An Inducible RNAi system for the functional dissection of genes in the mouse

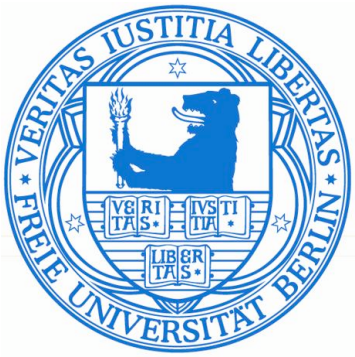
Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy

of the Freie Universität Berlin



by

Joana Alves Vidigal

from Lisbon, Portugal

December 2010

The work here described was performed at the

Max-Planck Institute for Molecular Genetics in Berlin

between February 2007 and December 2010

under the supervision of Prof. Dr. Bernhard G. Herrmann and Dr. Markus Morkel

1st Reviewer:	Prof. Dr. Bernhard G. Herrmann Max-Planck Institute for Molecular Genetics Ihnestrasse 73 D-14195 Berlin
2nd Reviewer:	Prof. Dr. Stephan Sigrist Institute of Biology - Freie Universität Berlin Takustr. 6, 14195 Berlin
Date of defence:	11/02/2011

Aos meus pais

e

aos meus irmãos

Table of Contents

1 INTRODUCTION

1.1 An introduction to the mouse as a model system

In the early 1900s, the rediscovery of Mendel's laws, and the question of whether they applied to organisms other than plants marked the beginning of the field of vertebrate genetics and the establishment of the mouse as a model system. For these early geneticists, the mouse was a valuable companion. It was small and its husbandry relatively facile and inexpensive. It had a reasonably short breeding time (around 3 weeks) and each litter could yield as many as 10 to 15 offspring. This meant that breeding schemes could be fairly easily set up to determine the inheritance pattern of phenotypic traits (Paigen, 2003a).

Importantly, the mouse shared with humans its basic biology, from gross anatomy and physiology to the spectrum of diseases that it succumbed to. Soon, mouse genetic research was guided towards the understanding of human pathologies, and in particular cancer (Tyzzer, 1909). This led to the development of a very sophisticated genetic system with the establishment of hundreds of inbred strains (Staats, 1980), derivation of new mutants, drawing of a dense genetic map (Lyon, 1990), and development of gene-mapping resources such as recombinant inbred strains (Bailey, 1971). The research tools developed in these early years helped shape the field of genetics and established the mouse as one of its most important model systems. After eight decades of classical genetic research, the first transgenic mice were created. Together with the relatively recent sequencing of the mouse genome (Waterston et al., 2002), transgenic techniques have made it possible to query virtually any genomic sequence, and to understand its role in mammalian biology - both in its normal and pathological state.

Because the development of transgenic mouse models is the main focus of the work described in this thesis, I will next give a short historical perspective and overview of the techniques and tools currently available.

1.2 The molecular revolution of mouse genetics

In the early 1980s, the emerging field of molecular biology came together with that of mouse genetics when five independent groups reported the generation of adult mice carrying foreign genes (Brinster et al., 1981; Costantini and Lacy, 1981; Gordon et al., 1980; Harbers et al., 1981; Wagner et al., 1981). These first transgenic mice were derived by direct injection of DNA into the male pronucleus of a fertilized mouse oocyte and showed that the DNA not only integrated into the genome but could also be transmitted to every cell of the developing mouse - including those in the germline – and could therefore be passed to its progeny. Most importantly, Wagner and collegues showed that, when introducing the intact sequence coding for the rabbit β-globin gene, this gene was expressed in the adult mouse in the correct tissue, i.e. the erythrocytes. This finding was significant for two reasons: first, it meant that a gene could be properly expressed away from its usual chromosomal environment, and secondly, it meant that the regulatory systems ensuring proper gene expression throughout development and adult life were basically conserved between rabbit and mouse, raising the fascinating possibility that regulatory conservation extended also to other animal orders including *Primata*. If that was true, then human gene regulation, and perhaps function, could be dissected in an organism as amenable to experimental research as the mouse. However, for the true molecular revolution of mouse genetics to begin, two major obstacles still had to be overcome. The first was determining a way to specifically modify the mouse genome and the second was how to bring this modification to the germ line.

1.3 The “knockout mouse”

In 2007, the Nobel Prize in Physiology or Medicine was awarded to Mario Capecchi, Martin Evans and Oliver Smithies for solving the problem of bringing specific genetic modifications into the mouse. The technique, commonly known as gene targeting, was seminal in shaping today’s knowledge about gene function and regulation, and is now standardly used in molecular genetics laboratories.

1.3.1 Isolation of Embryonic Stem Cells

Pluripotent cells are cells that are able to differentiate into all derivatives of the three germ layers, i.e., ectoderm, mesoderm and endoderm. By 1981, when the first transgenic mice were derived, two sources of pluripotent cells were well documented: the teratomas (Martin and Evans, 1974) and the inner cell mass (ICM) of a blastocyst (Lovell-Badge and Evans, 1980). While the isolation of cell lines from teratomas – the embryonal carcinoma or EC cells – proved to be fairly easy (Evans, 1972), isolation of pluripotent cell lines from a blastocyst (thought to be the normal counterpart of EC cells) proved to be slightly more demanding. This was mainly because the isolated cells would differentiate upon cultivation under standard conditions. In July 1981 Martin Evans, followed in December of that year by Gail Martin, reported the isolation of such cells and the term embryonic stem (ES) cell was coined (Evans and Kaufman, 1981; Martin, 1981). ES cells passed every pluripotency test: they formed teratomas *in vivo*, differentiated *in vitro*, expressed the expected surface antigen markers, and were able to contribute to the germline of a chimeric mouse when injected into the cavity of a developing blastocyst (Bradley et al., 1984; Evans, 2001). The success of these early experiments led Evans to propose the use of ES cells for gene modification studies. Indeed, in 1986, his group showed in another landmark paper, that ES cells infected with a recombinant retrovirus gave rise to chimeric mice that could transmit the viral DNA to their offspring, proving that germline transmission of foreign DNA

through modified ES cells was possible (Robertson et al., 1986). In the next step, the Evans laboratory introduced a mutated copy of an endogenous gene into the genome. The gene coding for a mutated version of hypoxanthine phosphoribosyltransferase (HPRT), often disrupted in the Lesch-Nyhan syndrome, was introduced by retroviral infection into cultured cells and was shown to be transmitted to the offspring of the resultant chimeric mice (Kuehn et al., 1987). For the first time, a mouse model for a human disease had been created by modification of ES cells. Although these experiments performed by Martin Evans and colleagues represent an enormous technological breakthrough, their applicability was somewhat limited in that the foreign DNA integrated randomly and often at multiple loci in the genome. This lack of precision made the targeted mutation of endogenous genes impossible. This problem was addressed by Mario Capecchi and Oliver Smithies, who independently discovered that the mammalian genome could be specifically modified by co-opting the cellular recombination machinery.

1.3.2 Genetic engineering by homologous recombination.

Recombination between homologous DNA molecules had been known to occur in bacteria since the mid-1900's, and in the 70s a similar process was reported to occur in eukaryotic cells between homologous chromosomes during meiosis (Paigen, 2003b). However, only in 1982 homologous recombination was confirmed to occur in somatic eukaryotic cells when Mario Capecchi observed that DNA microinjected into mammalian cells would integrate in concatemerized patterns (Folger et al., 1982). Despite these and other observations showing that eukaryotic cells could mediate homologous recombination, the harnessing of the enzymatic machinery involved in this process for the purpose of specifically modifying an endogenous genomic sequence was largely dismissed. This was due to the fact that a recombination event requiring foreign DNA to find its matching sequence within a host genome was viewed as extremely unlikely (Capecchi, 2001).

Despite these early disbeliefs, by 1986, both Capecchi's and Smithies' laboratories reported the successful targeting of the mouse genome (Smithies et al., 1985; Thomas et al., 1986), and Capecchi's group further demonstrated that such recombination events occurred at relatively high frequencies. Over the next several years, both groups, in collaboration with Martin Evans, showed that ES cells could be targeted by homologous recombination without losing their pluripotency (Doetschman et al., 1987; Thomas and Capecchi, 1987), and strategies for the selection of homologous recombinants were developed (Gossler et al., 1986; Mansour et al., 1988). The race to produce the first genetically engineered mouse was finally realized in 1989 with the generation of 'knockout' mouse lines (so called because they carried a disrupted gene) by several independent laboratories (Koller et al., 1989; Thompson et al., 1989; Zijlstra et al., 1989). A generalized overview of a gene targeting approach can be seen below (Figure 1.1).

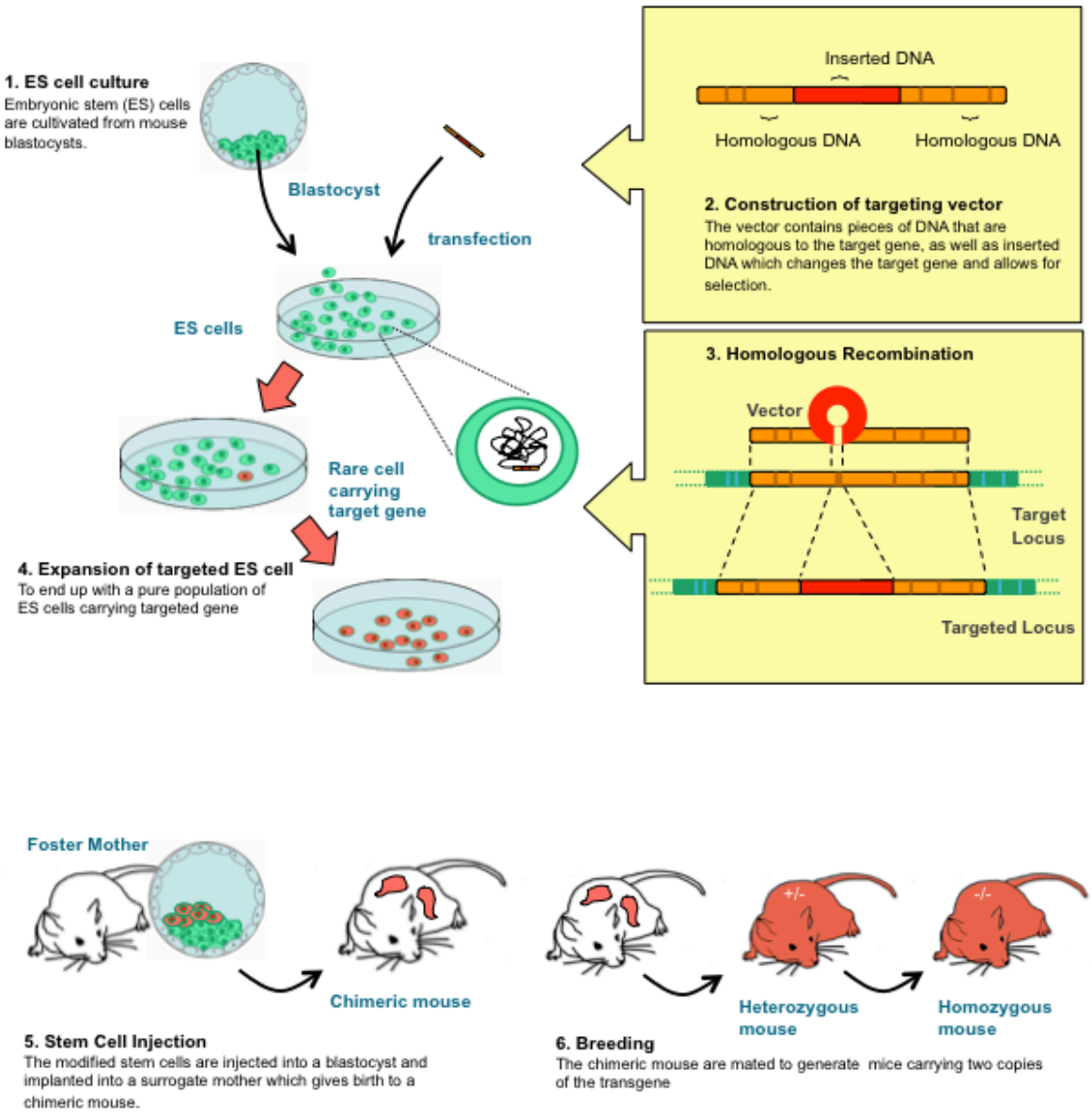


Figure 1.1. Gene Targeting by Homologous Recombination

(1) Embryonic stem cells, derived from the inner cell mass of a blastocyst-stage mouse embryo can be cultivated *in vitro*. The modification of these cells by gene targeting techniques requires the construction of a targeting vector that contains arms with homology to the target locus, flanking a DNA sequence that changes the locus and allows for selection (2). Upon transfection of the ES cells with the targeting vector, some cells will undergo recombination through the homologous sequences (3) and can be selected and expanded *in vitro* (4). To bring the targeted mutation into the genome of a living mouse, the modified ES cells can be injected into the cavity of a wild-type mouse blastocyst. This chimeric blastocyst can be transferred into the uterus of a foster mother where it will develop into a chimeric mouse carrying targeted cells in some of its tissues (5). If the targeted mutation is present in the germline, breeding of the chimeric mouse will result in heterozygous offspring carrying one mutated copy in all cells. Further breeding will give rise to homozygous mice, in which both copies of the gene are targeted.

1.4 Further developments in transgenic techniques

Since the initial establishment of gene targeting techniques, several modifications have been reported that have broadened the scope of its application. Described below are some of the most prominent developments.

1.4.1 Site-specific recombination systems

An ingenious improvement to the traditional "knockout" strategy came from exploiting the Cre-*lox* recombination system from the P1 bacteriophage (Sternberg et al., 1981). This system comprises two components: the Cre recombinase enzyme and the DNA sequences it recognizes, the *loxP* sites. The simplicity of this system made it a very attractive tool in the field of molecular genetics.

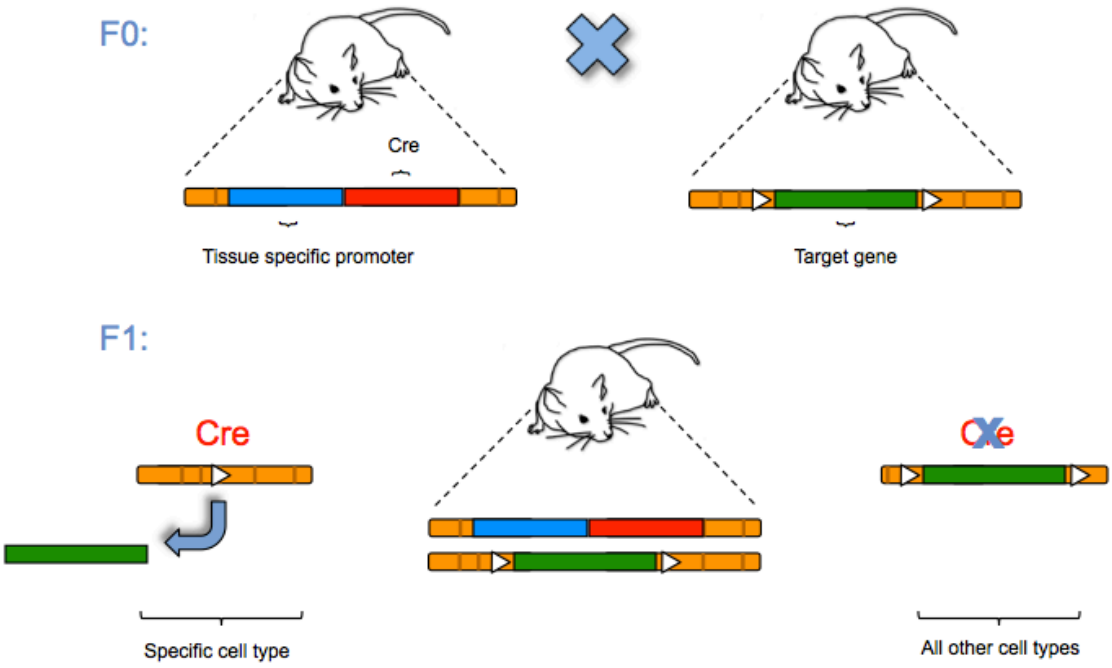


Figure 1.2. Cre-*lox* recombination system in the mouse

Site-specific recombination systems can be used to delete genomic sequences in a subset of cell types. When a mouse expressing Cre recombinase under the control of a tissue-specific promoter is crossed with a mouse containing *loxP* sites flanking a gene, the resulting progeny will carry copies of both the recombinase and the 'floxed' allele. In the tissues where Cre is expressed, recombination will lead to excision of the sequence flanked by the *lox* sites, while in all other cells the floxed allele will remain unchanged. White triangles represent *loxP* sites.

A major advantage of the Cre/*lox* system over traditional knockout strategies is that it can be used to disrupt a gene in a specific set of cells or in a temporally controlled manner. For this purpose, *loxP* sites are introduced in genome in the same orientation, flanking either whole genes or critical exons. Because *loxP* sites are only 34 base pairs (bps) long, such integration generally causes no detrimental effects to gene expression. When mice carrying these sequences are crossed with mice expressing the Cre recombinase, the progeny will undergo a genomic recombination wherever Cre is expressed, leading to the excision of the flanked DNA segment (Nagy, 2000; Sauer and Henderson, 1988). A similar site-specific recombination system relies on the recognition of FRT sites by the Flp recombinase (Branda and Dymecki, 2004). Both Cre and Flp can be expressed under tissue-specific promoters leading to recombination in only a subset of cells as depicted in Figure 1.2 (Gu et al., 1994). In addition these enzymes can be expressed under inducible promoters, allowing not only for spatial but also temporal control over the deletion (Kuhn et al., 1995).

1.4.2 Inducible expression systems

Although homologous recombination is an extremely powerful tool, its application is hampered when the disruption of a gene leads to embryonic lethality, since this precludes the study of its functions in later developmental stages or adulthood. In addition, overexpression of a transgene such as Cre can occasionally have unwanted physiological or toxic effects (Schmidt-Supprian and Rajewsky, 2007). The development of inducible expression systems has provided means to overcome these issues. Several inducible expression systems have been described. Tetracycline (Kuhn et al., 1995), type I-Interferon (Whyatt and Rathjen, 2001) and tamoxifen (Zhang et al., 1996) have all been used to produce drug-inducible systems. Here however, I will turn my attention to the first system described, and the one most pertinent to the work presented here, the tetracycline system (Figure 1.3).

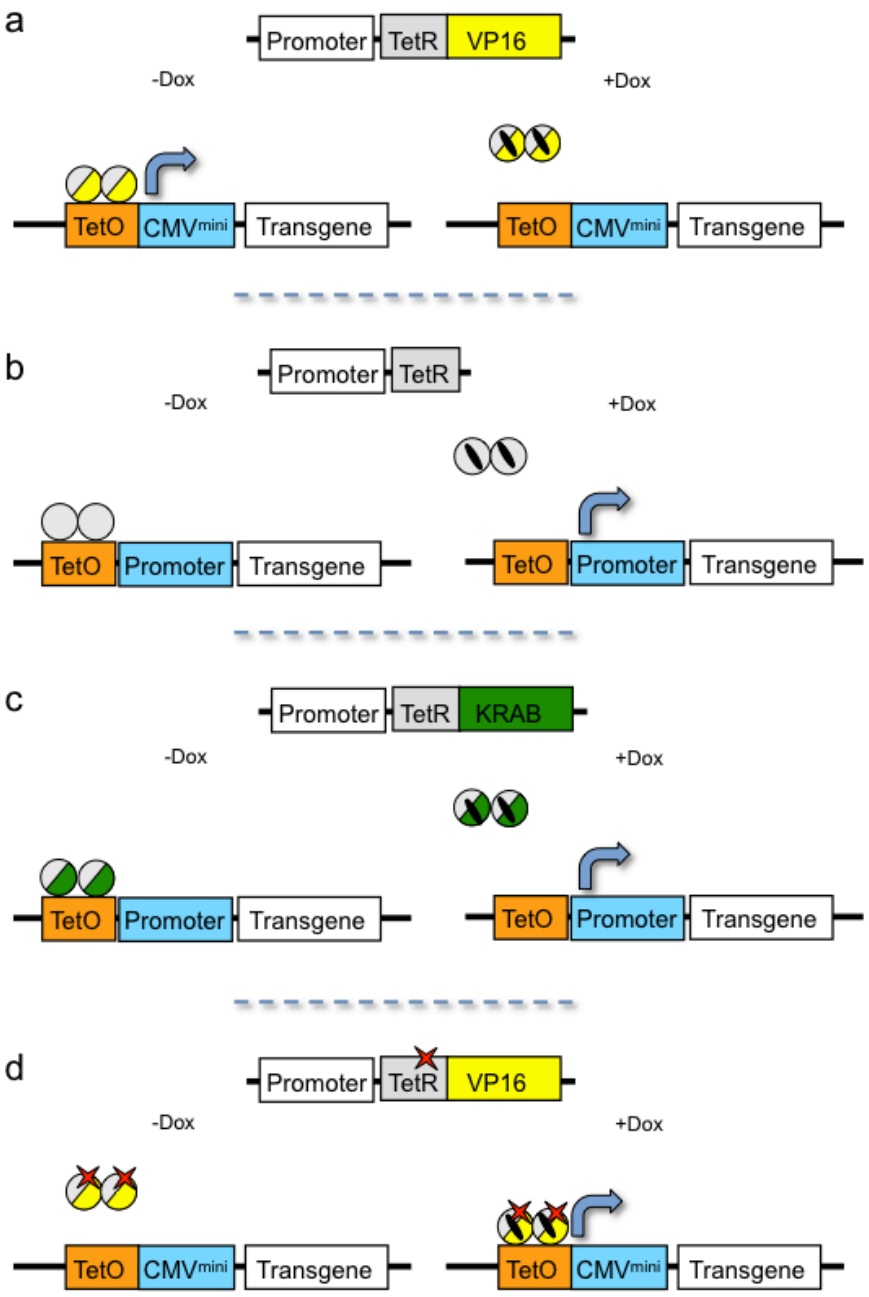


Figure 1.3. The tetracycline (Tet) system for mammalian transgenesis

(a) In the "Tet-off" system, the tTA – a fusion protein between the tet repressor (tetR) DNA binding domain and the transactivation domain of VP16 – binds to the tetracycline operator sites (TetO) and drives expression of a transgene only in the absence of doxycycline (dox). In the "Tet-on" system (b-d) transcription occurs only in the presence of dox. In the absence of dox, both the tetR (b) and the tTS (c) – a fusion protein between the tetR DNA binding sites and KRAB transcriptional repressor domain – bind to the TetO sites and prevent trancription. The rtTA on the other hand, a mutated version of tTA, binds only in the presence of dox, leading to transcriptional activation through the TetO sites (further detais can be found in the text).

Tet expression systems are modified from the Tc resistance operon from *Escherichia coli* and rely on the activity of a transcriptional activator or repressor that can be regulated by the presence of tetracycline (Tc) or of a tetracycline derivative such as doxycycline (dox) (Zhu et al., 2002). These systems exist in two complementary variants, “Tet-on” and “Tet-off”. In the first, expression of the transgene occurs in the presence of Tc or dox; in the second, transcription is only active in the absence of these compounds.

The “Tet-off” system depends on the activity of a tetracycline-controlled transactivator (tTA). tTA is a fusion protein between the Tet-repressor DNA binding domain (tetR) from the Tn*10* operon and the strong transactivating domain of VP16 from the Herpes simplex virus (Gossen and Bujard, 1992). In the absence of Tc or dox, tTA binds to the Tet responsive element (TRE) and drives expression of a transgene under its control. The TRE is composed of tet operator (tetO) concatemers followed by a minimal promoter. In the presence of Tc or dox, tTA cannot bind and the transgene remains silent.

The “Tet-on” system allows the usage of three different transcriptional regulators. The simplest of them is the tetR (Yao et al., 1998), the transcriptional repressor encoded in the original Tn*10* operon. In the absence of Tc or dox, tetR binds to tetO sequences and leads to repression of transcription from a promoter placed near the operator by preventing endogenous transcription factors from binding. This system has the inherent problem of leakiness (Meyer-Ficca et al., 2004). A second regulator is the tetracycline-controlled transcriptional silencer (tTS), which is a fusion protein between the tetR and the human Krueppel-associated Box (KRAB) transcriptional repressor (Freundlieb et al., 1999). In the absence of Tc, tTS binds to the tetO sequences and represses transcription from a nearby promoter by recruiting histone deacetylases and methyltransferases and inducing the formation of heterochromatin. In the presence of Tc, neither tetR nor tTS can bind tetO and transcription from nearby promoters is active. Finally, a “Tet-on” system can elicited by activity of a modified version of the tTA, the reverse-tTA or rtTA. rtTA differs from the tTA by only 4 amino acids in the DNA binding domain

of tetR (Gossen et al., 1995), but these changes are sufficient to alter the DNA binding characteristics of the protein so that the transactivator can only bind to the tetO sequences in the *presence* of Tc or dox.

Inducible expression systems such as these have been used extensively to temporally control the expression of transgenes integrated into the mouse genome. One of their most important applications is the control over the onset of expression of Cre recombinase in Cre-*lox* mice. Since in such settings the timing of recombination between *loxP* sites is controlled, embryonic lethality can be avoided thus allowing the study of later function of a particular gene. Because they allow for temporally and for spatially controlled gene disruption, animal models carrying Cre-*lox*/Tet or analogous systems are commonly called 'conditional knockouts'.

1.4.3 RNAi: an alternative to traditional gene targeting

The recent discovery that short double-stranded RNA (dsRNA) can trigger the silencing of genes in mammalian cells by a mechanism called RNA interference (RNAi) resulted in new advances in the field of functional genomics. As a tool to generate loss-of-function phenotypes, RNAi has several attractive features. First, genes can be conveniently silenced in mammalian cell lines by transient transfection of RNA triggers designed against a specific sequence. Such experiments are currently standardly used in cell culture and frequently opted for over traditional gene disruption because they produce similar results for a fraction of the time and cost. Secondly, unlike traditional or conditional 'knockouts', gene silencing by RNAi does not modify the targeted locus. This means that when coupled with inducible expression systems such as the ones described above, RNAi allows one to turn an endogenous gene on and off at will, which cannot be done using 'knockout' models. Finally, because the abundance of the target mRNA depends on the efficiency and abundance of the RNAi trigger, RNAi allows not only recapitulation of complete loss-of-function phenotypes, but also the study of the biological functions of proteins at varying physiological concentrations. Such studies can be extremely

important and informative in situations where different levels of a protein lead to distinct outcomes, as it is often the case in fields like developmental biology.

Because a major part of the work described in this thesis deals with the development of RNAi models to study mammalian gene function, in the next Chapters I will summarize what is currently known in this field. In particular, I will describe the endogenous triggers of gene silencing, their mechanism-of-action and how they can be exploited for experimental purposes. Finally, I will give an overview of the available systems to silence genes via RNAi with a focus on their merits and limitations.

1.5 RNA interference and its triggers

Few fields of biology have changed in the past decade as much as that of RNA. This rapid development was mainly driven by the discovery of small noncoding RNAs of around 20-30 nucleotides in length, which were shown to regulate gene expression (Elbashir et al., 2001b; Fire et al., 1998; Montgomery et al., 1998; Tuschl et al., 1999). The significance of this discovery led to the award of the 2006 Nobel Prize in Physiology or Medicine, to Andrea Fire and Craig Mello. Gene regulation by small RNAs occurs at many levels including, but not limited to, RNA stability, RNA translation, and regulation of chromatin structure. Although the mechanisms employed are diverse, they mainly have a negative effect on gene expression and so are collectively summed up under the *RNA silencing* or *RNA interference* banners. Two other themes are common to all the RNA silencing mechanisms. The first is that in all cases the noncoding RNA serves as a specificity molecule, guiding the effector machinery to the target nucleic acid by Watson-Crick base pairing. The second is that the central protein of the silencing machinery is always a member of the Argonaute superfamily.

Several classes of small RNAs involved in RNA silencing have been described, and it is certainly likely that many more will emerge in the near future. Although the borders between them become fuzzier with each new member added, small RNAs can be coarsely grouped in three main categories: small interfering RNAs (siRNAs), microRNAs (miRNAs) and Piwi interacting RNAs (piRNAs). Among the three classes, siRNAs and miRNAs have the most in common. They are both processed from double stranded precursors, while piRNAs are probably processed from single stranded precursors, though not much is known about their biogenesis. Both si/miRNAs act in somatic as well as germ line tissues and their action has been obseved in a wide variety of eukaryotic organisms. piRNAs, on the other hand, were found only in animals and their action is restricted to the germ line. Finally, si/miRNAs associate with members of the Argonaute clade of the Argonaute superfamily while piRNAs associate – as their name suggests – with members of the Piwi clade.

The differences between siRNAs and miRNAs are not easily delineated. Most of the processing steps are common to both molecules, although functional specificity for the proteins involved has been reported in some organisms (Lee et al., 2004b; Liu et al., 2004). Additionally, although in principle they silence genes through distinct mechanisms, some exceptions have been reported. Still, miRNAs are generally considered to be expressed from endogenous genes and to be involved in the fine-tuning of physiological processes by controlling the mRNA levels of endogenous genes through imperfect base-pair binding to their 3'UTR. siRNAs, on the other hand, are commonly seen as a product of exogenous triggers of RNA silencing (such as viruses, transposons or transgenes) and act as protectors of the genome against these foreign nucleic acids, leading to their cleavage upon perfect base-pair binding.

1.6 MicroRNAs

A true understanding of RNAi and how to use it in gene silencing is not possible without an understanding of the biogenesis and function of miRNAs. miRNAs are non-coding RNAs expressed by all metazoan eukaryotes. In higher eukaryotes, miRNAs account for 1% of the genome, with over 200 members per specie (Bartel, 2004). In humans, they have been estimated to control the expression of up to one third of the genes (Cullen, 2004). miRNAs have been implicated in a multitude of processes, from cell division (Hatfield et al., 2005), to differentiation (Makeyev et al., 2007) and apoptosis (Jovanovic and Hengartner, 2006) and their aberrant expression has been shown to be relevant in the context of diseases such as cancer (He et al., 2005; Lu et al., 2005; O'Donnell et al., 2005).

1.6.1 MicroRNA structure and expression

MicroRNAs were initially thought to be transcribed by the RNA polymerase III (pol III) since this is the polymerase that usually transcribes small RNAs such as ribosomal and transfer RNAs. Although this was found to be true for some miRNAs, the large majority is thought to be transcribed by via RNA polymerase II (pol II) (Lee et al., 2004a). This theory is supported by several lines of evidence. First, miRNAs are not transcribed as short RNAs, but rather as long primary transcripts or pri-miRNAs, often several kilobases long, that must go through a series of processing steps until they reach their mature form (Lee et al., 2002). These transcripts, like those of protein-coding genes, have both a 7-methylguanosine cap and a polyadenylation (polyA) signal (Cai et al., 2004). Secondly, several miRNAs have been described containing stretches of uracils (Lee et al., 2002), which would lead to termination of transcription by pol III (Bogenhagen and Brown, 1981; Cozzarelli et al., 1983). Thirdly, transcription of miRNAs has been shown to be sensitive to α-amanitin at concentrations that specifically inhibit pol II but do not disturb transcription by pol I or III (Lee et al., 2004a). Finally, chromatin immunoprecipitation

experiments have shown direct association of the pol II to the promoter of mir-23a2~27a~24-2 (Lee et al., 2004a).

Transcription by the pol II machinery means that miRNA expression can be controlled by pol II-associated transcription factors, allowing elaborate control of their expression both temporally and spatially. Indeed, expression profiles of miRNAs have shown tight regulation during developmental processes as well as tissue-specific expression, analogous to other genes under pol II control.

1.6.2 miRNA Maturation

MicroRNAs are transcribed as long precursor single-stranded RNAs (ssRNAs), the pri-miRNAs. These primary transcripts are usually several kbs long and have, at least in metazoans, typically a hairpin-like structure with a stem of 33 bp (Lee et al., 2002). The first step in the maturation of a miRNA occurs in the nucleus where pri-miRNAs are trimmed into shorter precursors, the pre-miRNAs (Figure 1.4). In mice, this step is catalyzed by Drosha, a class 2 RNase III enzyme together with its cofactor, DiGeorge syndrome critical region 8 (Dgcr8) (Lee et al., 2003). Together, Drosha and Dgcr8 form a large complex of around 65 kilodaltons (kDa) called the Microprocessor (Filippov et al., 2000; Gregory et al., 2004; Wu et al., 2000) (Fortin et al., 2002). Interestingly, although microRNAs seem to be widespread among eukaryotes, both Drosha and Dgcr8 are conserved only in animals (Fortin et al., 2002), pointing to a divergence of the early steps of miRNA biogenesis about 800 million years ago (Douzery et al., 2004).

Substrate recognition occurs at the level of Dgcr8. This protein binds the pri-miRNA near the base of the hairpin stem, in a region where the transcript is double-stranded as well as in another where it is single-stranded (Han et al., 2006). Upon binding to the Dgcr8/pri-miRNA complex, Drosha cleaves the RNA approximately 11 bp away from the ssRNA-dsRNA junction, producing a stem-loop molecule of around 69 bp in length with a 2 nt long 3' overhang, the pre-miRNA (Zeng and Cullen, 2005). This is the first step in the processing of a miRNA and defines

one of the ends of the final product. The second processing step occurs in the cytoplasm by Dicer and defines the second end of the microRNA.

Translocation of the pre-miRNA from the nucleus to the cytoplasm is mediated by Exportin-5, a RanGTP-dependent dsRNA-binding nuclear export receptor that specifically binds to pre-miRNAs in a sequence independent manner (Lund et al., 2004). Once in the cytoplasm, the pre-miRNA is bound by Dicer. The PAZ domain of Dicer binds to the pre-miRNA and defines the distance from the initial cut by Drosha (see Figure 1.4). Cleavage is then mediated by the two RNase III domains of Dicer leading to the production of a ~22 bp miRNA duplex with 2 nt 3' overhangs at both ends (Ketting et al., 2001; Knight and Bass, 2001). Because Dicer functions as a molecular ruler, cleaving the pre-miRNA at a predefined distance from the initial cut by Drosha, the nuclear processing of the miRNAs determines not only the 5' end of the mature miRNA but also its final sequence (Figure 1.4). After cleavage, Dicer remains bound to the miRNA duplex and the duplex is unwound by an as yet unidentified RNA helicase (Khvorova et al., 2003). Once unraveled, one of the strands, the mature miRNA, is incorporated into an RNA-induced or miRNA-induced silencing complex (RISC or miRISC). Strand selection seems to be accomplished on the basis of thermodynamic stability, whereby the strand with the most unstable base pair at the 5' end (e.g. G:U versus G:C) is incorporated into the complex, while the other is degraded (Gregory et al., 2005).

1.6.3 Gene silencing by miRNAs

RISC is a multiprotein complex and several of its components have already been identified, although very few have been characterized at the molecular level. The best studied of these components are the Dicer and the Argonaute (Ago) proteins, though only Argonaute proteins have been consistently found in all RISC complexes (Sontheimer, 2005). The mouse genome encodes four Argonaute proteins, (Ago1-4), and all of them have been shown to associate with miRNAs and lead to miRNA-mediated gene silencing (Su et al., 2009). Argonaute

proteins have four domains: the amino-terminal domain, PAZ, MID and PIWI. Binding of Argonaute proteins to miRNAs is mediated via two domains: the 5'end of the miRNA attaches to a binding pocket in the MID domain while the 3' nucleotides are anchored to a pocket of the PAZ domain (Wang et al., 2007). Upon binding, the miRNA guides the Ago protein to its target mRNA. With a few exceptions, animal miRNAs imperfectly base-pair to the 3'UTR of the messenger RNA (mRNA) and lead to inhibition of protein accumulation through translational inhibition or by destabilization of the mRNA through deadenylation and decapping (Figure 1.4) (Eulalio et al., 2008).

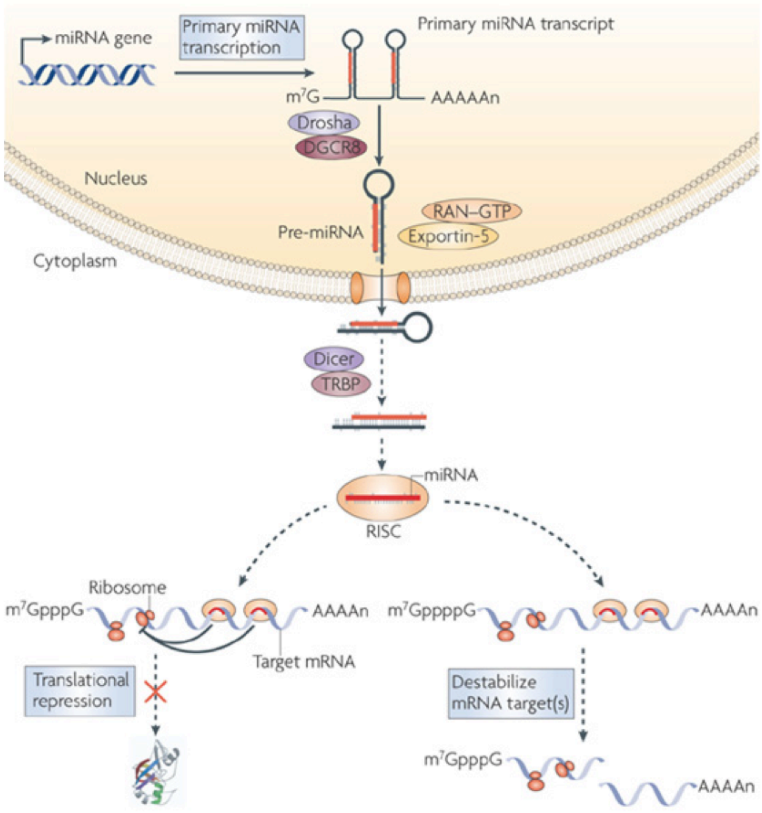


Figure 1.4. microRNA biogenesis and function in mammalian cells

microRNA genes are transcribed as long polyadenylated and capped primary transcripts. They are trimmed in the nucleus by the microprocessor complex, comprised of Drosha and Dgcr8, to a hairpin-shaped RNA called pre-miRNA. The pre-miRNA is exported to the cytoplasm with the help of Exportin-5, where it is further processed by Dicer into a microRNA duplex containing the miRNA (in red) and a passenger strand (in blue). Once incorporated into the RISC complex, the passenger strand is discarded and the mature miRNA guides the silencing complex to the target mRNA. Binding of the active RISC complex to the mRNA leads to gene silencing either by translational repression or by destabilization of the mRNA through deadenylation and decapping. For simplicity not all components of the miRNA biogenesis pathway are shown figure from (Lodish et al., 2008).

1.7 siRNAs

Small interfering RNAs are small single-stranded RNAs usually produced as a cellular response to exogenous dsRNAs. They were first discovered in plants, and later in worms, where introduction of long double stranded transcripts led to posttranscriptional silencing of endogenous genes with sequence complementarily to the double stranded trigger (Fire et al., 1998; Hamilton and Baulcombe, 1999). Silencing was shown to be accompanied by the generation of short (around 22 nt) RNA duplexes with 2 nt long 3' overhangs produced from the processing of the longer double stranded transcript by Dicer (Elbashir et al., 2001b; Hamilton and Baulcombe, 1999). Like in the miRNA silencing pathway, knockdown of genes was shown to be mediated by an Argonaute-containing multiprotein complex, the siRISC, that depends on the siRNA to find and silence its target mRNA.

