MASTERARBEIT

Titel der Masterarbeit

„The impact of *Airn* lncRNA expression on enhancer-promoter interactions"

verfasst von

Markus Muckenhuber, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2014

Table of Contents

4

1.   Zusammenfassung

Ein kleiner Teil von Säugetiergenen wird, abhängig vom elterlichen Ursprung, transkribiert und dieses epigenetische Phänomen wird genomische Prägung genannt. Solche Gene findet man oft in Cluster und in Verbindung mit langen nicht-kodierenden RNS (lnkRNS). Die Genestilllegung von diesen Clustern wird kontrolliert von Methylierungen an einem einzigen geprägten Kontrollelement (ICE), welches oft als Promoter für eine lnkRNS dient. Im *Igf2r* Genecluster von Mäusen wird die makro lnkRNS ausschließlich vom väterlichen Allele hergestellt und die Überlappung mit den protein-kodierenden Genes *Igf2r* führet zu dessen Stilllegung durch Störung der Transkription. Das mütterlichen ICE ist methyliert und verhindert so die Herstellung von *Airn* was stattdessen zur Produktion von *Igf2r* führet. *Igf2r* zeigt dieses Herstellungsmuster in fast allen Geweben, während zwei andere Gene in diesem Cluster, *Slc22a2* und *Slc22a3*, nur in extra-embryonal Geweben diese *Airn*-abhängige geprägte Expression vorweisen. Der genau Mechanismus, wie diese nicht von *Airn* überlappten Genen reguliert werden, ist unbekannt. In unsere Arbeitsgruppe wurde ein Model erarbeitet, welche die väterliche Stilllegung von *Slc22a2/a3* durch das Blockieren von gewebsspezifischen Verstärkerelementen (Enhancer) durch die Transkription von *Airn* vorschlägt. Diese Hypothese wird unterstützt von frühere Chromosome Conformation Capture (3C) und ChiP-Seq Experimente mit *Airn*-Mutanten. Mein Beitrag zu diesem Teil des Projekts war die Überprüfung dieser früheren Ergebnisse mit 3C Proben einer verkürzten Version von *Airn* um zu analysieren welche Auswirkungen *Airn* Transkription auf diese Verstärker-Promoter Interaktionen hat. Ich konnte eine Zunahme von Interaktionen in der Abwesenheit von *Airn* zeigen, was die früheren Ergebnisse unterstützt. Um diese Hypothese weiter zu teste, wurden Deletionen und Duplikationen vom *Airn*-Gene in Mäusen mit Chromosom Engineering hergestellt. So wurden eine 28kb Deletion/Duplikation des ersten Drittels vom *Airn*-Gene und zusätzlich eine 270kb Deletion/Duplikation, welche den gesamten 118kb Genekörper von *Airn* beinhaltet, hergestellt. Ich war verantwortlich für die Entwicklung von Southern Blot und PCR Assays für die Identifizierung der neuen Allele. Diese neune Allele ermöglich uns eine tieferen Einblick in den Mechanismus dieser gewebsspezifischen geprägten Stilllegung von unseren Zielgenen durch die *Airn* Transkription zu erlangen.

