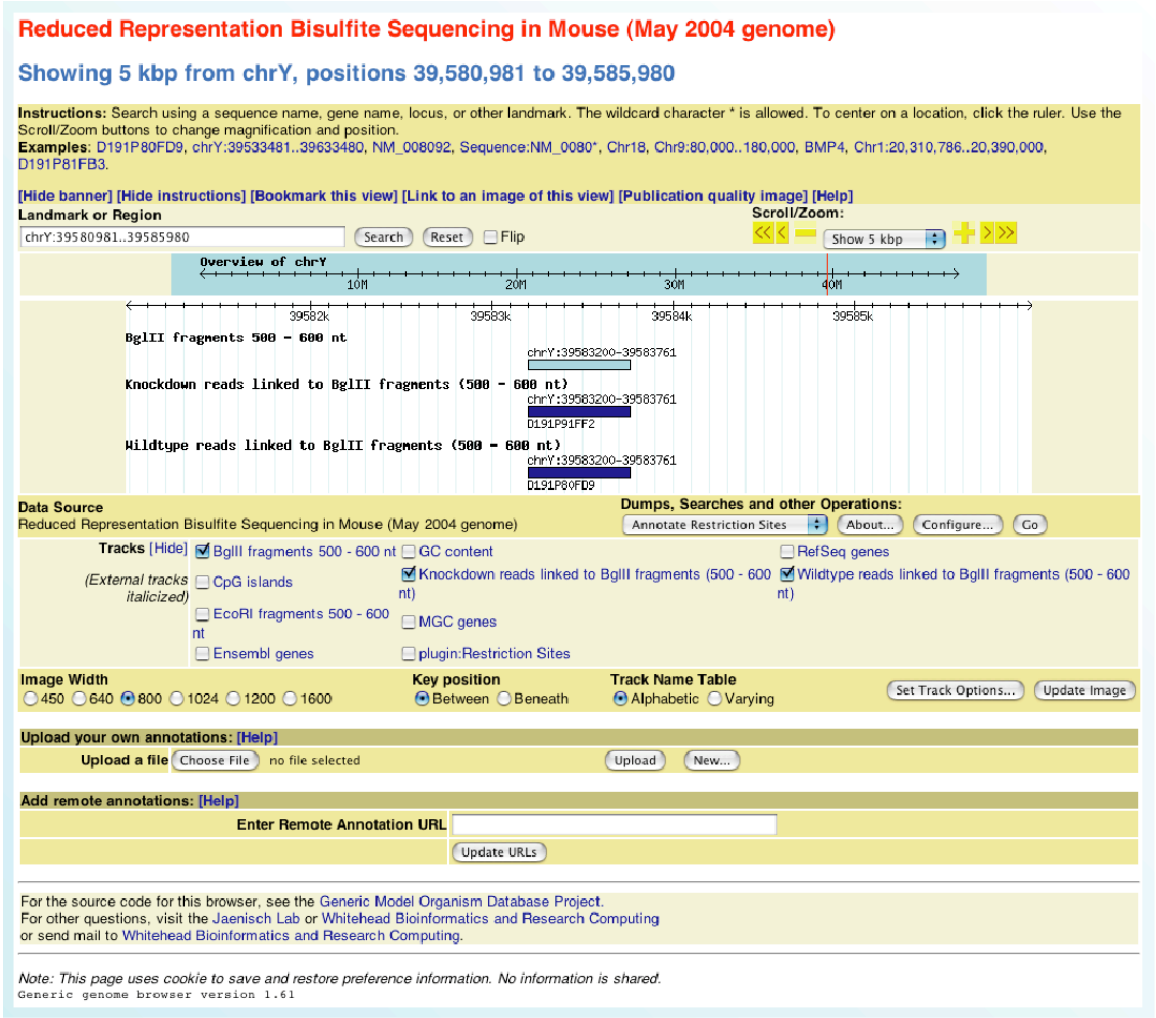

Figure 22: RRBS Genome Browser. The main window of the RRBS genome browser is shown. http://frodo.wi.mit.edu/cgi-bin/jaenisch_rrbs/gbrowse.cgi/mouse_may04
username = jaenisch/ password = rrbs.
The browser works like most common genome browsers, with zoom in/out and different tracks that can be displayed. In addition, the sequence for each fragment can be viewed. The knockdown and wildtype fragments display the respective bisulfite converted sequence. The corresponding sequence in the *Bgl*II track provides the uncoverted sequence, which can be used for individual alignments.

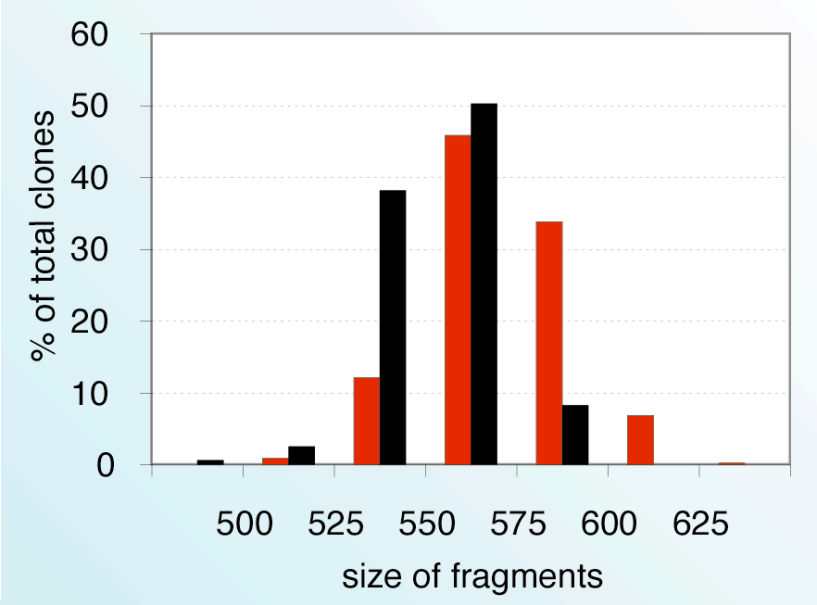

Figure 23: Size distributions of the sequenced clones from each library. RRBS reads from wild-type ES cells (black) had a mean of 553 bp and an SD of 17 bp. Dnmt[1^{kd},$3a^{-/-}$,$3b^{-/-}$] reads were (570 ± 20) bp in size (grey bars). The size distributions of the two libraries were overlapping but not identical.

Table 5: Fraction (in per cent) of various types of sequences in the mouse reference genome, the 500–600 bp BglII reduced representation thereof (RR genome) and RRBS sequences from Dnmt-deficient and wild-type ES cells.

	Genome[a]	RR genome	RRBS *Dnmt*	RRBS *wt*
GC content	42.0	41.5	43.7	43.1
CpG islands[b]	0.4	0.1	0.1	0.0
ENSEMBL genes	34.3[c]	35.0[c]	41.9[d]	35.3[d]
Promoter	5.0[e]	5.0[e]	7.0[f]	4.7[f]
SINEs	8.2	2.7	2.6	2.3
LINEs	19.2	10.2	10.7	11.3
LTR elements	9.9	2.9	2.9	4.3
MER DNA elements	0.9	0.2	0.1	0.2

[a]Repeat and GC content were taken from (Waterston et al., 2002)
[b]CpG islands were taken from the mm6 mouse genome assembly on the UCSC genome browser.
[c]Fraction of genome sequence that falls within gene bounds of non-overlapping ENSEMBL gene models.
[d]Fraction of RRBS sequences with significant hits to the ENSEMBL gene fraction of the genome.
[e]Fraction of genomesequence that falls within 5 kb upstream of the transcription start site of ENSEMBL gene models.
[f]Fraction of RRBS sequences with significant hits to regions 5 kb upstream of transcription start sites.

To determine whether these RRBS libraries were generally representative we compared the GC content, the representation of CpG islands, transcripts, promoter regions and different classes of repeat elements between the entire mouse genome (Waterston et al., 2002), the 500-600 bp *Bgl*II fraction thereof and the genome sequences hit by the RRBS clones (Table 5). While reducing the representation introduced a noticeable bias, in particular a reduction of repeats, bisulfite conversion, PCR amplification, cloning and sequencing did not. The GC content of loci covered by RRBS sequences ranged from 32 to 63%, indicating satisfactory performance of our protocol over a wide range of GC content. Likewise, the distribution of the sequenced clones in the genome did not show conspicuous hot or cold spots (see Figs 20 and 21). Taken together, our data suggest that RRBS libraries are sufficiently random and representative of the genome fraction used to make them.

Reducing the complexity by size fractionation of a limit digest with BglII (recognition site AGATCT) is expected to bias somewhat against GC-rich regions of the genome. Pooling two single digests with compatible enzymes such as BglII and BamHI (GGATCC) before the size selection would sample the genome more evenly and increase the complexity of the RRBS libraries.

Comparison of wild-type and Dnmt-deficient ES cells

The RRBS sequences revealed the methylation status of 66,212 cytosines in *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] ES cells (Table 4, bottom half). Only 38 of these were inferred to be methylated, 35 of them in CpG and three in CpT dinucleotide context. Considering the non-random distribution of mC among the four dinucleotides, it unlikely that all of them were caused by incomplete bisulfite conversion or PCR or sequencing errors. Moreover, the 35 mCpGs are ~1% of all bisulfite-sequenced CpGs, which is close to the 2% mCpG level determined by NNA (Fig. 19d). By comparison, 90% of CpGs were methylated in wild type ES cells. We also observed a considerable difference in the level of non-CpG methylation [(mCpA+mCpT)/C], which was >250-fold reduced in the Dnmt-deficient ES cells.