Even though siRNAs and miRNAs are very similar in terms of structure and of the enzymes involved in their pathways, these molecules silence genes through distinct mechanisms.

1.7.1 Gene silencing by siRNAs

All four mammalian Argonaute proteins have been show to associate with siRNAs, but only Ago2 can lead to siRNA-mediated gene silencing. Ago2 is unique in that it is the only protein from its family that has catalytic activity and this was found to be critical for gene knockdown by siRNAs (Liu et al., 2004). This means that, although siRNAs can, in principle, be loaded into RISC complexes containing any of the four Argonaute family members, only when loaded into a Ago2-containing RISC will the siRNA produce an active complex.

In the siRNA pathway, unlike what happens during miRNA biogenesis, both strands of the siRNA duplex are loaded into the siRISC. Once in the complex, one of the strands becomes the first substrate of Ago2 (Rand et al., 2005). As in the miRNA pathway, strand selection seems to

depend on relative thermodynamic stabilities. The cleavage and dissociation of one of the strands produces a functional RISC able to silence its cognate target upon binding.

siRNAs bind to their targets through perfect base-pairing, and their binding is not restricted to the 3'UTR region. Once bound, Ago2 cleaves the mRNA specifically 10 nt away from the 5' end of the siRNA leading to mRNA instability and degradation (Martinez and Tuschl, 2004).

1.7.2 siRNA triggers in mammals

In mammals and other vertebrates, unlike plants and worms, RNAi cannot be experimentally triggered with long dsRNAs, since these would activate the interferon response, leading to activation of the dsRNA-dependent protein kinase R (PKR) and global inhibition of protein synthesis (Platanias, 2005). There are however options to elicit a specific RNAi response. These use artificial molecules that mimic the structure of distinct RNA intermediates from the miRNA pathway (Figure 1.5), and therefore have to undergo similar processing steps.

The first described approach to mammalian RNAi involved the transfection of synthetic siRNA duplexes into cultured cells (Elbashir et al., 2001a). The siRNA duplex is similar in structure to the molecules produced in plants and worms after injection of long dsRNAs, and it is similar to the miRNA duplexes produced during miRNA biogenesis. This approach leads to efficient downregulation of target genes but it has the disadvantages that it is relatively expensive and transient, meaning that it can only be applied *in vitro*. A second method uses vectors encoding short-hairpin RNAs under the control of pol III promoters (Brummelkamp et al., 2002; Paddison et al., 2002). These RNA molecules mimic the structure of pre-miRNA hairpins and, since they are expressed from plasmids, can be stably integrated into the genome and brought into living mice. However, since pol III promoters do not lend themselves to controlled regulation, this approach does not easily allow for tissue specific expression or temporal control over the RNAi trigger. In addition, shRNAs have been shown to inefficiently dowregulate their target mRNAs when expressed from a single-copy transgene (Silva et al., 2005).

A third experimental trigger of RNAi in mammals depends on the transcription of a pre-miRNA mimetic. These mimics, the shRNA-mirs, are transcribed as longer single stranded RNAs that, like pre-miRNAs, form an internal hairpin-like structure due to sequence complementarity. shRNA-mirs undergo the same biogenesis steps as pre-miRNA, but result in the production of an siRNA. Because they mimic the structure of a miRNA primary transcript, shRNA-mirs can be transcribed under pol II promoters, meaning that they are more amenable to regulatory control. In addition, shRNA-mir based vectors have been shown to lead to a higher level of siRNA expression as well as a more dependable and efficient knockdown of target genes when compared to shRNAs expressed from the same promoter (Silva et al., 2005).

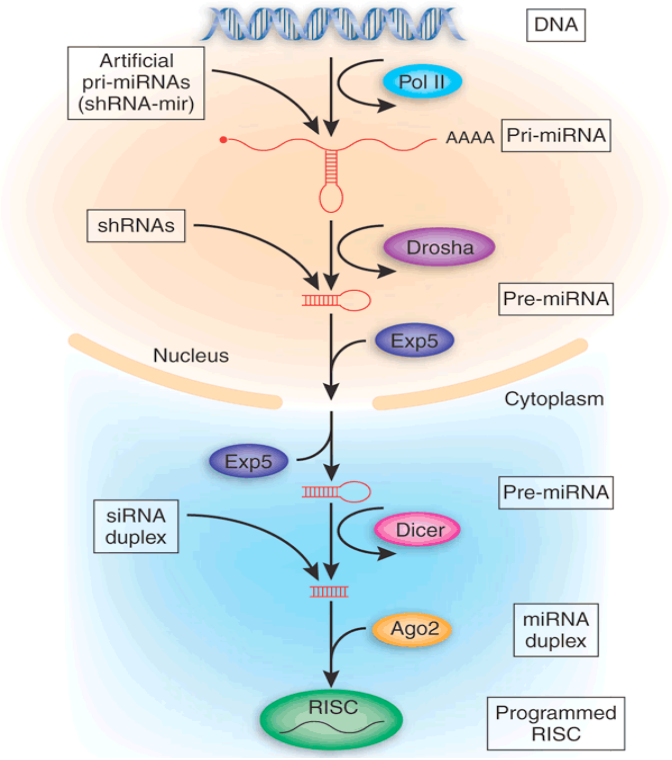


Figure 1.5. Entry points for experimental RNAi

To experimentally silence genes by RNAi three different triggers can be employed. shRNA-mirs mimic the structure of primary microRNA transcripts, and undergo all the processing steps of the miRNA biogenesis pathway. shRNAs are similar in structure to Drosha cleavage products and enter the pathway one step later. Finally, siRNA duplexes mimic the structure of miRNA duplexes and only undergo loading into the RISC complex. For simplicity not all factors involved in miRNA biogenesis are shown. Pol II, polymerase II, Exp5, Exportin-5, Ago2, Argonaute 2. Figure from (Cullen, 2005)

1.8 RNAi in the mouse embryo

Although RNA interference has become a useful transgenic technique for *in vitro* studies, there is a paucity of systems which reproducibly achieve efficient *in vivo* knockdown in mice. The application of RNAi as a tool has been a challenge, particularly in the mouse embryo, with only a handful of studies showing its feasibility. The first report of RNAi in the mouse embryo came in 2000 when the group of Zernicka-Goetz microinjected dsRNA targeting E-cadherin into single cell zygotes. This was shown to lead to dowregulation of E-cadherin protein levels when assayed at the morula stage by immunofluorescence and western blot (Wianny and Zernicka-Goetz, 2000). Two years later, RNAi was show to be possible in post-implantation embryos when siRNAs against GFP were co-electroporated ex-utero with a plasmid encoding this reporter into the forebrain of a 10 days post-coitum (dpc) mouse embryo (Calegari et al., 2002). Although both of these studies are relatively simple in the techniques they employed, they were significant in that they showed that both pre- and post-implantation mammalian embryos can sustain an RNAi response.

The first studies to truly unfold the full potential of RNAi as a tool to dissect gene function in the mouse embryo used a lentiviral-based vector, encoding shRNAs driven by the H1 RNA pol III promoter, to target ES cells (Kunath et al., 2003; Lickert et al., 2005). Targeted ES cell clones were then used in a tetraploid complementation assay to generate mouse embryos. In these

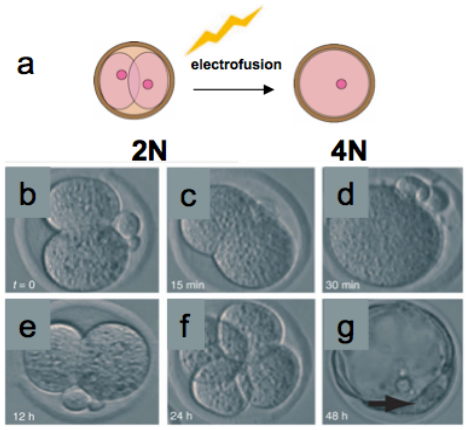


Figure 1.6. Generation of tetraploid embryos

(a) Schematic representation of tetraploid fusion. Wild-type two-cell stage embryos are subjected to a weak electric pulse. As a consequence the two blastomeres fuse giving rise to a one-cell-staged embryo with a tetraploid genome (b-d). A few hours after the electrofusion, the embryos will resume development (e-g) and give rise to blastocyst-stage embryos that lack a proper inner cell mass (arrow). 2N, diploid; 4N, tetraploid. Adapted from (Eakin and Hadjantonakis, 2006).

assays, ES cells are put in close contact with morula-stage tetraploid mouse embryos, generated by electrofusion of two blastomeres of a two-cell-stage embryo (Figure 1.6). Upon *in vitro* cultivation, these ES cells and tetraploid embryos develop into a chimeric blastocyst, whose inner-cell-mass is composed entirely of ES derived cells. When transferred into the uterus of a pseudopregnant mouse, these chimeric embryos will develop so that all the cells of the embryo proper are ESC-derived, while most of the extra-embryonic tissues will originate from the wild-type tetraploid embryos Figure 1.7 (George et al., 2007). Because all cells of the embryo proper carry the modification introduced into the ES cells, they can be directly used to study RNAi phenotypes.

When Rossant and colleagues analyzed embryos derived by this technique, carrying shRNAs against *Grsf1* or *Fragilis2*, they saw not only that protein levels were significantly downregulated when compared to wild-type embryos, but also that embryos showed several abnormalities, which implicated both genes as downstream components of the Wnt response during axial elongation (Kunath et al., 2003; Lickert et al., 2005). These studies represented breakthroughs in the field of RNAi technology, however, the techniques they used to deliver the RNAi trigger have, as described above, their inherent problems. First, the usage of lentiviral vectors means that integration of the knockdown transgene occurs randomly in the genome and at multiple loci, which can give rise to variability between clones due to integration site. Secondly, this approach does not incorporate a mechanism for inducibility. This can be problematic, especially when analyzing phenotypes in embryos derived by tetraploid complementation. Because phenotypes are studied in full ES cell- derived embryos it is important to have a control for ES cell integrity, to be sure that any abnormalities present are a consequence of the expression of the knockdown construct, and not a reflection of loss of pluripotency due to previous manipulation. Thirdly, use of pol III promoters leads to expression of the transgene at very high levels. Together with the integration of multiple copies this can lead to toxicity by subverting the RNAi machinery from its normal regulatory functions. In accordance

with this, some of the phenotypes reported using this system showed a higher severity than the corresponding knockout.

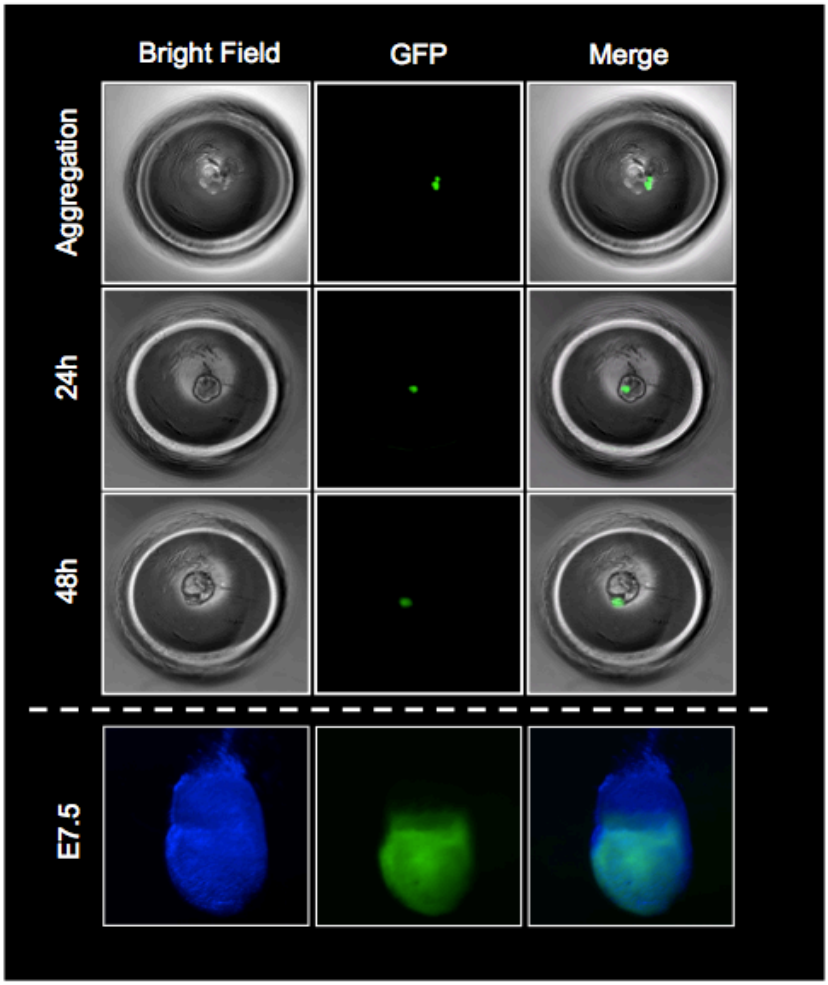


Figure 1.7. Generation of chimeric mice by tetraploid complementation.

Microscopy images showing various stages of a tetraploid complementation assay. A clump of ES cells modified to express GFP (green) is put together with two morula-stage tetraploid embryos (first row). 24h later, ES cells and tetraploid embryos have developed together into a late morula (second row). Two days after the aggregation has been set up the chimeric embryo has developed into a blastocyst, with the modified GFP-expressing cells making up the inner cell mass (third row). This chimeric blastocyst can then be transferred into a surrogate mother for further development. Analysis of the embryos at E7.5 shows that the ES cells have given rise to the embryo proper (bottom), while the tetraploid embryos have originated most of the extra-embryonic tissues.

1.9 Goal

Several systems have been developed previously to perform *in vivo* knockdown of genes in the mouse, however, lack of control over integration site, time or tissue of knockdown, or even over integrity of the cell lines have rendered them impractical. The work described in this thesis aimed at setting up a system that would allow us to overcome these limitations. In the following Chapters I describe the generation of animals using a vector system that allows fast and reliable knockdown of genes from a single and defined genomic locus, avoiding variability caused by integration site. The system is inducible, permitting control over the onset of gene silencing by RNAi and providing a control for the integrity of the transgenic ES cell clones used in the subsequent experiments. In addition, I show that this system can accommodate RNAi vectors of differing strengths, leading to the generation of mutants of graded phenotypes – from hypomorphic to complete loss-of-function. Finally, as evidence for the merit of this system, I describe novel roles for the gene brachyury, which have been discovered by studying the *in vivo* RNAi models described herein.

2 MATERIALS AND METHODS

2.1 Mouse strains and animal husbandry

The mouse strains used for this study were all purchased from the Harlan Laboratories (Harlan Winkelmann GmbH, Borchen, Germany). The outbred strain CD1® was derived from the Swiss-Webster mice at the beginning of the 20th century, and a stock established at the Institute for Cancer Research (ICR) in Philadelphia, PA (USA). Offspring from this strain were transferred to the Charles River Breeding Laboratories, in Willington, MA (USA), from which a breeding stock was obtained by the Harlan Sprague Dawley, Inc. (*Hsd:ICR(CD-1®)*). The outbred NMRI strain was derived from Swiss mice and established at the US Naval Medical Research Institute (NMRI). They were obtained in 1955 by the Federal Research Institute for Virus Diseases of Animals in Tübingen (Germay) and moved to the Central Institute for Laboratory Breeding, Hannover (Germany) in 1958, from where Harlan UK obtained a breeding stock is 1994 (*HsdHan:NMRI*). The strain 129 was originally derived from a cross of coat color stocks and a chinchilla stock at the beginning of the 20th century. The Harlan inbred strain, obtained from the Jackson Laboratory in Bar Harbor, ME (USA) represents a substrain of 129. 'S' refers to the origin from a congenic strain made by outcrossing the *steel* mutation. The number following this letter distinguishes the different 129 parental strains within the lineage. 'Sv' indicates the name of the Stevens lab (*129S2/SvHsd*). The strain C57BL is the most commonly used inbred strain. It was derived in 1921, and strain 6 (C57BL/6) was separated from strain 10. C56BL/6 was acquired by Harlan UK, formally OLAC (Ola), from the Jackson Laboratory (J) and then moved to Harlan Netherlands. It is therefore referred to as C67BL/6JOlaHsd.

The mice were housed at specific pathogen free (SPF) conditions at the animal facility of the Max Planck Institute for Molecular Genetics and kept under a 12h cycle of light and dark at 22°C and a relative humidity of 55±10%. They were fed a pelleted, irradiated diet (ssniff M-Z®, Soest, Germany) composed of 22% raw protein, 4.5% raw fat, 3.9% raw fiber and 6.8% raw ashes. Distilled water alone, or supplemented with 4 mg ml^{-1} and 1% sucrose – for transgene induction studies – was provided *ad libitum*. For timed matings, day of plug was assumed to be

E0.5. All animal experiments were approved by the Berlin State Office for Safety at Work, Health protection and Technical Safety (Landsamt für Arbeitsschutz und technische Sicherheit, LAGetSi) and carried out in accordance with the german animal welfare act (Tierschutzgesetz, TSchG).

2.2 Generation of the RNAi vector system

2.2.1 Recipient Constructs

The recipient transgenes were cloned by Dr. Phillip Grote (Max-Planck Institute for Molecular Genetics, Berlin, Germany) into the XbaI site of pROSA26-1 (Soriano, 1999) and used to target the *Gt(ROSA)26Sor* locus as described in subsection 2.3.1.1. The recipient transgene was assembled in pBluescript SK (Stratagene) and contained the following elements (5′ to 3′): a SV40 splice acceptor fused to a murinized version of tTS or rtTA (Clontech) with a SV40 polyadenylation signal, the 1.2-kb chicken β-globin insulator (5′HS4), a PGK-hygromycin selection cassette flanked by opposing loxP and lox5171 sites (Hoess et al., 1982; Lee and Saito, 1998), and a promoter-less neomycin resistance gene with bidirectional polyadenylation signal, followed by another chicken β-globin insulator.

2.2.2 Exchange Vectors

shRNA-mirs against genes of interest were identified and purchased from Invitrogen (Block it kit). After *in vitro* validation of their knockdown efficiency (see subsection 2.2.3), selected shRNA-mirs were amplified as described in subsection 2.4.3.2, using the mir155_SalI and mir155_MluI PCR primers (see Table 6) -containing recognition sequences for the SalI and the MluI restriction enzymes respectively - and cloned into exchange vectors. Exchange vectors were assembled in pBluescript SK and contained the following elements (5′ to 3′): optimized tetS

binding sites (TRE-tight, Clontech), CAGGS promoter, a transgene containing either one or two shRNAs-mir in a murine mir155 context (Invitrogen) followed by the thymidine kinase polyadenylation signal or the coding sequence of EGFP, one shRNAmir flanked by splice donor and acceptor sites and the rabbit β-globin polyadenylation signal, the chicken β-globin core insulator and a PGK promoter that complements the promoter-less neomycin resistance gene of the recipient locus upon integration. Exchange vectors were flanked by loxP and lox5171 sites.

2.2.3 *In vitro* validation of shRNA-mirs

For in vitro validation of gene knockdown, four individual shRNA-mirs in *pCDNA6.2GW/EmGFP miR* (Block it kit; Invitrogen) were tested per gene. Briefly, human SW480 cells were plated in 12-well-plates (Corning) in a density of $7x10^4$ cells per well and cultured overnight (o.n.) in Dulbecco's Modified Eagle's Medium (DMEM; BioWhittaker) containing 10% (v/v) of fetal calf serum (gibco) and 2mM of L-Glutamine (BioWhittaker) at 37°C in a humified 5% CO_2 incubator (HERAcell 150; Heraeus). The next day, cells were transfected with *pCDNA6.2GW/EmGFP miR* plasmids together with an expression vector carrying the murine target gene cDNA, in a concentration ratio of 4:1 (1μg of knockdown plasmid to 0.25μg of expression vector). Transfections were done with Lipofectamine 2000 (Invitrogen) according to manufacturers' instructions. Two days after transfection, total RNAs were extracted and abundance of target mRNA quantified by qPCR as described in subsection 2.5.2. The most efficient shRNA-mirs were used in subsequent *in vivo* experiments, and their sequences are listed below.

Table 1 Sequences of shRNA-mirs used in *in vivo* experiments

Target Gene	shRNA-mir	sequence 5' to 3'
T	267	TGC TGT TAG TTA GCT CCT TGA AGC GCG TTT TGG CCA CTG ACT GAC GCG CTT CAG AGC TAA CTA A
T	1457	TGC TGT AGA AGA TCC AGT TGA CAC CGG TTT TGG CCA CTG ACT GAC CGG TGT CAT GGA TCT TCT A
Foxa2	294	TGC TGT AGG ATG ACA TGT TCA TGG AGG TTT TGG CCA CTG ACT GAC CTC CAT GAA TGT CAT CCT A
Foxa2	584	TGC TGT GTA CGA GTA GGG AGG TTT GGG TTT TGG CCA CTG ACT GAC CCA AAC CTC TAC TCG TAC A
Noto	422	TGC TGT AAG GTA GCT CTG ATC CTG TGG TTT TGG CCA CTG ACT GAC CAC AGG ATG AGC TAC CTT A

2.3 Generation of Modified ES cells

All experimental procedures were based on the protocols described in (Nagy, 2003).

2.3.1 ES cell procedures

2.3.1.1 Electroporation

The procedures were performed under sterile conditions in a laminar flow hood (HERAsafe ; Heraeus). 2-5x10^6 ES cells were seeded onto a monolayer of mitotically inactivated primary embryonic fibroblasts, i.e., feeder cells (3x10^6/plate), in a gelatin-coated 10 cm cell culture dish (Corning) and incubated at 37°C in a humified 5% CO_2 incubator (HERAcell 150; Heraeus). The cells were grown in ES cell medium composed of Dulbecco's Modified Eagle's Medium (DMEM containing 4,500 mg/ml glucose, without sodium pyruvate; Sigma-Aldrich), 15% (v/v) ES cell-qualified, heat-inactivated fetal bovine serum (FBS; Gibco), 2 mM L-glutamine (Sigma-Aldrich), 50 U/ml penicillin (Sigma-Aldrich), 50 µg/ml streptomycin (Sigma-Aldrich), 1% 100× non-essential amino acids (Sigma-Aldrich), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 1% 100x nucleosides (Sigma-Aldrich). 1000 U/ml murine leukemia inhibitory factor (LIF; Chemicon) were added to keep the ES cells in an undifferentiated state. The medium was exchanged daily until round colonies were visible. Before trypsinization, the ES cells were grown in fresh medium for at least 2 h. The medium was aspirated, and the cells were carefully washed twice with cell-culture grade D-PBS (Lonza). 1 ml trypsin/EDTA solution (Gibco) was added, and the cells were incubated at 37°C for 10 min in order to disrupt cell-cell contacts. The enzyme was inactivated by the addition of 2 ml ES cell medium before pipetting vigorously up and down to produce a single cell suspension. The cell density was determined with a hemocytometer (Neubauer; Roth), and the cells were collected by centrifugation for 5 min at 200xg. 2-5x10^6 ES cells were then resuspended in 800 µl D-PBS and transferred to an electropration cuvette (0.4 cm gap; Bio-Rad). The respective linearized DNA (pROSA26A and pROSA26S knock-in

plasmids; linearized with PacI) was added and mixed with the ES cells. The cells were electroporated with a GenePulser (Bio-Rad) using 240 V and 500 µF. After electroporation, the cell suspension was immediately transferred to a 15 ml tube containing 10 ml ES cell medium, the cells were collected by centrifugation as above, and the pellet was resuspended in 10 ml ES cell medium. The cells were seeded on a monolayer of neomycin-resistant feeder cells ($1x10^6$/plate) in 6 cm cell culture dishes and incubated at 37°C in ES cell medium. Selection for clones that had integrated the respective DNA fragment was started 36 h after electroporation with selection medium containing 125 µg ml^{-1} Hygromycin B (Invitrogen). The medium was exchanged daily until ES cell colonies became visible (approximately 1 week after electroporation).

2.3.1.2 RMCE

To integrate exchange vectors into the *ROSA26* locus via recombination mediated cassette exchange, $3x10^5$ ROSA26A/S ES cells were plated in a gelatinized well of a 6 well plate (Corning) containing feeders, and cultured as before (2.3.1.1). After overnight (o.n.) incubation, 1.5 ml of fresh medium were added and ES cells transfected with 5.0 µg of exchange vector and 0.1µg of Cre recombinase expression vector using Lipofectamine 2000 (Invitrogen). DNAs were mixed with 125 µl of Opti-MEM and 25 µl of Lipofectamine were mixed with 110 µl of Opti-MEM and both mixtures were left to incubate for 5 min at room temperature (RT). Afterwards, both mixtures were combined and left to incubate at RT for 15 more minutes. After this incubation period, the combined solution was added to the cells and left to act for 3-5 hours at 37°C in a humified 5% CO_2 incubator (HERAcell 150; Heraeus). Cells were then split as described before (2.3.1.1) into 3x6 cm culture dishes (Corning) containing neomycin resistant feeders in the following ratios 1/6, 2/6 and 3/6. Selection with 250 µg ml^{-1} of Geneticin (Gibco) started the following day and lasted approximately 7 days, until resistant colonies were visible. Throughout the whole procedure ES medium was exchanged daily.

2.3.1.3 Picking

Fresh ES cell medium was added to the cells 3-4 h prior to picking. The cells were washed twice with D-PBS before covering the colonies with a layer of fresh D-PBS. Individual colonies were picked using disposable 10 µl pipette tips under a stereo microscope (MZ8; Leica) and transferred to the wells of a round-bottomed 96-well plate (Corning) filled with 50 µl cold trypsin/EDTA solution. After all colonies had been picked, the 96-well plate was placed in the 37°C incubator for 10 min. 100 µl ES cell medium were added per well to inactivate the trypsin. The colonies were disaggregated with a multi-channel pipette and transferred to the wells of a gelatinized, flat-bottomed 96-well plate (Corning) containing a monolayer of feeder cells ($1x10^6$ feeder cells/plate) and grown in regular ES cell medium. The next day, the medium was replaced by selection medium.

2.3.1.4 Splitting and Freezing

After the cells had been grown in selection medium for 2 days, they were washed twice with D-PBS before incubating them at 37°C with 70 µl trypsin/EDTA for 10 min. The trypsinization was stopped by adding 140 µl bicarbonate-free DMEM (Sigma-Aldrich) supplemented with 10 mM HEPES (Sigma-Aldrich) and 20% FBS (v/v). The cells in the so-called 'DNA Original Plate' were disaggregated, and 70 µl of the 210 µl were transferred to the wells of a round-bottomed 96-well plate (Corning) containing 70 µl 2x- concentrated ES cell freezing medium (bicarbonate-free DMEM (Sigma-Aldrich), 10 mM HEPES, 20% FBS, 20% DMSO). The contents of this so-called 'Master Plate' were mixed well by pipetting. Another 70 µl of the remaining 140 µl cell suspension were transferred to the wells of a gelatinized, flat-bottomed 96-well plate (Corning) containing 200 µl ES cell medium ('DNA Replica Plate'), and the contents were again mixed. The 'Master Plate' was sealed, placed inside a styrofoam box, and frozen at 80°C. 200 µl ES cell medium were added to the remaining 70 µl cell suspension in the 'DNA Original Plate' and the contents were mixed. The cells in the 'DNA Original Plate' and the 'DNA

Replica Plate' were grown to confluency at 37°C. DNA was isolated and subsequently used for analysis by Southern blot (see section 2.3.2).

2.3.2 Screening of ES cell clones by Southern blot

2.3.2.1 Genomic DNA isolation

Southern blot analysis was performed in order to identify ES cell clones that had successfully integrated in the *ROSA26* locus by either homologous recombination or RMCE. The procedure followed the protocol described in (Ramirez-Solis et al., 1993). ES cells were grown to confluency in the wells of a 96-well plate (in the so-called 'DNA Original Plate' and the 'DNA Replica Plate' (2.3.1.4) in a humidified incubator at 37°C and 5% CO_2 (HERAcell 150; Heraeus). The cells were carefully washed twice with D-PBS and 50 µl prewarmed lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM EDTA pH 8.0, 10 mM NaCl, 0.5% sarcosyl) containing 1 mg/ml proteinase K (Roche) were added per well. The plate was placed inside a humified chamber and incubated o.n. at 60°C. The next day, 100 µl ice-cold 75 mM NaCl/100% EtOH were added without mixing. The plate was allowed to stand on the bench for 30 min to precipitate the DNA as a filamentous network on the bottom of the wells. The plate was then carefully inverted to discard the solution, and excess liquid was blotted on a paper towel. The wells were rinsed 3 times by addition of 200 µl 70% EtOH. After the final wash, the precipitated DNA was allowed to dry on the bench. The 'DNA Replica Plate' was sealed and stored at -20°C. The 'DNA Original Plate' was used for restricition enzyme digestion.

2.3.2.2 DNA digestion and electrophoresis

A restriction digest mix was prepared containing 1 U/µl BamHI (Promega), for identification of correct integration of recipient contructs, or 1 U/µl BamHI and HindIII (Promega), for identification of correct RMCE, plus 1xBuffer B (Roche), 1 mM spermidine (Promega), 100 µg/ml BSA (Promega), 100 µg/ml RNase A (Promega). The mix was prewarmed to 37°C, and 30 µl

were added to each well without mixing. The reaction was incubated at 37°C for 2-4 h in a humified chamber before mixing the content of the wells. The 37°C incubation was continued o.n. in a humidified chamber. The next day, 6 µl 6x gel loading buffer were added to each well, and the DNA was electrophoretically separated in a 0.7% TBE agarose gel in 1x TBE electrophoresis buffer for 12-18 h at 30 V. The next day, the gel was documented with the GelDoc 2000 system (Bio-Rad).

2.3.2.3 DNA blotting

After electrophoretic size separation, the gel was pretreated in order to facilitate transfer of large DNA fragments. First, the DNA was partially depurinated by soaking the gel twice in 0.25 N HCl for 10 min. The gel was washed in ddH_2O for 5 min before denaturing the DNA by placing the gel in a bath of 0.5 N NaOH on a moving platform for 40 min at RT. The DNA was then blotted o.n. onto a preequilibrated nylon Zeta-Probe GT membrane (Bio-Rad) by capillary transfer. The DNA was UV-crosslinked to the membrane using 5000µJ/cm_2 radiation (Stratalinker 2400; Stratagene). The membrane was washed twice in 2xSSC pH 7.0 and either air-dried and stored at -20°C in a sealed plastic bag or directly hybridized.

2.3.2.4 Probe labeling

For the generation of the radioactively labeled probes, 25 ng of the respective template DNA fragments (complementary to the 5' homology arm of the *ROSA26* locus or to the neomycin resistence cassette for cells targeted as described in subsections 2.3.1.1 and 2.3.1.2, respectively) were diluted in 45 µl TE buffer and denatured for 3 min at 95°C. After the sample was cooled down on ice for 2 min, it was added to a reaction tube containing a dried mix of dATP, dGTP, dTTP, Klenow enzyme, and random primers (Rediprime II Random Prime Labelling System; Amersham), and the components mixed. All following procedures were performed in an isotope laboratory facility according to the manufacturer's instructions. 5 µl of [α-$_{32}$P]dCTP (Redivue; Amersham) with a specific activity of 3000 Ci/mmol were added, and the

mixture was incubated for 10 min at 37°C to allow for the labeling reaction catalyzed by the Klenow fragment of the DNA polymerase I. The radioactive sample was pipetted onto a G-25 MicroSpin Column (GE) to separate the labeled probe from unincorporated radioactive nucleotides, according to manufacturers' instructions. The labeled DNA probe was denatured for 3 min at 95°C, briefly put on ice, and added to the hybridization buffer (see below).

2.3.2.5 Hybridization and detection

The DNA-blotted membrane was prehybridized in a glass bottle with 5 ml prewarmed ExpressHyb Hybridization Solution (Clontech) for 30 min in a hybridization oven (Hybaid Shake'n'Stack; Thermo Scientific) at 68°C with constant rotation. After prehybridization, the denatured, labeled probe was added to the solution, and the membrane was hybridized o.n. at 60°C with constant rotation. The next day, the membrane was washed in 2x-concentrated SSC pH 7.0 for 5 min rocking at RT, in preheated 2x SSC/1% SDS for 30 min at 60°C, and – if higher stringency was required – in 0.1x SSC/0.1% SDS for up to 30 min. A rinse in 0.1x SSC at RT followed before sealing the membrane inside a plastic bag. The radioactively labeled DNA was exposed in a phosphoimager (Amersham) for 24h before detection.

2.4 Generation of Transgenic Mice

After identification of the clones that had undergone successful homologous recombination or RMCE, the respective cells were thawed, transferred to 15 ml tubes containing 5 ml ES cell medium and spun down at 200xg for 5 min. The supernatant was carefully aspirated, the pellet resuspended in ES cell medium, and the cells plated on a monolayer of feeder cells. Cells were grown in ES cell medium at 37°C for propagation. They were subsequently used for the generation of mutant mouse lines via ES cell- diploid embryo complementation (for live birth), or ES cell-tetraploid embryo complementation (for the generation of full ES cell derived embryos) techniques. These procedures were performed either by me or Karol Macura (Max Planck

Institute for Molecular Genetics, Berlin, Germany) and followed the protocols described in (Eakin and Hadjantonakis, 2006).

2.4.1 Generation of full-ES cell derived embryos by tetraploid complementation

For generation of full ES cell-derived mice, two-cell stage host embryos (E1.5) were collected by flushing the oviducts of wild-type females with M2 medium (Sigma-Aldrich), warmed up to 37°C. Embryos were collected in 3.5-cm dishes containing KSOM (Millipore) micro-culture drops, covered with mineral oil (Sigma-Aldrich), and kept in a humidified incubator at 37°C and 5% CO_2 (HERAcell 150; Heraeus). Tetraploidy was induced by electrofusing the plasma membranes of the two-cell diploid embryos. Briefly, embryos were washed twice in 0.3M mannitol solution (Sigma-Aldrich) and then subjected to a short electric pulse using a CF-150B pulse generator (repeat 1; voltage 35V; pulse duration 35 μs). Embryos were immediately washed three times in M2 medium, and incubated in a KSOM micro-culture drop at 37°C with 5% CO_2. Embryos were periodically monitored for successful electrofusion. Embryos that failed to fused were discarded. The following day, four-cell stage embryos were washed in M2 medium and incubated briefly in Tyrode's Acid solution (Sigma-Aldrich) to remove the *zona pelucida*. This step makes the embryos ready for aggregation. In the next step, two tetraploid embryos were put in close contact with an ES cell colony, containing 8-12 cells, in a KSOM micro-culture drop and the group was let to incubated for two days drop at 37°C 5% with CO_2. To obtain feeder-free colonies of modified ES cells for the aggregation step, ES cells were thawed two days in advance into a 6-cm culture dish (Corning) containing feeder cells and cultured as before (2.3.1.1). On the day of the aggregation the cells were washed twice with PBS (Lonza) and then incubated briefly with warm trypsin/EDTA solution (Gibco) until the ES cell colonies, but not the feeders, started detaching from the bottom. The enzyme was inactivated by adding 2 ml ES cell

medium, and the medium - containing the ES cell colonies ready for aggregation - was transferred into a clean 6 cm dish (Corning). Two days after setting up the aggregation, the blastocyst stage chimeric embryos were transferred into the uterus of a pseudopregnant female as described in (Nagy, 2003).

2.4.2 Generation of adult chimeric mice by diploid complementation

In the case of diploid complementation technique, the resulting animal will be comprised of a mixture of ES-derived cells along with a high proportion of host embryo-derived cells. The method follows similar steps to those described above (2.4.1) with a few modifications: host embryos were collected at four-cell stage (E2.5) and used immediately for aggregation with ES cell colonies without going through the electrofusion step. Additionally, aggregation was preformed using only one host embryo instead of two. All remaining steps were done as before.

2.4.3 Animal Genotyping by PCR

2.4.3.1 DNA isolation

Genomic DNA was isolated according to the protocol described in (Laird et al., 1991) as follows. Tail biopsies of 0.5-1 cm or yolk sacs from embryos were incubated o.n. at 56°C with gentle agitation in 0.5 ml or 0.15 ml of lysis buffer respectively (200 mM NaCl, 100 mM Tris-HCl pH 8.5, 5 mM EDTA, 0.2% SDS) containing 150 µg/ml proteinase K (Roche). The lysates were centrifuged for 10 min at 16,100xg, and the supernatants poured into new tubes containing an equal volume of isopropanol. The samples were mixed until the DNA was completely precipitated, and the precipitates fished out using disposable tips and transferred to new tubes containing 0.5 ml or 0.15 ml TE buffer respectively. The DNA was dissolved for several hours at 37°C under agitation and stored at 4°C.

2.4.3.2 PCR reaction and program

For PCR analysis of the DNA described above, a reaction mix was prepared on ice. The master mix contained PCR buffer (200 mM Tris pH 8.4, 500 mM KCl), $MgCl_2$, dNTP mix (dATP, dCTP, dGTP, dTTP), Taq DNA polymerase (all from Invitrogen), and primers amplifying a region in the neomycin resistance cassette (see Table 6). A reaction mix was assembled with the following concentrations: PCR buffer (1x); $MgCl_2$ (1.5mM); dNTPs (200μM); primers (0.2μM each), Taq DNA polymerase (0.05 U/μl).1 μl of genomic DNA was used as template in a 50 μl reaction. A negative control without template was included in each run. The PCR program used for genotyping is shown on Table 2.

Table 2. Genotyping PCR program

Step	Temperature	Time
1. Initial denaturation	95°C	10 min
2. Denaturation	95°C	15 sec
3. Elongation	55°C	45 sec
	go to step 2.	35 times
4. Melt curve	95°C	15sec
	60°C	15 sec
	95°C	15 sec
6. Cooling	4°C	∞

2.4.3.3 Gel electrophoresis

The PCR products were analyzed on a 2% (w/v) agarose gel. UltraPure agarose (Invitrogen) was mixed with the appropriate volume of 1xTAE buffer in a glass beaker, microwaved until dissolved, and cooled to about 60°C. Ethidium bromide was added at a concentration of 0.2 μg/ml, and the solution was well swirled before pouring it into a gel tray holding a comb. The comb was removed after solidification of the gel, and the tray was placed in a gel chamber (PerfectBlue Gel System; Peqlab) filled with 1xTAE electrophoresis buffer. One well was loaded with the 1 Kb Plus DNA Ladder (Invitrogen). 2 μl of 6x gel loading buffer were added to 10 μl of PCR mix before loading the samples. Following electrophoretic size

separation, the PCR products were visualized under transilluminescent UV light (λ=302 nm) and documented using the GelDoc 2000 system (Bio-Rad).