2.   Abstract

A small subset of mammalian genes is transcribed in a parent-of-origin specific manner and this epigenetic phenomenon is called genomic imprinting. These genes are often found clustered and in association with long noncoding RNAs (lncRNA). Imprinted gene silencing of a gene cluster is controlled by methylation of a single imprint control element (ICE), which is often a promoter for a lncRNA that initiates gene silencing. In the well-characterized *Igf2r* imprinted cluster the macro lncRNA *Airn* is exclusively expressed from the paternal allele, overlaps and silences the protein-coding gene *Igf2r in cis* via transcription interference. On the maternal allele the methylation of the ICE inhibits the transcription of *Airn* and consequently *Igf2r* is expressed. *Igf2r* shows this expression pattern in nearly all tissues whereas two other genes, *Slc22a2/a3,* show imprinted expression restricted to extra-embryonic tissues. The macro lncRNA *Airn* does not overlap these upstream lying genes and the exact mechanism, how this silencing process is controlled remains unclear. In the lab a model was proposed that *Slc22a2/a3* are silenced on the paternal allele by *Airn* transcription blocking tissue specific enhancers. Previous chromosome conformation capture (3C) and ChIP-seq experiments using *Airn* mutants support this hypothesis. My contribution to this part of the project was to validate these previous findings with 3C samples of *Airn* truncation samples and to analyze the effects of *Airn* transcription on enhancer-promoter interactions. I was able to detect a gain of interaction in the absence of *Airn*, which confirms that this lncRNA effects the chromosomal conformation in this region. To further test the enhancer-promoter interference model large deletions and duplications of the *Airn* gene in mice using chromosome engineering were created. Using this approach we generated a 28kb deletion/duplication of the first third of the *Airn* gene, and a 270kb deletion/duplication that includes the entire 118kb *Airn* gene. I designed new Southern blot and PCR assays to screen for these new alleles. As part of this, I was also involved in the maintenance and breeding strategy of this approach. At the moment we have all four strains maintained. In summary, chromosome engineering resulting in deletion or duplication of regions of the *Airn* gene and putative enhancer regions should give insights into the tissue specific mechanism of imprinted silencing by this lncRNA.

3.   Introduction

3.1.   Genomic imprinting, an epigenetic phenomenon

Epigenetic phenomena cause reversible alterations in gene expression levels without changing the underlying DNA sequences and can be mitotically inherited. DNA methylation (Ehrlich and Wang, 1981) is the essential requirement for this maintenance process. Once established the modified DNA can be recognized by different histone modifiers (Marmorstein and Zhou, 2014); (Cheng, 2014) or chromatin remodelers (Rippe et al., 2007). Histone modifications can be altered during the process of development, thereby promote changes in the chromosomal environment (Bannister and Kouzarides, 2011) and affect gene expression pattern. Normally, those changes are found in diploid organisms, like mouse or human, on both copies of an allele (biallelic). In mouse and humans more than 99% of all protein coding genes are regulated in a biallelic manner. In these mammals the remaining protein-coding genes (<1%) show different expression of each allele in a parent-of-origin-dependent manner that is called imprinted expression (Barlow et al., 1991). This epigenetic process, called genomic imprinting, can lead to either maternal or paternal monoallelic expression or to biased expression. A generally accepted level of biased expression to define imprinted expression is still open to debate (Barlow, 2011). Once established the *cis*-acting mechanism of imprinting has to be maintained in somatic cells throughout the life of the carrier and independent of its sex. However, the imprinted marks have to be dynamic and reversible, because during meiotic division in gametes the parental imprinted marks have to be erased and new marks have to be set in a sex-dependent manner (Jones, 2012).

Interestingly, genomic imprinting has evolved several times independently in different plants and animals. The first examples of genomic imprinting were discovered in Sciaridae (flies) and Coccidae (scale) where it is involved in sex determination. In fertilized eggs the paternal genome is completely silenced to create a haploid-like situation, which is required to start male development. With both parental genomes active, female development is initiated (Brown et

al., 1991); (Crouse et al., 1971); (Khosla et al., 2006). Imprinting expression is also present in seed-bearing plants like Arabidopsis (Kinoshita et al., 2008) and maize (Waters et al., 2013). Finally, genomic imprinting is linked to mammals, where it is an essential mechanism in during development (Barlow and Bartolomei, 2014).

3.2. Genomic imprinting in mammals

Imprinted expressed genes can be found in placental mammals as well as in marsupials (Stringer et al., 2014). For the small group of egg-laying mammals (platypus; echidna) not many studies have been performed so far, but it seems that these mammals do not have imprinted expression (Renfree et al., 2009). Apart from this subgroup, in mammals genomic imprinting is involved in essential process like development, X-inactivation, regulation of embryonic and placental growth (Cooper et al., 1971); (Isles and Holland, 2005); (Reik and Lewis, 2005); (Ideraabdullah et al., 2008).