In the *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] RRBS sequences, 25,020 bases were covered 2- or 3-fold, comprising 4,669 cytosines including 217 CpGs. Overlapping RRBS sequences agreed for most loci. In two cases, only one sequenced *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] clone

displayed a methylcytosine. At another discordant site, the two reads agreed at one mCpG but disagreed at another.

To address the issue of heterogeneity, we selected ten loci with mCpGs and ten loci without methylation and designed specific PCR primers to bisulfite re-sequence them in a targeted fashion. Multiple clones were sequenced for each locus in wild type, *Dnmt*[3a$^{-/-}$,3b$^{-/-}$] and the *Dnmt*[1^{kd},3a$^{-/-}$,3b$^{-/-}$] cells. In all but one case, at least one re-sequenced clone matched the previously determined mCpG pattern precisely, and the overall level of methylation for each region was similar in all cases (Fig. 24 and data not shown). Thus, as a rule, a single clone from the RRBS library provides a good indication of the general methylation pattern at any given site. This is in line with the predominantly bimodal methylation profiles observed previously (reviewed in Ref. (Bird, 2002)). For example, >80% of the loci in the HEP survey of the MHC were either hypermethylated or hypomethylated (Rakyan et al., 2004).

Four representative examples are shown in Fig. 24. For the two loci on chromosome 4 and 15, respectively, all clones, including the clone from the RRBS library, indicated complete absence of methylation in *Dnmt*[1^{kd},3a$^{-/-}$,3b$^{-/-}$] cells. The sister cell line with normal Dnmt1 levels (see Fig. 18) was also considerably demethylated at these sites compared to wild-type ES cells. The two other loci maintained more mCpGs in the methylation-impaired cell lines. The two CpGs on chromosome 17 that were most consistently methylated in *Dnmt*[3a$^{-/-}$,3b$^{-/-}$] cells showed also residual methylation in the

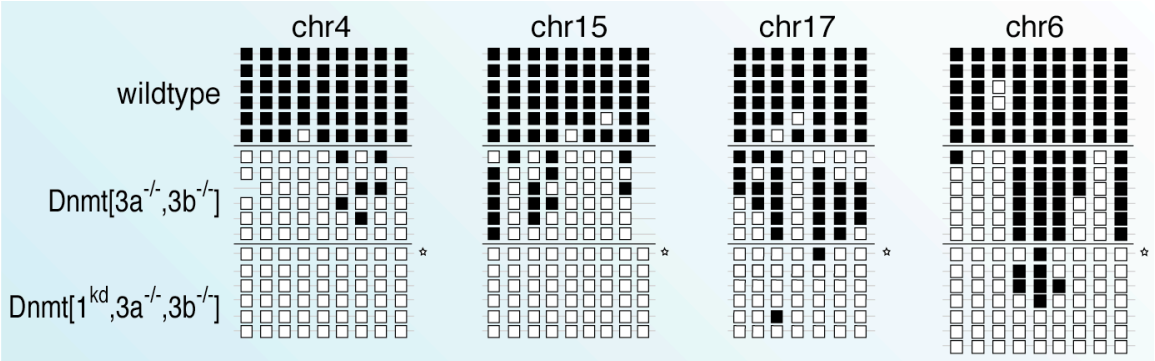

Figure 24: Targeted bisulfite sequencing of specific loci. Ten loci for which RRBS sequencing indicated mCpGs in Dnmt-deficient cells and 10 loci that were devoid of methylation were bisulfite re-sequenced using specific primers in wild-type (top), 3a/b double knockout (middle) and Dnmt[1^{kd},3a$^{-/-}$,3b$^{-/-}$] cells (bottom). Shown are two examples of each set. Each row represents a single sequenced molecule. Filled squares are methylated CpGs and empty ones indicate unmethylated sites. The asterisk indicates the original clone sequenced from the library.

Dnmt[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells. One of these two mCpGs was detected in the RRBS clone. Targeted resequencing detected methylation at the second CpG. This pattern is consistent with passive random loss of CpG methylation in *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells.

Chapter 4- Discussion

Much of our current research is focused on understanding the mechanisms that govern nuclear reprogramming through various approaches including nuclear transfer. This is ultimately aimed at generating patient-specific uncommitted stem and progenitor cells that may be useful for cell replacement therapies. The goal of the studies described here was to establish a set of methods that try to increase the understanding as well as the efficiency of epigenetic reprogramming. The potential applications of the tools and data presented here are manifold and will therefore be discussed in a more general way.

Conditional RNA interference

Since the development of gene targeting technologies in ES cells (Thomas and Capecchi, 1987), the gold standard for the analysis of gene function in mammals has been the creation of knock-out mice. Improvements to this technology have allowed for a more refined analysis of gene function at specific developmental stages or in specific tissues. These refined techniques are based on conditional knock-out strategies that are controlled by Cre-lox-regulated recombination (Van Dyke and Jacks, 2002). Despite significant technical improvements over the last decade, however, the creation of loss-of-function alleles in mice remains time consuming and costly. The recent demonstration that the RNA pol III-driven expression of shRNAs can be used to functionally silence gene expression in transgenic mice suggests that RNAi-based technologies might be a convenient alternative to gene targeting through homologous recombination (Carmell et al., 2003; Kunath et al., 2003; Rubinson et al., 2003).

A major limitation of current approaches for transgenic RNAi is the inability to regulate the expression of shRNA. Instead, approaches result in constitutive gene silencing in all tissues. The lentiviral vectors described here overcome this limitation.

The compact nature of RNA polymerase III promoters (Paule and White, 2000) prevents the use of a conventional Lox-STOP-lox strategy to achieve Cre-inducible shRNA expression. Some investigators have recently tried to circumvent this problem by placing the lox-STOP-lox cassette in the loop region of the shRNA (Fritsch et al., 2004; Kasim et al., 2004). However, this approach results in the transcription of the residual loxP site as part of the shRNA, resulting in the synthesis of a longer dsRNA that

processed less efficiently (Fritsch et al., 2004) and might elicit non-specific, off-target effects or an IFN response (Stark et al., 1998). By using a mutant lox site that contains a functional TATA box in its spacer sequence, we were able to obtain Cre-regulated transcription and efficient processing of a normal-length shRNA.

The potential applications of RNAi-based technology was further extended with the creation of the lentiviral vector pSicoR, in which constitutive shRNA expression can be terminated by a Cre-mediated recombination event. As demonstrated for Dnmt1, this vector can be used to determine the functional consequences of gene reactivation and will facilitate rescue experiments *in vivo*. In addition, by mimicking the action of small-molecule drugs designed to activate the proteins or pathways controlled by human disease genes (e.g., tumor suppressor genes), this strategy could be used to identify promising new targets for drug development.

Because preparation of conditional RNAi constructs only requires the cloning of short synthetic DNA sequences, a large number of conditional knock-down strains can be generated in parallel by a single investigator. This approach is thus ideally suited for large-scale projects aimed at the characterization of genetic pathways or the validation of candidate target genes identified through gene profiling screenings. For example, gene expression profiling in murine cancer models typically yields numerous genes that distinguish tumor from normal tissue. Using conventional or conditional knock-out strategies, it is practical to examine the functional relevance of only a small fraction of these genes. In contrast, shRNA conditional systems, such as pSico, greatly reduce the time, cost and effort required to perform such large-scale experiments.

It is important to note that although pSico and pSicoR were used in this work to control the expression of ‘‘artificial’’ shRNAs, they might also be used to spatially and temporally regulate the expression of naturally occurring microRNAs -- an approach that could help unravel the biological functions of this abundant class of small RNAs (Bartel, 2004).

The two lentiviral vectors reported here allow for greater control over gene inactivation compared with constitutive shRNA expression systems, an advance that expands the number of potential applications of RNAi-based technologies. The pSicoR

vector has been successfully used to test a proposed modification of the current nuclear transfer procedure called altered nuclear transfer.

Altered nuclear transfer

The ethical controversy surrounding nuclear transplantation arises from the unavoidable destruction of the reconstructed human blastocyst in order to obtain embryonic stem cells useful for biomedical research and therapy. The available evidence suggests that after nuclear transfer, the reconstructed embryos lack the potential to develop into normal human beings with any acceptable or practical efficiency (Jaenisch, 2004). Despite the incompatibility of this approach with normal human development, the utility and promise of nuclear transfer lies in the development of embryonic stem cells that have the same biological and molecular characteristics and the same therapeutic potential as those derived from fertilized embryos (Brambrink et al., 2006; Jaenisch, 2004). Altered nuclear transfer further cripples the already compromised blastocyst and eliminates the potential for the blastocyst to implant into the uterus and establish the fetal-maternal connection (Hurlbut, 2005). The genetic manipulations of the somatic donor cells that are required to generate this inherently abnormal blastocyst are simple and straightforward. Our data indicate that the removal of the *Cdx2* gene from somatic cells prior to nuclear transfer results in the formation of a blastocyst that lacks the functional cells of the trophectoderm lineage. This finding is consistent with previous results with embryos from mutant animals (Strumpf et al., 2005). Because the *Cdx2* gene is expressed prior to the blastocyst stage (Strumpf et al., 2005), *Cdx2*-deficient clones are abnormal at the pre-blastocyst stages, before an overtly abnormal phenotype becomes apparent. By reversing the *Cdx2* deficiency we demonstrate that fully competent ES cells can be derived from the inherently abnormal product of nuclear transfer using *Cdx2*-deficient donor cells.