2.5 RNA Molecular Biology Techniques

2.5.1 Northern Blot Analysis of small RNAs

2.5.1.1 RNA isolation

Total RNA was extracted from mouse embryos or ES cells using Trizol Reagent (Invitrogen) according to manufacturers' instructions. Trizol is a monophasic solution of phenol and guanidine isothiocyanate, which allows the extraction of whole RNA populations from cells and tissues with no bias for size. Importantly, small RNA isolation is not possible with protocols involving regular RNA purification columns since their size exclusion is too large. Briefly, embryos were collected in 1ml of Trizol and homogenized using a douncer. Cells were also collected in Trizol and homogenized by vigorous pipetting. Homogenized samples were incubated for 5 minutes at room temperature to allow for full dissociation of nucleoprotein complexes. Following homogenization, 0.2ml of chloroform (Roth) was added to each sample. Samples were centrifuged at 12,000g for 15 minutes at 4°C and upper aqueous phase transferred to a fresh tube. RNA was precipitated by mixing 0.5ml of isopropanol (Merk) per ml of Trizol followed by centrifugation at 12,000g for 10 minutes at 4°C. RNA pellet was washed with 70% EtOH. To minimize DNA contaminations, all samples were treated with RNase-free DNase I (Qiagen) according to manufactures instructions. After DNaseI treatment, RNA samples were cleaned by phenol-chloroform extraction as follows: reaction volume was set to 300μl with nuclease-free water and sodium acetate (pH 5.2) was added to a final concentration of 0.3M. An equal volume of a 1:1 mixture of Phenol (Roth) and Chloroform was added to each sample. Samples were mixed by shaking and centrifuged at 12,000g for 5 min at room temperature. The

upper phase was transferred to a fresh tube and an equal volume of chloroform was added to each sample and mixed by shaking. Samples were centrifuged once more at 12,000g for 5 minutes at room temperature and the upper phase collected in a new tube. RNA was precipitated by adding 2.5 volumes of 100% EtOH (Merk) in the presence of 1μl of Polyacryl Carrier (MRC) and centrifuging samples at 12,000g for 10 minutes at room temperature. Following centrifugation, RNA pellets were washed with 70% EtOH and ressuspended in 20μl of nuclease-free water. RNA was quantified by UV spectrophotometry (ND-1000 Nanodrop).

2.5.1.2 Gel Preparation and Electrophoresis

For detection of processed siRNAs, 30 μg of total RNA were loaded on a 15% acrylamid (Roth) TBE denaturing gel, containing 7M Urea (Sigma-Aldrich). Before loading, samples were mixed with an equal volume of loading dye (8M Urea, 25 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue), heated at 80°C for 5 min and spun down. The samples were ran for 2h at 250 V (Consort E831), until the dye reached the end of the gel.

2.5.1.3 RNA blotting

After electrophoretic size separation, the RNA was blotted o.n. with TBE, onto a preequilibrated Gene Screen Plus membrane (PerkinElmer) by capillary transfer. The next day, the RNA was UV-crosslinked to the membrane using 5000μJ/cm2 radiation (Stratalinker 2400; Stratagene) and the membrane either air-dried and stored at -20°C in a sealed plastic bag or used directly for hybridization.

2.5.1.4 Probe labeling

An oligonucleotide, complementary to the predicted siRNA, was ^{32}P-end labeled using the following reaction:

Table 3. Probe end-labeling reaction

oligonucleotide (10 μM)	1 μl
10x Kinase buffer (Promega)	2,5 μl
H_2O	14,5 μl
T4 polynucleotide kinase (Promega)	2 μl
γ 32^P-ATP	5 μl

The reaction was left to incubate for 2h at 37°C, before the labeled probe was cleaned using a G-25 MicroSpin Column (GE), according to manufactures' instructions.

2.5.1.5 Hybridization and detection

Hybridization was performed in Ultra-hyb Oligo buffer (Ambion). Briefly, membrane was pre-hybridized in this solution at 42°C for 2h, before the labeled probe was added and let to hybridized o.n.. The next day, the membrane was washed 3x for 10 min in 6xSSC/1% SDS, and then exposed for 48h in a phosphoimager (Amersham) before detection.

2.5.2 Quantitative Real Time PCR

2.5.2.1 RNA isolation

Total RNA was isolated from whole embryos or cells using the RNeasy Mini or Midi Kit (Qiagen) according to the manufacturers' instructions. All volumes were adjusted accordingly. Briefly, fresh tissue was homogenized in RLT Buffer with a T8 Ultra-Turrax disperser (IKA-Werke) and centrifuged at full speed for 3 min. The supernatant was transferred into a new tube, and 1 vol. 70% EtOH was added to the lysate, and the contents mixed by pipetting. The sample was transferred to an RNeasy column placed in a collection tube and centrifuged, and the flow-through discarded. RW1 Buffer was added, and the column spun down again. An on-column

DNase digestion step was performed: RNase-free DNase I was mixed with RDD Buffer and pipetted onto the spin column membrane, which was then incubated on the bench- top for 15 min. RW1 Buffer was added, and the column was centrifuged. The flow-through was discarded. The membrane was washed twice with RPE Buffer before eluting the RNA with RNase-free H2O into a new collection tube. The RNA concentration was measured by UV spectrophotometry (ND-1000 Nanodrop).

2.5.2.2 Reverse transcription (RT)-PCR

The SuperScript III First-Strand Synthesis System for RT-PCR from Invitrogen was used for reverse transcription of RNA into cDNA. For this, 8 µg RNA were diluted in a total volume of 8 µl DECP-H_2O. 1 µl random hexamers and 1 µl dNTP mix (10 mM) were added per reaction, and the samples were incubated at 65°C for 5 min before placing them on ice for 1 min. A reaction mixture was prepared containing 2 µl 10× RT Buffer, 4 µl $MgCl_2$ (25 mM), 2 µl 0.1 M DTT, 1 µl RNaseOUT per sample. It was added to the RNA mixture and incubated at 25°C for 2 min. 1 µl SuperScript III (50 U/µl) reverse transcriptase was added, and the samples were incubated successively at 25°C for 10 min, at 42°C for 50 min, and at 70°C for 15 min. The reactions were cooled on ice before adding 1 µl RNase H to each tube and incubating them for 20 min at 37°C. The cDNA samples were stored at -20°C until further use.

2.5.2.3 Quantitative real-time PCR

Quantitative real-time PCR was performed with the above yielded cDNA using the Power SYBR Green PCR Master Mix (Applied Biosystems). This mix contains polymerase, dNTPs, buffer, and SYBR Green I Dye, which binds to double-stranded DNA. Thus, the fluorescent signal reflects the amount of double-stranded PCR product that is generated during the reaction. For each reaction, a mixture of 10 µl Power SYBR Green PCR Master Mix, 0.5 µl forward primer (20 µM), 0.5 µl reverse primer (20 µM), and 7 µl DECP-H_2O were transferred to the wells of a MicroAmp Fast Optical 96-Well Reaction Plate (Applied Biosystems) before adding 2 µl of

1:20 diluted cDNA sample. Each sample was run in technical triplicates. Negative controls without cDNA were included in each run. The reaction was run on a StepOne Real-Time PCR System according to the program listed below (Table 2.9). The results were analyzed using the StepOnePlus Software v2.0.2. The housekeeping gene *Gapdh* was used as internal control for normalization of each cDNA sample. The sequences of the primers used in this assay are listed in Table 6.

2.5.2.4 PCR program

Table 4. Quantitative real time PCR program

Step	Temperature	Time
1. Initial denaturation	95°C	10 min
2. Denaturation	95°C	15 sec
3. Elongation	60°C	30 sec
	go to step 2.	39 times
4. Melt curve	95°C	15sec
	60°C	15 sec
	95°C	15 sec
6. Cooling	30°C	∞

2.5.3 cDNA microarray analysis

The Illumina TotalPrep RNA Amplification Kit (Ambion) was used for generating biotinylated, amplified cRNA for hybridization on Illumina Sentrix arrays. All steps were performed according to the manufacturer's instructions.

2.5.3.1 First strand cDNA synthesis

DNase-digested total RNA was prepared with the RNeasy Mini or Midi Kit (Qiagen) as described under subsection 2.5.2.1. For a single reaction, 500 ng RNA were diluted in a total volume of 11 µl RNase-free H_2O in a sterile, RNase-free tube. For each sample, a Reverse Transcription Master Mix was assembled at RT as follows: 1 µl T7 Oligo (dT) Primer, 2 µl 10x First Strand Buffer, 4 µl dNTP Mix, 1 µl RNase Inhibitor, and 1 µl ArrayScript were mixed well, and 9 µl of the Master Mix were transferred to each RNA sample. The samples were thoroughly

mixed by pipetting, briefly centrifuged to collect the reaction at the bottom of the tube, and placed in a 42°C incubator for 2 h. After the incubation, the tubes were immediately put on ice.

2.5.3.2 Second strand cDNA synthesis

80 µl Second Strand Master Mix was prepared on ice (63 µl RNase-free H_2O, 10 µl 10x Second Strand Buffer, 4 µl dNTP Mix, 2 µl DNA Polymerase, 1 µl RNase H) and transferred to each sample. The contents were mixed well, briefly centrifuged and incubated at 16°C for 2 h. After the incubation, the tubes were immediately placed on ice.

2.5.3.3 cDNA purification

250 µl cDNA Binding Buffer were added to each sample, the samples were mixed, and the reaction was pipetted onto the center of the cDNA Filter Cartridge. The cartridges were centrifuged at 10,000xg for 1 min at RT. The flow-through was discarded, and 500 µl Wash Buffer were applied to each cartridge before centrifuging them again as above. The flow-through was discarded, and the cartridges were again spun to remove trace amounts of buffer. The cartridges were placed in a cDNA Elution Tube, and 10 µl preheated (55°C) RNase-free H_2O were applied to the center of the filter. After a 2-min incubation at RT, the cartridges were again spun. Additional 9 µl preheated RNase-free H_2O were pipetted onto the filter, and the cartridges were centrifuged at 10,000xg for 2 min at RT. The eluted, purified cDNA samples were either stored at -20°C or directly used to proceed with the protocol.

2.5.3.4 cRNA transcription

The IVT Master Mix was prepared at RT (2.5 µl T7 10× Reaction Buffer, 2.5 µl T7 Enzyme Mix, 2.5 µl Biotin-NTP Mix), and 7.5 µl IVT Master Mix were added to each purified cDNA sample. The samples were thoroughly mixed by pipetting, briefly centrifuged to collect the reaction at the bottom of the tube, and placed in a 37°C incubator for 14 h. The reaction was stopped by the addition of 75 µl RNase-free H_2O.

2.5.3.5 cRNA purification

350 µl cRNA Binding Buffer and 250 µl 100% EtOH were added to each sample. The contents were mixed by pipetting and directly transferred onto the center of a cRNA Filter Cartridge. The cartridges were centrifuged at 10,000xg for 1 min at RT, and the flow-through was discarded. 650 µl Wash Buffer were applied to each cartridge, which were again spun as above. The flow-through was discarded, and the cartridges again centrifuged to remove trace amounts of buffer. The cartridges were placed in cRNA Collection Tubes, and 70 µl preheated (55°C) RNase-free H_2O were applied to the center of the filter. After a 2-min incubation at RT, the purified cRNA was eluted by spinning the cartridges at 10,000xg for 2 min at RT. The cRNA concentration was measured by UV spectrophotometry (ND-1000 Nanodrop)., and the cRNA quality was assessed on a 1% (w/v) agarose gel.

2.5.3.6 Hybridization

To assess genome-wide expression profiles, 750 ng purified cRNA in a total volume of 5 µl RNase-free H_2O were provided for each sample. The samples were hybridized to MouseRef-8 Expression BeadChips (Illumina) and scanned on the Illumina BeadArray Reader. All these procedures were performed by Aydah Sabah (Service Department, Max Planck Institute for Molecular Genetics, Berlin, Germany) according to the manufacturer's instructions.

2.5.3.7 Data analysis

Expression data was processed using Bead Studio GX software (Illumina), using the cubic spline normalization method, and background subtraction. Marginally expressed genes were removed for analysis by a detection value cutoff of one in at least four out of eight profiles from the respective tissue domain. Deregulated genes were selected by more than 2-fold change in expression with a P-value <0.05 in an unpaired Welch t-test, without multiple testing correction. Comparison of randomized expression profiles was between 18 sets that contained equal numbers of profiles of induced and non-induced forebrains.

For analysis of off-target regulation, Sylamer was used as suggested by the authors (see www.ebi.ac.uk/enright/sylamer/sylamer.html). In short, identifiers of expressed genes (detection cutoff as above) in caudal end or forebrain data sets were ordered by induced versus non-induced fold change expression. Next, Sylamer was used to determine the distribution of occurrence of 6-, 7- and 8-mer sequences complimentary to the experimental seeds (shRNAmir 1: TAACTA, AACTAA, TAACTAA, CTAACTA, GCTAACTA, CTAACTAA; shRNAmir2: CTTCTA, TCTTCT, TCTTCTA, ATCTTCT, GATCTTCT, ATCTTCTA) in these ordered lists, by analyzing a genome-wide set of murine 3′ UTR sequences (van Dongen et al., 2008). Both these analysis were performed by Dr. Markus Morkel (Max Planck Institute for Molecular Genetics, Berlin, Germany).

2.6 Histology

2.6.1 Standard histology procedures

2.6.1.1 Embryo preparation

Mice were euthanized via cervical dislocation and embryos harvested in cold PBS and fixed o.n. at 4°C in 4% Paraformaldehyde (PFA; Sigma-Aldrich) in PBS. The next day, embryos up to E14.5 were washed twice with PBS and then processed manually through an ethanol series: 30%, 50%, 70%, before further processing in a MICROM STP 120 processor (MICROM) according to the program 1 (Table 5). E16.5 old fetuses were fixed in 4% PFAPBS o.n. at 4°C and then incubated for two days in 2% PFA/PBS with 0.25M EDTA to decalcify. Subsequently, fetuses were washed twice with PBS and processed manually through the ethanol series: 30%, 50%, 70%, before further processing in a MICROM STP 120 processor (MICROM) according to the program 2 (Table 5).The specimens were embedded in metal molds with paraffin (Histowax; Leica) using an EC 350-1 embedding station (MICROM) and placed onto the cooling plate until

the paraffin block was solidified. 5 µm-thick sections were cut on a rotary microtome (HM 355 S; MICROM), transferred onto adhesion microscope slides (SuperFrost; Menzel), and dried o.n. at 37°C. All slides were stored in a desiccated slide box at 4°C until further use.

Table 5. MICROM STP 120 processing programs

Program 1

Solution	Time	Agitation
80% EtOH	120 min	2
96% EtOH	120 min	2
100% EtOH	60 min	2
100% EtOH	60 min	2
100% EtOH	60 min	2
100% Xylene	90 min	2
100% Xylene	90 min	1
100% Parafin	120 min	1
100% Parafin	120 min	1

Program 2

Solution	Time	Agitation
70% EtOH	120 min	2
70% EtOH	110 min	2
80% EtOH	100 min	2
90% EtOH	120 min	2
96% EtOH	120 min	2
100% EtOH	60 min	2
100% EtOH	60 min	1
100% EtOH	60 min	1
100% Xylene	90 min	2
100% Xylene	90 min	2
100% Parafin	120 min	2
100% Parafin	120 min	1

2.6.1.2 Hematoxylin and Eosin (H&E) staining

PFA-fixed, paraffin-embedded sections were deparaffinized in xylene (2x for 8 min) and rehydrated (100% EtOH, 2x for 5 min; 90% EtOH, 80% EtOH, ddH2O, each 1x for 5 min). The samples were stained with Mayer's Hematoxylin Solution (Sigma-Aldrich) for 3 min, washed under running tap water for 15 min, dipped in 0.25% HCl/EtOH for 3 sec followed by a 5-min incubation in tap water, stained with alcoholic Eosin Y Solution (Sigma-Aldrich) for 5 min, and briefly washed in ddH_2O for 30 sec. The sections were then dehydrated (dip into 90% EtOH, 100% EtOH 2x for 5 min, xylene 1x for 5-10 min), immediately mounted in Entellan (Merck), and covered with glass cover slips (Roth).

2.6.1.3 Alcian Blue staining on sections

PFA-fixed, paraffin-embedded sections were deparaffinized and rehydrated as before (2.6.1.2). The samples were stained in 1% Alcian Blue (Sigma-Aldrich) for 25 min, then washed in taped water and counterstained in Eosin Y Solution (Sigma-Aldrich) for 5 sec. The slides were

washed again in tap water for 5 min and then dehydrated thought an EtOH series (dip into 70% EtOH, 80% EtOH, 96% EtOH and 100%). Slides were then incubated in xylene 1x for 5-10 min, mounted in Entellan (Merck) and covered with glass cover slips (Roth).

2.6.1.4 Imaging

Tissue sections were photographed with an AxioStar plus microscope (Zeiss) and a DFC320 camera (Leica) using the FireCam V3.0 software (Leica).

2.6.2 Whole-mount in situ hybridization

The procedures for the whole-mount in situ hybridization followed the online available protocols of the Molecular Anatomy of the Mouse Embryo Project (mamep) of the Max Planck Institute for Molecular Genetics (Berlin, Germany; http://mamep.molgen.mpg.de/). For all steps, RNase-free solutions were used.

2.6.2.1 Fixation of mouse embryos

Timed-pregnant mice were euthanized via cervical dislocation. The uterus was removed, and the embryos were dissected in cold PBS and fixed o.n. at 4°C in 4% PFA/PBS solution on a roller mixer. The next day, the fixative was removed by two washes with cold PBS for 10 min, and the embryos were dehydrated through a graded methanol series (25% MetOH/PBS, 50% MetOH/PBS, 75% MetOH/PBS, 1x for 10 min each, 100% MetOH 2x for 10 min). All steps were performed at 4°C on a roller mixer with pre-cooled solutions. The fixed embryos were stored in 100% MetOH at -20°C until further use.

2.6.2.2 Processing of mouse embryos

The desired number of fixed embryos of one stage were pooled. If not stated otherwise, all subsequent steps comprised a 10 min incubation at 4°C on a roller mixer. The embryos were rehydrated (1x 75% MetOH/PBST, 1x 50% MetOH/PBST, 1x 25% MetOH/PBST, 2x PBST), bleached in a 6% H_2O_2/PBST solution (10 min for E8.5, 20 min for E9.5, 30 min for E10.5, 45

min for E11.5), and washed 3x in PBST. The specimens were digested with 10 µg/ml proteinase K/PBST (7 min for E8.5, 10 min for E9.5, 13 min for E10.5, 17 min for E11.5) in order to allow for better penetration of the labeled probe. The digestion process was stopped by an incubation with 2 mg/ml glycine/PBST solution followed by two washes with PBST. The embryos were refixed in 0.2% glutaraldehyde/4% PFA/PBST for 30 min at RT while rolling, washed twice with PBST at RT, and preincubated with hybridization solution (50% formamide, 5x SSC pH 5.0, 1% SDS, 0.05 µg/ml yeast RNA (Sigma-Aldrich), 0.05 µg/ml heparin (Sigma-Aldrich), in RNase-free H2O) for 15 min at RT. The solution was exchanged with fresh hybridization solution, and the embryos were prehybridized for 2 h at 68°C to reduce unspecific background staining. Prehybridized embryos were either directly used for hybridization or stored at -20°C in hybridization solution.

2.6.2.3 Preparation of labeled probes

All probes used are listed in the Molecular Anatomy of the Mouse Embryo Project (mamep) (Max Planck Institute for Molecular Genetics; Berlin, Germany; http://mamep.molgen.mpg.de/), and were prepared as described in the database.

2.6.2.4 Hybridization

Wells from a 12-well plate (Corning) were filled with 2 ml preheated hybridization solution (68°C). A clean netwell (15 mm diameter, 74 µm mesh; Corning) was placed in each well, and the prehybridized embryos were sorted into the wells. The embryos were prewarmed to 68°C for 30 min. Meanwhile, a fresh 12-well plate with 2 ml hybridization solution containing 23.5 µl of hydrolyzed probe per well was prepared. The probes were denatured at 80°C for 10 min by placing the plate inside a sealed plastic box in a hot water bath. Right after denaturation, the netwells with the embryos were transferred to the plate containing the probes. The plate was placed inside a humid plastic box and incubated o/n in an oven with rocking function (BFED 053; Binder) at 68°C for hybridization of the probe to its complementary mRNA.

2.6.2.5 Antibody incubation

The netwells were transferred back to a 12-well plate filled with hybridization solution and incubated at 68°C for 30 min. The netwells were transferred into a netwell reagent tray (Corning) and washed twice at 68°C for 30 min with approximately 90 ml of preheated Solution 1 (50% formamide, 5x SSC pH 5.0, 1% SDS) followed by two washes for 30 min and two washes for 60 min at 68°C with Solution 3T (50% formamide, 2× SSC pH 5.0, 0.1% Tween-20) inside a humified box. The box was removed from the oven, and the embryos were washed three times with TBST for 15 min at RT. During the washing procedure, the antibody solution was prepared as follows: 6 ml TBST were mixed with a pinch of embryo powder (for preparation of the embryo powder: see below) and heat-inactivated for 30 min at 70°C in a waterbath. After cooling the solution on ice, 60 µl heat-inactivated lamb serum (Gibco) and 12.5 µl of α-dig-AP antibody (Roche) were added, and the solution was incubated at 4°C for at least 1 h rolling in the dark. The solution was spun down at 1,500x g for 10 min at 4°C, and the supernatant was transferred into a tube containing 19 ml 1% lamb serum/TBST and mixed well. The antibody solution was stored at 4°C in the dark until use. Following the last washing steps, the embryos were blocked in a fresh 12-well plate filled with 2 ml 10% lamb serum/TBST per well for 2-3 h at RT. The embryos were then transferred to a fresh 12-well plate containing 2 ml antibody solution per well and incubated o/n rocking at 4°C in the dark. The next day, the embryos were washed twice for 15 min, twice for 30 min, and at least six times for 1 h at RT in the dark with approximately 90 ml TBST. To reduce background staining, the specimens were incubated o.n. at 4°C in TBST in the dark.

2.6.2.6 Staining

The next day, the embryos were washed four times at RT for 15 min with 90 ml freshly prepared NTMT (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 50 mM MgCl2, 0.1% Tween-20). Meanwhile, the staining solution was prepared by adding 112.5 µl NBT (75 µl/µg; Roth) and 87.5 µl BCIP (50 µg/µl; Roth) to 25 ml ice-cold NTMT. The solution was filter-sterilized using a 0.45

µm syringe filter (Schleicher&Schuell) before filling 2 ml of it into the wells of a fresh 12-well plate. The embryos were then transferred into the staining solution and incubated rocking in the dark at RT. The staining intensity was monitored periodically under a binocular. Once an appropriate staining was obtained, the reaction was stopped by washing the embryos once in NTT (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 0.1% Tween-20) and several times in PBST at RT. The stained embryos were postfixed in 4% PFA/PBST and stored in the dark at 4°C.

2.6.2.7 Imaging

Whole-mount specimens were photographed with a SteREO Discovery.V12 microscope (Zeiss) and an AxioCam Color camera (Zeiss) using the AxioVision 4.6 software (Zeiss).

2.6.3 Immunohistochemistry on sections

PFA-fixed, paraffin-embedded tissue sections were deparaffinized in xylene (3× for 5 min) and rehydrated (100% EtOH, 2x for 5 min; 90% EtOH, 80% EtOH, ddH_2O, each once for 5 min). The slides were then incubated in 1% H_2O_2 for 10 min at RT in order to quench endogenous peroxidase activity. The sections were washed three times for 5 min in PBS before unmasking the epitopes. Unmasking was done by, boiling the slides in a glass beaker containing a buffer with moderately acidic pH (pH 6.0, 10 mM sodium citrate; pH adjusted with 1 M citric acid) for 20 min. After this heat-induced antigen retrieval, the glass beaker was placed in an ice bucket, and the buffer was cooled down to approximately 40°C. The slides were washed 3x for 5 min in PBS and blocked o.n. at 4°C with 2.5% horse serum (ImmPRESS REAGENT kit; Vector Laboratories). The next day, the sections were incubated with the primary antibody against T (Kispert and Herrmann, 1994), diluted 1:1000 in 2.5% horse serum, for 2 h at RT, inside a humidified chamber. Uncoupled antibody was removed by five washes in PBS for 5 min. The sections were then incubated with ImmPRESS anti-rabbit Ig reagent for 60 min at RT; This reagent contains micropolymers of a very active peroxidase coupled to affinity purified

secondary antibodies. The slides were washed three times for 3 min in PBS and incubated for 15 min at RT in peroxidase substrate solution (Vector NovaRED; Vector Laboratories). The slides were rinsed in tap water at RT for 5 min before counterstaining them for 5 sec in hematoxylin (Mayer's Hematoxylin Solution; Sigma-Aldrich). The sections were again rinsed in tap water for 5 min, dehydrated (dip into 90% EtOH, 100% EtOH 2x for 5 min, xylene once for 5-10 min), immediately mounted in Entellan (Merck) and covered with glass cover slips (Roth).

Tissue sections were photographed as before (2.6.1.4).

2.6.4 Skeleton staining

The procedure followed the protocol described in (Kessel et al., 1990). Mice were euthanized by CO_2 inhalation. The skin and all visceral organs were removed before fixing the carcasses for four days in 100% EtOH at RT on a roller mixer. They were incubated in 100% acetone for three days, rinsed in water, and incubated in staining solution (1 vol. 0.3% Alcian Blue 8 GX (Sigma-Aldrich) in 70% EtOH, 1 vol. 0.1% Alizarin Red S (Sigma-Aldrich) in 95% EtOH, 1 vol. 100% acetic acid, 17 vol. 100% EtOH) for 10 days on a roller mixer. Excess soft tissue was removed by incubation with Clearing Solution (20% glycerol, 1% KOH). For storage, the specimens were transferred into 50%, 80%, and finally 100% glycerol.

2.7 Primers, general chemicals and solutions

2.7.1 PCR Primers

Table 6. List of primers used and their sequence

primer name	gene	purpose	sequence 5' to 3'
neo_f	neomycin	genotyping	GAATGAACTGCAGGACGAGGC
neo_r	neomycin	genotyping	GATCCCCTCAGAAGAACTCGTC
T_f	T	qPCR	CAGCTGTCTGGGAGCCTGG
T_r	T	qPCR	TGCTGCCTGTGAGTCATAAC
GAPDH_f	GAPDH	qPCR	TCAAGAAGGTGGTGAAGCAG
GAPDH_r	GAPDH	qPCR	ACCACCCTGTTGCTGTAGCC
Foxa1_f	Foxc2	qPCR	GACATACCGACGCAGCTACA
Foxa1_r	Foxc2	qPCR	GGCACCTTGAGAAAGCAGTC
Noto_f	Noto	qPCR	ACCTGACGGAGAATCAGGTG
Noto_r	Noto	qPCR	ACTGCCAATTCCCAACTCAG
GFP_f	GFP	qPCR	CAAAGACCCCAACGAG
GFP-r	GFP	qPCR	CGGCGGCGGTCAACGAACT
mir155_SalI	mir155	subcloning	AGGTAGTAGGTCGACCAGTGGATC
mir155 MluI	mir155	subcloning	GCTAACGCGTCTCGAGTGCGGCCAGATCTG

2.7.2 General solutions

Table 7. Description of general solutions used in this work

Name	Composition
Gel loading buffer, 6×	1× TAE, 60 mM EDTA pH 8.0, 50% glycerol, 0.1% bromophenol blue
PBS(T)	137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH2PO4, (0.1% Tween-20)
SSC, 20×	300 mM trisodium citrate dihydrate, 3 M NaCl, pH adjusted with 1 M citric acid
TAE, 50×	2 M Tris, 950 mM acetic acid, 62.5 mM EDTA
TBE, 50×	4.5 M Tris, 4.5 M boric acid, 125 mM EDTA
TBS(T)	140 mM NaCl, 2.7 mM KCl, 25 mM Tris-HCl pH 7.5, (0.1% Tween-20)
TE	10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0
Tris-HCl	1 M Tris, pH adjusted with HCl (37%)

2.7.3 General Chemicals

Table 8. List of Chemicals

Name	Supplier
0.3 M mannitol	Sigma-Aldrich
Acetic acid (100%)	Merck
Acetone	Merck
Acrylamid/Bisacrylamid	Roth
Ammonium acetate	Merck
Chloramphenicol	Sigma-Aldrich
Chloroform	Roth
Citric acid	Merck
DMSO	Sigma-Aldrich
DTT	Sigma-Aldrich
EDTA	Sigma-Aldrich
Ethanol	Merck
Ethidium bromide (10%)	Roth
Formaldehyde (37%)	Fluka
Formamide	Merck
Genticin	Gibco
Glutaraldehyde	Sigma-Aldrich
Glycerol	Merck
Glycine	Merck
Glycogen	Peqlab
H_2O_2 (30%)	Roth
HCl (37%)	Merck
HEPES	Sigma-Aldrich
Hygromycin	Invitrogen
Isopropanol	Merck
KCl	Sigma-Aldrich
KOH	Merck
KSOM	Milipore
M2	Sigma-Aldrich
Methanol	Merck
$MgCl_2$	Roth
Na_2 HPO_4	Sigma-Alrich
NaCl	Merck
NaOH	Merck
Paraformaldehyde	Sigma-Aldrich
Phenol	Roth
SDS	Fluka
Sodium acetate	Sigma-Aldrich
Tris	Sigma-Aldrich
Trisodium citrate dihydrate	Merck
Tween-20	Sigma-Aldrich
Tyrode's Acid solution	Sigma-Aldrich
Urea	Sigma-Aldrich
Xylene	Roth

3 RESULTS - PART I

An Inducible RNAi system for the functional dissection of genes

Contributors

Joana A. Vidigal

Markus Morkel

Lars Wittler

Antje Brouwer-Lehmitz

Phillip Grote

Karol Macura

Bernhard G. Herrmann

3.1 Experimental contributions

Markus Morkel supervised the project, cloned the *KD4*, *KD2*, and *KD1-T* vectors, did the *in vitro* induction tests, validated the *T* shRNA-mirs and performed the transcriptome analysis.

Lars Wittler microdissected the embryos for expression profiling and helped with phenotyping.

Antje Brouwer-Lehmitz did the RMCE, cloned the *Foxa2* and *Noto* exchange vectors and validated their shRNA-mirs.

Phillip Grote generated the recipient cell lines.

Karol Macura performed some of the tetraploid complementation assays.

Bernhard G. Herrmann conceived and supervised the project.

All other experiments were done by me, including experimental design, promoter and insulation tests, cloning of exchange vectors, production of embryos by tetraploid complementation, whole mount *in situ* hybridization experiments, northern blots, and embryo phenotyping.

3.2 An integrated vector system to target the *ROSA26* locus by RMCE

3.2.1 Overview of the vector system

The *Gt*(*ROSA*)*26Sor* (*ROSA26*) locus has been widely used to generate knock-in mice since it allows strong ubiquitous expression of transgenes throughout embryonic development and adulthood (Friedrich and Soriano, 1991; Zambrowicz et al., 1997), its disruption does not lead to any overt phenotype (Zambrowicz et al., 1997), and homologous recombination occurs at a high frequency using available targeting vectors.

To achieve single-copy integration of RNAi constructs in the mouse genome, a bipartite system was designed to allow recombination of transgenes into the *ROSA26* locus by recombinase-mediated cassette exchange (RMCE) (Feng et al., 1999). RMCE in this locus is

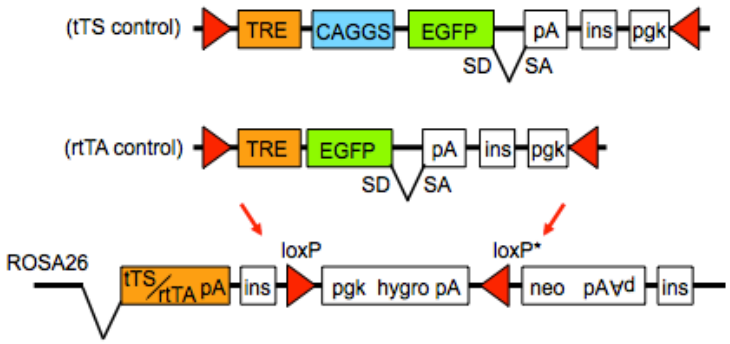


Figure 3.1. A generalized diagram of the vector system to target the *ROSA26* locus by RMCE

Top, example of exchange vectors compatible with tTS (up) or rtTA (down) control. Vectors can recombine into the recipient locus via asymmetric *lox* sites (*loxP* and *loxP**). Upon positive recombination, the PGK promoter, located at the 3' end of the exchange vector is positioned nearby the promoter-less neomycin resistance cassette (neo) allowing for positive selection. Bottom, the recipient construct is integrated into the *ROSA26* locus and encodes for either tTS or rtTA. In the presence of Doxycycline these transcriptional regulators, ubiquitously expressed by the ROSA26 promoter, lead to the expression of a transgene located nearby the TRE (Tetracycline responsive element; comprised of TetO sites in tTS controlled vectors and TetO sites and a minimal CMV promoter in rtTA controlled vectors). CAGGS, ubiquitous promoter; EGFP, enhanced green fluorescent proterin; SD, splice donor, SA, splice acceptor, pA, polyadenylation signal; ins, insulator sequence.

accomplished by integrating, through homologous recombination, a recipient vector carrying two opposing asymmetric *lox* sites flanking a hygromycin resistance cassette (see Figure 3.1 a generalized diagram of the vector system). In the presence of Cre and an exchange vector carrying reciprocal *lox* sites, the recombinase will favor an integrative recombination over a excisive reaction (Albert et al., 1995). Such event leads to the replacement of the hygromycin cassette by the exchange vector. Upon correct integration, the PGK promoter in the 3' region of the exchange vector is placed upstream of a promoter-less neomycin resistance cassette, allowing for easy positive selection of recombined clones (see Figure 3.1).

In the knockdown exchange vectors, the RNAi trigger is transcribed via an shRNA-mir transgene (see Figure 3.10). Like protein-coding genes, shRNA-mirs can be expressed by pol II promoters, allowing for the integration of inducibility by placing the transgene under the control of an rtTA/tTS responsive element (TRE; Figure 3.1). Expression of the transgene is controlled either by the tTS or the rtTA transcriptional regulators, located in the 5' region of the recipient vector, and expressed ubiquitously through the *ROSA26* promoter (Friedrich and Soriano, 1991). Upon recombination, tTS and rtTA can bind to the TRE in the exchange vector, leading to expression of the transgene only in the presence of doxycycline (dox).

3.2.2 Initial targeting of the *ROSA26* locus

The recipient transgene was assembled in the pBluescript plasmid and contained the following elements (5' to 3'): an SV40 splice acceptor fused to a murinized version of tTS or rtTA with a SV40 polyadenylation signal, the 1.2kb chicken β-globin insulator, a PGK-hygromycin selection cassette flanked by opposing *loxP* and *lox5171* sites (Hoess et al., 1982) and a promoter-less neomycin resistance gene with a bidirectional polyadenylation signal followed by another chicken β-globin insulator (Figure 3.2a). To target the *ROSA26* locus, the recipient transgene was inserted into the *XbaI* site of pROSA26-1 plasmid (Soriano, 1999).

This recipient targeting construct was linearized with *PacI* and electroporated into G4 ES cells (George et al., 2007). Following hygromycin selection, colonies that had undergone homologous recombination were identified by Southern-blot analysis of genomic DNA digested with BamHI, using a probe external to the 5' homology arm. Positive recombination led to detection of a 4.7 Kb band, as compared to a 5.8 kb band for the wt allele (Figure 3.2b). Positive clones were expanded and are referred to as *ROSA26S* – if targeted with a vector containing the tTS – or *ROSA26A* – if targeted with a vector containing the rtTA. *ROSA26S* and *ROSA26A* clones gave rise to highly chimeric, healthy and fertile mice by diploid aggregation, and to

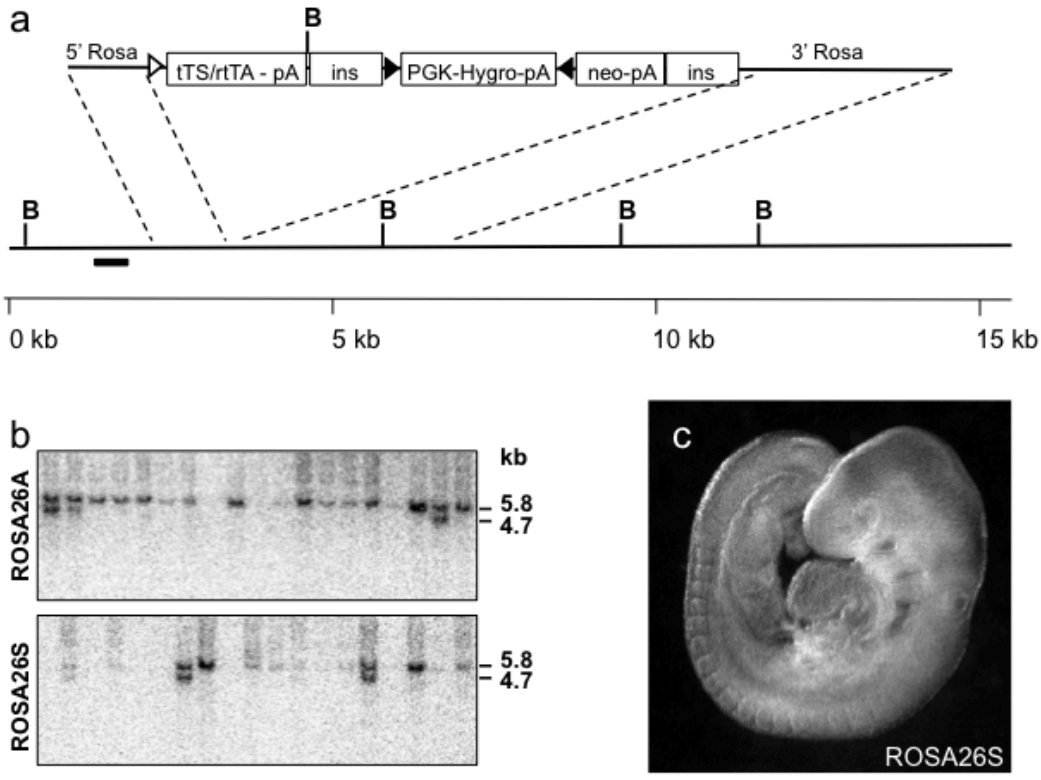


Figure 3.2. Targeting of the *ROSA26* locus

(a) Top, schematic representation of *ROSA26A*/*ROSA26S* targeting constructs. Trans-splicing sequence is shown a white triangle. Asymmetric loxP sites are shown as black triangles. The probe used for Southern blot analysis is depicted as a black box. Bottom, restriction map of the locus showing BamHI cleavage sites. (b) Southern blot analysis of clones targeted with ROSA26A (top) and ROSA26S (bottom) constructs and digested with *BamHI*. Correct targeting results in detection of a 4.7 kb band, as compared to a 5.8 kb band for the wt alelle. (c) Correctly targeted clones gave rise to phenotypically normal E9.5 embryos by tetraploid complementation. tTS, tetracycline-controlled transcriptional silencer; rtTA, reverse tetracycline-controlled transcriptional activator; pA, polyadenylation signal; ins, insulator sequence; Hygro, hygromycin resistance gene; neo, neomycin resistance gene; 3'/5' Rosa, homology arms to targeted *ROSA26* locus; B, *BamHI* cleavage site; kb, kilobases; PGK, PGK promoter.

phenotypically normal embryos by tetraploid aggregation (Figure 3.2c).