In the early 1980s nucleus transplantation experiments in mice demonstrated that normal embryogenesis only occurs with a maternal and a paternal genome copy. Two maternal (gynogenetic) or two paternal (androgenetic) copies are embryonic lethal (McGrath and Solter, 1984); (Barton et al., 1984) (**Figure 1**). To exclude genetic factors, inbred mouse lines with identical parental chromosomes were used to show that an epigenetic mechanism is required to establish imprinting (Surani et al., 1984); (Cattanach and Kirk, 1985). In 1991 the first autosomal imprinted expressed genes were described. First of all, the insulin-like growth factor 2 receptor (*Igf2r*) was characterized as maternally expressed imprinted gene (Barlow et al., 1991). Shortly after, the insulin-like growth factor 2 (*Igf2*), a paternally expressed imprinted gene, was identified (DeChiara et al., 1991) and later on in the same year the lncRNA *H19* (Bartolomei et al., 1991) was discovered as a maternally expressed imprinted gene in the same cluster. In mice the essential role of many imprinted genes in development, growth and behavior has been demonstrated by loss of function studies, although fewer studies have

demonstrated an imprinted phenotype for single genes (Ferguson-Smith, 2011). In terms of behavioral phenotypes, imprinted genes are also involved in postnatal development where they affect suckling and metabolism of newborns (Yotova et al., 2008).

Egg   **Sperm**

Zygote

Mat   Pat

2x Mat   **1x Mat**   **1x Pat**   **2x Pat**

Gynogenetic   **Wild-type**   **Androgenetic**

lethal   viable   *lethal*
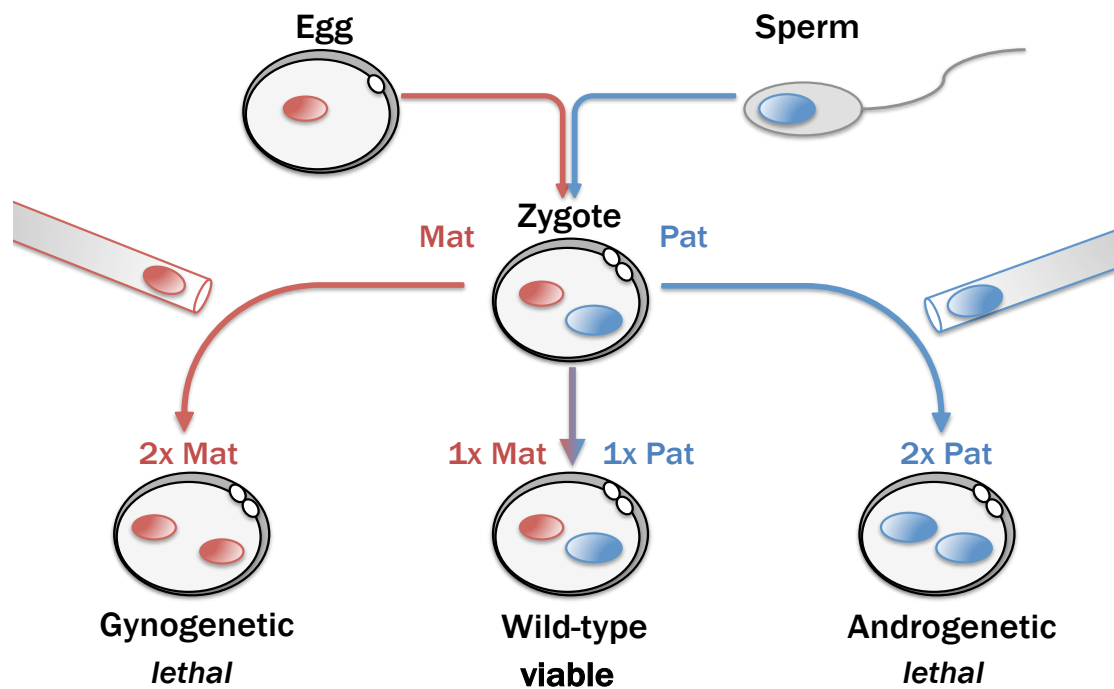
Figure 1: Nucleus transplantation experiments in mouse

In mammals a copy of the maternal and the paternal genome (biparental) are necessary for embryogenesis. Two maternal (gynogenetic) as well as two paternal (androgenetic) pronuclei are embryonic lethal at early stages (adapted from (Barlow and Bartolomei, 2014)).