If ANT was ever contemplated as an approach for the generation of human ES cells by nuclear transfer, the following issues need to be considered. First, although *CDX2* is expressed in the trophectoderm of human blastocysts (Adjaye et al., 2005) and derivatives of hES cells (Hyslop et al., 2005), its expression pattern in the human fetus has not been determined. As such, it is unknown whether the effect of *CDX2* on placentation will be the same in humans as in mice. Because the effect of gene inhibition

on human placentation cannot be tested, surrogate assays such as *in vitro* differentiation of hES cells are required to assess the effect of *CDX2* deficiency on human trophoblast development. In addition, with the use of retroviral vectors for gene transduction (Pfeifer et al., 2002) comes the risk of insertional mutagenesis and thus the activation of oncogenes that can cause leukemia (Hacein-Bey-Abina et al., 2003). However, this probably does not represent a serious problem in ANT because, in contrast to the gene therapy trials using retroviral infection of bone marrow cells, viral integration into the fibroblasts does not lead to a proliferative advantage and selective outgrowth of infected cells due to an activated oncogene. Finally, since all nuclear transfer-derived ES cells are clonal, simple DNA analysis would indicate whether proviral integration occurred in the vicinity of an oncogene.

The results reported in this study provide proof of principle that inhibition of genes important for trophoblast function can prevent placentation without interfering with ES cell potency, and may thus provide a scientific way to side-step the ongoing debate surrounding the nuclear transfer technology. However, because the *Cdx2*-deficient embryo is not obviously abnormal before the onset of *Cdx2* expression, this approach may not solve the ethical dilemma. Moreover, research with primate or human cells will be required to assess whether *Cdx2* is an optimal target for human application. Finally, we wish to emphasize that ANT is simply a modification and not an alternative to nuclear transfer. As the approach requires additional manipulation of the donor cells, it is likely to complicate the logistics of production and safety assessment of therapeutic patient-specific ES cell lines.

Some critics have raised doubts about whether an essential role of *CDX2* in human placentation could ever be established, thus making it impossible to assess the potential effectiveness of the ANT approach in humans (Solter, 2005). While *CDX2*'s role in human placentation cannot be studied in mutant embryos, a simple *in vitro* assay could be developed. For example, human ANT blastocysts could be explanted in culture to verify whether *CDX2* deficiency prevents the generation of trophoblast cells – the cells that generate the placental lineage - as it does in the mouse. Although such surrogate assays cannot provide 100% certainty regarding the effect of *CDX2* deficiency on human placentation, a positive *in vitro* result would provide sufficient confidence that inhibition

of the *CDX2* gene would abrogate the reproductive cloning potential of ANT-derived human blastocysts. It has also been argued that because of the dependence of siRNA expression on genomic integration, numerous human ANT embryos would have to be tested to assure sufficient *CDX2* inhibition. However, linking siRNA with GFP expression provides a simple way to isolate donor cells that contain a high level of CDX2 siRNA prior to nuclear transfer. Finally, although ANT will probably not persuade everyone who opposes nuclear cloning, and attempts to find technological solutions to contentious ethical issues may represent a diversion of the scientific process (Melton et al., 2004), we also find merit in another argument. Our work is supported by public funds. As a result, it could be argued that we have an obligation to take the public debate on scientific and ethical issues seriously and to contribute to possible solutions. As long as the experiments performed in pursuit of this goal are scientifically sound, it is a moot point whether such efforts would lead to a resolution of the controversy.

As described in Chapter 1, a great deal of current research is focused on the de-differentiation and/or reprogramming of somatic cells with the ultimate goal of generating less differentiated cells that will be useful for patient therapy. The cloning of fully differentiated cells has demonstrated that no genetic information is lost during development and that nuclear totipotency is retained. This suggested that epigenetic mechanisms of gene regulation and differentiation are responsible for keeping somatic cells in their differentiated state. In order to better define the epigenome of different cell types, we have developed an approach for large-scale high resolution DNA methylation analysis.

Genome-wide high resolution DNA methylation analysis

In this work, we explored the feasibility of large-scale shotgun bisulfite sequencing for genome-wide analysis of DNA methylation. We have shown that bisulfite sequencing libraries can be constructed that are largely unbiased and representative of the genome. These libraries display few false-positive methylcytosines caused by incomplete cytosine to uracil conversion or PCR and sequencing errors.

Insert sizes of the libraries were kept very small (500-600 bp) for two reasons. First, the bisulfite reaction requires relatively high temperatures (50-60°C) and a low pH (pH5) -- conditions that are known to cause depurination and strand breakage. In

addition, smaller molecules are less prone to damage and require fewer PCR cycles to recover intact molecules suitable for cloning compared with larger ones, thereby minimizing the risk of a skewed representation. Finally, larger-insert clones would require sequencing of both strands, and C-poor strands have proven difficult to sequence in our hands.

We used limit digestion with *Bgl*II and size fractionation to reduce the complexity of the DNA. The resulting RRBS libraries cover a small but reproducible fraction of the genome and are therefore suitable for large-scale comparative methylation studies across different strains, tissues or cell types. Based on the overall success rate (72%) and insert-size distributions encountered during this pilot study (Fig. 23), we expect that for a pair-wise comparison, sequencing 100 x 384 RRBS clones from each DNA sample will produce 4.0 Mb of high-quality, overlapping bisulfite sequence with 2- to 3-fold coverage in each library of fragments within 1 SD of the mean size. Assuming that improvements in sequencing of C-poor strands (85% success rate) and better libraries with congruent insert-size distributions can be made, the same sequencing effort would yield ~5.8 Mb of pair-wise comparative sequence which, of course, is still only a tiny fraction of the genome.

At this level of genome coverage, differential methylation at most individual sites in the genome, including many functionally important ones, is likely to escape detection. However, we expect the coverage to be sufficient to generate methylation variable position markers for future bisulfite SNP "epigenotyping" (Murrell et al., 2005). A genome-wide set of comparative bisulfite sequences may prove useful to train computer algorithms for predicting methylation patterns. RRBS sequencing may be sufficient to detect genomic imprints (or the loss thereof), tissue-specific regulated methylation domains or long-range methylation gradients along a chromosome. We also envision RRBS applications in epigenetic cancer profiling and bio-marker discovery.

Despite the essential role of the known DNA methyltransferases in mouse development (Li et al., 1992; Okano et al., 1999), DNA methylation and the enzymes responsible for its establishment and maintenance appear to be largely dispensable in undifferentiated ES cells. *Dnmt*1-deficient ES cells retain approximately 20% CpG methylation, likely due to continuous *de novo* methylation by Dnmt3a and Dnmt3b.

Although early passage Dnmt3a/b double mutant ES cells show almost wild-type levels of CpG methylation (Chen et al., 2003; Jackson et al., 2004), they progressively lose methylation with <1% remaining after 75 passages (Jackson et al., 2004). This gradual loss may reflect the infidelity of the maintenance enzyme Dnmt1.

Our data showed that ES cells that lack the DNA methyltransferases Dnmt3a and 3b and have greatly reduced levels of Dnmt1 were viable with 1-2% CpG methylation remaining after only six passages. The extremely low rate of false-positive methylcytosines allowed us to identify and inspect some of the rare sites that retained methylation. There were no obvious hotspots for residual mCpGs in the genome (Figures 20 and 21). Also, there was no correlation between the numbers of CpGs and the residual methylation at a given site. The distance to CpG islands or to known genes also appeared to be random and none of the loci was notably conserved across species. Finally, no specific motif was detected upstream or downstream of the residual mCpG dinucleotides (data not shown). Thus these findings provide no evidence of specific maintenance of residual mCpG by an unidentified DNA methyltransferase. Rather, *Dnmt*[1^{kd},$3a^{-/-}$,$3b^{-/-}$] cells seem to lose residual CpG methylation in a random fashion over time.

Only 3 of the 25,505 sequenced CpT dinucleotides were inferred to be methylated in Dnmt-deficient cells, and no methylated CpA was detected. By comparison, wild-type cells showed 0.7% CpT and 2.4% CpA methylation in agreement with previous observations (Dodge et al., 2002; Ramsahoye et al., 2000). Previous experiments have also shown that the presence of Dnmt1 is not required for non-CpG methylation (Ramsahoye et al., 2000). In contrast, non-CpG methylation becomes almost undetectable in ES cells lacking Dnmt3a and Dnmt3b (Dodge et al., 2002). Both global nearest neighbor data and our bisulfite-sequencing data therefore suggest that the *de novo* DNA methyltransferases 3a and/or 3b are responsible for asymmetric CpA and CpT methylation in murine ES cells.

In this pilot study we have employed a combination of RNAi-induced knock-down and complete knockout of DNA methyltransferases to generate murine ES cells that were almost completely devoid of DNA methylation. These cells had only 1-2% residual CpG methylation left after a few passages, and non-CpG methylation was over 250-fold reduced compared to wild-type ES cells.

Unamplified, nearly methylation-free genomic DNA is an ideal substrate to optimize and test conditions for genome-wide bisulfite conversion, PCR amplification and library construction for future genomic shotgun bisulfite sequencing of mammalian genomes. We have shown that essentially complete bisulfite conversion can be achieved without undue adverse effects on library complexity and sequence representation.

Perspectives

RNA interference has become one of the predominant tools to assess gene function across all species. In the near future, it is likely that a catalog of functional siRNAs will be available for every known gene in most organisms. And further studies comparing the mechanisms of endogenous miRNAs and synthetic siRNA should increase the efficiency of rational RNAi design. The discovery and optimization of RNAi will undoubtedly lead to advances in medical treatment, either directly as therapy or by revealing new gene targets for therapy.