3.2.3 Integration of transgenes in the modified *ROSA26* by RMCE

To test the efficiency of RMCE into the modified *ROSA26* locus, an exchange vector carrying an enhanced green fluorescence protein (EGFP) reporter (Chalfie et al., 1994), expressed under the ubiquitous CAGGS promoter (Niwa et al., 1991) was co-transfected with a Cre expression vector into *ROSA26S* cells (Figure 3.3a). Transfectants were selected with neomycin and resistant colonies screened for positive recombination by Southern blot using a neomycin probe on genomic DNA digested with *BamHI*/*HindIII*. Positive recombination led to detection of a 2.7 kb band as opposed to the 4.1 kb band detected in the parental cell line (Figure 3.3b), and occurred at a high frequency among resistant clones.

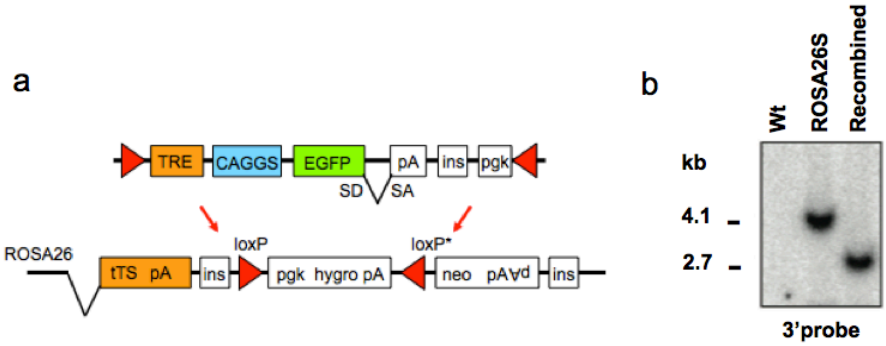


Figure 3.3. Integration of a reporter into the ROSA26 locus by RMCE

(a) Schematic representation of the vector system. The exchange vector (top), expressing an EGFP reporter, can integrate into the recipient locus (bottom) by recombination through asymmetric *lox* sites (*loxP* and *loxP**). (b) positive recombination leads to a band shift from 4.1 to 2.7kbs as detected by Southern blot with neomycin probe. Abbreviations as before.

3.2.4 *In vitro* expression of a transgene from the *ROSA26* locus

To determine whether expression of integrated transgenes was responsive to tetracycline analogs in the medium, targeted cells, carrying the EGFP transgene, were cultured in the presence or absence of dox. In the presence of this drug, cells produced EGFP protein as determined by fluorescence microscopy (Figure 3.4a), while in the absence of dox they did not. Additionally, the transgene could be reversibly induced *in vitro* for several rounds of induction/de-induction (Figure 3.4b).

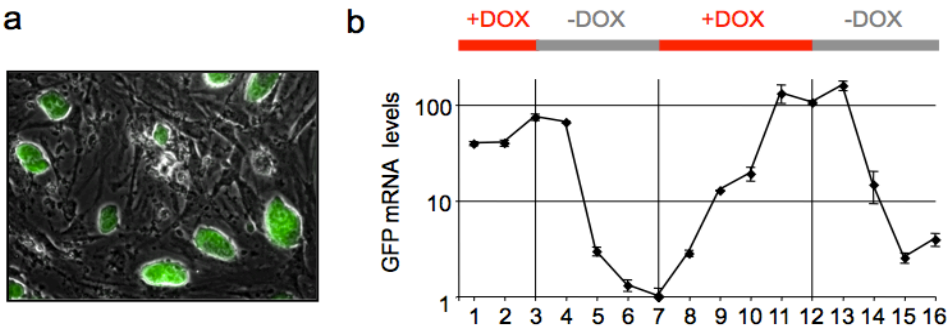


Figure 3.4. *In vitro* expression of a reporter integrated in the *ROSA26* locus by RMCE

(a) Visualization of EGFP in ES cell clones cultured in the presence of Dox. (b) Reversible induction of transgene expression *in vitro*. ES cells carrying exchange vector expressing EGFP under the CAGGS promoter and under tTS control were subjected to cycles of repression (−Dox) and re-induction (+Dox) for 16 days. Levels of EGFP transgene expression were measured by qRT-PCR. Error bars represent technical replicates.

3.2.5 *In vivo* expression of transgenes from the *ROSA26* locus

Next, to test the expression of transgenes *in vivo*, ES cells carrying the EGFP reporter were used to generate embryos in a tetraploid complementation assay. In this assay, all tissues from the embryo proper derive from the modified ES cells (see Figure 1.6) and therefore, expression of a transgene can be directly analyzed with no need to establish a mouse line. At the blastocyst stage, chimeric embryos - composed of an ES cell-derived inner cell mass and

tetraploid visceral endoderm and trophoectoderm - were transferred into the uterus of pseudopregnant females, and expression of the transgene induced by addition of dox to the drinking water.

At E9.5, induced embryos carrying EGFP under the CAGGS promoter showed expression of the reporter in all tissues (Figure 3.5). These results are in line with earlier reports demonstrating that transgenes integrated into *ROSA26* are expressed ubiquitously throughout development (Soriano, 1999). Additionally, transgene activation levels were similar among induced littermates (Figure 3.6a). When transgene expression was not induced, levels of EGFP remained undetectable both by fluorescence microscopy and quantitative PCR (see Figure 3.5 and Figure 3.6a) showing that the system is tight.

To test whether transgenes could also be expressed in a tissue-specific manner, ES cells

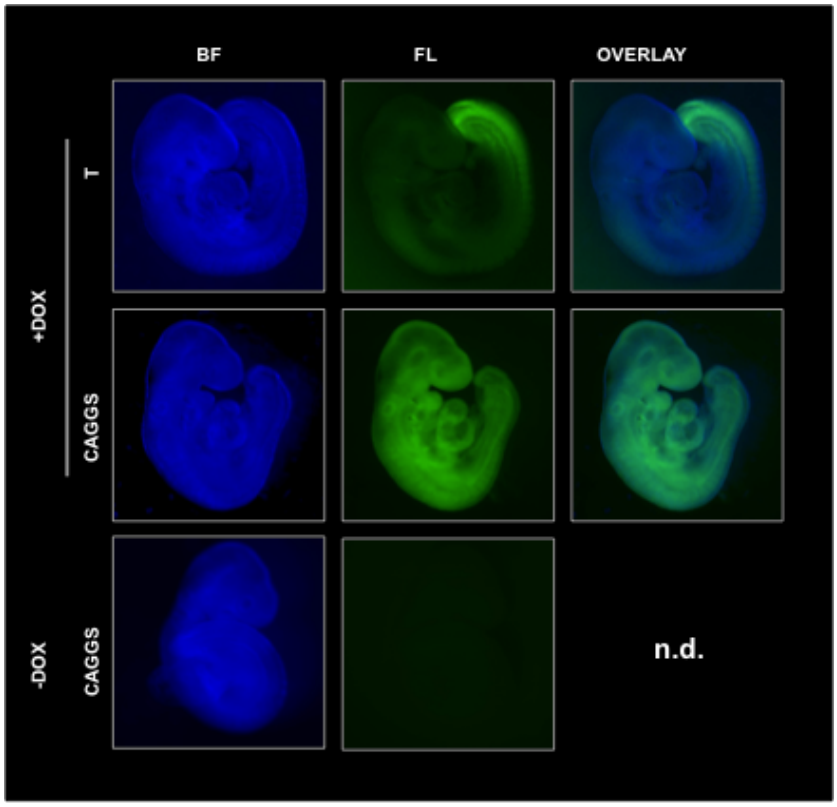


Figure 3.5. *In vivo* expression of a reporter from the ROSA26 locus

Fluorescence images of E9.5 control (−Dox) or induced (+Dox) embryos showing ubiquitous EGFP expression under control of the CAGGS promoter and tissue-specific expression under control of the T-streak promoter. BF, bright field; FL, fluorescence; n.d. not displayed.

carrying an EGFP reporter under the control of the T-streak promoter were derived and used to generate embryos by tetraploid complementation as before. Once again, non-induced embryos showed absence of reporter expression, whereas Induced embryos had detectable expression of EGFP in the caudal end (Figure 3.5), consistent with the expression pattern described for this promoter fragment (Perantoni et al., 2005).

Importantly, not all tested ubiquitous promoters gave a satisfactory expression pattern *in vivo*. In particular, expression of an EGFP reporter by the cytomegalovirus (CMV) promoter (Niwa et al., 1991) resulted in a patchy expression pattern in E9.5 induced embryos (Figure 3.6b), inconsistent activation of the transgene among induced littermates and low activation levels in induced embryos (Figure 3.6c).

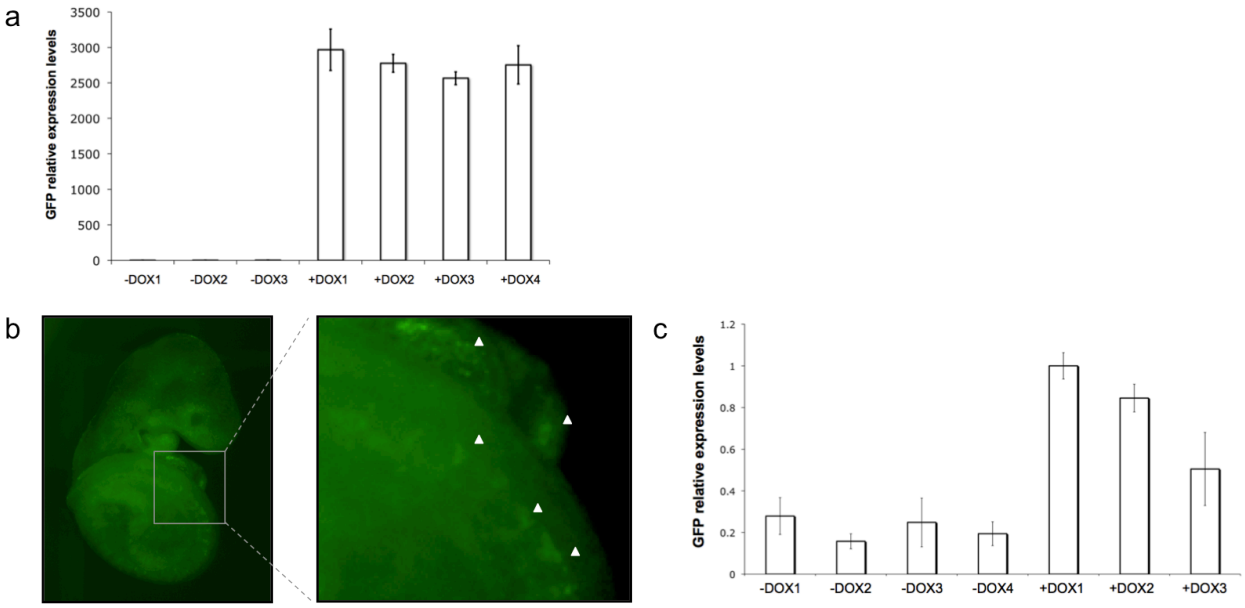


Figure 3.6. *in vivo* expression of EGFP under the control of CAGGS and CMV promoters

(a and c) Quantification of EGFP mRNA levels by qPCR in induced (+Dox) and un-induced (-Dox) embryos expressing EGFP under the control of the CAGGS (a) or the CMV (c) promoter. Each bar represents a biological replicate. Error bars represent technical replicates. (b) Fluorescent images of an E9.5 induced embryo showing patchy reporter expression under control of the CMV promoter. A close-up of the boxed region is shown on the right, with examples of expression hotspots marked with arrow heads.

Taken together these results show that the RMCE system in the *ROSA26* locus allows for the expression of single copy transgenes in both ubiquitous and tissue-specific manners and that expression of the transgene is responsive to the presence of doxycycline.

3.2.6 Insulation of transgenes in the *ROSA26* locus.

Importantly, tissue-specific expression like the one shown in Figure 3.5 could only be achieved when transgenes were insulated from surrounding genomic sequences. Expression of genes in eukaryotic organisms often relies on regulation by enhancer sequences which are recognized and bound by transcription factors that help initiating transcription from the promoter. Enhancers can act independently of their distance or location to a gene giving them the potential

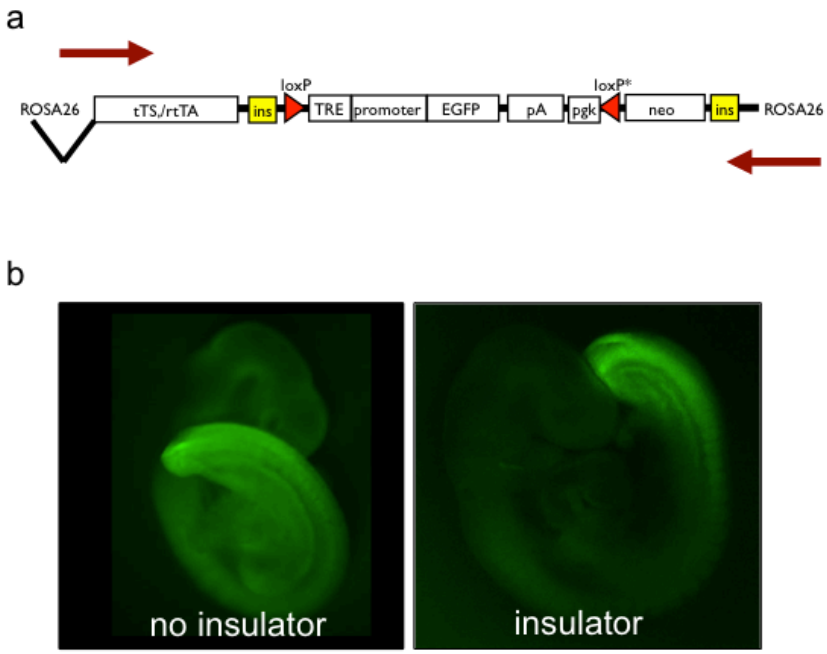


Figure 3.7. Effects of insulators on the tissue-specific expression of a transgene

(a) Schematic representation of the ROSA26 locus after RMCE. Insulator (ins) sequences at the 5' and 3' regions of the recipient locus prevent transcriptional interference from the genomic locus (red arrows). (b) Fluorescent images of E9.5 induced embryos showing the expression of the EGFP reporter under the T-streak promoter in the absence (left) and presence (right) of insulator sequences.

to be promiscuous and leading to improper activation of transgenes (Gaszner and Felsenfeld, 2006). Insulator sequences can be used to avoid influences from genomic regulatory elements on transgene transcription since they act as transcriptional buffers by interfering with enhancer-promoter interactions when located between the two (Mongelard and Corces, 2001).

To test the effects of insulation on the expression of a reporter in the *ROSA26* locus, ES cells expressing EGFP under the tissue-specific T-streak promoter were generated, which either carried or not insulator sequences flanking the RMCE site (Figure 3.7). Embryos were derived by tetraploid complementation and transgene expression induced by dox treatment as described earlier. In the absence of insulation, EGFP was ectopically expressed throughout the embryo (Figure 3.7b; left). The addition of insulator sequences up- and downstream of the expression cassette eliminated this unspecificity, leading to correct EGFP expression only at the caudal end of the embryo at E9.5 (Figure 3.7b; right).

3.3 Knockdown of genes from the *ROSA26* locus

3.3.1 Proof of principle: knockdown of brachyury

The results described above show that the *ROSA26* locus can be used for RMCE of transgenes via asymmetric *lox* sites, and that expression of transgenes in this locus can be controlled both temporally and spatially. To determine whether this system is also suitable for the expression of RNAi constructs, a panel of exchange vectors targeting brachyury, a key player in mouse development, was designed and tested for efficient *in vivo* knockdown.

Around E6.0 of mouse development, the embryo becomes asymmetric along both its proximal-distal and its anterior-posterior axes through processes of cell differentiation, migration, and embryonic patterning. Anterior-posterior symmetry is lost with the migration of cells from the distal visceral endoderm to the anterior region of the embryo, forming the 'anterior visceral endoderm'. After this initial loss of symmetry, the primitive streak forms at the opposing end, defining the posterior region of the embryo (Figure 3.8). During gastrulation, epiblast cells

Figure 3.8. The expression and function of brachyury during mouse development

(a) Schematic representation of a sagittal section through a E7.0 mouse embryo showing the primitive streak (purple) in the posterior region of the embryo, the epiblast (pink), the primitive endoderm (light green), the anterior visceral endoderm (red) plus the extra-embryonic visceral endoderm and extra-embryonic ectoderm in dark green and light purple respectively. Modified from (Arnold and Robertson, 2009). (b) Expression of *brachyury* visualized by whole mount *in situ* hybridization. Brachyury is expressed in the primitive streak from the onset of gastrulation (left; E7.5 embryo). Later, its expression becomes restricted to the tailbud (tb) and the notochord (nt; right; E9.5 embryo). (c) Mutant embryos lacking brachyury generate insufficient mesoderm leading to truncation of posterior structures at E9.5. Images courtesy of Pedro Rocha (b, left) and Markus Morkel (c) or adapted from the MAMEP database (b, right; http://mamep.molgen.mpg.de/). Prox, proximal; Dist. Distal; Ant/A, anterior; Post/P, posterior.

undergo an epithelial to mesenchymal transition, ingress through the primitive streak and give rise to two new germ layers, the mesoderm and the definitive endoderm (Arnold and Robertson, 2009). This process is controlled by multiple signals including fibroblast growth factors (Fgfs), Wnts, and transforming growth factors (TGF) (Kimelman, 2006). Brachyury, a transcription factor encoded by the *T* gene, located on chromosome 17, is one of the best characterized direct Wnt targets during primitive streak formation (Yamaguchi et al., 1999b). *T* is expressed in the primitive streak from the onset of gastrulation (Figure 3.8b; left). Later, its expression becomes restricted to the caudal region of the embryo and to the notochord (Figure 3.8b; right), a mesoderm-derived rod-shaped structure located ventral of the neural tube that functions as a signaling center during development and serves as a template for the formation of the vertebral column (Stemple, 2005).

Brachyury was shown to be required for both mesoderm formation and notochord differentiation. Homozygous embryos lacking Brachyury activity do not form sufficient mesoderm due to cell migration defects (Wilson et al., 1993), which result in truncation of posterior structures. In addition, the notochord is not established, and the allantois, another mesoderm derived structure fails to fuse to the chorion leading to embryonic lethality around E10.0 due to absence of a functional placenta (Wilson et al., 1993)(Figure 3.8c).

Interestingly, rescue experiments have shown a direct correlation between the levels of brachyury expression and the extension of the anterior-posterior axis (Stott et al., 1993), which is in line with the phenotypes observed for heterozygous mutants, that survive embryonic development but are born with shortened tails (Wilson et al., 1993). This makes brachyury a very appealing choice as a proof of principle gene for the generation of RNAi loss-of-function mutants, since the effectiveness of its *in vivo* knockdown can be easily assessed by the degree of axis truncation it causes.

3.3.2 *In vitro* tests of shRNA-mirs targeting brachyury

To assess their knockdown efficacy *in vitro*, four individual shRNA-mirs targeting brachyury were identified by queering the *Block it* database (Invitrogen), and co-transfected with the murine *T* cDNA into human SW480 cells, and the levels of T transcript detected by qPCR 48 hours later.

All four shRNA-mirs led to a significant silencing of the brachyury mRNA when expressed from the pCDNA6.2GW vector (T-267: knockdown of 90.8%, T-527: 89.4%, T-709: 87.7%, T-1457: 91.5%) while mir155 sequences alone or containing an siRNA against an unrelated gene had no effect (Figure 3.9a). The two most efficacious shRNA-mirs, T-267 and T-1457 were used for further *in vivo* experiments. These shRNA-mirs target the brachyury mRNA in its first exon or in its 3'UTR (for T-267 and T-1457 respectively; see Figure 3.9b for a diagram).

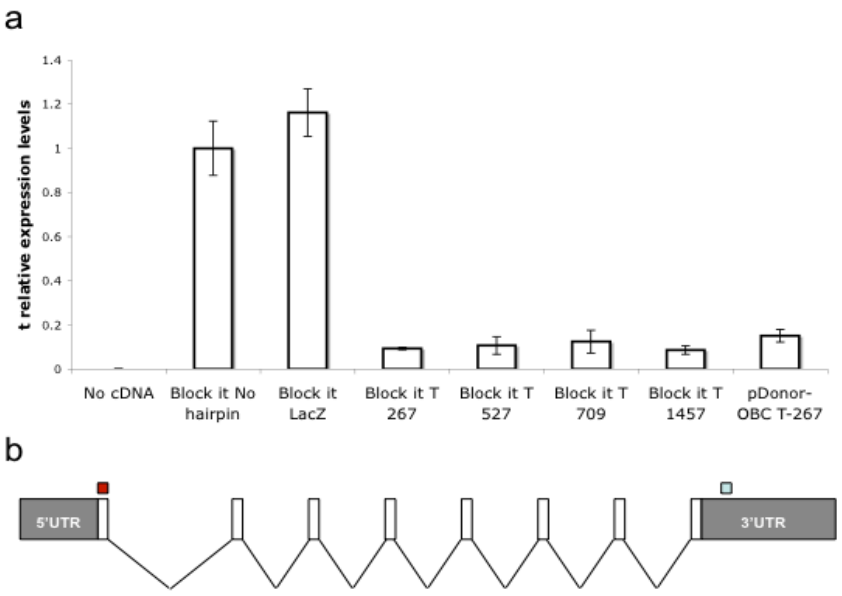


Figure 3.9. In vitro efficacy of shRNA-mirs targeting brachyury

(a) pCDNA6.2GW (Block it) or exchange (KD3) vectors containing either no hairpin sequence (empty), sequences directed against unrelated genes (*LacZ*) or sequences against *T* were co-transfected with an expression vector for *T*. Target mRNA was quantified by qRT-PCR and normalized to transfections containing unrelated shRNA-mirs. Error bars give range of technical replicates. (b) Schematic representation of the brachyury mRNA showing recognition regions for siRNA sequence 267 (red) and 1457 (blue). Untranslated regions are depicted in grey, exons in white and introns as connective lines between exons.

3.3.3 Integration of *T* knockdown constructs in *ROSA26S*

The shRNA-mirs T-267 and T-1457 were used to clone a panel of exchange vectors under the control of tTS (KD1 to KD3; see Figure 3.10). Importantly, this cloning did not compromise the knockdown efficacy of individual shRNA-mirs (Figure 3.9b, far right). These exchange vectors contained either both shRNA-mirs (KD1), the T-267 shRNA-mir alone (KD2) or an intronic T-267 shRNA-mir and the coding sequence of EGFP (KD3), all under the control of the CAGGS promoter. In this setup, the tTS was ubiquitously expressed under control of the *ROSA26* promoter and therefore represses the RNAi constructs until the block is released by exposure of the cells/embryos to doxycycline.

Transient transfection of *ROSA26S* cells with *KD1-KD3* vectors and a Cre recombinase expression vector efficiently produced clones with a complete combined transgene system. Positive ES cell clones were named *KD1-T*, *KD2-T* or *KD3-T* depending on the vector used for recombination.

3.3.4 Expression and processing of shRNA-mirs from the *ROSA26* locus

The T-267 and T-1457 shRNA-mirs were designed to mimic the structure of the murine microRNA mir155 (Moffett and Novina, 2007). They were generated by substituting the sequence corresponding to the mature miRNA, with the sequence of the artificial siRNAs (Figure 3.10b). In order to successfully silence their target mRNA, they have to undergo a series of processing steps which result in the production of a 21nt long siRNA (see Figure 1.5 for an overview of the biogenesis pathway).

To determine if T-267 and T-1457 were correctly processed from the exchange vectors following integration into the mouse genome, *KD1-T* cells, containing both shRNA-mirs were cultured with or without addition of doxycyline to the medium. After three days of induction, total

RNA was isolated, and the presence of processed siRNA confirmed by small RNA Northern blot. When cells were cultured with dox, both siRNAs could be detected co-migrating with oligonucleotides corresponding to their mature sequence (Figure 3.10c). In contrast, RNA from cells cultured in the absence of dox had undetectable levels of both shRNA-mirs. These results show that processing of the primary shRNA-mir transcript occurs correctly after integration of the RNAi constructs in the *ROSA26* locus, and that the RNAi triggers are not expressed in the absence of dox.

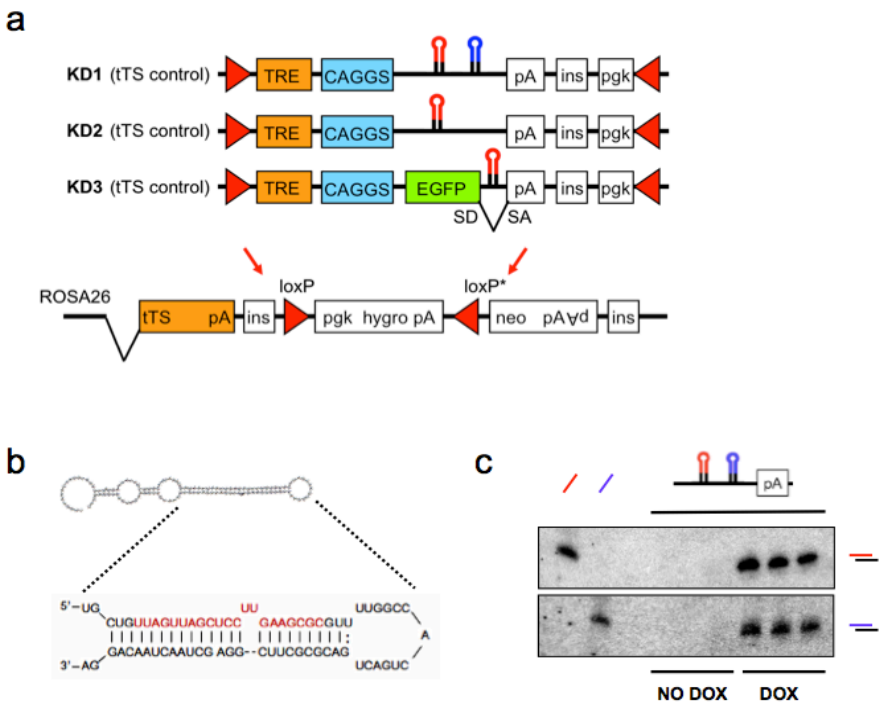


Figure 3.10. Exchange vectors for inducible *in vivo* knockdown of brachyury

(a) Schematic representation of the recipient locus and exchange vectors KD1-KD3 used for RNAi. T-267 and T-1457 shRNA-mirs are represented as red and blue hairpins respectively. (b) Secondary structure of shRNA-mir T-267, as pedicted by *RNAfold* (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). Native mir155 sequences are shown in black, T-267 siRNA sequence is shown in red. (c) Expression and processing of T-267 and T-1457 shRNA-mirs in induced (DOX) and non-induced (NO DOX) ES cells. Detected bands co-migrated with a 21mer oligonucleotide, indicating correct processing of the siRNA from the shRNA-mir precursors.

3.3.5 *In vivo* knockdown of brachyury

To test the knockdown efficiency achieved by each of the exchange vectors, ES cells carrying the *KD1* to *KD3* constructs targeting brachyury were used to generate embryos by tetraploid complementation. As before, activation of the transgene was done by giving dox to the foster mothers in the drinking water. For all three RNAi constructs, control embryos derived in the absence of dox were phenotypically normal (128/128 embryos; Figure 3.11).

Embryos induced with doxycycline, however exhibited various malformations. Embryos expressing the *KD1-T* construct were indistinguishable from *T/T* loss of function mutants (98/98; compare embryos in Figure 3.11 and Figure 3.8c). When analyzed at E9.5, these embryos showed formation of only a few somites - discrete segments of mesoderm along both sides of the neural tube that give rise to dermis, skeletal muscle and vertebrae – marked by *Uncx*, as well as an early arrest of posterior development. In addition, these embryos did not differentiate a notochord (marked by *Shh*) and failed to generate a functional placenta, leading to embryonic death around E10.0. Embryos expressing the *KD2-T* construct showed either the loss-of-function

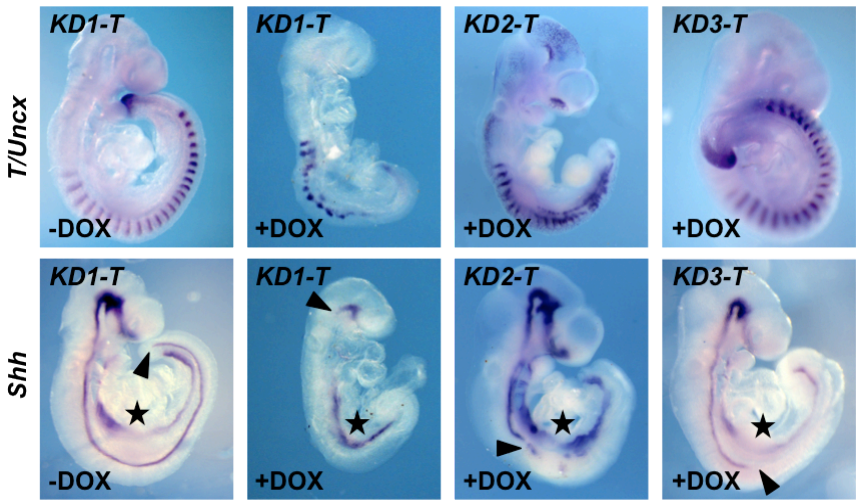


Figure 3.11. Analysis of phenotypes obtained by *in vivo* knockdown of *T*

Whole mount *in situ* hybridizations of E9.5-10.0 transgenic embryos carrying *T* knockdown constructs. Upper panel shows simultaneous detection of *T* and *Uncx* transcripts, lower panel shows localization of *Shh* transcripts. RNAi constructs are indicated at the top, induction schemes on the bottom of each panel. Arrow heads mark the caudal-most point of the notochord, asterisks mark endoderm staining by the *Shh* specific probe.

phenotype of *KD1-T* embryos (18/33) or a weaker phenotype (15/33) characterized by axial elongation up to the tailbud, formation of more though still irregular somites and notochord differentiation up to the forelimb bud (Figure 3.11). An even milder phenotype was observed in *KD3-T* induced embryos (23/23), in which notochord formation was also arrested at the level of the forelimb bud but somite formation was unperturbed. A portion of these embryos showed proper allantoic development, and their caudal ends displayed brachyury transcripts, which were not detectable in *KD1-T* and *KD2-T* embryos (Figure 3.11).

In all embryos of this series tail outgrowth failed or was abnormal, showing a *T* RNAi phenotype. Thus, in addition to the loss of function phenotype, embryos with hypomorphic phenotypes where generated, which are rarely observed in standard knockout alleles.

3.3.6 *T* shRNA-mir expression and processing in embryos

To see whether shRNA-mirs were correctly expressed and processed *in vivo*, RNA was isolated from control and induced *KD1-T* to *KD3-T* embryos at embryonic day 9.5 and probed by small RNA northern blot for the presence of mature T-267 siRNA.

In the absence of doxycycline, no T-267 siRNA was detected (Figure 3.12). In contrast, when embryos were derived in the presence of dox, all three constructs led to the expression of T-267. Importantly, the detected bands co-migrated with a 21mer oligonucleotide, indicating

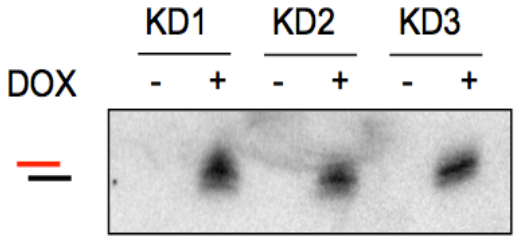


Figure 3.12. Expression and processing of T-267 shRNA-mir in mouse embryos
Total RNA was isolated from induced (+DOX) and control (-DOX) *KD1-T*, *KD2-T* and *KD3-T* embryos and probed by small RNA Northern blot for the presence of processed T-267 siRNA.

correct processing of the siRNA from the shRNA-mir precursor transcript. Surprisingly, although *KD2-T* and *KD3-T* constructs result in distinct phenotypes *in vivo* (Figure 3.11), no differences in processing or expression levels were detected (Figure 3.12).

3.3.7 System validation: *in vivo* knockdown of *Foxa2*

The forkhead box A2 (*Foxa2*) gene, codes for a member of the winged-helix family of transcription factors (Lai et al., 1991; Weigel and Jackle, 1990), and is expressed in the node, notochord, floor plate of the neural tube and gut of mouse embryos (Figure 3.13b). Like the notochord, both the node and the floor plate are signaling centers involved in embryonic patterning. The node is located immediately anterior to the primitive streak (Figure 3.15b), and secretes differentiation factors important for gastrulation and establishment of left-right asymmetry (Wolpert, 2007). The floor plate is a structure located at the ventral midline of the neural tube (Figure 3.13a) and functions as an organizing center by inducing the ventralization of tissues and guiding neuron differentiation along the dorsoventral axis of the neural tube (Wolpert, 2007).

Foxa2 is essential for the formation of both node and notochord. Mice lacking Foxa2 die

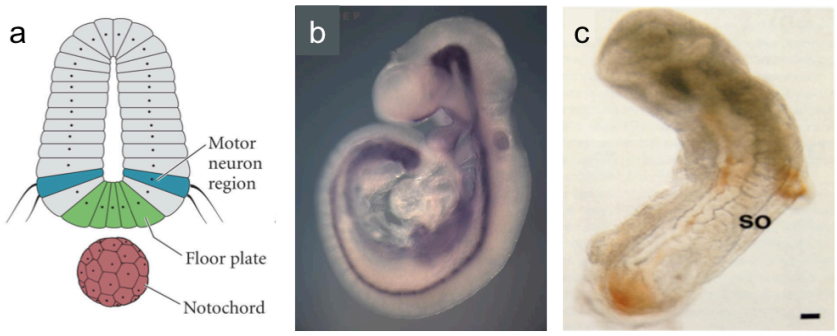


Figure 3.13 Expression and function of F*oxa2* during mouse development

(a) Schematic representation of a transverse section through a mouse embryo, showing both the floor plate (green) and the notochord (red) where *Foxa2* is expressed. The motor neuron region is patterned by the floor plate and is shown in blue. (b) *Foxa2* expression at E9.5 visualized by whole mount *in situ* staining. (c) Phenotypic abnormalities in *Foxa2* loss-of-function mutants. Figures were adapted from (Gilbert, 2006) (a), the MAMEP database (b; http://mamep.molgen.mpg.de/) and (Ang and Rossant, 1994) (c).

around E10.5 showing generalized patterning defects, which affect the somites and the neural tube and lead to embryos with truncation of anterior structures (Ang and Rossant, 1994; Weinstein et al., 1994), see Figure 3.13c. To see if such phenotypical abnormalities could be phenocopied by a *Foxa2* RNAi mutant, shRNA-mirs were selected in vitro as before (Figure 3.16a). Two efficacious hairpins targeting its mRNA in the first and second exons (Figure 3.16b, top) were identified, and used to generate *KD1-Foxa2* and *KD3-Foxa2* exchange vectors as before (see Figure 3.10 for schematic representation of vectors).

ES cells containing the vectors *KD1-Foxa2* and *KD3-Foxa2* were generated by RMCE with *ROSA26S* cells and used to generate transgenic embryos as before. Again, control embryos raised in the absence of doxycycline were phenotypically normal (9/9). Induction of the *KD1-Foxa2* construct resulted in major malformations in a fraction of embryos (4/10), resembling those described for *Foxa2* loss-of-function mutants (Figure 3.14).

Induction of the *KD3-Foxa2* transgene resulted in a milder phenotype, with a fraction of embryos showing only truncation of the forebrain at E9.5 (2/7;Figure 3.14). This phenotype has not been reported previously for *Foxa2* mutants, but it is in agreement with studies showing that its expression in axial mesoderm is required for proper forebrain specification through genetic interaction with *Otx2* (Hallonet et al., 2002; Kimura-Yoshida et al., 2007). Thus, *KD3-Foxa2*

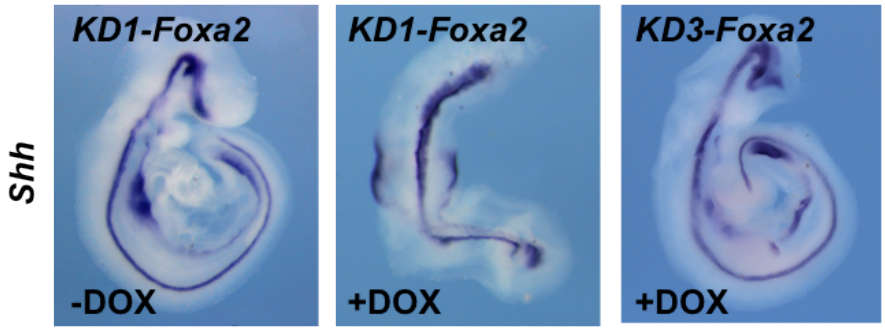


Figure 3.14. Analysis of embryonic phenotypes obtained by *in vivo* knockdown of Foxa2

Detection of *Shh* transcripts by whole mount *in-situ* hybridization of E9.5 transgenic embryos carrying *Foxa2* knockdown constructs. RNAi constructs are indicated on the top, induction schemes on the bottom

induced embryos give a hypomorphic phenotype for this gene, and should be a useful model to study the role of *Foxa2* in the differentiation of anterior structures during mouse development.

3.3.8 System validation: *in vivo* knockdown of *Noto*

Notochord homolog (Noto) is a homeobox transcription factor expressed in the node and in the nascent notochord. It acts downstream of the organizer genes *Foxa2* and *T* and it was shown to be required for the formation of the caudal part of the notochord, as well as node morphogenesis and ciliogenesis (Beckers et al., 2007). In the mouse, the node cilia rotate in a clockwise direction and this movement generates a leftward flow that leads to asymmetric distribution of morphogens between the left and right sides of the body (Nonaka et al., 2002) Mutations that affect cilia development, like that described for *Noto*, result in disruption of this

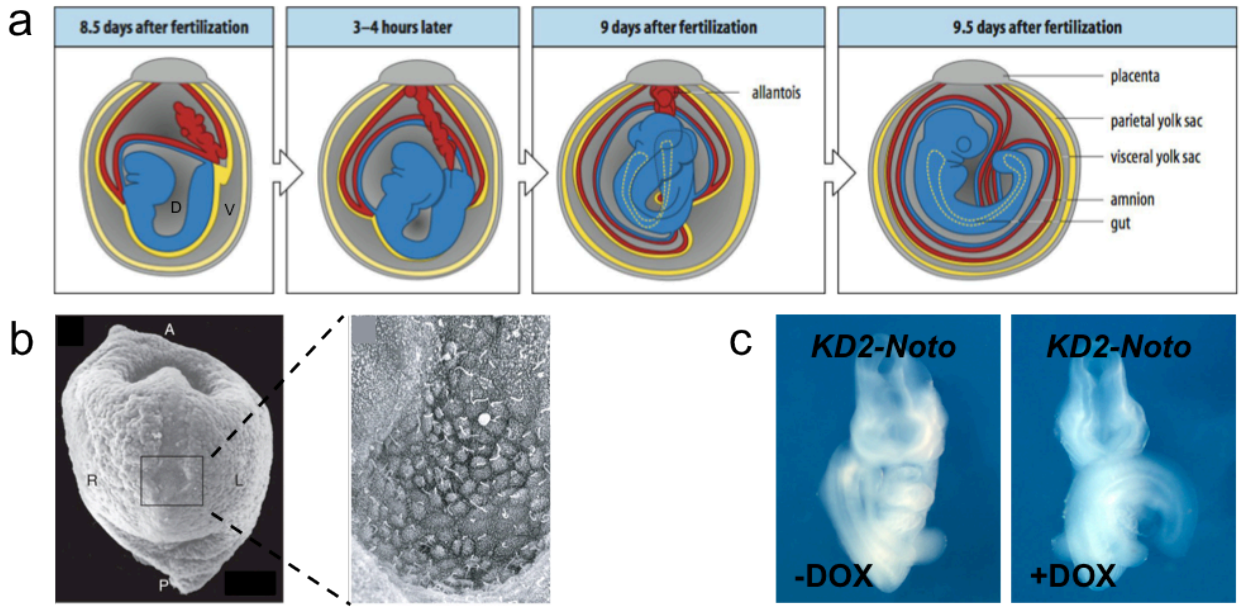


Figure 3.15 Lack of *Noto* leads to absence of nodal cilia and randomized embryonic turning

(a) Turning in the mouse embryo. Between E8.5 and E9.5 the embryo rotates on its axis and becomes surrounded by its extra-embryonic membranes. (b) Scanning electron micrographs of the mouse node at E7.5; ventral view. A magnification of the boxed region on the right shows the nodal cilia, involved in the establishment of left-right asymmetry. (c) Embryonic phenotypes obtained by knockdown of *Noto*. Induction of the *KD2-Noto* transgene (+Dox) leads to abnormal embryonic turning, detected by the position of the tail at E9.5 (ventral view). Turning in control embryos (-Dox) occurs normally. D, dorsal; V, ventral; A, anterior; P, posterior; R, right; L, left. Figures (a) and (b) adapted from (Wolpert, 2007) and (Hirokawa et al., 2009) respectively.

flow and randomization of the body situs, i.e., the position and orientation of organs in the body (Nonaka et al., 1998; Okada et al., 1999). Left-right asymmetry is first apparent at embryonic day 8.5, in the orientation of the heart looping. Later, around E9.0, asymmetry defines the direction of the embryonic turning, a process during which the embryo rotates on its axis so that the dorsal surface - initially facing inwards - ends up facing out, and the embryo becomes surrounded by its extra-embryonic membranes (Figure 3.15a-b). The direction of turning determines the position of the tail at E9.5 which, in wild-type embryos, is always positioned to the right. When left-right asymmetry is disrupted, as in *Noto* mutants, the turning occurs randomly resulting in a fraction of the embryos with the tail to the left (Beckers et al., 2007).

In order to have an independent validation of the RNAi system, an exchange vector targeting *Noto* was designed. The resulting transgene, *KD2-Noto*, carries an efficacious hairpin that targets the first exon of the mRNA, leading to an 80% reduction of its abundance when tested *in vitro* (Figure 3.16 and Figure 3.10 for an overview of the vectors). As before, cells carrying this construct were used to derive embryos by tetraploid aggregation in the presence or absence of dox. At embryonic day 9.5, control embryos looked phenotypically normal (13/13). Induced embryos on the other hand showed defects in embryonic turning (2/14; Figure 3.15c), in line with the *Noto* loss-of-function mice, which also show turning defects with incomplete penetrance (Beckers et al., 2007).

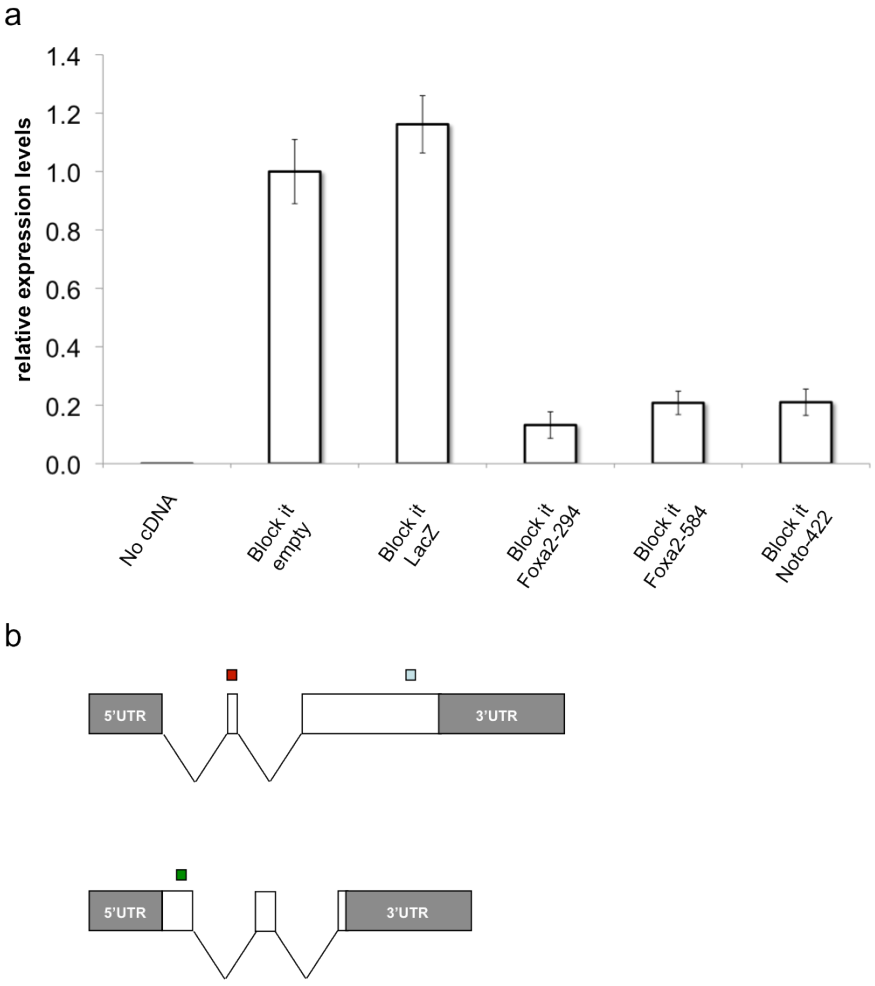


Figure 3.16. *In vitro* validation of shRNA-mir against *Foxa2* and *Noto*

(a) pCDNA6.2GW vectors (Block it) containing either no hairpin sequence (empty), sequences directed against unrelated genes (*LacZ*) or sequences against *Foxa2* or *Noto* were co-transfected with expression vectors for these genes. Target mRNA was quantified by qRT-PCR and normalized to transfections containing unrelated shRNA-mirs. Error bars give range of technical replicates. (b) Schematic representation of the *Foxa2* (top) and *Noto* (bottom) mRNAs showing recognition regions for siRNA sequence 294 (red), 584 (blue) or 422 (green). Untranslated regions are depicted in grey, exons in white and introns as connective lines between exons.

Taken together, the successful *in vivo* knockdown of *T*, *Foxa2* and *Noto* show that RNAi via the shRNA-mir transgene system described here is robust and well-suited for generating phenocopies of loss of function as well as hypomorphic phenotypes in mouse embryos.

3.4 Temporally controlled *in vivo* gene silencing

3.4.1 tTS leads to irreversible transgene silencing during mouse development

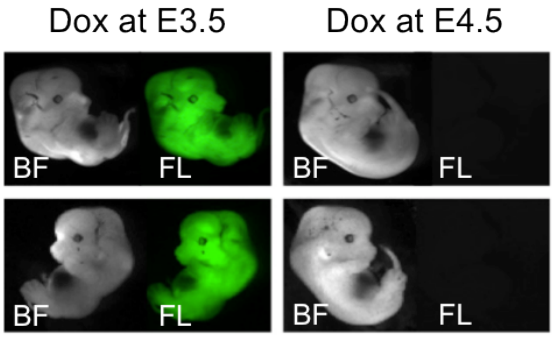


Figure 3.17. Irreversible transgene silencing in KD3-T embryos

Images of E12.5 induced embryos. Administration of Doxycycline until E3.5 results in activation of the transgene in all animals whereas at E4.5 it is unable to relieve transcriptional repression by tTS. BF, bright field; FL, fluorescence.

Although induction of the knockdown constructs under the control of tetracycline-dependent transcriptional silencer (tTS) was not problematic up to E3.5, administration of doxycycline to transgenic embryos at later time points did not result in transgene activation (Figure 3.17). This is a well-documented consequence of transcription repression by the KRAB fusion domain in tTS, which induces irreversible gene silencing through promoter DNA methylation when allowed to act during mouse post-implantation development (Wiznerowicz et al., 2007). For this reason, to achieve temporal control, knockdown constructs were integrated in *ROSA26A* ES cells, expressing the reverse tetracycline-dependent transcriptional activator (rtTA) under the ubiquitous *ROSA26* promoter.

3.4.2 rtTA allows temporal control over knockdown

Two exchange vectors, carrying hairpins against brachyury, were cloned and recombined into *ROSA26A* cells. KD4-T ES cells contained shRNA-mirs T-267 and T-1457 whereas KD5 cells expressed an intronic T-267 shRNA-mir and the coding sequence of EGFP (Figure 3.18a).

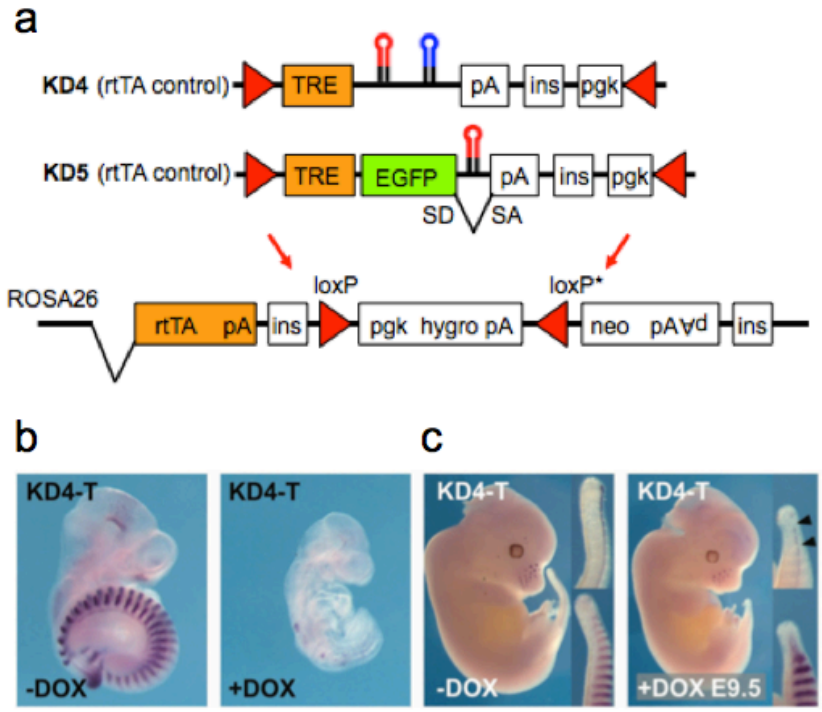


Figure 3.18. Exchange vectors and *T* RNAi phenotypes under rtTA control

(a) Schematic representation of the recipient locus and exchange vectors KD4 and KD5 used for RNAi. T-267 and T-1457 shRNA-mirs are represented as red and blue hairpins respectively. (b) E10.0 transgenic control embryos (−Dox) or embryos induced for KD4 expression (+Dox) were hybridized *in situ* with *T* plus *Uncx*. Treatment with Doxycycline (+Dox) throughout development leads to a phenotype indistinguishable from *T* loss of function mutatnts. (c) E12.5 control (left) and mutant embryo (right) induced for RNAi at E9.5 using the rtTA controlled KD4-T construct. Inlays show *Shh* (upper) or *Uncx* (lower) expression detected by *in situ* hybridization. Arrowheads mark discontinuous *Shh* expression.

Non-induced *KD4-T* embryos, derived by tetraploid complementation and isolated at E9.5 or E12.5 were phenotypically normal (17/17 embryos). Embryos induced to express *KD4-T* throughout development displayed a phenotype, which was indistinguishable from E9.5 *T* loss of function or *KD1-T* embryos (16/16 embryos; Figure 3.18b). However, when *KD4-T* was induced at E9.5 of development, embryos survived mid-gestation and the phenotype was confined to the tail. Embryos examined at E12.5 showed an arrest of mesoderm formation and notochord development at mid-tail, indicating that loss of *T* function induced by *KD4-T* RNAi commenced one day after the doxycycline treatment was initiated (12/12 embryos; Figure 3.18c).

Induced embryos derived by tetraploid complementation of *KD5-T* ES cells and dissected at E10.5 showed a milder phenotype, identical to *KD2-T* induced embryos (8/8), while control embryos for were phenotypically normal (11/11; Figure 3.19). Significantly, EGFP reporter

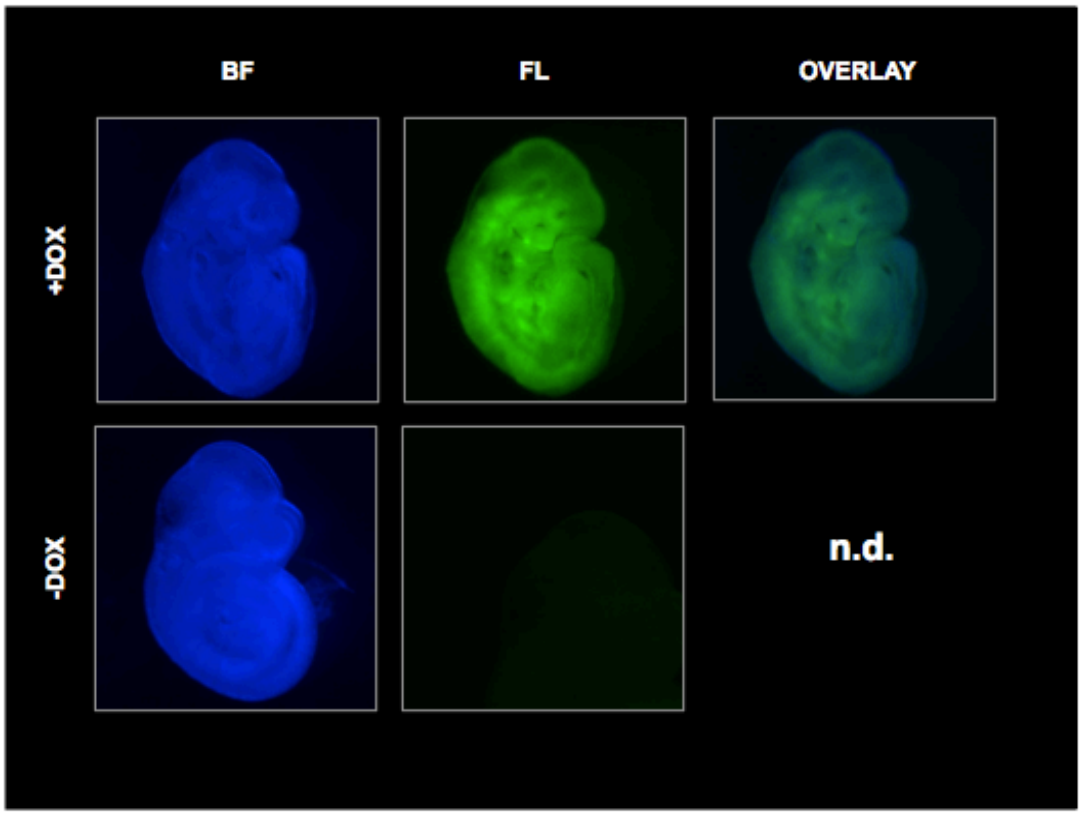


Figure 3.19 Analysis of KD5-T embryos

Fluorescent images of E10.5 control (-Dox) or induced (+Dox) embryos. Induced embryos show ubiquitous EGFP expression under the control of the rtTA. In the absence of Doxycyclne, no EGFP could be detected. BF, bright field; FL, fluorescence; n.d. not displayed.

expression was detected throughout *KD5-T* induced embryos, but absent in controls, suggesting that this system can be applied to all tissues of the developing embryos with low background activity. Together, these results demonstrate the suitability of the rtTA vector system for temporally controlled RNAi in the embryo.

3.5 Transcriptome analysis of the *KD1-T* RNAi model

3.5.1 Deregulated genes in *KD1-T* caudal ends

To determine if the specificity of the *KD1-T* phenotype was also reflected by deregulation of genes involved in the processes controlled by Brachyury, RNA from caudal ends of E8.5 embryos (6-7 somites) was analyzed by microarray. This stage was chosen because it is the time point when the *T* loss of function phenotype first becomes apparent, but with a still low compositional imbalance between mesoderm and neural tissue.

Expression profiles from induced and control embryos were compared (Figure 3.20), and statistical analysis identified 37 significantly deregulated genes (>2-fold and $P < 0.05$; 9 up-, 28 down-regulated). Among the downregulated genes, were the experimental target brachyury, *Tbx6* which was shown to require Brachyury for maintenance of its expression (Chapman DL

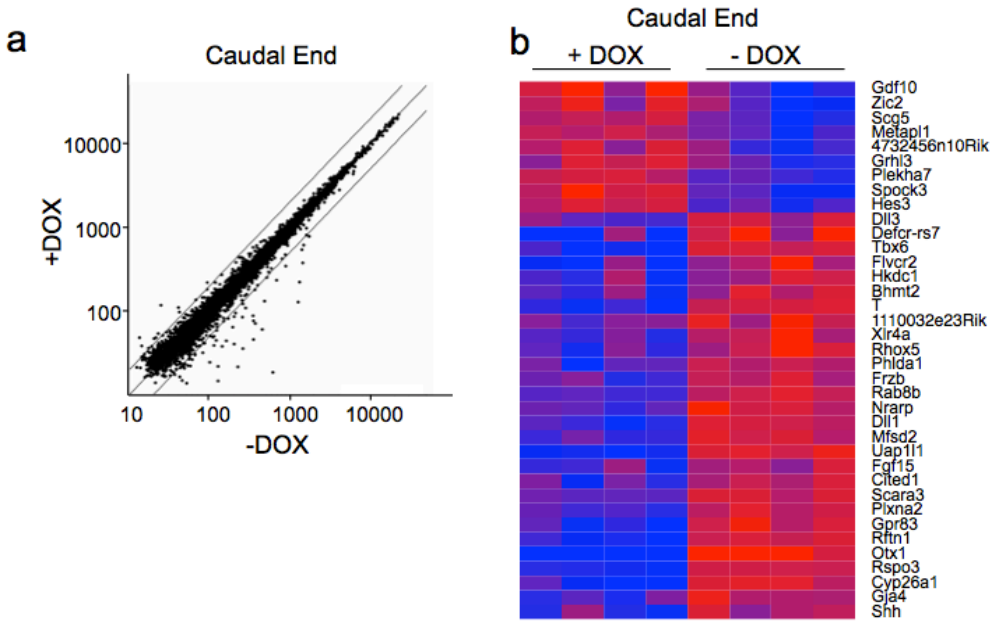


Figure 3.20. Expression analysis of *KD1-T* caudal ends

(a) Scatter plot analysis of gene expression in the caudal ends of *KD1-T* transgenic E8.5 control (−Dox) or induced (+Dox) embryos, as determined by microarray hybridization. Diagonal lines represent 2-fold deregulation. (b) Heat maps of genes found to be deregulated more than 2-fold ($P < 0.05$) in caudal ends. Each lane represents a biological replicate (red, upregulated; blue, downregulated; gene symbols on the right).

1996), the notochord marker *Shh*, as well as components from the Wnt (Martin and Kimelman, 2008), Fgf (Schulte-Merker and Smith, 1995) and Dll/Notch (Hofmann et al., 2004) signaling pathways, known to be involved in mesoderm formation and affected by the absence of T (Figure 3.21).

These data confirm the specificity of the RNAi phenotype generated by *KD1-T* expression and the suitability of analysis of RNAi for the functional analysis of genes in embryonic development.

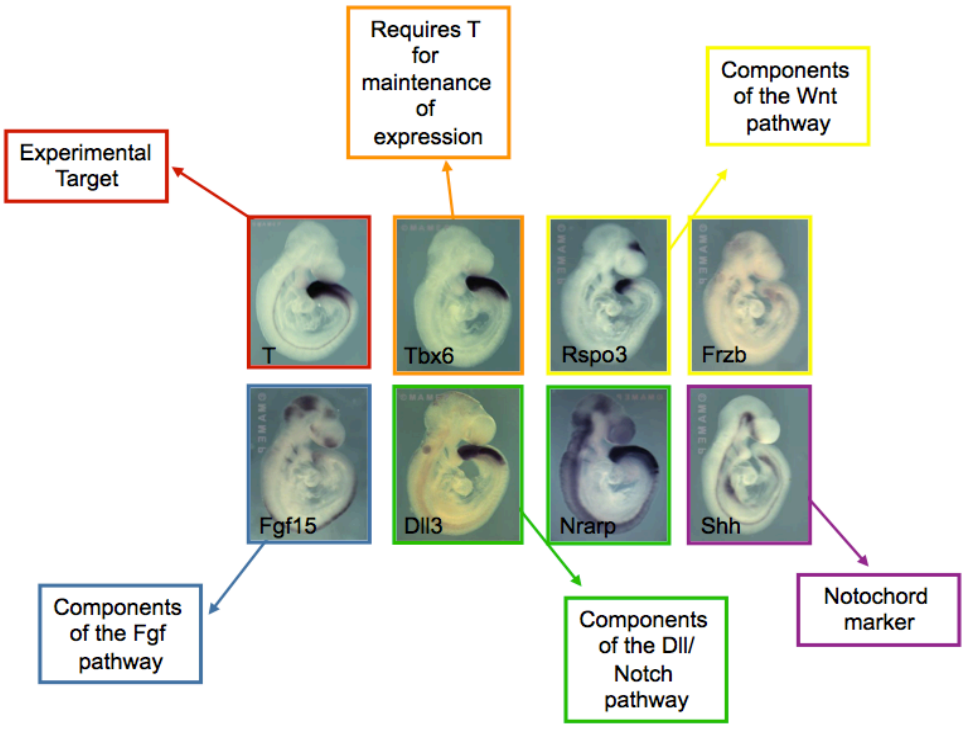


Figure 3.21. Downregulated genes are involved in the processes controlled by T

Expression patterns of selected downregulated genes as visualized by whole mount *in situ* hybridization. Boxed texts highlight the relevance of these genes in the context of the processes controlled by T. Images taken from the MAMEP database (http://mamep.molgen.mpg.de/).

3.5.2 Knockdown occurs in the absence of off-target effects

One of the biggest concerns when using experimental RNAi is the occurrence of off-target effects, i.e. the unwanted modulation of mRNAs through unspecific RNA interference. To

determine if off-target effects could be detected in *KD1-T* embryos, expression profiles from induced and control forebrains of E8.5 embryos were analyzed by microarray (Figure 3.22).

The forebrain is an embryonic structure, which does not express *T*, and consequently should not be affected by the expression of the *KD1-T* knockdown construct. Statistical analysis between control and induced forebrain expression profiles yielded three deregulated genes (>2-fold and $P < 0.05$; 1 up-, 2 downregulated), which is less than in comparisons made between randomized groups (4.7 deregulated; >2-fold and $P < 0.05$).

As an independent method, expression profiles were analyzed for underrepresentation of mRNAs containing motifs complementary to the seed sequences of the experimental shRNA-mirs, which could be an indicator of off-target gene regulation. Using the Sylamer tool (van Dongen et al., 2008), such a bias could not be detected in genome-wide forebrain or caudal end data sets (Figure 3.23).

Collectively, these data show that the RNAi system can generate loss-of-function models in the absence of detectable off-target effects.

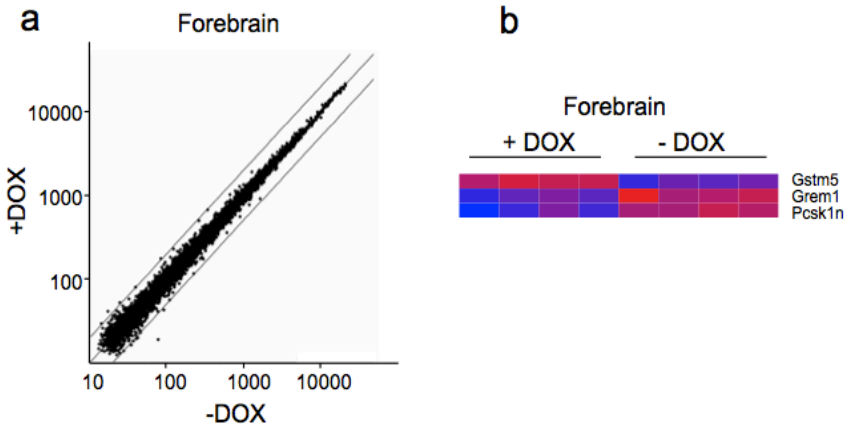


Figure 3.22. Expression analysis of *KD1-T* forebrains

(a) Scatter plot analysis of gene expression in the forebrains of *KD1-T* transgenic E8.5 control (−Dox) or induced (+Dox) embryos, as determined by microarray hybridization. Diagonal lines represent 2-fold deregulation. (b) Heat maps of genes found to be deregulated more than 2-fold ($p < 0.05$) in forebrains. Each lane represents a biological replicate (red, upregulated; blue, downregulated; gene symbols on the right).

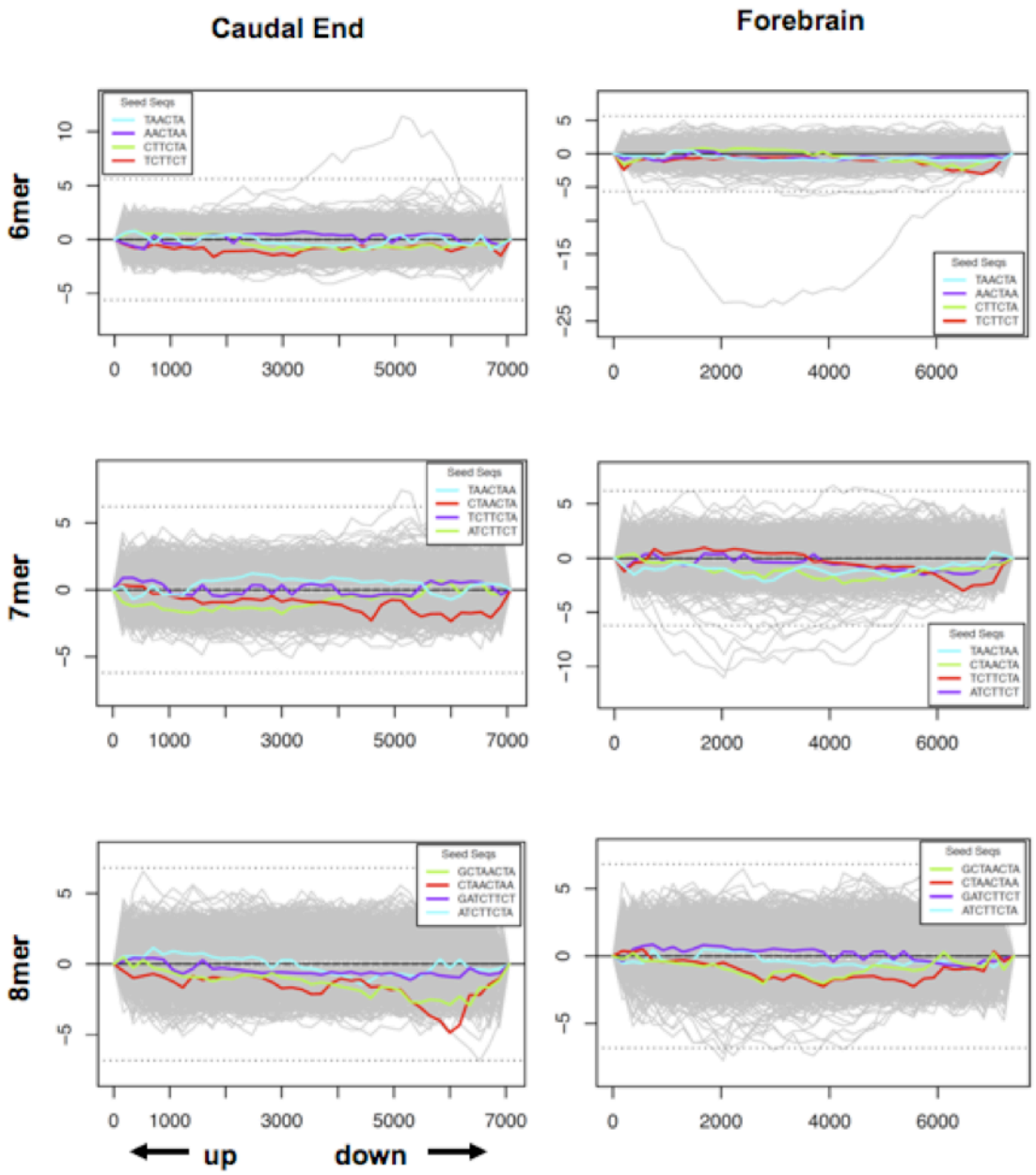


Figure 3.23. Analysis of off-target effects in KD1-T embryos using Sylamer.

Graphs show the distributions of seed sequence occurrence in caudal end- (left) and forebrain- (right) derived ordered gene lists. Genes that are overrepresented in induced *KD1-T* samples to the left (up), genes underrepresented in induced samples to the right (down). Motif lengths of 6 (top), 7 (middle) and 8 (bottom panels) nucleotides corresponding to the experimental shRNA-mir seed sequences were used. Cyan, purple, green, red: 6-,7- and 8-mer complimentary to *T* shRNA-mirs; grey: occurrence of unrelated 6-,7- and 8-mers; dashed lines: confidence thresholds.

3.6 Summary and Conclusion

In summary, I present here a novel single-copy, single-locus integrated transgenic system for shRNA-mir mediated RNAi in the mouse. By use of this system, I demonstrate for the first time, conditional RNAi in the mouse embryo that phenocopies genetic null models. I present a panel of vectors allowing the generation of various hypomorphs of different strengths, which have not been possible to obtain with loss of function alleles. This set of tools is suitable for investigating the function of genes playing multiple roles in development in more detail than previously possible with knockout alleles.

The system presented here has distinct advantages in generating null and hypomorphic mutants with minimal off-target regulation. In particular, the possibility to derive genetically identical experimental and control embryos ensures compatibility with high-throughput applications, such as expression profiling or massive parallel sequencing.

4 RESULTS - PART II

KD3-T hypomorphic embryos as a model for uro-rectal-caudal-syndrome

Contributors

Joana A. Vidigal

Lars Wittler

Judith Proske

Bernhard G. Herrmann

4.1 Experimental contributions

Lars Wittler supervised the work on the uro-rectal phenotype.

Judith Proske did the dissection of embryos with anal atresia and the H&E sections.

Bernhard G. Herrmann supervised the project.

All other experiments were done by me, which includes generation of the RNAi model, and dissection/analysis of embryos at E12.5 and E13.5.

4.2 *KD3-T* induced embryos survive mid-gestation and develop an ectopic neural tube

The *KD3-T* RNAi model described in Chapter 3 showed a hypomorphic phenotype for brachyury, with truncation of notochord development but not of axis length or somite development at E9.5 (Figure 3.11). In a fraction of these embryos, allantoic development was not compromised, suggesting they could develop beyond mid-gestation. In order to study later developmental defects in embryos with reduced Brachyury levels, a mouse line for this RNAi model was established. Transgenic animals raised in the absence of doxycycline were viable, healthy and fertile.

To determine later phenotypes in *KD3-T* mice, transgenic males were mated with wild-type females and doxycycline administered at the day of the plug. E12.5 embryos were isolated and genotyped by PCR with primers binding to the neomycin coding sequence. Mutant embryos were recovered at non-Mendelian ratios - reflecting the lethality of the *KD3-T* transgene in a portion of the embryos – and displayed a clear axis truncation, with development of only a short

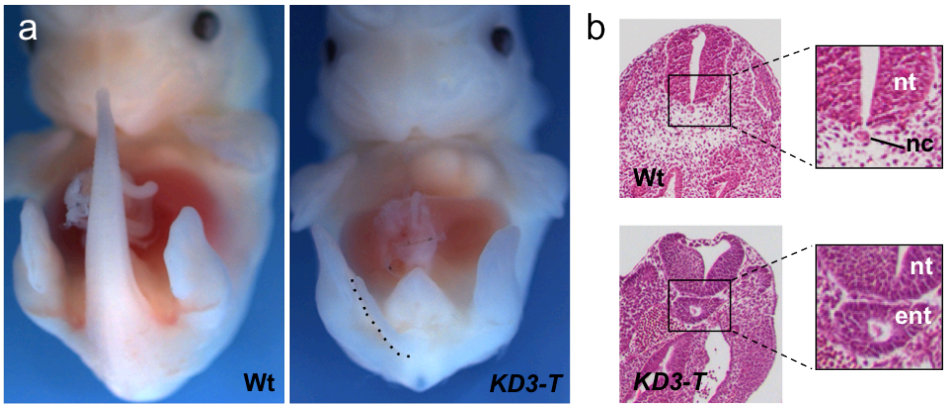


Figure 4.1.*KD3-T* induced embryos develop a tail filament and an ectopic neural tube

(a)Ventral view of wild-type (Wt; left) and induced *KD3-T* (right) littermates at E12.5. The tail filament in the RNAi mutant is delineated by a doted line. (b) Transverse sections of wild-type (Wt; top) and induced *KD3-T* embryos (bottom) at E10.5. Magnification of the boxed areas on the right shows the notochord (nc) below the neural tube (nt) in the wild-type embryo. In the mutant, an ectopic neural tube (ent) forms in the place of the notochord.

tail filament (Figure 4.1a; Right). Wild-type littermates were phenotypically normal (Figure 4.1a; left). Other mutations that lead to a caudally truncated phenotype such as in *Wnt3a* (Yamaguchi et al., 1999b), *Fgfr1* (Ciruna et al., 1997) and *Cyp26a1* (Abu-Abed et al., 2001; Sakai et al., 2001) mutants, often lead to the development of secondary neural tubes and these have been previously described in *T* loss-of-function mutants (Yamaguchi et al., 1999b). To determine if such structure was also present in induced *KD3-T* mice, embryos were dissected and sectioned at E10.5. Hematoxylin and Eosin (H&E) staining of mutant embryos revealed the formation of an ectopic tubular epithelial structure immediately ventral to the neural tube, which develops in place of the notochord (Figure 4.1b). Thus, in the upon knockdown of brachyury, an ectopic neural tube-like structure forms in a region with reduced axial mesoderm, emphasizing the similarities between *T* and other genes involved in posterior mesoderm formation, and providing further evidence of the specificity of the knockdown model.

4.3 KD3-T mutants have spina bifida

At E8.5 of mouse development the process of neural tube closure starts, and by E9.5 it is complete along the entire body axis with the exception of the caudal neuropore, where closure occurs continuously as the axis elongates. The process of neural tube closure involves tight regulation of processes of proliferation, migration and patterning, during which the folds of the neural plate elevate, come together at the dorsal midline and fuse to give rise to a hollow tube, known as the neural tube (Figure 4.2a) (Copp et al., 2003). The process of closure does not occur simultaneously along the entire axis, but starts at three closure points. Failure to close the tube in any of these regions results in neural tube defects, which are called *spina bifida* when arising in the lumbosacral neural tube posterior of closure point 1 (Figure 4.2b) (Copp and Greene, 2009).

Analysis of *KD3-T* embryos at E13.5 revealed incomplete neural tube closure in the spinal region of all embryos, resulting in *spina bifida* (Figure 4.2d). This seems to be a defect

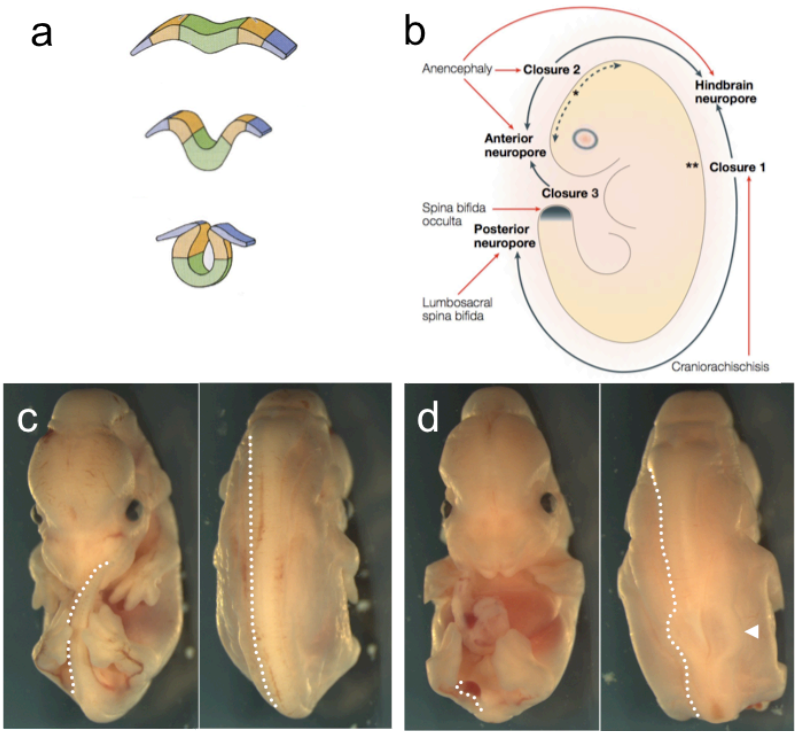


Figure 4.2. Neural tube closure in *KD3-T* mutants

(a) Representation of the process of neural tube closure. Neural folds elevate, become apposed at the midline and fuse resulting in a neural tube that is covered by epithelial ectoderm (in blue). Modified from (Gilbert, 2000). (b) Initial points of neural tube closure in the mouse embryo. Black arrows indicate closure directions after initial contact. Diseases caused by failure of neural tube closure are also shown and their designation depends on the site of failure, represented in red. Adapted from (Copp et al., 2003). (c-d) Ventral and dorsal views of wild-type (c) and KD3-T mutant embryos at E13.5. Tail and neural tube are delineated with a dotted line. Arrow head points to a region with spina bifida, resulting from improper neural tube closure in the lumbar region.

secondary to the notochord truncation (Figure 3.11; Figure 4.1b) since this structure is known to be involved in induction and patterning of the neural plate (van Straaten et al., 1988) and was also shown to play a role in the closure of the neural tube in amphibians (Jacobson, 1984).