In the 50 years since Briggs and King first reported the nuclear transfer technique, NT has become a powerful research tool that holds great promise for therapy in the years to come. The prospect of deriving customized ES cells has significant implications for medical therapy, and, more immediately, for the use of patient-derived ES cells as disease models. For instance, stem cells derived from a patient that suffers from Parkinson's disease would provide limitless material to study the physiology of the affected cells. Typically, by the time Parkinson's disease is diagnosed, most of the cells have been lost. Stem cells derived from patients could be differentiated into neurons *in vitro* and then used for mechanistic studies. These cells might also provide a means of screening candidate drugs and compounds with the potential to counter neurodegeneration. The long-term goal, however, remains patient-tailored cell replacement therapy. This goal will require longer-term clinical trials and the resolution of numerous safety issues before such a therapy could be implemented in patients.

One of the next steps in studying stem cell biology will be to decipher the epigenome of stem cells and compare it to differentiated cells. As discussed, large-scale random bisulfite sequencing complements existing directed bisulfite sequencing strategies, which are well suited to analyze a limited number of gene promoters and regulatory sequence elements in a large number of samples. One advantage of sequencing

clone libraries in a random fashion is that no target-specific PCR or sequencing primers are needed. Once the library is made, the method is amenable to automation and is scaleable. Since the bisulfite reads are not assembled but merely aligned to the reference genome sequence, we expect this method to work well in combination with highly parallel sequencing technologies that produce single reads of approximately 100 bases in length (Margulies et al., 2005). Finally, in principle, bisulfite-converted libraries can be constructed from randomly sheared DNA for future whole-genome bisulfite sequencing.

References

Adjaye, J., Huntriss, J., Herwig, R., Benkahla, A., Brink, T., Wierling, C., Hultschig, C., Groth, D., Yaspo, M. L., Picton, H.*, et al.* (2005). Primary differentiation in the human blastocyst: Comparative molecular portraits of inner cell mass and trophectoderm cells. Stem Cells.

Altshuler, D., Pollara, V. J., Cowles, C. R., Van Etten, W. J., Baldwin, J., Linton, L., and Lander, E. S. (2000). An SNP map of the human genome generated by reduced representation shotgun sequencing. Nature *407*, 513-516.

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell *122*, 553-563.

Baguisi, A., Behboodi, E., Melican, D. T., Pollock, J. S., Destrempes, M. M., Cammuso, C., Williams, J. L., Nims, S. D., Porter, C. A., Midura, P.*, et al.* (1999). Production of goats by somatic cell nuclear transfer. Nat Biotechnol *17*, p456-461.

Barberi, T., Klivenyi, P., Calingasan, N. Y., Lee, H., Kawamata, H., Loonam, K., Perrier, A. L., Bruses, J., Rubio, M. E., Topf, N.*, et al.* (2003). Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice. Nat Biotechnol *21*, 1200-1207.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bedell, J. A., Budiman, M. A., Nunberg, A., Citek, R. W., Robbins, D., Jones, J., Flick, E., Rholfing, T., Fries, J., Bradford, K.*, et al.* (2005). Sorghum genome sequencing by methylation filtration. PLoS Biol *3*, e13.

Berns, K., Hijmans, E. M., Mullenders, J., Brummelkamp, T. R., Velds, A., Heimerikx, M., Kerkhoven, R. M., Madiredjo, M., Nijkamp, W., Weigelt, B.*, et al.* (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature *428*, 431-437.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bertwistle, D., Sugimoto, M., and Sherr, C. J. (2004). Physical and functional interactions of the Arf tumor suppressor protein with nucleophosmin/B23. Mol Cell Biol *24*, 985-996.

Bestor, T. H. (2000a). The DNA methyltransferases of mammals. Hum Mol Genet *9*, 2395-2402.

Bestor, T. H. (2000b). The DNA methyltransferases of mammals. Hum Mol Genet *9*, 2395-2402.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006). The Epigenome Influences Efficiency of Reprogramming by Nuclear Transfer. (submitted).

Blelloch, R. H., Hochedlinger, K., Yamada, Y., Brennan, C., Kim, M., Mintz, B., Chin, L., and Jaenisch, R. (2004). Nuclear cloning of embryonal carcinoma cells. Proc Natl Acad Sci U S A *101*, 13985-13990.

Boese, Q., Leake, D., Reynolds, A., Read, S., Scaringe, S. A., Marshall, W. S., and Khvorova, A. (2005). Mechanistic insights aid computational short interfering RNA design. Methods Enzymol *392*, 73-96.

Boiani, M., Eckardt, S., Scholer, H. R., and McLaughlin, K. J. (2002). Oct4 distribution and level in mouse clones: consequences for pluripotency. Genes Dev *16*, 1209-1219.

Bortvin, A., Eggan, K., Skaletsky, H., Akutsu, H., Berry, D. L., Yanagimachi, R., Page, D. C., and Jaenisch, R. (2003). Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. Development *130*, 1673-1680.

Bourc'his, D., Le Bourhis, D., Patin, D., Niveleau, A., Comizzoli, P., Renard, J. P., and Viegas-Pequignot, E. (2001a). Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. Curr Biol *11*, 1542-1546.

Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B., and Bestor, T. H. (2001b). Dnmt3L and the establishment of maternal genomic imprints. Science *294*, 2536-2539.

Bradley, A., Evans, M., Kaufman, M. H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Brambrink, T., Hochedlinger, K., and Jaenisch, R. (2006). Gene expression in embryonic stem cells from cloned and fertilized embryos. Proc Natl Acad Sci U S A *103*,933-8.

Briggs, R., and King, T. J. (1952). Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. Proc Natl Acad Sci U S A *38*, 455-463.

Briggs, R., and King, T. J. (1957). Changes in the nuclei of differentiating endoderm cells as revealed by nuclear transplantation. J Morphol *100*, 269-311.

Bromhall, J. D. (1975). Nuclear transplantation in the rabbit egg. Nature *258*, 719-722.

Campbell, K. H., Alberio, R., Choi, I., Fisher, P., Kelly, R. D., Lee, J. H., and Maalouf, W. (2005). Cloning: eight years after Dolly. Reprod Domest Anim *40*, 256-268.

Campbell, K. H., McWhir, J., Ritchie, W. A., and Wilmut, I. (1996). Sheep cloned by nuclear transfer from a cultured cell line. Nature *380*, 64-66.

Carmell, M. A., Zhang, L., Conklin, D. S., Hannon, G. J., and Rosenquist, T. A. (2003). Germline transmission of RNAi in mice. Nat Struct Biol *10*, 91-92.

Cezar, G. G., Bartolomei, M. S., Forsberg, E. J., First, N. L., Bishop, M. D., and Eilertsen, K. J. (2003). Genome-wide epigenetic alterations in cloned bovine fetuses. Biol Reprod *68*, 1009-1014.

Chapman, V., Forrester, L., Sanford, J., Hastie, N., and Rossant, J. (1984). Cell lineage-specific undermethylation of mouse repetitive DNA. Nature *307*, 284-286.

Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J., and Beck, F. (2004). Cdx2 is essential for axial elongation in mouse development. Proc Natl Acad Sci U S A *101*, 7641-7645.

Chawengsaksophak, K., James, R., Hammond, V. E., Kontgen, F., and Beck, F. (1997). Homeosis and intestinal tumours in Cdx2 mutant mice. Nature *386*, 84-87.

Chen, T., Ueda, Y., Dodge, J. E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol *23*, 5594-5605.

Chesne, P., Adenot, P. G., Viglietta, C., Baratte, M., Boulanger, L., and Renard, J. P. (2002). Cloned rabbits produced by nuclear transfer from adult somatic cells. Nat Biotechnol *20*, 366-369.

Chung, Y., Klimanskaya, I., Becker, S., Marh, J., Lu, S. J., Johnson, J., Meisner, L., and Lanza, R. (2006). Embryonic and extraembryonic stem cell lines derived from single mouse blastomeres. Nature *439*, 216-219.

Chung, Y. G., Ratnam, S., Chaillet, J. R., and Latham, K. E. (2003). Abnormal regulation of DNA methyltransferase expression in cloned mouse embryos. Biol Reprod.

Cibelli, J. B., Stice, S. L., Golueke, P. J., Kane, J. J., Jerry, J., Blackwell, C., Ponce de Leon, F. A., and Robl, J. M. (1998). Cloned transgenic calves produced from nonquiescent fetal fibroblasts. Science *280*, 1256-1258.

Colombo, E., Marine, J. C., Danovi, D., Falini, B., and Pelicci, P. G. (2002). Nucleophosmin regulates the stability and transcriptional activity of p53. Nat Cell Biol *4*, 529-533.

Cowan, C. A., Atienza, J., Melton, D. A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309*, 1369-1373.

Cross, S. H., and Bird, A. P. (1995). CpG islands and genes. Curr Opin Genet Dev *5*, 309-314.

Dean, W., Bowden, L., Aitchison, A., Klose, J., Moore, T., Meneses, J. J., Reik, W., and Feil, R. (1998). Altered imprinted gene methylation and expression in completely ES

cell- derived mouse fetuses: association with aberrant phenotypes. Development *125*, 2273-2282.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., and Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci U S A *98*, 13734-13738.

Dennis, K., Fan, T., Geiman, T., Yan, Q., and Muegge, K. (2001). Lsh, a member of the SNF2 family, is required for genome-wide methylation. Genes Dev *15*, 2940-2944.

Do, J. T., and Scholer, H. R. (2004). Nuclei of embryonic stem cells reprogram somatic cells. Stem Cells *22*, 941-949.