4.4 *KD3-T* mutants display urorectal defects

The results described so far emphasize the role of Brachyury during posterior trunk development, not only at the level of axial extension but also as a player in inductive events by controlling the formation of the notochord. The differentiation of the caudal trunk, the region affected by low Brachyury levels, involves the formation of two systems: the urogenital and the anorectal. These are summed together under the name 'urorectal system', due to their common embryological etiology. At E10.5 in the mouse, corresponding to week 5 in humans, the developing urogenital and anorectal systems come together in a common outlet called the cloaca (Figure 4.3, right). As development of the urorectal system progresses, a mesoderm-derived structure, the urorectal septum, expands towards the cloaca leading to its separation into the anal and caudal urethral orifices around E12.5 (Figure 4.3) (Stephens, 1988). Defects in the

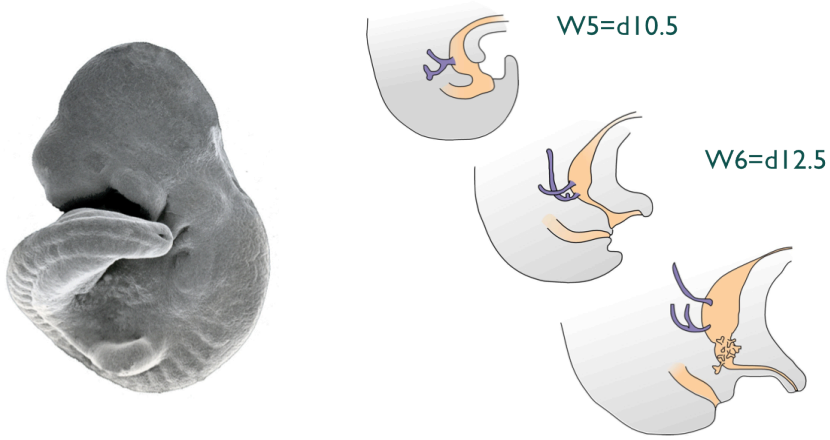


Figure 4.3 Development of the urorectal system

(Left) Scanning electron micrograph of a E9.5 mouse embryo. (Right) a diagram showing the development of the urorectal system throughout day 10.5 to 12.5 of mouse development (d10.5; d12.5), corresponding to week 5 and 6 of human development (W5; W6). Courtesy of Dr. Lars Wittler.

development of the urorectal septum are thought to lead to urorectal malformations, which include anal atresia, meaning that the anus is either not formed or it is in the wrong position (Kluth, 2010).

The observation that the *KD3-T* RNAi model exhibited several hallmarks of the mouse uro-rectal-caudal syndrome (Gruneberg, 1952), including truncation of posterior structures and neural tube defects, prompted the closer examination of development of the uro-rectal system. *KD3-T* males were mated to wild-type females and knockdown of brachyury induced in the progeny as described previously. A fraction of induced *KD3-T* fetuses showed the formation of a small anal pit in place of the anal orifice at E17.5 (Figure 4.4a). Sectioning and H&E staining of these embryos showed absence of a connection between the anal pit and the gastrointestinal

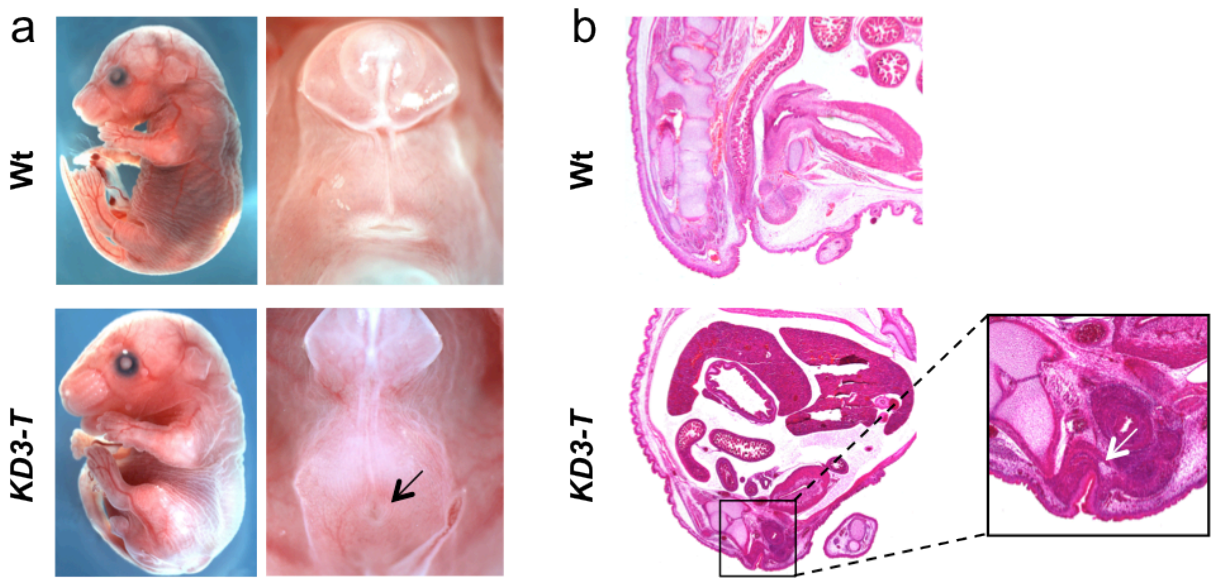


Figure 4.4. Anal atresia in KD3-T induced mice

(a) Induced *KD3-T* fetuses (bottom) and wild-type littermates (top) at E17.5. Magnification on the right shows a close up of the uro-rectal region. Arrow points to the anal pit in mutant animals. (b) H&E stainings of sagittal sections of embryos shown in (a). Arrow shows absence of a connection between the anal pit and the gastrointestinal tract.

tract (Figure 4.4b).

Together, these results show that *KD3-T* induced animals recapitulate key features of the uro-rectal-caudal syndrome, including anal atresia, and can therefore be used as a model to analyze the morphological and molecular aspects of this disease.

4.5 Summary and Conclusion

In summary, *KD3-T* induced embryos, displaying a hypomorphic phenotype for brachyury, exhibit an imperforate anus, neural tube closure defects, and posterior axis truncation. These are all defining features of the mouse uro-rectal-caudal syndrome and involve brachyury in the differentiation of the uro-genital and ano-rectal systems.

Importantly, the genetic and molecular causes behind ano-rectal malformations are poorly understood, mainly due to lack of animal models. The *KD3-T* RNAi mouse, showing a highly reproducible phenotype, is therefore a welcomed model with which to study the morphogenesis of ano-rectal malformations in comparison to normal development.

5 RESULTS – PART III

Late induction of *T* knockdown reveals a novel role of Brachyury during skeletogenesis

Contributors

Joana A. Vidigal

Bernhard G. Herrmann

5.1 Experimental contributions

Bernhard G. Herrmann supervised the project.

All experiments were planed and done by me.

5.2 Brachyury is required for the differentiation but not maintenance of the notochord

The notochord is a transient structure during vertebrate development. It is formed by mesodermal cells that migrate through the primitive streak and serves two main functions: it is a source of midline signals that patterns surrounding tissues, and serves as a key skeletal element during embryonic development (Stemple, 2005). Brachyury is expressed in notochordal cells from the moment they ingress through the primitive streak and its expression has been detected in the midline of mouse embryos as late as E17.5 (Herrmann and Kispert, 1994; Wilkinson et al., 1990). Brachyury is required for notochord differentiation (Figure 5.1; Top) but the role of its continuous expression in the midline remains unclear. To address this question, a mouse line for the *KD4-T* RNAi model was established. This transgene allows the study of Brachyury loss-of-function phenotypes in a temporally controlled manner and can be used to circumvent early embryonic lethality (Figure 3.18). Mutant animals, carrying the *KD4-T* transgene and raised in the absence of doxycycline were viable, healthy and fertile.

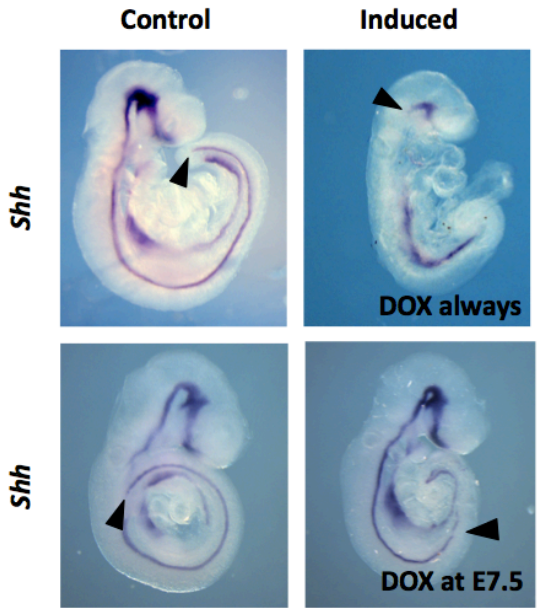


Figure 5.1 brachyury is required for differentiation but not maintenance of the notochord

Whole mount *in-situ* hybridizations for *Shh* in control and induced transgenic embryos. Induction of the knockdown at the beginning of embryogenesis (DOX always) leads to absence of notochord differentiation at E9.5 (top, right). When the knockdown is induced at E7.5, the notochord is truncated at the point of mesoderm formation arrest, but *Shh* transcripts can still be detected in more anterior regions (bottom, right). Arrow head marks the caudal end of the notochord.

In order to dissect the role of brachyury's notochordal expression during late stages of development, transgenic males were mated to wild-type females and the RNAi construct induced in their progeny by doxycycline treatment at E7.5, a time point when notochord development is already initiated. Dissection of embryos at E9.5 showed mild truncation of the body axis in mutant embryos, consistent with abrogation of brachyury's mRNA at E8.5 (6/6). Wild-type littermates were phenotypically normal (5/5). Analysis of notochord development by whole mount *in situ* hybridization for *Shh*, an independent notochordal marker, revealed truncation of this structure at the point of mesoderm formation arrest. Notochord anterior to this region however, was unaffected (Figure 5.1; Bottom). These results show that, although Brachyury is required for the differentiation of the notochord, it is not necessary for its maintenance.

5.3 Gross morphology of E16.5 *KD4-T* embryos induced at E8.5

To determine the effects of knockdown of brachyury in older animals, the *KD4-T* construct was induced in mutant embryos at E8.5. At E16.5, *KD4-T* fetuses could be isolated at normal Mendelian ratios, but were significantly smaller than their wild-type littermates (Figure 5.2a-c; mean length from head to tail base – wt: 1.78 ± 0.08cm; *KD4-T*: 1.55±0.10 cm; $p=1.7x10^{-3}$; mean weight – wt: 7.6±0.46g; *KD4-T*: 6.3±0.57g; $p=1.4x10^{-3}$). In addition, mutant embryos displayed a truncation of the tail (mean tail length – wt: 87±4mm; *KD4-T*: 44±11mm; $p=3.3x10^{-5}$), consistent with arrest of mesoderm formation and axial elongation around day 9.5 of embryogenesis.

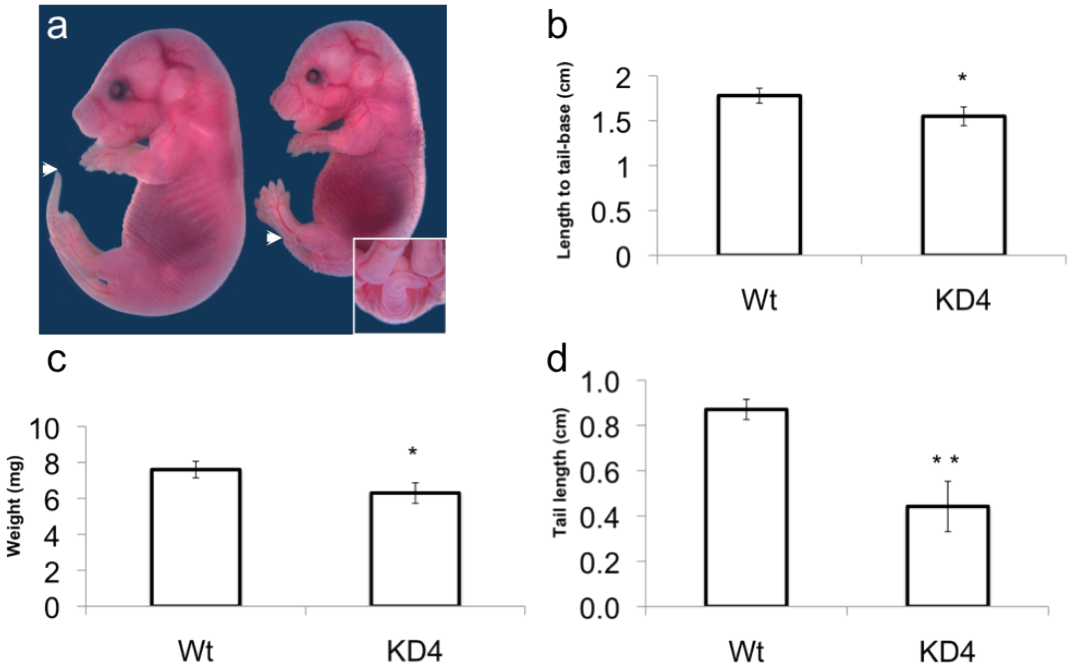


Figure 5.2. Gross appearance of E16.5 *KD4-T* embryos induced at E8.5

(a) Gross appearance of *KD4-T* embryos (right), induced at E8.5, and wild-type litter mates (left). Mutant embryos are smaller in weight and length and display a truncated tail. Arrow heads point to the end of the tail. Inset shows ventral view of mutant tail. (b-d) Quantification of weight and length of mutant and wild-type mice. Data are mean ± standard deviations from 5-6 mice at E16.5. * and ** indicate statistical significance at $p < 2.0x10^{-3}$ and $p < 3.5x10^{-5}$ respectively.

5.4 Skeletal defects in *KD4-T* embryos

Skeletal preparations of wild-type and mutant fetuses at E18.5 showed that, apart from a truncation of the caudal vertebrae, *KD4-T* mutants, induced at E8.5, formed an intact skeleton (Figure 5.3), characterized by abnormalities in bone mineralization. In the skull, *KD4-T* mice displayed an absent or rudimentary supraoccipital bone, as well as reduced mineralization of the parietal and interparietal bones (Figure 5.4a). In addition, mutant embryos had no sign of hyoid ossification (Figure 5.4d), a process that should be completed by day 15.5 of embryogenesis. The clavicles, were also significantly smaller in mutant fetuses (Figure 5.4b). This was not a mere reflection of the generalized size difference between *KD4-T* and wild-type mice, since the ratio between the size of clavicles and femurs - a bone that seems unaffected by the knockdown of brachyury - was statistically different between mutant and wild-type animals (Figure 5.4c- wt: 0.84 ± 0.14; *KD4-T*: 0.68 ± 0.02; p= 0.04).

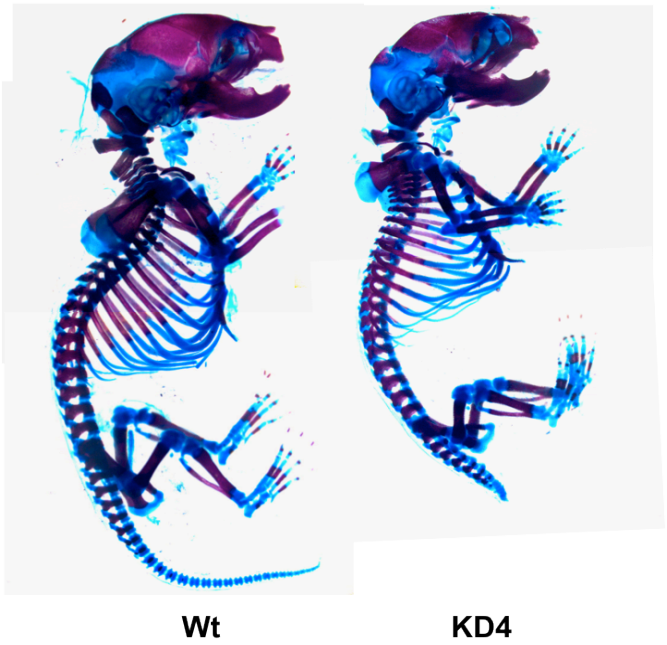


Figure 5.3. *KD4-T* embryos have an intact skeleton

Skeletal preparations of *KD4-T* (KD4) and a wild-type littermate (Wt) at E18.5. Skeletons are stained with alcian blue for cartilage and alizarin red for bone.

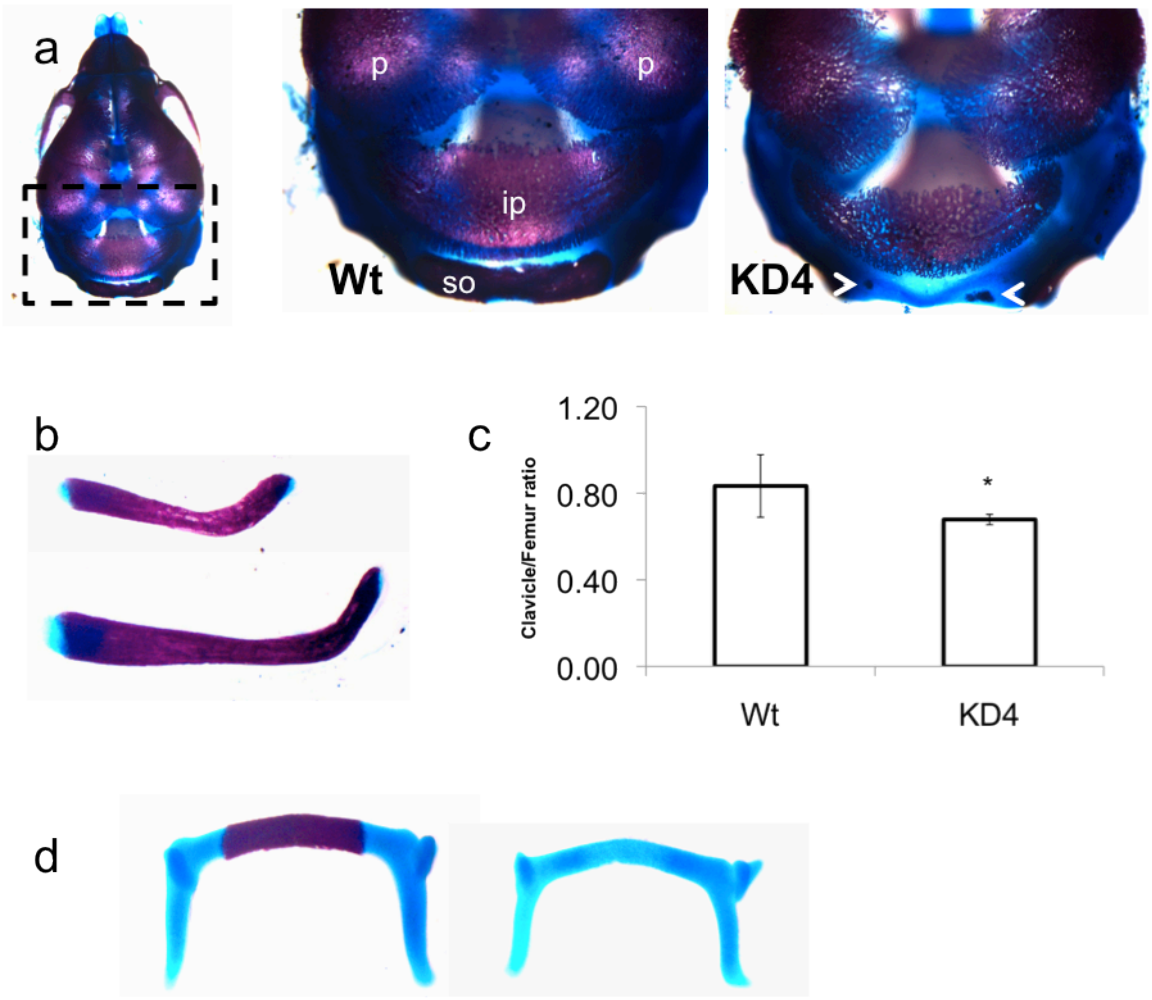


Figure 5.4. Skeletal defects in *KD4-T* mutant fetuses induced at E8.5

(a) dorsal view of a wild-type (Wt) and mutant (KD4) skulls. A wild-type skull on the left shows the position of the magnified region. P, parietal bone; ip, interparietal bone; so, supraoccipital bone. Arrow heads shows rudimentary supraoccipital bones in *KD4-T* mutant skull. (b) clavicles of mutant (top) and wild-type (bottom) mice. (c) Ratios between clavicle and femur lengths in wild-type (Wt) and mutant (KD4) embryos. Data are mean ± standard deviations from 3-4 mice at E18.5. * indicates statistical significance at $p < 0.05$. (d) Hyoid bones of wild-type (left) and *KD4-T* (right) embryos.

In other regions of the skeleton, knockdown of brachyury resulted in delayed ossification of the vertebral bodies of the cervical vertebrae and irregular ossification of vertebral bodies of lumbar vertebrae (Figure 5.5a). Additionaly, ossification of the sternal bars of the rib cage showed mild to severe misalignement in *KD4-T* mice, but not in wild-type littermates (Figure 5.5 b,c), which was accompanied by desaligned rib pairs.

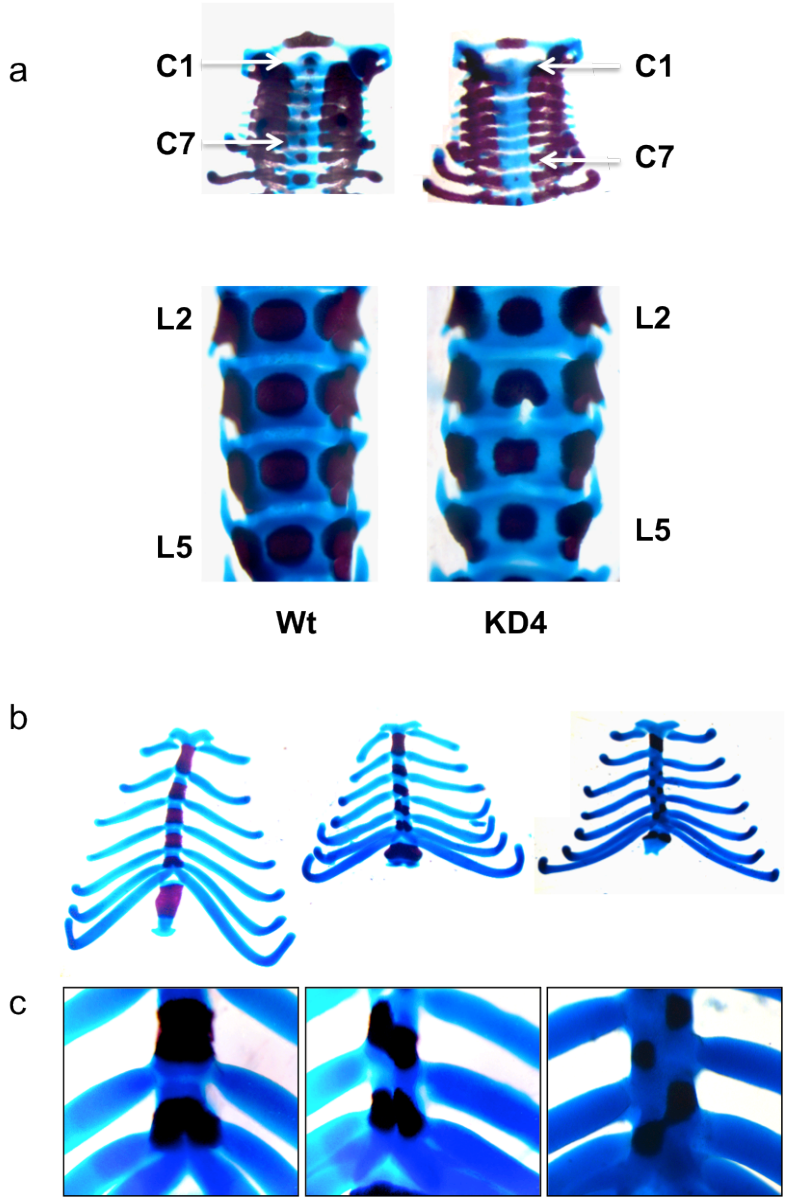


Figure 5.5. Skeletal defects of *KD4-T* fetuses induced at E8.5

(a) Ventral view of cervical (top) and lumbar (bottom) vertebrae from wild-type (wt) and mutant (KD4) mice. (b) Representative rib cage bones from wild-type (left) and *KD4-T* (center an right) animals. (c) Detail view of the sternal bars from the rib cages shown in b.

5.5 Loss of Brachyury affects neural crest cell migration

The formation of axial skeleton, comprised of skull, spine, sternum and ribs, and affected in *KD4-T* mutants, depends on two mesenchymal cell lineages: neural crest cells , derived from the roof plate of the neural tube, and paraxial mesoderm cells, either from the segmented trunk (sclerotome) or from the unsegmented head mesenchyme (Karaplis, 2002). In order the contribute to skeletogenesis, cells from these lineages undergo an epithelial-to-mesenchymal transition, migrate to the site of bone formation and differentiate into either chondrocytes, which give rise to cartilage, or osteoblasts, that give rise to mineralized bone (Hall and Miyake, 2000).

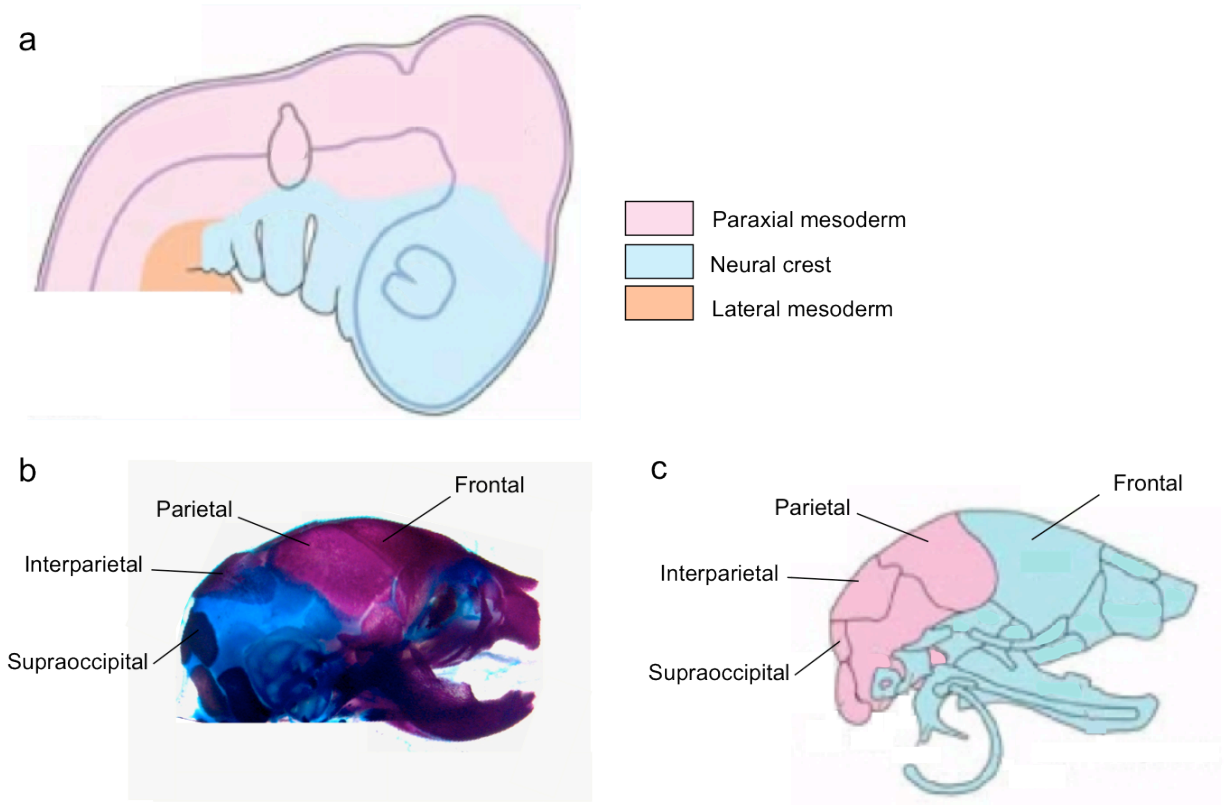


Figure 5.6. Origin of skull bones in the mouse

(a) Representation of the location of the neural crest/mesoderm border in the head after the initial migration of neural crest cells. (b) Lateral view of a mouse skull stained with alcian blue for cartilage and alizarin red for bone. (c) Representation of the neural crest and paraxial mesoderm contribution for the formation of the bones in the skull. In pink are represented bones of paraxial mesoderm origin, in blue those derived from neural crest cells. Adapted from (Noden and Schneider, 2006).

In the mouse skull, neural crest cells (NCCs) give rise to craniofacial bones (Le Douarin et al., 1993) and the paraxial mesoderm to bones positioned more dorsally, with the interface between the two populations corresponding roughly to the site of the coronal suture between the frontal and the parietal bones (Jiang et al., 2002) (Figure 5.6). The phenotypes observed in *KD4-T* mutants, with absence of hyoid ossification – a neural crest derived structure – and compromised ossification of the supraoccipital, parietal and interparietal bones – derived from the head mesenchyme - show that skull skeletogenesis via both these cell lineages is affected.

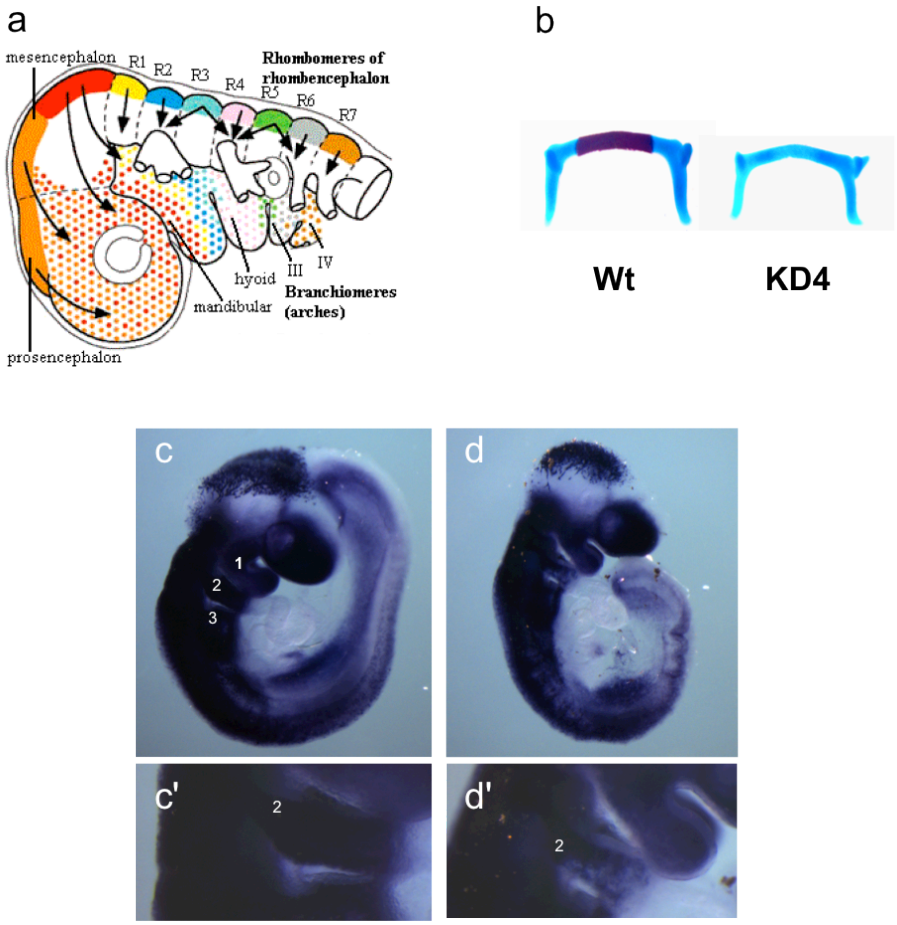


Figure 5.7 Effects of *T* knockdown on neural crest cell migration

(a) Schematic representation of the head of a E9.5 mouse embryo, showing the origin and direction of migration of neural crest cell populations. Adapted from (Couly et al., 2002) (b) Representative hyoid bones from E18.5 wild-type (Wt) and *KD4-T* mutant (KD4) animals. (c-d) whole mount i*n situ* hybridization for crabp1 in E9.5 wid-type (c) and *KD4-T* embryos induced at E7.5 (d) A detailed view of the second brachial arch is shown in c' and d' for wild-type and mutant embryos respectively. 1, 2 and 3 indicate the position of the first, second and third branchial arches respectively.

During craniofacial development, NCCs delaminate from the dorsal region of the neural tube into the branchial arches and the frontonasal process to give rise to cartilage and bone (Noden, 1988). NCCs delaminating from different levels of the anteroposterior axis migrate into specific regions of the head and give rise to defined structures of the facial skeleton (Lumsden et al., 1991) (Figure 5.7a). The hyoid bone, whose differentiation is affected in *KD4-T* mutants (Figure 5.7b), is formed by NCCs delaminating from the sixth and fourth rhombomeres into the third and second branchial arch respectively (Jiang et al., 2002). To trace the migration of NCC upon knockdown of brachyury, a whole mount *in-situ* hybridization against the *Crabp1*, a NCC migration marker (Ruberte et al., 1991), was done. While analysis of wild-type embryos at E9.5 showed complete colonization of the second branchial arch by neural crest cells (Figure 5.7c and c'), in *KD4-T* mutants, induced at E7.5, colonization of this structure was incomplete (Figure 5.7d and d'). Together, the absence of ossification of the hyoid bone at E18.5 and the compromised colonization of the second brachial by NCCs at E9.5 - which will give rise to the superior region of the hyoid body where mineralization is compromised - suggest that expression of brachyury in the notochord is required for proper neural crest cell migration.

The remaining skeletal defects observed upon brachyury knockdown occur in bones derived from paraxial mesoderm, either originating from the unsegmented head mesenchyme (parietal, interparietal and supraoccipital bones) or from the sclerotome. Several transcription factors such as *Pax1*, *Foxc2* and have been implicated in the differentiation and migration of these cells, and their disruption leads to skeletal abnormalities similar to those described here for brachyury. These include absence of ossification of the vertebral bodies in the cervical region, irregular ossification of vertebral bodies in the lumbar and sacral regions, abnormal mineralization of parietal and interparietal bones as well as absence of supraoccipital bone mineralization (Furumoto et al., 1999; Iida et al., 1997; Reinhold et al., 2006). These similarities might be an indication that, like neural crest cells, paraxial mesoderm cells require Brachyury for correct specification.

5.6 Analysis of chondrogenic markers in *KD4-T* mutants

Sox9 - a member of the Sox gene family characterized by a high-mobility-group-box DNA binding motif related to the one in the sex determining factor SRY - is one of the earliest transcription factors acting during skeletal development. *Sox9* is expressed in the skeletal precursors that will give rise to both chondrocytes and osteoblasts, and its activity is required for chondrocyte lineage specification (Akiyama, 2008), one of the first events during skeletal differentiation (Karsenty, 2008). After entering the chondrogenic lineage, mesenchymal cells differentiate into chondroblasts, cells with an extracellular matrix rich in collagens and proteoglycans. This differentiation step is also controlled by Sox9, together with two of its direct downstream targets: *Sox5* and *Sox6* (Han and Lefebvre, 2008; Lefebvre and Smits, 2005).

To determine whether embryos could properly activate *Sox9* and its downstream target

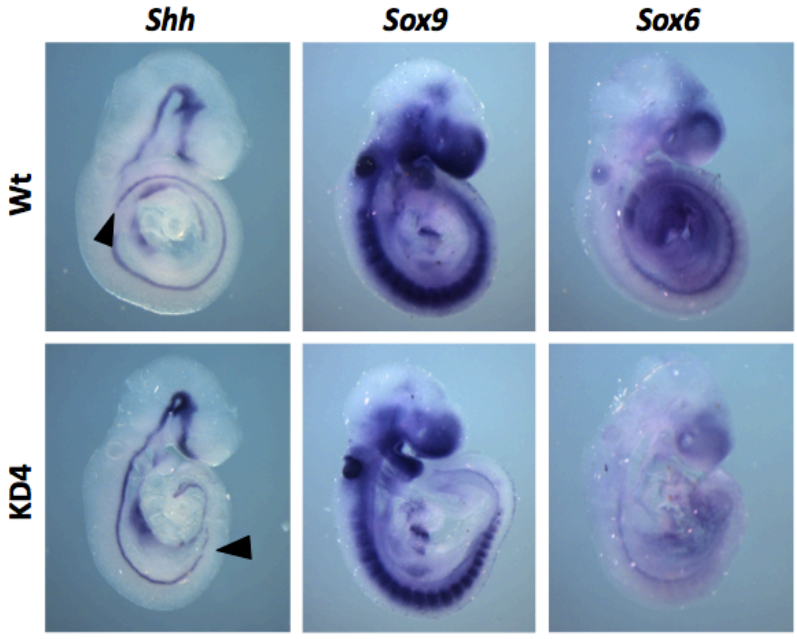


Figure 5.8. *Sox6*, a direct target of Sox9 during bone formation, is deregulated in *KD4-T* mutants

Whole mount *in situ* hybridizations of E9.5 *KD4-T* mutants, induced at E7.5, and wild-type littermates. Probes used in the hybridizations are indicated on top of each panel. Arrow heads in embryos stained for *Shh* transcripts mark the caudal end of the notochord.

Sox6 after brachyury knockdown, the expression of these factors was assayed by whole mount *in-situ* hybridization in E9.5 *KD4-T* and wild-type embryos treated with dox from E7.5 on. While the expression of *Sox9* was grossly comparable between mutant and wild-type embryos, the expression of *Sox6* was severely reduced upon knockdown of brachyury suggesting that the earliest steps of skeletal differentiation might be compromised.

5.7 Brachyury is required for nucleus pulposus differentiation

The role of the notochord in axial skeleton development goes beyond inductive events. It also directly participates in the formation of the verteral column. Around E11.0, signals from the notochord induce sclerotome migration, condensation and differentiation around the notochord and neural tube. The resulting structure, the perinotochordal tube, has a segmented patterned formed by alternating zones of high and low condensation (Figure 5.9; left). Between E12.5 and E15.5 the less condensed zones develop into vertebrae, and the notochordal cells, initialy forming a rod-like structure at the center of the perinotochordal tube, are displaced to the intervertebral regions were they undergo hypertrophy to form the nuclei pulposi of the intervertebral discs (Aszodi et al., 1998; Rufai et al., 1995) (Figure 5.9; middle).

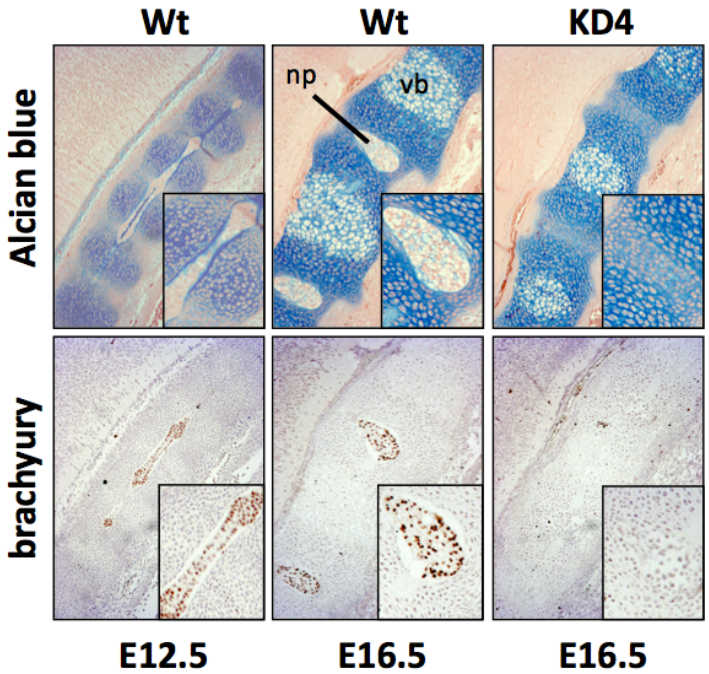


Figure 5.9 Brachyury's function and expression during vertebrae differentiation

Top, alcian blue and eosin staining of sagittal sections from E12.5 and E16.5 wild-type (Wt) and E16.5 *KD4-T* (KD4) embryos induced at E8.5. Bottom, immunohistochemistry for Brachyury, showing expression in wild-type animals in notochordal cells at E12.5 and in the nucleus pulposus at E16.5. In *KD4-T* animals, Brachyury expression was not detected. Insets show magnification of structures where Brachyury is expressed. np, nucleus pulposus; vb, vertebral body.