Dodge, J. E., Ramsahoye, B. H., Wo, Z. G., Okano, M., and Li, E. (2002). De novo methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. Gene *289*, 41-48.

Doench, J. G., Petersen, C. P., and Sharp, P. A. (2003). siRNAs can function as miRNAs. Genes Dev *17*, 438-442.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res *19*, 4008.

Dykxhoorn, D. M., Novina, C. D., and Sharp, P. A. (2003). Killing the messenger: short RNAs that silence gene expression. Nat Rev Mol Cell Biol *4*, 457-467.

Eads, C. A., and Laird, P. W. (2002). Combined bisulfite restriction analysis (COBRA). Methods Mol Biol *200*, 71-85.

Eggan, K., Akutsu, H., Loring, J., Jackson-Grusby, L., Klemm, M., Rideout, W. M., 3rd, Yanagimachi, R., and Jaenisch, R. (2001a). Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. Proc Natl Acad Sci U S A *98*, 6209-6214.

Eggan, K., Baldwin, K., Tackett, M., Osborne, J., Gogos, J., Chess, A., Axel, R., and Jaenisch, R. (2004). Mice cloned from olfactory sensory neurons. Nature *428*, 44-49.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Elbashir, S. M., Harborth, J., Weber, K., and Tuschl, T. (2002). Analysis of gene function in somatic mammalian cells using small interfering RNAs. Methods *26*, 199-213.

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

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

Feinberg, A. P. (2004). The epigenetics of cancer etiology. Semin Cancer Biol *14*, 427-432.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fogg, M. J., Pearl, L. H., and Connolly, B. A. (2002). Structural basis for uracil recognition by archaeal family B DNA polymerases. Nat Struct Biol *9*, 922-927.

Fritsch, L., Martinez, L. A., Sekhri, R., Naguibneva, I., Gerard, M., Vandromme, M., Schaeffer, L., and Harel-Bellan, A. (2004). Conditional gene knock-down by CRE-dependent short interfering RNAs. EMBO Rep *5*, 178-182.

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., and Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A *89*, 1827-1831.

Fujita, N., Shimotake, N., Ohki, I., Chiba, T., Saya, H., Shirakawa, M., and Nakao, M. (2000). Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. Mol Cell Biol *20*, 5107-5118.

Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H., and Nakao, M. (1999). Methylation-mediated transcriptional silencing in euchromatin by methyl- CpG binding protein MBD1 isoforms. Mol Cell Biol *19*, 6415-6426.

Galli, C., Lagutina, I., Crotti, G., Colleoni, S., Turini, P., Ponderato, N., Duchi, R., and Lazzari, G. (2003). Pregnancy: a cloned horse born to its dam twin. Nature *424*, 635.

Garnes, J., Ciancio, M. and Gnirke, A. (2002). Construction of large-insert bacterial clone libraries. In Genomic Mapping and Sequencing (UK, Horizon Scientific Press).

Gasaryan, K. G., Hung, N. M., Neyfakh, A. A., and Ivanenkov, V. V. (1979). Nuclear transplantation in teleost Misgurnus fossilis L. Nature *280*, 585-587.

Gaudet, F., Hodgson, J. G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J. W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. Science *300*, 489-492.

Gaudet, F., Rideout, W. M., 3rd, Meissner, A., Dausman, J., Leonhardt, H., and Jaenisch, R. (2004). Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing. Mol Cell Biol *24*, 1640-1648.

Gerdes, T., and Wabl, M. (2004). Autoreactivity and allelic inclusion in a B cell nuclear transfer mouse. Nat Immunol *5*, 1282-1287.

Grunau, C., Clark, S. J., and Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res *29*, E65-65.

Gurdon, J. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J Embryol exp Morph *10*, 622-640.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E.*, et al.* (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415-419.

Haines, T. R., Rodenhiser, D. I., and Ainsworth, P. J. (2001). Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. Dev Biol *240*, 585-598.

Hakelien, A. M., Gaustad, K. G., and Collas, P. (2004). Transient alteration of cell fate using a nuclear and cytoplasmic extract of an insulinoma cell line. Biochem Biophys Res Commun *316*, 834-841.

Hakelien, A. M., Landsverk, H. B., Robl, J. M., Skalhegg, B. S., and Collas, P. (2002). Reprogramming fibroblasts to express T-cell functions using cell extracts. Nat Biotechnol *20*, 460-466.

Hall, V. J., Ruddock, N. T., and French, A. J. (2005). Expression profiling of genes crucial for placental and preimplantation development in bovine in vivo, in vitro, and nuclear transfer blastocysts. Mol Reprod Dev *72*, 16-24.

Han, Y. M., Kang, Y. K., Koo, D. B., and Lee, K. K. (2003). Nuclear reprogramming of cloned embryos produced in vitro. Theriogenology *59*, 33-44.

Hannon, G. J., and Rossi, J. J. (2004). Unlocking the potential of the human genome with RNA interference. Nature *431*, 371-378.

Hashimshony, T., Zhang, J., Keshet, I., Bustin, M., and Cedar, H. (2003). The role of DNA methylation in setting up chromatin structure during development. Nat Genet *34*, 187-192.

He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J., and Hammond, S. M. (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828-833.

Hennet, T., Hagen, F. K., Tabak, L. A., and Marth, J. D. (1995). T-cell-specific deletion of a polypeptide N-acetylgalactosaminyl-transferase gene by site-directed recombination. Proc Natl Acad Sci U S A *92*, 12070-12074.

Hochedlinger, K., Blelloch, R., Brennan, C., Yamada, Y., Kim, M., Chin, L., and Jaenisch, R. (2004). Reprogramming of a melanoma genome by nuclear transplantation. Genes Dev *18*, 1875-1885.

Hochedlinger, K., and Jaenisch, R. (2002a). Monoclonal mice generated by nuclear transfer from mature B and T donor cells. Nature *415*, 1035-1038.

Hochedlinger, K., and Jaenisch, R. (2002b). Nuclear transplantation: lessons from frogs and mice. Curr Opin Cell Biol *14*, 741-748.

Howell, C. Y., Bestor, T. H., Ding, F., Latham, K. E., Mertineit, C., Trasler, J. M., and Chaillet, J. R. (2001). Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell *104*, 829-838.

Humpherys, D., Eggan, K., Akutsu, H., Friedman, A., Hochedlinger, K., Yanagimachi, R., Lander, E. S., Golub, T. R., and Jaenisch, R. (2002). Abnormal gene expression in cloned mice derived from embryonic stem cell and cumulus cell nuclei. Proc Natl Acad Sci U S A *99*, 12889-12894.

Humpherys, D., Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W. M., 3rd, Biniszkiewicz, D., Yanagimachi, R., and Jaenisch, R. (2001). Epigenetic instability in ES cells and cloned mice. Science *293*, 95-97.

Hurlbut, W. B. (2005). Altered nuclear transfer as a morally acceptable means for the procurement of human embryonic stem cells. Perspect Biol Med *48*, 211-228.

Hyslop, L. A., Stojkovic, M., Armstrong, L., Walter, T., Stojkovic, P., Przyborski, S., Herbert, M., Murdoch, A., Strachan, T., and Lako, M. (2005). Downregulation of NANOG induces differentiation of Human Embryonic Stem Cells to Extraembryonic Lineages. Stem Cells.

Inoue, K., Wakao, H., Ogonuki, N., Miki, H., Seino, K., Nambu-Wakao, R., Noda, S., Miyoshi, H., Koseki, H., Taniguchi, M., and Ogura, A. (2005). Generation of cloned mice by direct nuclear transfer from natural killer T cells. Curr Biol *15*, 1114-1118.

Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003). Expression profiling reveals off-target gene regulation by RNAi. Nat Biotechnol *21*, 635-637.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol Cell Biol *24*, 8862-8871.

Jaenisch, R. (1997). DNA methylation and imprinting: why bother? Trends Genet *13*, 323-329.

Jaenisch, R. (2004). Human cloning - the science and ethics of nuclear transplantation. N Engl J Med *351*, 2787-2791.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet *33 Suppl*, 245-254.

Jaenisch, R., and Wilmut, I. (2001). Developmental biology. Don't clone humans! Science *291*, 2552.

Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem *3*, 274-293.

Jones, P. A., and Baylin, S. B. (2002). The fundamental role of epigenetic events in cancer. Nat Rev Genet *3*, 415-428.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-191.

Kang, Y. K., Koo, D. B., Park, J. S., Choi, Y. H., Chung, A. S., Lee, K. K., and Han, Y. M. (2001a). Aberrant methylation of donor genome in cloned bovine embryos. Nat Genet *28*, 173-177.

Kang, Y. K., Koo, D. B., Park, J. S., Choi, Y. H., Kim, H. N., Chang, W. K., Lee, K. K., and Han, Y. M. (2001b). Typical demethylation events in cloned pig embryos. Clues on species- specific differences in epigenetic reprogramming of a cloned donor genome. J Biol Chem *276*, 39980-39984.

Kang, Y. K., Park, J. S., Koo, D. B., Choi, Y. H., Kim, S. U., Lee, K. K., and Han, Y. M. (2002). Limited demethylation leaves mosaic-type methylation states in cloned bovine pre-implantation embryos. Embo J *21*, 1092-1100.

Kasim, V., Miyagishi, M., and Taira, K. (2004). Control of siRNA expression using the Cre-loxP recombination system. Nucleic Acids Res *32*, e66.

Kato, Y., Tani, T., Sotomaru, Y., Kurokawa, K., Kato, J., Doguchi, H., Yasue, H., and Tsunoda, Y. (1998). Eight calves cloned from somatic cells of a single adult. Science *282*, 2095-2098.

Kawasaki, H., and Taira, K. (2004). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature *431*, 211-217.

Kishigami, S., Mizutani, E., Ohta, H., Hikichi, T., Thuan, N. V., Wakayama, S., Bui, H. T., and Wakayama, T. (2006). Significant improvement of mouse cloning technique by treatment with trichostatin A after somatic nuclear transfer. Biochem Biophys Res Commun *340*, 183-189.

Koralov, S. B., Novobrantseva, T. I., Hochedlinger, K., Jaenisch, R., and Rajewsky, K. (2005). Direct in vivo VH to JH rearrangement violating the 12/23 rule. J Exp Med *201*, 341-348.

Kunath, T., Gish, G., Lickert, H., Jones, N., Pawson, T., and Rossant, J. (2003). Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. Nat Biotechnol *21*, 559-561.

Laird, P. W. (2003). The power and the promise of DNA methylation markers. Nat Rev Cancer *3*, 253-266.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., and Reik, W. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis *35*, 88-93.

Lanza, R. P., Cibelli, J. B., Diaz, F., Moraes, C. T., Farin, P. W., Farin, C. E., Hammer, C. J., West, M. D., and Damiani, P. (2000). Cloning of an endangered species (Bos gaurus) using interspecies nuclear transfer. Cloning *2*, 79-90.

Lee, B. C., Kim, M. K., Jang, G., Oh, H. J., Yuda, F., Kim, H. J., Shamim, M. H., Kim, J. J., Kang, S. K., Schatten, G., and Hwang, W. S. (2005). Dogs cloned from adult somatic cells. Nature *436*, 641.

Lee, K. Y., Huang, H., Ju, B., Yang, Z., and Lin, S. (2002). Cloned zebrafish by nuclear transfer from long-term-cultured cells. Nat Biotechnol *20*, 795-799.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195-3205.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet *3*, 662-673.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Lin, H., Yamada, Y., Nguyen, S., Jackson-Grusby, L., Meissner, A., Linhart, H., and Jaenisch, R. (2006). Suppression of intestinal neoplasia by deletion of Dnmt3b. Mol Cell Biol *(in press)*.

Lippman, Z., Gendrel, A. V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., Lavine, K., Mittal, V., May, B., Kasschau, K. D.*, et al.* (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471-476.

Lippman, Z., Gendrel, A. V., Colot, V., and Martienssen, R. (2005). Profiling DNA methylation patterns using genomic tiling microarrays. Nat Methods *2*, 219-224.

Loi, P., Ptak, G., Barboni, B., Fulka, J., Jr., Cappai, P., and Clinton, M. (2001). Genetic rescue of an endangered mammal by cross-species nuclear transfer using post-mortem somatic cells. Nat Biotechnol *19*, 962-964.

Lucifero, D., Mertineit, C., Clarke, H. J., Bestor, T. H., and Trasler, J. M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. Genomics *79*, 530-538.

Mann, M. R., Chung, Y. G., Nolen, L. D., Verona, R. I., Latham, K. E., and Bartolomei, M. S. (2003). Disruption of Imprinted Gene Methylation and Expression in Cloned Preimplantation Stage Mouse Embryos. Biol Reprod.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z.*, et al.* (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Matzke, M. A., and Birchler, J. A. (2005). RNAi-mediated pathways in the nucleus. Nat Rev Genet *6*, 24-35.

McCaffrey, A. P., Nakai, H., Pandey, K., Huang, Z., Salazar, F. H., Xu, H., Wieland, S. F., Marion, P. L., and Kay, M. A. (2003). Inhibition of hepatitis B virus in mice by RNA interference. Nat Biotechnol *21*, 639-644.

McGrath, J., and Solter, D. (1983). Nuclear transplantation in the mouse embryo by microsurgery and cell fusion. Science *220*, 1300-1302.

McGrath, J., and Solter, D. (1984). Inability of mouse blastomere nuclei transferred to enucleated zygotes to support development in vitro. Science *226*, 1317-1319.

McManus, M. T., and Sharp, P. A. (2002). Gene silencing in mammals by small interfering RNAs. Nat Rev Genet *3*, 737-747.

Meissner, A., and Jaenisch, R. (2006). Generation of nuclear transfer-derived pluripotent ES cells from cloned Cdx2-deficient blastocysts. Nature *439*, 212-215.

Melton, D. A., Daley, G. Q., and Jennings, C. G. (2004). Altered nuclear transfer in stem-cell research - a flawed proposal. N Engl J Med *351*, 2791-2792.

Miller, V. M., Xia, H., Marrs, G. L., Gouvion, C. M., Lee, G., Davidson, B. L., and Paulson, H. L. (2003). Allele-specific silencing of dominant disease genes. Proc Natl Acad Sci U S A *100*, 7195-7200.

Mittal, V. (2004). Improving the efficiency of RNA interference in mammals. Nat Rev Genet *5*, 355-365.

Munsie, M. J., Michalska, A. E., O'Brien, C. M., Trounson, A. O., Pera, M. F., and Mountford, P. S. (2000). Isolation of pluripotent embryonic stem cells from reprogrammed adult mouse somatic cell nuclei. Curr Biol *10*, 989-992.

Murrell, A., Rakyan, V. K., and Beck, S. (2005). From genome to epigenome. Hum Mol Genet *14 Spec No 1*, R3-R10.

Nan, X., Campoy, F. J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471-481.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Neumann, B., and Barlow, D. P. (1996). Multiple roles for DNA methylation in gametic imprinting. Curr Opin Genet Dev *6*, 159-163.

Niwa, H., Miyazaki, J., and Smith, AG. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. *24*, 372-6.

Novik, K. L., Nimmrich, I., Genc, B., Maier, S., Piepenbrock, C., Olek, A., and Beck, S. (2002). Epigenomics: genome-wide study of methylation phenomena. Curr Issues Mol Biol *4*, 111-128.

Ohgane, J., Wakayama, T., Kogo, Y., Senda, S., Hattori, N., Tanaka, S., Yanagimachi, R., and Shiota, K. (2001). DNA methylation variation in cloned mice. Genesis *30*, 45-50.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Okano, M., and Li, E. (2002). Genetic analyses of DNA methyltransferase genes in mouse model system. J Nutr *132*, 2462S-2465S.

Okano, M., Xie, S., and Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet *19*, 219-220.

Okano, M., Xie, S., and Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res *26*, 2536-2540.

Ota, A., Tagawa, H., Karnan, S., Tsuzuki, S., Karpas, A., Kira, S., Yoshida, Y., and Seto, M. (2004). Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. Cancer Res *64*, 3087-3095.

Paddison, P. J., Silva, J. M., Conklin, D. S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K.*, et al.* (2004). A resource for large-scale RNA-interference-based screens in mammals. Nature *428*, 427-431.

Paule, M. R., and White, R. J. (2000). Survey and summary: transcription by RNA polymerases I and III. Nucleic Acids Res *28*, 1283-1298.

Paulin, R., Grigg, G. W., Davey, M. W., and Piper, A. A. (1998). Urea improves efficiency of bisulphite-mediated sequencing of 5'-methylcytosine in genomic DNA. Nucleic Acids Res *26*, 5009-5010.

Pfeifer, A., Ikawa, M., Dayn, Y., and Verma, I. M. (2002). Transgenesis by lentiviral vectors: lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. Proc Natl Acad Sci U S A *99*, 2140-2145.

Polejaeva, I., Chen, S., Vaught, T., Page, R., Mullins, J., Ball, S., Dai, Y., Boone, J., Walker, S., Ayares, D.*, et al.* (2000). Cloned pigs produced by nuclear transfer from adult somatic cells [see comments]. Nature *407*, 86-90.

Prather, R. S., Sims, M. M., and First, N. L. (1989). Nuclear transplantation in early pig embryos. Biol Reprod *41*, 414-418.

Rabinowicz, P. D., Schutz, K., Dedhia, N., Yordan, C., Parnell, L. D., Stein, L., McCombie, W. R., and Martienssen, R. A. (1999). Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. Nat Genet *23*, 305-308.

Rakyan, V. K., Hildmann, T., Novik, K. L., Lewin, J., Tost, J., Cox, A. V., Andrews, T. D., Howe, K. L., Otto, T., Olek, A.*, et al.* (2004). DNA Methylation Profiling of the Human Major Histocompatibility Complex: A Pilot Study for the Human Epigenome Project. PLoS Biol *2*, e405.

Ramsahoye, B. H. (2002a). Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. Methods *27*, 156-161.

Ramsahoye, B. H. (2002b). Nearest-neighbor analysis. Methods Mol Biol *200*, 9-15.

Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A *97*, 5237-5242.

Ratnam, S., Mertineit, C., Ding, F., Howell, C. Y., Clarke, H. J., Bestor, T. H., Chaillet, J. R., and Trasler, J. M. (2002). Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. Dev Biol *245*, 304-314.

Razin, A., and Cedar, H. (1994). DNA methylation and genomic imprinting. Cell *77*, 473-476.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089-1093.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W. S., and Khvorova, A. (2004). Rational siRNA design for RNA interference. Nat Biotechnol *22*, 326-330.

Rideout, W. M., 3rd, Hochedlinger, K., Kyba, M., Daley, G. Q., and Jaenisch, R. (2002). Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. Cell *109*, 17-27.

Riggs, A. D., and Pfeifer, G. P. (1992). X-chromosome inactivation and cell memory. Trends Genet *8*, 169-174.