Sox6 and *Sox5* are co-expressed with *Sox9* in the developing vertebral column and are required to direct the fate of notochordal cells. In the absence of both these factors, induction of the sclerotome occurs normally but the development of the vertebral column is dramatically impared with degeneration of notochordal cells, absence of nuclei pulposi, and deficiencies in extracellular matrix sheat fromation (Smits and Lefebvre, 2003).

The severe reduction of *Sox6* expression upon brachyury knockdown prompted the analysis of the roles of T in later steps of vertebral column differentiation. Immunohistochemical analysis on sagittal sections of wild-type embryos showed that at E12.5 Brachyury is still expressed in the notochord (Figure 5.9; left). Four days later, at E16.5 - when the notochord remants have moved to the intervertebral regions - Brachyury could be specifically detected in the cells of the nuclei pulposi (Figure 5.9; middle). To investigate the consequences of brachyury knockdown in the development of this structure, *KD4-T* animals were treated with dox at E8.5 and analyzed at E16.5. Alcian blue staining of sagittal sections revealed that in such embryos, similarly to what happens in *Sox6*$^{-/-}$; *Sox5*$^{-/-}$ double mutants, differentiation of nuclei pulposi did not occur (Figure 5.9; right).

Together, these results show that expression of Brachyury in the notochord is required for the differentiation of nucleus pulposus during mouse vertebral development and that its functions could be mediated by *Sox6,* perhaps in conjunction with *Sox5*.

5.8 Summary and Conclusion

In summary, knockdown of brachyury at later stages of mouse development uncovered a novel and unexpected role for *T* during skeletal development. *T* is expressed in the midline as late as E17.5 and its function is required for correct mineralization of bones across the axial skeleton. Induction of *T* RNAi at E8.5 led to abnormal ossification of the parietal and interparietal bones, absence of ossification of the supraoccipital and hyoid bones as well as absence of ossification of the vertebral bodies of the cervical region and irregular ossification of the vertebral bodies of more caudal vertebra. In addition, mutant embryos showed missalignment of ossification of the vertebral bars.

Comparative molecular analysis between *T* knockdown embryos and their wild-type littermates revealed that brachyury regulates skeletogenesis via two distinct mechanisms. First, expression of Brachyury in the notochord is required for correct mesenchymal cell induction, as exemplified by abnormal neural crest cell migration which results in lack of hyoid ossification. Secondly, T is required for vertebral column development, specifically for diferentiation of nucleus pulposus, pherhaps in part by regulating the expression of *Sox6*.

6 DISCUSSION

6.1 An inducible system to knockdown genes in the mouse

I present here an inducible RNA interference system for *in vivo* gene knockdown in the mouse from a single and defined genomic locus. Of note, this is the first time that a conditional knockdown system has been reported in the mouse embryo. Using this system, I have successfully derived RNAi models for three key players in mouse development, T, Foxa2 and Noto, which phenocopy the genetic null mutants. In addition, I have generated hypomorphic RNAi models for both T and Foxa2, which should prove useful tools to dissect events during posterior and anterior development respectively.

Previously, constitutive high-dosage expression of shRNA transgenes under control of the H1 RNA polymerase III promoter has been used for functional RNAi in the mouse embryo (Lickert et al., 2005). This approach did not, however, allow for direct comparison of induced versus non-induced embryos, or to bypass early embryonic-lethal phenotypes. In addition, some experiments reported more severe phenotypes than the respective null mutants. This can be a direct consequence of the sustained, high-level expression of the RNAi trigger by the high-level H1 promoter (Myslinski et al., 2001), which can result in cellular toxicity and death by saturating the endogenous microRNA processing machinery (Grimm et al., 2006). In this respect, the expression of RNAi triggers from pol II promoters has proven to be an effective and safer alternative to achieve experimental gene silencing (Giering et al., 2008).

Other systems have been used in the adult mouse to achieve inducible knockdown under control of pol II promoters (Dickins et al., 2007; Szulc et al., 2006). These relied however on multiple and random genomic integrations by lentiviral/retroviral-based RNAi vectors. The appeal of site-specific chromosomal integration is that variability of transgene expression due to copy number or positional effects is eliminated (Feng et al., 1999). Moreover, it allows controlled integration of transgenes into loci whose disruption is innocuous. In the knockdown system

described here, the combination of RMCE and inducibility allows for derivation of genetically identical experimental and control embryos, ensuring compatibility with high-throughput applications, such as expression profiling or massive parallel sequencing. In addition, distinct RNAi mouse models are generated with minimal genetic differences.

While the results described here apply the RNAi vector system for dissection of gene functions in an embryonic context, such a tool should also prove useful to study gene function in the adult mouse, both during tissue homeostasis as well as in the context of diseases. Construction of exchange vectors requires only routine cloning steps which are standardized for all desired target genes. The subsequent steps of selection and screening are also standardized, making the derivation of loss-of-function mutants fast and efficient. Additionally, the possibility of concatemirizing shRNA-mirs (such as in the *KD1* construct) or co-expressing the shRNA-mir with the coding sequence of a gene (such as in the *KD3* construct) provides the opportunity to target several genes simultaneously, or to combine the silencing of one gene with the overexpression of another.

6.1.1 Ubiquitous and tissue specific knockdown

Two promoters were tested for driving ubiquitous transgene expression in tTS controlled vectors. While use of the CAGGS promoter (Niwa et al., 1991) resulted in uniform reporter expression in E9.5 embryos (Figure 3.5), reproducible transgene levels between induced littermates (Figure 3.6) and ultimately gene knockdown *in vivo* (Figure 3.11, Figure 3.14, Figure 3.15), expression under the control of the CMV promoter did not (Figure 3.6). Indeed, although CMV is amongst the strongest eukaryotic promoters - and therefore often used to drive transgene expression *in vitro* (Fritschy et al., 1996; van den Pol and Ghosh, 1998) - its expression *in vivo* is not consistent, with several reports showing its silencing in multiple mouse tissues (Fritschy et al., 1996; Mehta et al., 2009; Schmidt et al., 1990; Villuendas et al., 2001). Silencing of this promoter has been associated with extensive CpG and non-CpG DNA

methylation (Brooks et al., 2004), recruitment of HDACs by transcriptional repressors (Mulero-Navarro et al., 2006), reduced histone acetylation (Murphy et al., 2002) and H3K4 histone methylation (Mehta et al., 2009).

Similar to expression under the control of the CAGGS promoter in tTS expressing cells, expression through a Tet Responsive Element (TRE) in rtTA carrying cells also leads to uniform reporter expression in mid-gestation embryos and to reproducible gene knockdown *in vivo* (Figure 3.18, Figure 3.19). Interestingly, expression levels of transgenes activated by the rtTA seem to be higher than those obtained with the CAGGS promoter, since *KD5-T* mutants show more severe phenotypes than *KD3-T* mutants, even though both constructs differ only in the promoter fragment. This can be an advantageous feature when modifying vectors to achieve inducible tissue-specific RNA interference. Indeed, while tissue-specific EGFP expression was possible with single copy integration into the *ROSA26* locus (Figure 3.5), expression of a knockdown construct under the same promoter did not result in phenotypic abnormalities nor reduction of target mRNA abundance (not shown). This might be a consequence of generation of insufficient shRNA-mir levels, which result in a suboptimal siRNA/target mRNA ratio. A possible strategy to overcome this issue would be to express the rtTA in a tissue-specific manner. Because regulation by this protein leads to high transgene activation levels, this approach would likely lead to an amplification of the tissue-specific signal at the level of the rtTA-controlled shRNA-mir. This amplification might result in sufficient trigger abundance to induce efficient knockdown only in the cells where rtTA is expressed. In fact, a similar approach has already been applied to produce tissue-specific knockdown in adult mice (Dickins et al., 2007), although it relied on multiple integration of retroviral-based RNAi vectors. Further studies will help elucidate whether this strategy would also be efficient for driving tissue-specific gene silencing from single locus transgenes.

6.1.2 Why is the *KD3-T* phenotype milder than the *KD2-T* phenotype?

Both *KD2-T* and *KD3-T* constructs carry the same shRNA-mir against brachyury (Figure 3.10). Surprisingly, induction of these constructs leads to distinct outcomes *in vivo*. *KD2-T* induced embryos display a clear axis truncation and defects in somitic mesoderm at E9.5. Expression of *KD3-T* on the other hand, leads to a milder phenotype with truncation visible at the level of the notochord but not of the axis length or somite formation. A possible reason for these differences is that the intronic location of T-267 in *KD3-T* embryos might result in the production of lower amounts of mature siRNA. shRNA-mirs mimic the structure of the primary microRNA transcripts and therefore have to undergo the same processing steps in the nucleus and cytoplasm before they can be loaded into a functional RISC complex and lead to gene silencing (Figure 1.4). The location of an shRNA-mir in the intron of a coding sequence - as in the *KD3* construct – may lead to competition events between the intron-degradation processes and the processes involved in the shRNA-mir maturation. Surprisingly, northern blot analysis of RNAs from *KD1-KD3* transgenic embryos did not show detectable differences in expression levels or processing of T-267 between the two constructs. This could be a result of the low sensitivity of the method. Alternatively, the competition events could take place in only a subset of cells, which would not yield detectable differences in RNA extracts from whole embryos, or the dissimilarities between *KD2-T* and *KD3-T* embryos might be independent of processing events. Further studies are required to elucidate the cause of the observed differences.

6.1.3 Can knockdown be achieved in all tissues of a mouse embryo?

The observation that the RMCE system, described in Chapter 3, allows for ubiquitous expression of a transgene during mouse development suggests that it can be applied to

knockdown genes in any tissue of the embryo. Successful gene knockdown however, depends not only on the presence of an RNAi trigger, but also on the machinery that mediates gene silencing. Ago2 has been shown to be essential for siRNA-mediated gene silencing (Liu et al., 2004) and preliminary results suggest that its expression, as assayed by *in situ* hybridization on embryo sections, is excluded the from visceral endoderm during the establishment of the three germ layers (Figure 6.1). These results are supported by transcriptome data from microdissected E6.5 embryos (Markus Morkel and Walter Brichmeier; 2003, unpublished; personal communication), which show absence of Ago2 mRNA in this tissue. Although additional experiments are required to confirm these data, it seems likely that the visceral endoderm is not amenable to experimental RNAi. In this respect, it would be interesting to test if these results are recapitulated by experimental knockdown of a reporter gene expressed ubiquitously during mouse development. If so, then a strategy to circumvent this problem could involve co-expression of Ago2 with the RNAi trigger in the tissues where this enzyme is missing. Such an approach has already been tested *in vitro*, where co-delivery of Ago2 resulted in increased knockdown efficiency without adverse effects (Diederichs et al., 2008). However, since these

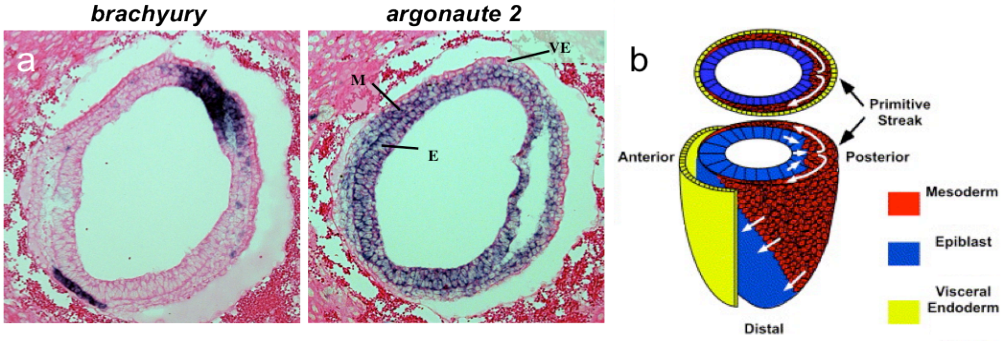


Figure 6.1 Expression of Argonaute 2 in E7.5 mouse embryos

(a) Expression patterns of brachyury (left) and argonaute 2 (right) in transversal sections of E7.5 deciduas, visualized by *in situ* hybridization. Staining for brachyury marks the primitive streak and nascent mesoderm. Staining for argonaute 2 shows expression in the ectoderm (E) and mesoderm (M) but not in the visceral endoderm (VE). (b) Schematic representation of an E7.5 embryo, showing a transverse view on the top. Mesoderm is represented in red, epiblast/ectoderm in blue and visceral endoderm in yellow. Adapted from (Zohn et al., 2006).

studies were limited to cultured cells, the consequences of Ago2 overexpression in the context of a whole animal remain unclear, and would have to be assessed as part of the validation experiments for the applicability of this strategy *in vivo*.

6.1.4 Argonaute 2 and the role of small noncoding RNAs in mouse development

The data discussed above raises the question of what role the restricted expression of Ago2 might play during embryogenesis. Recently, a new class of small noncoding RNAs has been described in mammals. Endogenous small interfering RNAs (endo-siRNAs) are produced from dsRNA precursors that can arise from various genomic sources, including structured loci, convergent and bidirectional transcription, and mRNAs paired to antisense pseudogene transcripts (Tam et al., 2008; Watanabe et al., 2008) (Figure 6.2). Endo-siRNAs are 21

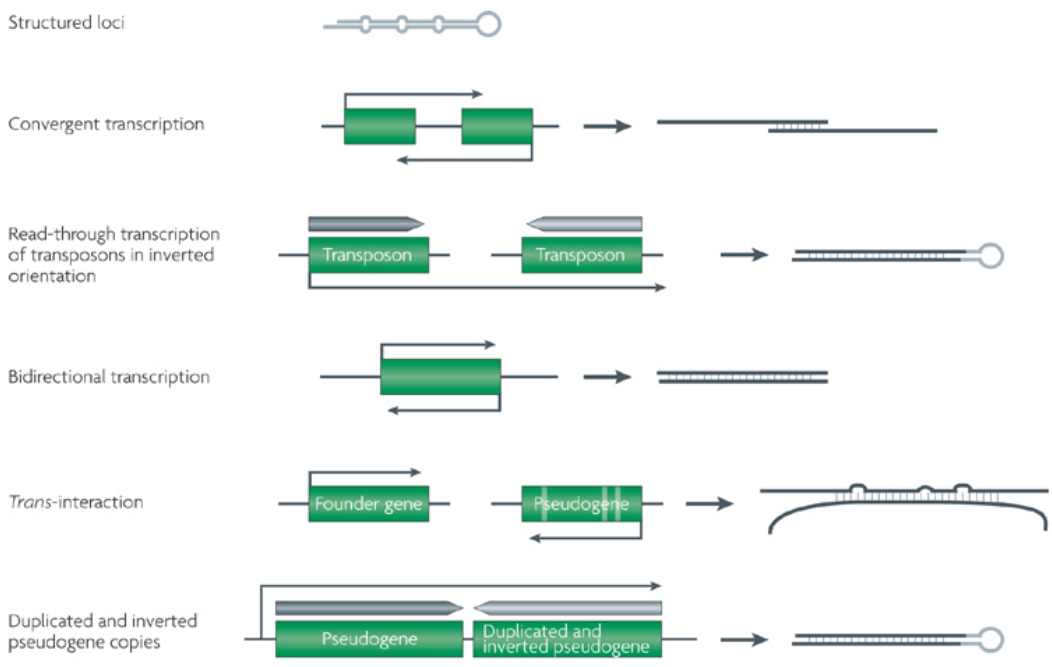


Figure 6.2. Genomic sources of mammalian endo-siRNAs

Endo-siRNAs derive from double stranded RNA precursors. These precursors can originate from structured loci that give rise to transcripts that pair intramolecularly, complementary overlapping transcripts, and bidirectionaly transcribed loci. Additionally, the precursors can be produced by base pairing of an mRNA from a protein coding gene, a peseudogene, or from transcription of regions of pseudogenes that form inverted repeat structures. Solid arrows indicate orientation; black arrows indicate transcription. From (Ghildiyal and Zamore, 2009).

nucleotides long and their maturation was shown to be dependent on Dicer processing in the cytoplasm, but to be independent of the microprocessor complex (Drosha and Dgcr8) in the nucleus (Babiarz and Blelloch, 2009) (Figure 6.3). They require Ago2 to silence genes (Watanabe et al., 2008) and their function was found to be essential during the early stages of mouse development. Indeed, the strong phenotypic abnormalities observed in the absence of Dicer, which include arrested development at gastrulation (in Dicer$^{-/-}$ embryos developing in the presence of maternally contributed Dicer (Bernstein et al., 2003)) or arrest at the first cell division (in the absence of zygotic *and* maternal Dicer (Tang et al., 2007)), have been recently shown to be a consequence of depletion of mature endo-siRNAs and not miRNAs as previously thought

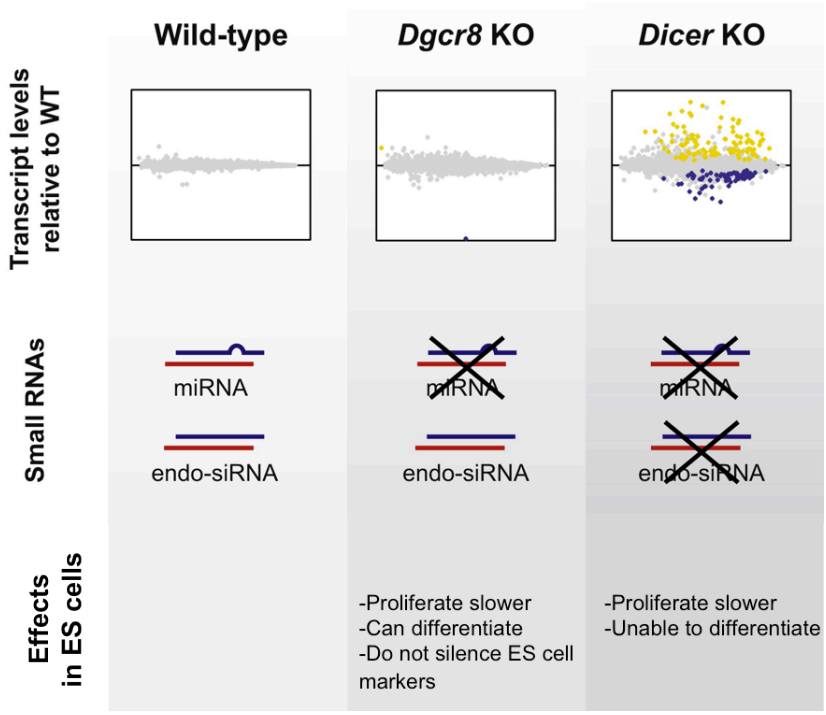


Figure 6.3 Effect of the absence of Dicer and Dgcr8 in the mouse

Top panel shows heat-maps of up-regulated (yellow) or down-regulated (blue) transcripts in Dgcr8$^{-/-}$ (Dgcr8 KO), and Dicer$^{-/-}$ (Dicer KO) relative to wild-type cells. A schematic representation in the middle shows which small RNAs are processed correctly in the absence of these enzymes. In ES cells, absence of *Dicer* results in slow proliferating cells that are unable to differentiate (Kanellopoulou et al., 2005). In the absence of *Dgcr8*, ES cells also proliferate slowly. They can differentiate to some extent, although they never silence stem cell markers (Wang et al., 2007). Figure adapted from (Suh et al., 2010).

(Ma et al., 2010; Suh et al., 2010). In addition, comparative analysis of ES cells lacking both miRNA and endo-siRNA functions, and ES cells lacking miRNA function alone (Figure 6.3) suggests that endo-siRNAs play important roles during cell differentiation (Kanellopoulou et al., 2005; Wang et al., 2007).

Similarly to Dicer, knockout of *Ago2* leads to arrested development at the two-cell-stage in the absence of both the maternal and zygotic enzyme (Lykke-Andersen et al., 2008), or at gastrulation when the maternal protein is present (Morita et al., 2007). These defects implicate Ago2 in several steps of mouse development. Although, to date, endo-siRNAs in the mouse have only been described in oocytes, it is tempting to speculate that the phenotypic defects seen at gastrulation in the absence of Ago2 are a reflection of defective gene regulation by these small RNAs - since the function of this protein in the miRNA pathway seems to be redundant with that of the remaining members of the Argonaute family (Su et al., 2009) – and that absence of *Ago2* expression in the visceral endoderm plays a role during the establishment of the body plan at gastrulation.

6.2 Analysis of *T* knockdown models

The results described in Chapters 4 and 5, obtained from the analysis of the *KD3-T* and *KD4-T* RNAi models respectively, exemplify the usefulness of applying inducible RNAi to dissect gene function. Activation of the *KD3-T* transgene *in vivo* results in a hypomorphic phenotype for brachyury, where mesoderm formation is compromised but not arrested. In this model, embryos can be analyzed far past the embryonic stage when absence of Brachyury becomes lethal. Hypomorphic phenotypes, like this one, have been previously shown to be useful in the study of several genes (Goldman et al., 2009; Greco et al., 1996; Rocha et al., 2010), but they are usually generated serendipititously. In this respect, a distinct advantage of the 'knockdown' versus 'knockout' technology is that the level of target mRNA depletion can be controlled, allowing the purposeful generation of mice with *lower* but not *absent* gene expression. The *KD4-T* mouse model, on the other hand, allows the study of T loss-of-function phenotypes in a temporally controlled manner. Using this model, I have uncovered a previously unknown role for Brachyury during skeletal formation. Importantly, the defects observed upon *T* knockdown occur at the level of the axial skeleton (skull, spine, sternum and ribs), with the appendicular skeleton (comprised of upper and lower limbs, pelvis, and pectoral girdles) developing normally. This provides further evidence of the specificity of the knockdown, which affects only the structures surrounding the notochord and does not give rise to generalized malformations as a consequence of ubiquitous expression of the RNAi trigger.

6.2.1 T and the uro-rectal-caudal syndrome

Congenital urorectal defects occur in humans with a frequency of approximately 1 in 2500 to 1 in 5000 live births (Cho et al., 2001; Cuschieri, 2001). Although a genetic cause has been established for some of these malformations (Kang et al., 1997), the molecular and morphological mechanisms behind them remain unclear. In the mouse, mutations in several

signaling pathways including Wnt (Yamaguchi et al., 1999a), Fgf (Petiot et al., 2005), and Bmp (Sasaki et al., 2004) lead to abnormal urorectal development. However, most of these mutations result in other severe developmental defects and early embryonic death. Therefore, animal models with reproducible malformations in the urorectal system that can be used to examine later steps in the development of these diseases are scarce. The analysis of the *KD3-T* mouse model, showing highly reproducible phenotypes consistent with the murine uro-rectal-caudal syndrome (Gruneberg, 1952), will help in the greater understanding of urorectal development, which is currently ill-defined, and will help identify new candidate genes that might be causative of the human disease.

6.2.2 Role of brachyury during skeletal development

Previous studies on Brachyury have focused mainly on its roles in early embryogenesis and established it as an essential factor during notochord differentiation. In Chapter 5, I have focused on T functions in later stages of notochord development, when this structure is involved in mesenchymal cell induction and directly contributes to the formation of the vertebral column. The phenotypes obtained upon knockdown of brachyury implicate it in several steps of skeletal formation. Firstly, expression of Brachyury seems to be required for correct induction of bone mesenchymal precursors and secondly, its presence is necessary for the differentiation of the nucleus pulposus of the vertebral column. These functions are summarized in Figure 6.4, and will be discussed further below.

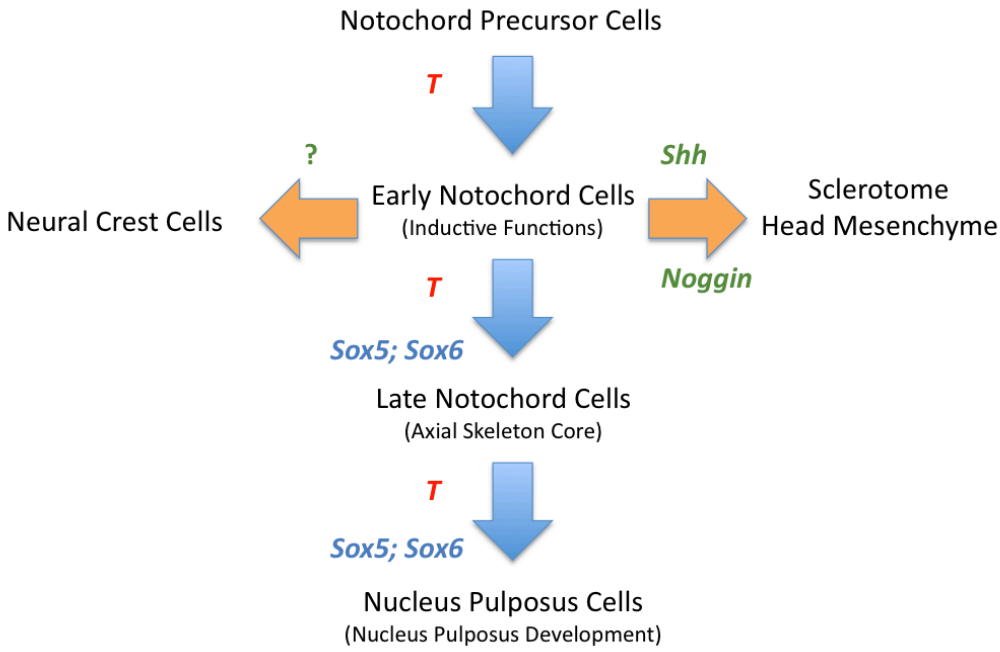


Figure 6.4. A model for Brachyury function during notochord development

Notochord development involves four major differentiation steps, which seem to be regulated by T. Precursor cells migrate through the primitive streak and require T to differentiate into the early notochord. This early structure is able to induce formation and differentiation of neighboring tissues via secreted signals, and this function is compromised by brachyury knockdown. The sclerotome and the head mesenchyme are know to respond to Shh and Noggin proteins emanating from the notochord, but a notochordal signal acting positively on neural crest cells has not been described so far. As development progresses, notochordal cells are integrated into the axial skeleton and ultimately give rise to the nucleus pulposus of the vertebral column. These steps also seem to require the presence of T, as well as of Sox5 and Sox6. Diagram modified from (Smits and Lefebvre, 2003).

6.2.3 T and the induction of neural crest cells

Knockdown of brachyury beginning at E8.5 leads to absence of hyoid bone ossification in E18.5 fetuses. This phenotype is fully penetrant, and seems to be caused by abnormal neural crest cell (NCC) migration, which results in incomplete colonization of the second branchial arch (Figure 5.7). This is unexpected for two reasons: first, signals from the notochord are know to *inhibit* the migration of NCCs, keeping them from crossing the midline, but a positive influence of the notochord on these cells has so far not been reported. Secondly, these migration defects seem to be specific to only a subset of NCCs, since most of the craniofacial structures, which also derive from cranial NCCs, appear to develop normally.

The notochord is known to be involved in both long- and short-range signaling to neighboring tissues throughout embryogenesis. During somite differentiation for example, Shh signals to nearby cells to induce differentiation into the sclerotome lineage, while in cells positioned further away - where the signal concentration is lower - markers of the dermomyotome lineage are upregulated giving rise to muscle and skin (Gilbert, 2006). It is possible that NCCs are also subject to a similar type of regulation, with high concentrations of one or more molecules inhibiting NCC migration and lower concentrations inducing it. Such long-range effects could have easily been overlooked in the *in vivo* transplantation experiments, where ectopically positioned notochords led to a striking absence of NCCs from the surrounding regions (Pettway et al., 1990). However, to date, no signaling molecule from this structure has been described that is able to positively influence NCCs specification/migration.

Another puzzling observation is that knockdown of brachyury seems to affect mainly cells that migrate to the second branchial arch. NCCs start delaminating from the roof plate of the neural tube as soon as this structure has formed (Figure 4.2a and b), which in the cranial region of the mouse occurs between E8.0 and late E9.0 (Copp et al., 1990), around the time point when, in *KD4-T* induced mice, *T* mRNA has been depleted from the notochord by RNAi. The impact of the *T* knockdown in only a subset of NCCs suggests that there is already some heterogeneity in the population prior to migration, which results in distinct responses to the depletion of *T*. A growing body of work suggests that not all pre-migratory NCCs are alike. When individual cells are analyzed *in vitro* following serial dilution, they form colonies with distinct cell types (Le Douarin and Dupin, 1993), which can respond differently to growth factors (Zhang et al., 1997). This is, in part, mediated by differential expression of membrane receptors among the pre-migratory population (Henion et al., 1995; Meyer and Birchmeier, 1995; Schatteman et al., 1992). These molecular differences have been shown to correlate with the timing of NCC migration from the neural tube (Ma et al., 1999), and it is possible that they also correlate with the axial level from which they originate. In the context of the phenotypes described in Chapter 5,

it plausible that presence of Brachyury in the notochord is required for the expression of an as yet unidentified signal (Figure 6.4), to which a subset of cells migrating from the fourth rhombomere to the second branchial arch can respond to, or that the timing of knockdown leads to a second arch bias.

6.2.4 T and the induction of mesoderm-derived mesenchymal cells

The majority of the remaining defects in *KD4-T* embryos point towards defective skeletogenesis mediated by paraxial mesodermal cells. Like NCCs, mesoderm-derived mesenchymal cells have to migrate to the site of bone formation and differentiate into the correct skeletal lineages. Several transcription factors have been shown to be important for these processes. Foxc2, a member of the forkehead family of transcription factors, is expressed in the sclerotome and in the head mesenchyme, and its function is required for normal development of the axial skeleton. Mice lacking Foxc2 do not form the supraoccipital bone, have a reduced interparietal bone, and have several vertebral defects which include absence of ossification centers in the cervical vertebrae and split ossification centers in thoracic, lumbar and sacral vertebrae (Winnier et al., 1997). These defects are strikingly similar to those observed in *KD4-T* mice (Figure 5.4, Figure 5.5). Another transcription factor, Pax1, is necessary for vertebral and rib differentiation. Pax1 is expressed in all sclerotomal cells at E8.5 (Deutsch et al., 1988). Later, its expression in the axial skeleton becomes confined to the mesenchymal condensations that will give rise to the intervertebral discs of the column (Wallin et al., 1994), the pharyngeal pouches (Wallin et al., 1996), and the developing sternum (Wilm et al., 1998). *Pax1*-null mutants show several skeletal abnormalities which also include absence of ossification in centra of the cervical vertebrae and split ossification centers in more posterior vertebrae as well as fused ossification of the sternebral bars of the rib cage (Wilm et al., 1998).

Both Pax1 and Foxc2 activate skeletogenesis by promoting differentiation of mesodermal cells (Kim et al., 2009; Rodrigo et al., 2003) and initiation of their expression is known to be controlled by signals emanating from the notochord (Furumoto et al., 1999). The overlap of defects observed in the absence of both Pax1 and Foxc2 with the ones in *KD4-T* mutant embryos might indicate that the skeletal defects seen in upon knockdown of brachyury are related to abnormal mesoderm induction due to impaired notochord signaling to the neighboring tissues. This would certainly be in line with the data discussed above, implicating Brachyury in the control of NCC migration. The question that then arises is which signaling factor is affected by the knockdown of *T*? A key molecule in notochordal inductivation events is Shh, a diffusible signal that can induce sclerotome over considerable distances (Fan, 1995). Mice lacking Shh have a significantly smaller sclerotome and express *Pax1* and other differentiation markers at drastically reduced levels, which results in severe skeletal abnormalities. These defects have been associated with a role of Shh in the maintenance/expansion of the population of committed sclerotome cells rather than specification of the sclerotomal fate (Chiang et al., 1996). A second factor secreted by the notochord, Noggin, seems to act at the level of sclerotome initiation. Noggin is expressed in the notochord at the time of sclerotome specification and homozygous *Noggin* mutants display a drastic delay in *Pax1* expression. Interestingly, Noggin can initiate *Pax1* expression in presomitic mesoderm explants independently of Shh activity (McMahon et al., 1998), suggesting that *in vivo* Noggin acts to initiate low levels of *Pax1* expression (and perhaps of other markers as well) and that Shh acts to maintain and augment these levels (McMahon et al., 1998).

In the case of *KD4-T* mutant mice, further experiments will help elucidate whether the skeletal defects observed are a consequence of abnormal specification or maintenance of the sclerotome. Whole-mount *in situ* hybridization experiments on *KD4-T* embryos induced at E7.5 show that upon knockdown of brachyury the notochord still expresses Shh. Nonetheless, the levels of expression may be reduced in comparison to wild-type embryos, leading to abnormal

activation of its downstream targets. Alternatively, Noggin expression may be impaired, but this point has not been experimentally addressed so far.

6.2.5 T in the development and disease of nuclei pulposi

This study has identified a critical role for Brachyury in the transcriptional regulation of the nucleus pulposus. Very little is known about the genetic program of these cells. They continue to express T, as well as Sox5 and Sox6, and all three transcription factors where shown to be required for their differentiation. In *Sox5*$^{-/-}$; *Sox6*$^{-/-}$ compound mutants, the notochord is formed and fulfils its inductive functions, but fails to become surrounded by an extracellular matrix sheet, and disintegrates after sclerotome cell migration via apoptosis. As a result, fetuses display a vertebral column devoid of nuclei pulposi (Smits and Lefebvre, 2003). This is a similar phenotype to that observed with brachyury RNAi (Figure 5.9). *Sox5*$^{-/-}$; *Sox6*$^{-/-}$ mutants continue to express *T*, but *KD4-T* mutants display a dramatic reduction in *Sox6* mRNA, suggesting that brachyury might play a role upstream of this transcription factor in the development of the nucleus pulposus. Interestingly, *in vitro* differentiation of *KD4-T* ES cells with or without dox, showed a dramatic decrease in expression of extracellular matrix proteins upon knockdown of *T* (Eun-ha Shin and Bernhard G. Herrmann; unpublished results), indicating that, like Sox5 and Sox6, Brachyury might also play a role in the formation of the notochordal extracellular sheet. Accordingly, expression profiles from chordomas (malignant neoplasms derived from remnants of the notochord, whose development in humans has been associated with a duplication of the brachyury locus (Yang et al., 2009)) also show a significant up-regulation of extracellular matrix genes (Fujita et al., 2005; Henderson et al., 2005; Vujovic et al., 2006). Chordomas are rare malignancies. They make up only 0.2% of all tumors arising in the central nervous system and about 3% of all primary bone neoplasms. The average age of diagnosis is 58, and there are currently only a handful of effective treatments and no cure, leading to death usually within 10 years of the initial diagnosis (U.S. National Institute of Health; www.cancer.gov). To date,

brachyury is the only gene to have been implicated in this disease, and immunohistochemical analysis has shown that its expression is highly specific for this type of tumor (Vujovic et al., 2006). Since both nuclei pulposi and chordomas arise from notochordal cells (Walmsley, 1953; Yamaguchi et al., 2004) and their development has been associated with Brachyury expression, it is likely that they rely at least partially on the same transcriptional program. The study of Brachyury function in the context of normal notochordal development, as well as during disease, will likely result in a better understanding of both processes and the molecular mechanisms that drive them. This may ultimately lead to the identification of new candidate targets for drug development.

7 SUMMARY

Functional analysis of genes is key to understanding gene regulatory networks controlling embryonic development. I have developed an integrated vector system for inducible gene silencing by shRNA-mir-mediated RNA interference in mouse embryos, as a fast method for dissection of mammalian gene functions. As a validation of the vector system, I have generated loss-of-function phenotypes for Brachyury, Foxa2 and Noto transcription factors - which play pivotal roles in embryonic development - as well hypomorphic phenotypes. Knockdown of genes in this system can be temporally controlled, thus allowing to bypass early embryonic lethality caused by loss-of-function of key developmental genes. Importantly, off-target effects in this system were not detectable at the transcriptome level.

As evidence for the merit of this system, I have used two RNAi models for brachyury to study its function during embryogenesis. Analysis of a hypomorphic mutant for brachyury established this gene as regulator of urorectal organogenesis during mouse trunk development. This model has now integrated a panel of few others, which can be used to understand the molecular mechanisms behind urorectal malformations. Additionally, temporal control over the onset of knockdown established brachyury as a key regulator of skeletogenesis, controlling mesenchymal cell induction by the notochord as well as differentiation of the vertebral column.

Taken together, the results described here show that this system allows the dissection of gene functions at unprecedented detail and speed.

8 ZUSAMMENFASSUNG

Für das Verständnis der genregulativen Netzwerke während der Embryonalentwicklung ist die Analyse von Genfunktion essentiell. Eine solche Funktionsanalyse kann durch die gezielte Inaktivierung von Genen erreicht werden. Um eine schnelle und einfache Inaktivierung von Genen zu ermöglichen, habe ich ein flexibles Transgensystem entwickelt, mit dem Gene während der Embryogenese mittels eines miRNA basierten shRNA-Knockdown aus und wieder angeschaltet werden können. Zur Validierung dieses Systems wurden Embryonen generiert, in denen die Transkriptionsfaktoren Brachyury, Foxa2 und Noto partiell (hypomorph) und vollständig inaktiviert wurden. Embryonen mit kompletten Knockdown phänokopieren beschriebene Knock-Out Phänotypen. Im Gegensatz zum konventionellen Knockout Experiment, ist der Knockdown in meinem Transgensystem zeitlich kontrollierbar, was die Analyse von Genen, die in der frühen Embryogenese eine essentielle Funktion haben und deren Funktionsverlust zu früher embryonaler Letalität führt, zu späteren Zeitpunkten der Entwicklung ermöglicht. Unspezifische Effekte der Expression der miRNAs, waren in diesem System nicht detektierbar.

Im zweiten Teil der Dissertation habe ich zur weiteren Validierung des Systems und zu einer genaueren Analyse der Funktion von Brachyury, zwei verschiedene Modelle für die Inaktivierung von brachyury verwendet. Zum einen zeigte die Analyse der hypomorphen Variante von brachyury, dass Brachyury eine wichtige Funktion bei der Etablierung des Urogenitaltraktes während der Rumpfentwicklung übernimmt. Hiermit wurde eines der ersten Mausmodelle für die Untersuchung von urorektalen Fehlbildungen, die auch in der menschlichen Entwicklung auftreten, etabliert. Zum Anderen konnte ich durch die komplette Inaktivierung von brachyury zu einem späteren Zeitpunkt der Embryogenese zeigen, dass Brachyury eine entscheidende Rolle bei der Skelettentwicklung hat. Brachyury kontrolliert hierbei die Induktion von mesenchymalen Zellen über den Notochord, als auch die Differentierung der Wirbelsäule.