Robl, J. M., Prather, R., Barnes, F., Eyestone, W., Northey, D., Gilligan, B., and First, N. L. (1987). Nuclear transplantation in bovine embryos. J Anim Sci *64*, 642-647.

Robertson, K. D. (2002). DNA methylation and chromatin - unraveling the tangled web. Oncogene *21*, 5361-5379.

Robertson, K. D., and Wolffe, A. P. (2000). DNA methylation in health and disease. Nat Rev Genet *1*, 11-19.

Rubinson, D. A., Dillon, C. P., Kwiatkowski, A. V., Sievers, C., Yang, L., Kopinja, J., Rooney, D. L., Ihrig, M. M., McManus, M. T., Gertler, F. B.*, et al.* (2003). A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. Nat Genet *33*, 401-406.

Sakatani, T., Kaneda, A., Iacobuzio-Donahue, C. A., Carter, M. G., de Boom Witzel, S., Okano, H., Ko, M. S., Ohlsson, R., Longo, D. L., and Feinberg, A. P. (2005). Loss of imprinting of Igf2 alters intestinal maturation and tumorigenesis in mice. Science *307*, 1976-1978.

Sandy, P., Ventura, A., and Jacks, T. (2005). Mammalian RNAi: a practical guide. Biotechniques *39*, 215-224.

Sanford, J. P., Clark, H. J., Chapman, V. M., and Rossant, J. (1987). Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. Genes Dev *1*, 1039-1046.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol *241*, 172-182.

Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.

Shin, T., Kraemer, D., Pryor, J., Liu, L., Rugila, J., Howe, L., Buck, S., Murphy, K., Lyons, L., and Westhusin, M. (2002). A cat cloned by nuclear transplantation. Nature *415*, 859.

Silva, J. M., Li, M. Z., Chang, K., Ge, W., Golding, M. C., Rickles, R. J., Siolas, D., Hu, G., Paddison, P. J., Schlabach, M. R.*, et al.* (2005). Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet *37*, 1281-1288.

Sims, M., and First, N. L. (1994). Production of calves by transfer of nuclei from cultured inner cell mass cells. Proc Natl Acad Sci U S A *91*, 6143-6147.

Solter, D. (2005). Politically correct human embryonic stem cells? N Engl J Med *353*, 2321-2323.

Song, E., Lee, S. K., Wang, J., Ince, N., Ouyang, N., Min, J., Chen, J., Shankar, P., and Lieberman, J. (2003). RNA interference targeting Fas protects mice from fulminant hepatitis. Nat Med *9*, 347-351.

Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H., and Schreiber, R. D. (1998). How cells respond to interferons. Annu Rev Biochem *67*, 227-264.

Strelchenko, N., Verlinsky, O., Kukharenko, V., and Verlinsky, Y. (2004). Morula-derived human embryonic stem cells. Reprod Biomed Online *9*, 623-629.

Strichman-Almashanu, L. Z., Lee, R. S., Onyango, P. O., Perlman, E., Flam, F., Frieman, M. B., and Feinberg, A. P. (2002). A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes. Genome Res *12*, 543-554.

Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Sun, X., Lewandoski, M., Meyers, E. N., Liu, Y.-H., Maxson, R. E., and Martin, G. R. (2000). Conditional inactivation of *Fgf4* reveals complexity of signalling during limb bud development. Nat Genet *25*, 83-86.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11*, 1553-1558.

Tanaka, M., Hadjantonakis, A. K., and Nagy, A. (2001). Aggregation chimeras. Combining ES cells, diploid and tetraploid embryos. Methods Mol Biol *158*, 135-154.

Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.

Taranger, C. K., Noer, A., Sorensen, A. L., Hakelien, A. M., Boquest, A. C., and Collas, P. (2005). Induction of dedifferentiation, genomewide transcriptional programming, and epigenetic reprogramming by extracts of carcinoma and embryonic stem cells. Mol Biol Cell *16*, 5719-5735.

Tesar, P. J. (2005). Derivation of germ-line-competent embryonic stem cell lines from preblastocyst mouse embryos. Proc Natl Acad Sci U S A *102*, 8239-8244.

Thomas, K. R., and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell *51*, 503-512.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tucker, K. L., Beard, C., Dausmann, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E., and Jaenisch, R. (1996). Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. Genes Dev *10*, 1008-1020.

Turker, M. S. (1999). The establishment and maintenance of DNA methylation patterns in mouse somatic cells. Semin Cancer Biol *9*, 329-337.

Van Dyke, T., and Jacks, T. (2002). Cancer modeling in the modern era: progress and challenges. Cell *108*, 135-144.

Ventura, A., Meissner, A., Dillon, C. P., McManus, M., Sharp, P. A., Van Parijs, L., Jaenisch, R., and Jacks, T. (2004). Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci U S A *101*, 10380-10385.

Wade, P. A. (2001). Methyl CpG-binding proteins and transcriptional repression. Bioessays *23*, 1131-1137.

Wade, P. A., Gegonne, A., Jones, P. L., Ballestar, E., Aubry, F., and Wolffe, A. P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat Genet *23*, 62-66.

Wagers, A. J., and Weissman, I. L. (2004). Plasticity of adult stem cells. Cell *116*, 639-648.

Wakayama, S., Ohta, H., Kishigami, S., Van Thuan, N., Hikichi, T., Mizutani, E., Miyake, M., and Wakayama, T. (2005). Establishment of male and female nuclear transfer embryonic stem cell lines from different mouse strains and tissues. Biol Reprod *72*, 932-936.

Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature *394*, 369-374.

Wakayama, T., Rodriguez, I., Perry, A. C., Yanagimachi, R., and Mombaerts, P. (1999). Mice cloned from embryonic stem cells. Proc Natl Acad Sci U S A *96*, p14984-14989.

Walsh, C. P., Chaillet, J. R., and Bestor, T. H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet *20*, 116-117.

Wang, Z., and Jaenisch, R. (2004). At most three ES cells contribute to the somatic lineages of chimeric mice and of mice produced by ES-tetraploid complementation. Dev Biol *275*, 192-201.

Wang, Z., Meissner, A. and Jaensich, R. (2004). Nuclear Cloning and Epigentic Reprogramming. In Handbook of Stem Cells, (ed. R. Lanza) Academic Press, USA.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P.*, et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.

Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L., and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet *37*, 853-862.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Willadsen, S. M. (1986). Nuclear transplantation in sheep embryos. Nature *320*, 63-65.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Woods, G. L., White, K. L., Vanderwall, D. K., Li, G. P., Aston, K. I., Bunch, T. D., Meerdo, L. N., and Pate, B. J. (2003). A mule cloned from fetal cells by nuclear transfer. Science *301*, 1063.

Yamada, Y., Jackson-Grusby, L., Linhart, H., Meissner, A., Eden, A., Lin, H., and Jaenisch, R. (2005). Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. Proc Natl Acad Sci U S A *102*, 13580-13585.

Yamada, Y., Watanabe, H., Miura, F., Soejima, H., Uchiyama, M., Iwasaka, T., Mukai, T., Sakaki, Y., and Ito, T. (2004). A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q. Genome Res *14*, 247-266.

Yokochi, T., and Robertson, K. D. (2002). Preferential methylation of unmethylated DNA by Mammalian de novo DNA methyltransferase Dnmt3a. J Biol Chem *277*, 11735-11745.

Zhang, Y., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev *13*, 1924-1935.

Zhou, Q., Renard, J. P., Le Friec, G., Brochard, V., Beaujean, N., Cherifi, Y., Fraichard, A., and Cozzi, J. (2003). Generation of fertile cloned rats by regulating oocyte activation. Science *302*, 1179.

Acknowledgements

I would like to thank my family and friends for their continous support and especially Marek Polewski for the cover design.

I am thankful to Konrad Hochedlinger and Laurie Jackson-Grusby for discussions and help throughout my PhD and for reading of manuscripts and this thesis. I am thankful to Caroline Beard, Chris Lengner and Grant Wilstead for comments on the thesis and in particular Heather van Epps for lots of editing.

I would like to thank Kevin Eggan, Jessie Dausman, Ruth Flannery and Gerry Kemske.

I also have to thank my main collaborators Andrea Ventura (Tyler Jacks, MIT CCR) and Andreas Gnirke (Eric Lander, Broad Institute) who contributed significantly to this work.

I am very thankful to my thesis advisers Jörn Walter and Uwe Walldorf.

Most importantly, I would like to thank Rudolf Jaenisch for giving me the opportunity to work in his lab and for letting me pursue all the different projects.

I would also like to thank my former mentors: Rabi Simantov, Claas Junghans, Hans von Doehren and Pamela Silver.

Finally, Boehringer Ingelheim Fonds for generous support of my PhD, in particular Hermann Fröhlich, Monika Beutelspacher and Claudia Walther.

Summary

The goal of the studies described here was to establish a set of methods aimed at ultimately enhancing the efficiency of epigenetic reprogramming. The use of RNA interference (RNAi) for studying and influencing gene function has become an essential part of biology. In this work, we have described two lentivirus-based vectors used for conditional, Cre-lox regulated RNAi in cells and in mice. One vector triggers Cre-dependent activation (pSico) and the other Cre-dependent termination (pSicoR) of shRNA expression. These vectors were used to conditionally and reversibly knock-down p53, Npm, and Dnmt1 expression in ES cells and in MEFs. As a proof of principle, pSico was used to generate conditional and tissue-specific knock-down mice. As outlined in Chapter 1 conditional depletion of various gene products by RNAi will provide better understanding of the factors involved in epigenetic reprogramming. This knowledge should ultimately lead to enhancing the efficiency of successful reprogramming. The pSicoR system was applied in later experiments to temporally suppress Cdx2 function in donor nuclei prior to nuclear transfer, a modification of the current procedure, termed altered nuclear transfer (ANT). Finally, deciphering the epigenome of different cell types is critical for understanding the regulation of both normal development and disease states. In the last part of this work we have devised a new strategy that permits high-resolution comparative DNA methylation analysis. The system was tested in ES cells depleted of DNA methylation by elimination of the DNA methyltransferases Dnmt1, Dnmt3a and 3b. Dnmt1 was knocked down in Dnmt3a and 3b double knockout ES cells using a pSicoR-Dnmt1 vector.