Zusammenfassend, konnte ich innerhalb der Dissertation darlegen, dass das von mir entwickelte Transgensystem eine schnelle, effektive und flexible Analyse von Genfunktion erlaubt.

9 ABBREVIATIONS

Ago	Argonaute
bp	base pair
cDNA	complementary DNA
CMV	cytomegalovirus
Dgcr8	DiGeorge syndrome critical region 8
DNA	desoxyribunucleic acid
dox	doxycycline
dsRNA	double-stranded RNA
E	embryonic day
EC cell	embryonal carcinoma cell
(E)GFP	(enhanced) green fluorescent protein
ES cell	embryonic stem cell
Fgf	fibroblast growth factor family
floxed	flanked by lox sites
H&E	hematoxylin and eosin
HPRT	hypoxanthine phosphoribosyltransferase
ICM	inner cell mass
Kb	kilo base
KRAB	Krueppel-associated Box
miRNA	microRNA
mRNA	messenger RNA
NCC	neural crest cell
nt	nucleotide
piRNA	Piwi interacting RNA
Pol	RNA polymerase
polyA	poly-adenylation
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA transcript
qPCR	quantitative polymerase chain reaction
RISC	RNA-induced silencing complex
RMCE	recombinase mediated cassette exchange
RNA	ribonucleic acid
RNAi	RNA interference
RT	room temperature
rtTA	reverse tTA
Shh	sonic hedgehog
shRNA	short hairpin RNA
siRNA	small interfering RNA
ssRNA	single strabded RNA
Tc	tetracycline
tetO	tet operator
tetR	tetracycline controlled transcriptional repressor
TGF	transforming growth factor
TRE	tetracycline responsive element
tTA	tetracycline controlled transactivator
tTS	tetracycline-controlled transcriptional silencer
UTR	untranslated region
Wnt	wingless-type MMTV integration site family
Wt	wild-type

10 ACKNOWLEDGMENTS

First of all, I want to thank Dr. Markus Morkel who supervised the biggest part of the work I present here, as well as Prof. Bernhard G. Herrmann, for supporting me as a PhD student in his department. I thank Prof. Sephan Sigrist from the Institute for Biology for serving as my doctoral advisor at the Freie Universität Berlin.

This work would not have been feasible without the help of numerous members of the Max Planck Institute for Molecular Genetics. First and foremost I thank Antje Brouwer-Lehmitz, an outstanding technician in our department, whose endless efforts made possible the establishment of the inducible knockdown system. I thank Ingo Voigt and Karol Macura for introducing me to the fascinating technique of tetraploid complementation, and Dr. Lars Wittler for his incessant help phenotyping and microdissecting mouse embryos. Most of all, I thank him for sharing his love for science. I thank also Dr. Ludger Hartmann from the Animal Facility and his team, in particular Sonja Banko and Eileen Jungnickel for excellent animal work.

I am grateful to all members of the Developmental Genetics Department for providing a great working atmosphere and for helping me solve the everyday problems in the lab. In particular, I thank Manuela Scholze, without whom the department would stop. I thank also Benedikt Schwartz, Marc Leushacke and Eun-ha Shin who started their PhDs with me, and with whom I shared most of the good and bad experiences associated to a doctoral thesis. I thank also Arica Beisaw, for her friendship and good humor and Dr. Heinrich Schrewe and Dr. Tracie Pennimpede for the help with the thesis. Most of all, I thank Tracie for making me laugh even in the most stressful moments and for always offering a helping hand with a smile.

None of the work presented here would have been possible without my family and friends in Portugal, who always receive me back with opened arms and remind me how good it feels to be home. I thank specially my parents and siblings, to whom I dedicate this thesis.

Above all others I thank Pedro. You are my rock in stormy weathers. With you I feel at home.

11 PUBLICATIONS

- **Vidigal JA,** Morkel M, Wittler L, Brouwer-Lehmitz A, Grote P, Macura K, Herrman BG (2010). An inducible RNA interference system for the functional dissection of mouse embryogenesis. Nucleic Acids Research; 38(11):e122

- **Vidigal JA** and Herrman BG. Skeletal abnormalities in brachyury loss-of-function mice. (in preparation).

12 CURRICULUM VITAE

For reasons of data protection, the Curriculum vitae is not published in the online version

13 BIBLIOGRAPHY

Abu-Abed, S., Dolle, P., Metzger, D., Beckett, B., Chambon, P. and Petkovich, M. (2001). The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. *Genes Dev* **15**, 226-40.

Akiyama, H. (2008). Control of chondrogenesis by the transcription factor Sox9. *Mod Rheumatol* **18**, 213-9.

Albert, H., Dale, E. C., Lee, E. and Ow, D. W. (1995). Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J* **7**, 649-59.

Ang, S. L. and Rossant, J. (1994). HNF-3 beta is essential for node and notochord formation in mouse development. *Cell* **78**, 561-74.

Arnold, S. J. and Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* **10**, 91-103.

Aszodi, A., Chan, D., Hunziker, E., Bateman, J. F. and Fassler, R. (1998). Collagen II is essential for the removal of the notochord and the formation of intervertebral discs. *J Cell Biol* **143**, 1399-412.

Babiarz, J. E. and Blelloch, R. (2009). Small RNAs - their biogenesis, regulation and function in embryonic stem cells. *StemBook (http://www.stembook.org/node/583*.

Bailey, D. W. (1971). Recombinant-inbred strains. An aid to finding identity, linkage, and function of histocompatibility and other genes. *Transplantation* **11**, 325-7.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-97.

Beckers, A., Alten, L., Viebahn, C., Andre, P. and Gossler, A. (2007). The mouse homeobox gene Noto regulates node morphogenesis, notochordal ciliogenesis, and left right patterning. *Proc Natl Acad Sci U S A* **104**, 15765-70.

Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V. and Hannon, G. J. (2003). Dicer is essential for mouse development. *Nat Genet* **35**, 215-7.

Bogenhagen, D. F. and Brown, D. D. (1981). Nucleotide sequences in Xenopus 5S DNA required for transcription termination. *Cell* **24**, 261-70.

Bradley, A., Evans, M., Kaufman, M. H. and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-6.

Branda, C. S. and Dymecki, S. M. (2004). Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev Cell* **6**, 7-28.

Brinster, R. L., Chen, H. Y., Trumbauer, M., Senear, A. W., Warren, R. and Palmiter, R. D. (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* **27**, 223-31.

Brooks, A. R., Harkins, R. N., Wang, P., Qian, H. S., Liu, P. and Rubanyi, G. M. (2004). Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *J Gene Med* **6**, 395-404.

Brummelkamp, T. R., Bernards, R. and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550-3.

Cai, X., Hagedorn, C. H. and Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**, 1957-66.

Calegari, F., Haubensak, W., Yang, D., Huttner, W. B. and Buchholz, F. (2002). Tissue-specific RNA interference in postimplantation mouse embryos with endoribonuclease-prepared short interfering RNA. *Proc Natl Acad Sci U S A* **99**, 14236-40.

Capecchi, M. R. (2001). Generating mice with targeted mutations. *Nat Med* **7**, 1086-90.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-5.

Chiang, C., Litingtung, Y., Lee, E., Young, K. E., Corden, J. L., Westphal, H. and Beachy, P. A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383**, 407-13.

Cho, S., Moore, S. P. and Fangman, T. (2001). One hundred three consecutive patients with anorectal malformations and their associated anomalies. *Arch Pediatr Adolesc Med* **155**, 587-91.

Ciruna, B. G., Schwartz, L., Harpal, K., Yamaguchi, T. P. and Rossant, J. (1997). Chimeric analysis of fibroblast growth factor receptor-1 (Fgfr1) function: a role for FGFR1 in morphogenetic movement through the primitive streak. *Development* **124**, 2829-41.

Copp, A. J., Brook, F. A., Estibeiro, J. P., Shum, A. S. and Cockroft, D. L. (1990). The embryonic development of mammalian neural tube defects. *Prog Neurobiol* **35**, 363-403.

Copp, A. J. and Greene, N. D. (2009). Genetics and development of neural tube defects. *J Pathol* **220**, 217-30.

Copp, A. J., Greene, N. D. and Murdoch, J. N. (2003). The genetic basis of mammalian neurulation. *Nat Rev Genet* **4**, 784-93.

Costantini, F. and Lacy, E. (1981). Introduction of a rabbit beta-globin gene into the mouse germ line. *Nature* **294**, 92-4.

Couly, G., Creuzet, S., Bennaceur, S., Vincent, C. and Le Douarin, N. M. (2002). Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development* **129**, 1061-73.

Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D. and Bogenhagen, D. F. (1983). Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell* **34**, 829-35.

Cullen, B. R. (2004). Transcription and processing of human microRNA precursors. *Mol Cell* **16**, 861-5.

Cullen, B. R. (2005). RNAi the natural way. *Nat Genet* **37**, 1163-5.

Cuschieri, A. (2001). Descriptive epidemiology of isolated anal anomalies: a survey of 4.6 million births in Europe. *Am J Med Genet* **103**, 207-15.

Deutsch, U., Dressler, G. R. and Gruss, P. (1988). Pax 1, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. *Cell* **53**, 617-25.

Dickins, R. A., McJunkin, K., Hernando, E., Premsrirut, P. K., Krizhanovsky, V., Burgess, D. J., Kim, S. Y., Cordon-Cardo, C., Zender, L., Hannon, G. J. et al. (2007). Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* **39**, 914-21.

Diederichs, S., Jung, S., Rothenberg, S. M., Smolen, G. A., Mlody, B. G. and Haber, D. A. (2008). Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites. *Proc Natl Acad Sci U S A* **105**, 9284-9.

Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S. and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* **330**, 576-8.

Douzery, E. J., Snell, E. A., Bapteste, E., Delsuc, F. and Philippe, H. (2004). The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc Natl Acad Sci U S A* **101**, 15386-91.

Eakin, G. S. and Hadjantonakis, A. K. (2006). Production of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. *Nat Protoc* **1**, 1145-53.

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**, 494-8.

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

Eulalio, A., Huntzinger, E. and Izaurralde, E. (2008). Getting to the root of miRNA-mediated gene silencing. *Cell* **132**, 9-14.

Evans, M. J. (1972). The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. *J Embryol Exp Morphol* **28**, 163-76.

Evans, M. J. (2001). The cultural mouse. *Nat Med* **7**, 1081-3.

Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-6.

Fan, C.-M. (1995). Long-Range Sclerotome Induction by Sonic Hedgehog: Direct Role of the Amino-Terminal Cleavage Product and Modulation by the Cyclic AMP Siganling Pathway. *Cell*, 457-465.

Feng, Y. Q., Seibler, J., Alami, R., Eisen, A., Westerman, K. A., Leboulch, P., Fiering, S. and Bouhassira, E. E. (1999). Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. *J Mol Biol* **292**, 779-85.

Filippov, V., Solovyev, V., Filippova, M. and Gill, S. S. (2000). A novel type of RNase III family proteins in eukaryotes. *Gene* **245**, 213-21.

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**, 806-11.

Folger, K. R., Wong, E. A., Wahl, G. and Capecchi, M. R. (1982). Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Mol Cell Biol* **2**, 1372-87.

Fortin, K. R., Nicholson, R. H. and Nicholson, A. W. (2002). Mouse ribonuclease III. cDNA structure, expression analysis, and chromosomal location. *BMC Genomics* **3**, 26.

Freundlieb, S., Schirra-Muller, C. and Bujard, H. (1999). A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *J Gene Med* **1**, 4-12.

Friedrich, G. and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5**, 1513-23.

Fritschy, J. M., Brandner, S., Aguzzi, A., Koedood, M., Luscher, B. and Mitchell, P. J. (1996). Brain cell type specificity and gliosis-induced activation of the human cytomegalovirus immediate-early promoter in transgenic mice. *J Neurosci* **16**, 2275-82.

Fujita, N., Miyamoto, T., Imai, J., Hosogane, N., Suzuki, T., Yagi, M., Morita, K., Ninomiya, K., Miyamoto, K., Takaishi, H. et al. (2005). CD24 is expressed specifically in the nucleus pulposus of intervertebral discs. *Biochem Biophys Res Commun* **338**, 1890-6.

Furumoto, T. A., Miura, N., Akasaka, T., Mizutani-Koseki, Y., Sudo, H., Fukuda, K., Maekawa, M., Yuasa, S., Fu, Y., Moriya, H. et al. (1999). Notochord-dependent expression of MFH1 and PAX1 cooperates to maintain the proliferation of sclerotome cells during the vertebral column development. *Dev Biol* **210**, 15-29.

Gaszner, M. and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* **7**, 703-13.

George, S. H., Gertsenstein, M., Vintersten, K., Korets-Smith, E., Murphy, J., Stevens, M. E., Haigh, J. J. and Nagy, A. (2007). Developmental and adult phenotyping directly from mutant embryonic stem cells. *Proc Natl Acad Sci U S A* **104**, 4455-60.

Ghildiyal, M. and Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**, 94-108.

Giering, J. C., Grimm, D., Storm, T. A. and Kay, M. A. (2008). Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic. *Mol Ther* **16**, 1630-6.

Gilbert, S. F. (2000). Principles of Development. *Sinauer Associates, Inc.,Publishers Sunderland; MA, USA*.

Gilbert, S. F. (2006). Developmental Biology, Eigth Edition. *Sinauer Associates, Inc.,Publishers Sunderland; MA, USA*.

Goldman, D. C., Bailey, A. S., Pfaffle, D. L., Al Masri, A., Christian, J. L. and Fleming, W. H. (2009). BMP4 regulates the hematopoietic stem cell niche. *Blood* **114**, 4393-401.

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A. and Ruddle, F. H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci U S A* **77**, 7380-4.

Gossen, M. and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **89**, 5547-51.

Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W. and Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766-9.

Gossler, A., Doetschman, T., Korn, R., Serfling, E. and Kemler, R. (1986). Transgenesis by means of blastocyst-derived embryonic stem cell lines. *Proc Natl Acad Sci U S A* **83**, 9065-9.

Greco, T. L., Takada, S., Newhouse, M. M., McMahon, J. A., McMahon, A. P. and Camper, S. A. (1996). Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. *Genes Dev* **10**, 313-24.

Gregory, R. I., Chendrimada, T. P., Cooch, N. and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631-40.

Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N. and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-40.

Grimm, D., Streetz, K. L., Jopling, C. L., Storm, T. A., Pandey, K., Davis, C. R., Marion, P., Salazar, F. and Kay, M. A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**, 537-41.

Gruneberg, H. (1952). The Genetics of the Mouse, 2nd Edition. *The Hague: Martinus Nijhof*, 650.

Gu, H., Marth, J. D., Orban, P. C., Mossmann, H. and Rajewsky, K. (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103-6.

Hall, B. K. and Miyake, T. (2000). All for one and one for all: condensations and the initiation of skeletal development. *Bioessays* **22**, 138-47.

Hallonet, M., Kaestner, K. H., Martin-Parras, L., Sasaki, H., Betz, U. A. and Ang, S. L. (2002). Maintenance of the specification of the anterior definitive endoderm and forebrain depends on the axial mesendoderm: a study using HNF3beta/Foxa2 conditional mutants. *Dev Biol* **243**, 20-33.

Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950-2.

Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., Sohn, S. Y., Cho, Y., Zhang, B. T. and Kim, V. N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**, 887-901.

Han, Y. and Lefebvre, V. (2008). L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer. *Mol Cell Biol* **28**, 4999-5013.

Harbers, K., Jahner, D. and Jaenisch, R. (1981). Microinjection of cloned retroviral genomes into mouse zygotes: integration and expression in the animal. *Nature* **293**, 540-2.

Hatfield, S. D., Shcherbata, H. R., Fischer, K. A., Nakahara, K., Carthew, R. W. and Ruohola-Baker, H. (2005). Stem cell division is regulated by the microRNA pathway. *Nature* **435**, 974-8.

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. et al. (2005). A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828-33.

Henderson, S. R., Guiliano, D., Presneau, N., McLean, S., Frow, R., Vujovic, S., Anderson, J., Sebire, N., Whelan, J., Athanasou, N. et al. (2005). A molecular map of mesenchymal tumors. *Genome Biol* **6**, R76.

Henion, P. D., Garner, A. S., Large, T. H. and Weston, J. A. (1995). trkC-mediated NT-3 signaling is required for the early development of a subpopulation of neurogenic neural crest cells. *Dev Biol* **172**, 602-13.

Herrmann, B. G. and Kispert, A. (1994). The T genes in embryogenesis. *Trends Genet* **10**, 280-6.

Hirokawa, N., Tanaka, Y. and Okada, Y. (2009). Left-right determination: involvement of molecular motor KIF3, cilia, and nodal flow. *Cold Spring Harb Perspect Biol* **1**, a000802.

Hoess, R. H., Ziese, M. and Sternberg, N. (1982). P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc Natl Acad Sci U S A* **79**, 3398-402.

Hofmann, M., Schuster-Gossler, K., Watabe-Rudolph, M., Aulehla, A., Herrmann, B. G. and Gossler, A. (2004). WNT signaling, in synergy with T/TBX6, controls Notch signaling by regulating Dll1 expression in the presomitic mesoderm of mouse embryos. *Genes Dev* **18**, 2712-7.

Iida, K., Koseki, H., Kakinuma, H., Kato, N., Mizutani-Koseki, Y., Ohuchi, H., Yoshioka, H., Noji, S., Kawamura, K., Kataoka, Y. et al. (1997). Essential roles of the winged helix transcription factor MFH-1 in aortic arch patterning and skeletogenesis. *Development* **124**, 4627-38.

Jacobson, A. G. (1984). Further evidence that formation of the neural tube requires elongation of the nervous system. *J Exp Zool* **230**, 23-8.

Jiang, X., Iseki, S., Maxson, R. E., Sucov, H. M. and Morriss-Kay, G. M. (2002). Tissue origins and interactions in the mammalian skull vault. *Dev Biol* **241**, 106-16.

Jovanovic, M. and Hengartner, M. O. (2006). miRNAs and apoptosis: RNAs to die for. *Oncogene* **25**, 6176-87.

Kanellopoulou, C., Muljo, S. A., Kung, A. L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D. M. and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**, 489-501.

Kang, S., Graham, J. M., Jr., Olney, A. H. and Biesecker, L. G. (1997). GLI3 frameshift mutations cause autosomal dominant Pallister-Hall syndrome. *Nat Genet* **15**, 266-8.

Karaplis, A. C. (2002). Embryonic Development of Bone and the Molecular regulation of Intramembranous and Endochondral Bone Formation. *Principles of Bone Biology, Second Edition* **1**.

Karsenty, G. (2008). Transcriptional control of skeletogenesis. *Annu Rev Genomics Hum Genet* **9**, 183-96.

Kessel, M., Balling, R. and Gruss, P. (1990). Variations of cervical vertebrae after expression of a Hox-1.1 transgene in mice. *Cell* **61**, 301-8.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. and Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev* **15**, 2654-9.

Khvorova, A., Reynolds, A. and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-16.

Kim, S. H., Cho, K. W., Choi, H. S., Park, S. J., Rhee, Y., Jung, H. S. and Lim, S. K. (2009). The forkhead transcription factor Foxc2 stimulates osteoblast differentiation. *Biochem Biophys Res Commun* **386**, 532-6.

Kimelman, D. (2006). Mesoderm induction: from caps to chips. *Nat Rev Genet* **7**, 360-72.

Kimura-Yoshida, C., Tian, E., Nakano, H., Amazaki, S., Shimokawa, K., Rossant, J., Aizawa, S. and Matsuo, I. (2007). Crucial roles of Foxa2 in mouse anterior-posterior axis polarization via regulation of anterior visceral endoderm-specific genes. *Proc Natl Acad Sci U S A* **104**, 5919-24.

Kispert, A. and Herrmann, B. G. (1994). Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Dev Biol* **161**, 179-93.

Kluth, D. (2010). Embryology of anorectal malformations. *Semin Pediatr Surg* **19**, 201-8.

Knight, S. W. and Bass, B. L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **293**, 2269-71.

Koller, B. H., Hagemann, L. J., Doetschman, T., Hagaman, J. R., Huang, S., Williams, P. J., First, N. L., Maeda, N. and Smithies, O. (1989). Germ-line transmission of a planned alteration made in a hypoxanthine phosphoribosyltransferase gene by homologous recombination in embryonic stem cells. *Proc Natl Acad Sci U S A* **86**, 8927-31.

Kuehn, M. R., Bradley, A., Robertson, E. J. and Evans, M. J. (1987). A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. *Nature* **326**, 295-8.

Kuhn, R., Schwenk, F., Aguet, M. and Rajewsky, K. (1995). Inducible gene targeting in mice. *Science* **269**, 1427-9.

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**, 559-61.

Lai, E., Prezioso, V. R., Tao, W. F., Chen, W. S. and Darnell, J. E., Jr. (1991). Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the Drosophila homeotic gene fork head. *Genes Dev* **5**, 416-27.

Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R. and Berns, A. (1991). Simplified mammalian DNA isolation procedure. *Nucleic Acids Res* **19**, 4293.

Le Douarin, N. M. and Dupin, E. (1993). Cell lineage analysis in neural crest ontogeny. *J Neurobiol* **24**, 146-61.

Le Douarin, N. M., Ziller, C. and Couly, G. F. (1993). Patterning of neural crest derivatives in the avian embryo: in vivo and in vitro studies. *Dev Biol* **159**, 24-49.

Lee, G. and Saito, I. (1998). Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene* **216**, 55-65.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-9.

Lee, Y., Jeon, K., Lee, J. T., Kim, S. and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* **21**, 4663-70.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H. and Kim, V. N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-60.

Lee, Y. S., Nakahara, K., Pham, J. W., Kim, K., He, Z., Sontheimer, E. J. and Carthew, R. W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69-81.

Lefebvre, V. and Smits, P. (2005). Transcriptional control of chondrocyte fate and differentiation. *Birth Defects Res C Embryo Today* **75**, 200-12.

Lickert, H., Cox, B., Wehrle, C., Taketo, M. M., Kemler, R. and Rossant, J. (2005). Dissecting Wnt/beta-catenin signaling during gastrulation using RNA interference in mouse embryos. *Development* **132**, 2599-609.

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L. and Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-41.

Lodish, H. F., Zhou, B., Liu, G. and Chen, C. Z. (2008). Micromanagement of the immune system by microRNAs. *Nat Rev Immunol* **8**, 120-30.

Lovell-Badge, R. H. and Evans, M. J. (1980). Changes in protein synthesis during differentiation of embryonal carcinoma cells, and a comparison with embryo cells. *J Embryol Exp Morphol* **59**, 187-206.

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A. et al. (2005). MicroRNA expression profiles classify human cancers. *Nature* **435**, 834-8.

Lumsden, A., Sprawson, N. and Graham, A. (1991). Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development* **113**, 1281-91.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. and Kutay, U. (2004). Nuclear export of microRNA precursors. *Science* **303**, 95-8.

Lykke-Andersen, K., Gilchrist, M. J., Grabarek, J. B., Das, P., Miska, E. and Zernicka-Goetz, M. (2008). Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. *Mol Biol Cell* **19**, 4383-92.

Lyon, M. F. (1990). L. C. Dunn and mouse genetic mapping. *Genetics* **125**, 231-6.

Ma, J., Flemr, M., Stein, P., Berninger, P., Malik, R., Zavolan, M., Svoboda, P. and Schultz, R. M. (2010). MicroRNA activity is suppressed in mouse oocytes. *Curr Biol* **20**, 265-70.

Ma, Q., Fode, C., Guillemot, F. and Anderson, D. J. (1999). Neurogenin1 and neurogenin2 control two distinct waves of neurogenesis in developing dorsal root ganglia. *Genes Dev* **13**, 1717-28.

Makeyev, E. V., Zhang, J., Carrasco, M. A. and Maniatis, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell* **27**, 435-48.

Mansour, S. L., Thomas, K. R. and Capecchi, M. R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348-52.

Martin, B. L. and Kimelman, D. (2008). Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. *Dev Cell* **15**, 121-33.

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**, 7634-8.

Martin, G. R. and Evans, M. J. (1974). The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. *Cell* **2**, 163-72.

Martinez, J. and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes Dev* **18**, 975-80.

McMahon, J. A., Takada, S., Zimmerman, L. B., Fan, C. M., Harland, R. M. and McMahon, A. P. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev* **12**, 1438-52.

Mehta, A. K., Majumdar, S. S., Alam, P., Gulati, N. and Brahmachari, V. (2009). Epigenetic regulation of cytomegalovirus major immediate-early promoter activity in transgenic mice. *Gene* **428**, 20-4.

Meyer, D. and Birchmeier, C. (1995). Multiple essential functions of neuregulin in development. *Nature* **378**, 386-90.

Meyer-Ficca, M. L., Meyer, R. G., Kaiser, H., Brack, A. R., Kandolf, R. and Kupper, J. H. (2004). Comparative analysis of inducible expression systems in transient transfection studies. *Anal Biochem* **334**, 9-19.

Moffett, H. F. and Novina, C. D. (2007). A small RNA makes a Bic difference. *Genome Biol* **8**, 221.

Mongelard, F. and Corces, V. G. (2001). Two insulators are not better than one. *Nat Struct Biol* **8**, 192-4.

Montgomery, M. K., Xu, S. and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **95**, 15502-7.

Morita, S., Horii, T., Kimura, M., Goto, Y., Ochiya, T. and Hatada, I. (2007). One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. *Genomics* **89**, 687-96.

Mulero-Navarro, S., Carvajal-Gonzalez, J. M., Herranz, M., Ballestar, E., Fraga, M. F., Ropero, S., Esteller, M. and Fernandez-Salguero, P. M. (2006). The dioxin receptor is silenced by promoter hypermethylation in human acute lymphoblastic leukemia through inhibition of Sp1 binding. *Carcinogenesis* **27**, 1099-104.

Murphy, J. C., Fischle, W., Verdin, E. and Sinclair, J. H. (2002). Control of cytomegalovirus lytic gene expression by histone acetylation. *EMBO J* **21**, 1112-20.

Myslinski, E., Ame, J. C., Krol, A. and Carbon, P. (2001). An unusually compact external promoter for RNA polymerase III transcription of the human H1RNA gene. *Nucleic Acids Res* **29**, 2502-9.

Nagy, A. (2000). Cre recombinase: the universal reagent for genome tailoring. *Genesis* **26**, 99-109.

Nagy, A. (2003). Manipulating the Mouse Embryo - A Laboratory Manual (3rd ed.). *Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press*.

Niwa, H., Yamamura, K. and Miyazaki, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193-9.

Noden, D. M. (1988). Interactions and fates of avian craniofacial mesenchyme. *Development* **103 Suppl**, 121-40.

Noden, D. M. and Schneider, R. A. (2006). Neural Crest Cells and the Community of Plan for Craniofacial Development: Historical Debates and Current Prespectives. *Neural Crest Induction and Differentiation*.

Nonaka, S., Shiratori, H., Saijoh, Y. and Hamada, H. (2002). Determination of left-right patterning of the mouse embryo by artificial nodal flow. *Nature* **418**, 96-9.

Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., Kido, M. and Hirokawa, N. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell* **95**, 829-37.

O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. and Mendell, J. T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839-43.

Okada, Y., Nonaka, S., Tanaka, Y., Saijoh, Y., Hamada, H. and Hirokawa, N. (1999). Abnormal nodal flow precedes situs inversus in iv and inv mice. *Mol Cell* **4**, 459-68.

Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. and Conklin, D. S. (2002). Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* **16**, 948-58.

Paigen, K. (2003a). One hundred years of mouse genetics: an intellectual history. I. The classical period (1902-1980). *Genetics* **163**, 1-7.

Paigen, K. (2003b). One hundred years of mouse genetics: an intellectual history. II. The molecular revolution (1981-2002). *Genetics* **163**, 1227-35.

Perantoni, A. O., Timofeeva, O., Naillat, F., Richman, C., Pajni-Underwood, S., Wilson, C., Vainio, S., Dove, L. F. and Lewandoski, M. (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* **132**, 3859-71.

Petiot, A., Perriton, C. L., Dickson, C. and Cohn, M. J. (2005). Development of the mammalian urethra is controlled by Fgfr2-IIIb. *Development* **132**, 2441-50.

Pettway, Z., Guillory, G. and Bronner-Fraser, M. (1990). Absence of neural crest cells from the region surrounding implanted notochords in situ. *Dev Biol* **142**, 335-45.

Platanias, L. C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol* **5**, 375-86.

Ramirez-Solis, R., Davis, A. C. and Bradley, A. (1993). Gene targeting in embryonic stem cells. *Methods Enzymol* **225**, 855-78.

Rand, T. A., Petersen, S., Du, F. and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **123**, 621-9.

Reinhold, M. I., Kapadia, R. M., Liao, Z. and Naski, M. C. (2006). The Wnt-inducible transcription factor Twist1 inhibits chondrogenesis. *J Biol Chem* **281**, 1381-8.

Robertson, E., Bradley, A., Kuehn, M. and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* **323**, 445-8.

Rocha, P. P., Scholze, M., Bleiss, W. and Schrewe, H. (2010). Med12 is essential for early mouse development and for canonical Wnt and Wnt/PCP signaling. *Development* **137**, 2723-31.

Rodrigo, I., Hill, R. E., Balling, R., Munsterberg, A. and Imai, K. (2003). Pax1 and Pax9 activate Bapx1 to induce chondrogenic differentiation in the sclerotome. *Development* **130**, 473-82.

Ruberte, E., Dolle, P., Chambon, P. and Morriss-Kay, G. (1991). Retinoic acid receptors and cellular retinoid binding proteins. II. Their differential pattern of transcription during early morphogenesis in mouse embryos. *Development* **111**, 45-60.

Rufai, A., Benjamin, M. and Ralphs, J. R. (1995). The development of fibrocartilage in the rat intervertebral disc. *Anat Embryol (Berl)* **192**, 53-62.

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J. and Hamada, H. (2001). The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. *Genes Dev* **15**, 213-25.

Sasaki, Y., Iwai, N., Tsuda, T. and Kimura, O. (2004). Sonic hedgehog and bone morphogenetic protein 4 expressions in the hindgut region of murine embryos with anorectal malformations. *J Pediatr Surg* **39**, 170-3; discussion 170-3.

Sauer, B. and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A* **85**, 5166-70.

Schatteman, G. C., Morrison-Graham, K., van Koppen, A., Weston, J. A. and Bowen-Pope, D. F. (1992). Regulation and role of PDGF receptor alpha-subunit expression during embryogenesis. *Development* **115**, 123-31.

Schmidt, E. V., Christoph, G., Zeller, R. and Leder, P. (1990). The cytomegalovirus enhancer: a pan-active control element in transgenic mice. *Mol Cell Biol* **10**, 4406-11.

Schmidt-Supprian, M. and Rajewsky, K. (2007). Vagaries of conditional gene targeting. *Nat Immunol* **8**, 665-8.

Schulte-Merker, S. and Smith, J. C. (1995). Mesoderm formation in response to Brachyury requires FGF signalling. *Curr Biol* **5**, 62-7.

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**, 1281-8.

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. and Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230-4.

Smits, P. and Lefebvre, V. (2003). Sox5 and Sox6 are required for notochord extracellular matrix sheath formation, notochord cell survival and development of the nucleus pulposus of intervertebral discs. *Development* **130**, 1135-48.

Sontheimer, E. J. (2005). Assembly and function of RNA silencing complexes. *Nat Rev Mol Cell Biol* **6**, 127-38.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21**, 70-1.

Staats, J. (1980). Standardized nomenclature for inbred strains of mice: seventh listing for the International Committee on Standardized Genetic Nomenclature for Mice. *Cancer Res* **40**, 2083-128.

Stemple, D. L. (2005). Structure and function of the notochord: an essential organ for chordate development. *Development* **132**, 2503-12.

Stephens, F. D. (1988). Embryology of the cloaca and embryogenesis of anorectal malformations. *Birth Defects Orig Artic Ser* **24**, 177-209.

Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M. and Hoess, R. (1981). Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harb Symp Quant Biol* **45 Pt 1**, 297-309.

Stott, D., Kispert, A. and Herrmann, B. G. (1993). Rescue of the tail defect of Brachyury mice. *Genes Dev* **7**, 197-203.

Su, H., Trombly, M. I., Chen, J. and Wang, X. (2009). Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes Dev* **23**, 304-17.

Suh, N., Baehner, L., Moltzahn, F., Melton, C., Shenoy, A., Chen, J. and Blelloch, R. (2010). MicroRNA function is globally suppressed in mouse oocytes and early embryos. *Curr Biol* **20**, 271-7.

Szulc, J., Wiznerowicz, M., Sauvain, M. O., Trono, D. and Aebischer, P. (2006). A versatile tool for conditional gene expression and knockdown. *Nat Methods* **3**, 109-16.

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M. et al. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534-8.

Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S. C., Sun, Y. A., Lee, C., Tarakhovsky, A., Lao, K. and Surani, M. A. (2007). Maternal microRNAs are essential for mouse zygotic development. *Genes Dev* **21**, 644-8.

Thomas, K. R. and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503-12.

Thomas, K. R., Folger, K. R. and Capecchi, M. R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419-28.

Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. and Melton, D. W. (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313-21.

Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* **13**, 3191-7.

Tyzzer, E. E. (1909). A Study of Inheritance in Mice with Reference to their Susceptibility to transplantable Tumors. *J Med Res* **21**, 519-73.

van den Pol, A. N. and Ghosh, P. K. (1998). Selective neuronal expression of green fluorescent protein with cytomegalovirus promoter reveals entire neuronal arbor in transgenic mice. *J Neurosci* **18**, 10640-51.

van Dongen, S., Abreu-Goodger, C. and Enright, A. J. (2008). Detecting microRNA binding and siRNA off-target effects from expression data. *Nat Methods* **5**, 1023-5.

van Straaten, H. W., Hekking, J. W., Wiertz-Hoessels, E. J., Thors, F. and Drukker, J. (1988). Effect of the notochord on the differentiation of a floor plate area in the neural tube of the chick embryo. *Anat Embryol (Berl)* **177**, 317-24.

Villuendas, G., Gutierrez-Adan, A., Jimenez, A., Rojo, C., Roldan, E. R. and Pintado, B. (2001). CMV-driven expression of green fluorescent protein (GFP) in male germ cells of transgenic mice and its effect on fertility. *Int J Androl* **24**, 300-5.

Vujovic, S., Henderson, S., Presneau, N., Odell, E., Jacques, T. S., Tirabosco, R., Boshoff, C. and Flanagan, A. M. (2006). Brachyury, a crucial regulator of notochordal development, is a novel biomarker for chordomas. *J Pathol* **209**, 157-65.

Wagner, T. E., Hoppe, P. C., Jollick, J. D., Scholl, D. R., Hodinka, R. L. and Gault, J. B. (1981). Microinjection of a rabbit beta-globin gene into zygotes and its subsequent expression in adult mice and their offspring. *Proc Natl Acad Sci U S A* **78**, 6376-80.

Wallin, J., Eibel, H., Neubuser, A., Wilting, J., Koseki, H. and Balling, R. (1996). Pax1 is expressed during development of the thymus epithelium and is required for normal T-cell maturation. *Development* **122**, 23-30.

Wallin, J., Wilting, J., Koseki, H., Fritsch, R., Christ, B. and Balling, R. (1994). The role of Pax-1 in axial skeleton development. *Development* **120**, 1109-21.

Walmsley, R. (1953). The development and growth of the intervertebral disc. *Edinb Med J* **60**, 341-64.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R. and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* **39**, 380-5.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T. et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539-43.

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**, 520-62.

Weigel, D. and Jackle, H. (1990). The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? *Cell* **63**, 455-6.

Weinstein, D. C., Ruiz i Altaba, A., Chen, W. S., Hoodless, P., Prezioso, V. R., Jessell, T. M. and Darnell, J. E., Jr. (1994). The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. *Cell* **78**, 575-88.

Whyatt, L. M. and Rathjen, P. D. (2001). Interferon-inducible ES cell expression systems. *Methods Mol Biol* **158**, 301-18.

Wianny, F. and Zernicka-Goetz, M. (2000). Specific interference with gene function by double-stranded RNA in early mouse development. *Nat Cell Biol* **2**, 70-5.

Wilkinson, D. G., Bhatt, S. and Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657-9.

Wilm, B., Dahl, E., Peters, H., Balling, R. and Imai, K. (1998). Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency. *Proc Natl Acad Sci U S A* **95**, 8692-7.

Wilson, V., Rashbass, P. and Beddington, R. S. (1993). Chimeric analysis of T (Brachyury) gene function. *Development* **117**, 1321-31.

Winnier, G. E., Hargett, L. and Hogan, B. L. (1997). The winged helix transcription factor MFH1 is required for proliferation and patterning of paraxial mesoderm in the mouse embryo. *Genes Dev* **11**, 926-40.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P. and Trono, D. (2007). The Kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *J Biol Chem* **282**, 34535-41.

Wolpert, L. (2007). Principles of Development, Third Edition. *Oxford University Press*.

Wu, H., Xu, H., Miraglia, L. J. and Crooke, S. T. (2000). Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. *J Biol Chem* **275**, 36957-65.

www.cancer.gov, U. S. N. I. o. H. *Division of Cancer Epidemiology & Genetics*.

Yamaguchi, T., Suzuki, S., Ishiiwa, H. and Ueda, Y. (2004). Intraosseous benign notochordal cell tumours: overlooked precursors of classic chordomas? *Histopathology* **44**, 597-602.

Yamaguchi, T. P., Bradley, A., McMahon, A. P. and Jones, S. (1999a). A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. *Development* **126**, 1211-23.

Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N. and McMahon, A. P. (1999b). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev* **13**, 3185-90.

Yang, X. R., Ng, D., Alcorta, D. A., Liebsch, N. J., Sheridan, E., Li, S., Goldstein, A. M., Parry, D. M. and Kelley, M. J. (2009). T (brachyury) gene duplication confers major susceptibility to familial chordoma. *Nat Genet* **41**, 1176-8.

Yao, F., Svensjo, T., Winkler, T., Lu, M., Eriksson, C. and Eriksson, E. (1998). Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Hum Gene Ther* **9**, 1939-50.

Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., Kerr, W. G. and Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A* **94**, 3789-94.

Zeng, Y. and Cullen, B. R. (2005). Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J Biol Chem* **280**, 27595-603.

Zhang, J. M., Hoffmann, R. and Sieber-Blum, M. (1997). Mitogenic and anti-proliferative signals for neural crest cells and the neurogenic action of TGF-beta1. *Dev Dyn* **208**, 375-86.

Zhang, Y., Riesterer, C., Ayrall, A. M., Sablitzky, F., Littlewood, T. D. and Reth, M. (1996). Inducible site-directed recombination in mouse embryonic stem cells. *Nucleic Acids Res* **24**, 543-8.

Zhu, Z., Zheng, T., Lee, C. G., Homer, R. J. and Elias, J. A. (2002). Tetracycline-controlled transcriptional regulation systems: advances and application in transgenic animal modeling. *Semin Cell Dev Biol* **13**, 121-8.

Zijlstra, M., Li, E., Sajjadi, F., Subramani, S. and Jaenisch, R. (1989). Germ-line transmission of a disrupted beta 2-microglobulin gene produced by homologous recombination in embryonic stem cells. *Nature* **342**, 435-8.