Zusammenfassung

Der Transfer eines differenzierten Zellkerns in eine entkernte Eizelle (Kerntransfer) ist einer von mehreren experimentellen Ansätzen um die Reprogrammierung von differenzierten Zellen zu erreichen. Unter Reprogrammierung vesteht man grundsätzlich die Erweiterung des Entwicklungspotentials einer differenzierten Zelle. Eines der Hauptziele von Kerntransfer-Experimenten ist es undifferenzierte Stamm- oder Vorläuferzellen hervorzubringen, welche für Zellersatztherapien genutzt werden können. Der Hauptvorteil von humanen Kerntransfer-Stammzellen liegt darin, dass sie patientenspezifisch sind und dadurch nicht vom Immunsystem als fremd erkannt würden.

Einer der Schwerpunkte unserer gegenwärtigen Forschung ist es, ein besseres Verständnis der Mechanismen und Faktoren, die in der Kernreprogrammierung involviert sind, zu gewinnen. Die Tatsache, dass vollständig differenzierte Zellen mittels Kerntransfer reprogrammiert werden können, zeigt, das im Verlauf der Entwicklung keine genetische Information verloren geht, d.h. differenzierte Zellkerne enthalten sämtliche Informationen um einen kompletten Organismus hervorzubringen. Diese Ergebnisse deuten darauf hin, dass die Regulierung der Differenzierung über epigenetische Mechanismen gesteuert wird. Epigenetische Modifikationen sind stabile Veränderungen der DNA oder des Chromatins, die jedoch die primäre DNA Sequenz nicht verändern. Das Ziel meiner Untersuchen ist es, Methoden und ein besseres Verständnis der involvierten Faktoren zu entwickeln, um den Vorgang der Reprogrammierung verbessern zu können.

Im ersten Teil meiner Arbeit beschreibe ich ein neues System zur Cre-Lox regulierbaren Gen Inhibierung durch RNA Interferenz. Die Effektivität und Funktionalität des Systems wurde für mehrere Gene *in vitro* und *in vivo* gezeigt. Neben vielen anderen nützlichen Anwendungen, wie konditionelle Regulierung von essentiellen Genen *in vivo*, was hier für das Tumorsuppressorgen p53 gezeigt wurde, erlaubt das System transiente Blockierung von Faktoren, die epigenetische Modifikationen regulieren, wie z.B. DNA Methyltransferase 1 (Dnmt1). Es konnte bereits gezeigt werden, dass eine Reduzierung der genomischen DNA Methylierung einen positiven

Einfluss auf die Effizienz des Reprogrammierens durch Kerntransfer hat. Allerdings ergeben sich aus der daraus bedingten Hypomethylierung der DNA auch negative Konsequenzen, wie z.B. vermehrtes Auftreten von Tumoren. Diese negativen Auswirkungen lassen sich durch zeitlich beschränkte Inhibierung des Enzyms vermindern, da nach dem Entfernen des RNAi Systems das endogene Gen wieder aktiv ist.

Im zweiten Teil beschreibe ich eine Modifikation der normalen Kerntransfer Technik. Dabei wird eine Gen, mittels des oben beschriebenen RNAi Systems blockiert, was unerlässlich ist für die Differenzierung in Trophectoderm, welches später die Plazenta formt. Dadurch wird kein funktioneller Embryo erzeugt, aber es lassen sich trotzdem Stammzellen gewinnen. Diese Experimente stellen eine wissenschaftliche Basis für die Diskussion über die Gewinnung von Stammzellen dar, und erlauben ausserdem weitere Analysen von essentiellen Faktoren die für die extraembryonalen Gewebe notwendig sind. Dies ist wichtig, da viele essentielle Gene im geklonten Embryo selbst, aber auch in seinen extraembryonalen Teilen, nicht korrekt reaktiviert werden.

Obwohl die DNA Sequenz zwischen Stammzellen und differenzierten Zellen identisch ist, sind sie epigentisch verschiedenen. Um diese Unterschiede im gesamten Genom besser untersuchen zu können, haben wir eine Methode entwickelt, die es erlaubt grosse Teile des Epigenoms zu analysieren und zwischen verschiedenen Zelltypen zu vergleichen.

Publications

Journals

Lin H, Yamada Y, Nguyen S, Jackson-Grusby L, **Meissner A**, Linhart H and Jaenisch R. *Suppression of intestinal neoplasia by deletion of Dnmt3b*. Mol Cell Biol. (2006) (in press)

Meissner A and Jaenisch R. *Generation of nuclear transfer-derived pluripotent ES cells from cloned Cdx2-deficient blastocysts*. Nature. 2006 Jan 12;439(7073):212-5. Epub 2005 Oct 16

Meissner A*, Gnirke A*, Ramsahoye B, Bell G, Lander ES, Jaenisch R. *Reduced Representation Bisulfite Sequencing for comparative high-resolution DNA methylation analysis.* Nucleic Acid Research. 2005 Oct 13;33(18):5868-77. Print 2005.

Yamada Y, Jackson-Grusby L, Linhart H, **Meissner A**, Eden A, Lin H, Jaenisch R. *Opposing effects of DNA hypomethylation and intestinal carcinogenesis.* Proc Natl Acad Sci U S A. 2005 Sep 20;102(38):13580-5. Epub 2005 Sep 8.

Ventura A*, **Meissner A***, Dillon CP, McManus M, Sharp PA, Van Parijs L, Jaenisch R, Jacks T. *Cre-lox-regulated conditional RNA interference from transgenes.* Proc Natl Acad Sci U S A. 2004 Jul 13;101(28):10380-5.

Gaudet F, Rideout WM 3rd, **Meissner A**, Dausman J, Leonhardt H, Jaenisch R. *Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing.* Mol Cell Biol. 2004 Feb;24(4):1640-8.

Chen WG, Chang Q, Lin Y, **Meissner A**, West AE, Griffith EC, Jaenisch R, Greenberg ME. *Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2.* Science. 2003 Oct 31;302(5646):885-9.

Submitted

Blelloch R, Wang Z, **Meissner A**, Pollard S, Smith AG, and Jaenisch R, *The Epigenome Influences Efficiency of Reprogramming by Nuclear Transfer*.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Hubert DJ, Cuff J, Fry B, **Meissner A**, Werning M, Jaensich R, Wagschal A, Feil R, Schreiber SL, and Lander ES. *Epigenetic landscape in embryonic stem cells*.

Reviews and Book chapters

Meissner A and Jaenisch R. *From frog cloning to customized human embryonic stem Cells*. BIF Futura (2006). In press.

Wang Z*, **Meissner A***, Jaenisch R. *Nuclear Cloning and Epigenetic Reprogramming*. Essentials of Stem Cell Biology, Elsevier 2005

Wang Z*, **Meissner A***, Jaenisch R. *Nuclear Cloning and Epigenetic Reprogramming*. Handbook of Stem Cells, Elsevier 2004

Curriculum Vitae

Personal Details

Dipl-Ing. Alexander Meissner, born on August 31, 1975 in Berlin Germany.

Nine Cambridge Center, Cambridge, MA 02142, USA

Phone: 001-617-6696898 Email: meissner@wi.mit.edu

Education

2002-current PhD Biology, Whitehead Institute/ Massachusetts Institute of Technology

1996-2002 Diplom Engineer Biotechnology, Technological University of Berlin (GPA: 1.4 very good)

1989-1995 Walther-Rathenau-Gymnasium, Abitur

Additional Scientific Training

2001 Harvard Medical School, Boston
Identification of topoisomerase II targets by ChIP on Chip

2000 Mologen AG, Berlin
Chromatography, validation and purification of genetherapy vectors

1999 Cancer Research Center Hawaii, Honolulu
PCR, cloning, protein purification for pRB, cyclin D1 and Cdk4

1999 Weizmann Institute of Science, Rehovot
Immunocytochemistry, tissue culture, investigating the neurotoxicity of MDMA

1998 Connex GmbH, Munich
Immunological work, ELISA, lateral flow test development (H. Pylori)

1998 MetaGen GmbH (Schering AG), Berlin
Tumor genome analysis, PCR, DNA-sequencing

1997 Löwenbräu AG, Munich
Basic training in microbiology

Awards

2006 Keystone Symposia Scholarship

2003 Boehringer Ingelheim PhD fellowship

1999 Weizmann Institute of Science, Karyn Kupcinet Summer fellowship

Miscellaneous

2006 Poster presenter, Keystone Symposium: Stem cells, Canada

2005 Invited speaker, Keystone Symposium: Stem cells, Senescence and Cancer, Singapore

2005 Invited discussant, BIF International Titisee Conference: Stem cells

2005 Cold Spring Habor Laboratory, "Science-Get it across" workshop

2005 Poster presenter, Gordon Conference Epigenetics, USA

2004 Invited discussant, BIF International Titisee Conference: RNA interference

2004 Poster presenter, Keystone Symposium: Epigenetics, USA

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Ort, Datum

(UNTERSCHRIFT)