Dysfunction in the regulation of specific imprinted expression genes in human can be linked to different types of genetic diseases like Angleman syndrome (Angelman, 1965), Prader-Willi syndrome (Cassidy et al., 2012), Beckwith-Wiedermann syndrome (Shuman et al., 1993), Silver-Russell syndrome (Bartholdi et al., 2009) and also to cancer (Uribe-Lewis et al., 2011). For molecular biology imprinted gene clusters are powerful toolboxes to understand epigenetic control mechanisms. The presence of both the active and repressive allele in the same cellular environment represents a wonderful model system to understand, how epigenetic gene regulation is controlled.

3.3.   Hypotheses about the origin of genomic imprinting

From an evolutionary point of view, genomic imprinting seems to be paradox, as the protection against expression of recessive mutated allele in a biallelic situation is lost. While in a haploid organism the expression of a mutated allele could cause a phenotype, in a diploid situation the expression of the wildtype-like allele from the sister chromosome could rescue the mutant phenotype (Perrot et al., 1991). Haploidy offers the benefit of a reduced mutational load, because of lower number of potential mutation targets (Otto and Goldstein, 1992). On the one hand only one copy of an imprinted gene is expressed, meaning that the benefits of being diploid is lost for this set of genes. On the other hand the mutant load is not lower, because the second chromosome set can still be mutated. It seems, that genomic imprinting does not promote the fitness of the host. Why should evolution promote such a mechanism? What advantages are provided by imprinted expression? More than twenty years ago the first theories were set up to answer these questions.

The "Parental Conflict Hypothesis" (Moore and Haig, 1991) also called kinship theory of genomic imprinting, argues that a genetic conflict of interest between maternal and paternal genomes on evolutionary fitness for the next generation lead to the establishment of a new regulation system, genomic imprinting. There is a strong difference in the contribution of parents to their offspring. Therefore imprinted genes are mainly involved in the transfer of maternal resources to embryos and developmental growth. This hypothesis is supported by examples such as the growth promoting *Igf2* showing paternal imprinted expression, while the growth-repressing gene *Igf2r* shows maternal imprinted expression.

The "Ovarian Time Bomb" or "trophoblast defense" hypothesis (Varmuza and Mann, 1994) see genomic imprinting as an protection mechanism against parthenogenesis (asexual reproduction), which does not naturally occur in placental mammals (Kaufman, 1983). The presence of imprinted expression and the absence of parthenogenesis are both restricted to mammals,

indicating that genomic imprinting may be a protection mechanism against parthenogenesis, which can lead to ovarian trophoblast disease (Solter, 1988).

All these hypotheses can at least partially explain the function of imprinting expression and how the evolution of such a mechanism could be driven. However, it is likely a combination of some of these theories may explain the parent-of-origin specific expression pattern different imprinted gene clusters. This point is also supported by the fact, that regulation mechanisms of imprinted cluster are highly variable, although commonly used features can be found.

3.4.   Features of genomic imprinting

First of all, imprinted expression is an epigenetic phenomena and a consequence of inheritance not sex. The imprinted cycle (Autuoro et al., 2014) (**Figure 2**) shows the changes of imprinted expression pattern during development. The modification of identical alleles in two different ways can only happen when the maternal and paternal chromosomes are separated. Such a situation occurs in the parental germline. Two different modification systems have to be present to establish specific imprinted marked alleles in ovum and sperm (Li et al., 2004). After fertilization, both of these alleles are in the same nuclear environment. At this point, the different patterns have to be maintained throughout the entire adulthood of an organism in somatic cells. In contrast, the marks have to be erased before gametogenesis to establish parent-of-origin specific marks (Davis et al., 2000); (Barlow and Bartolomei, 2014). Afterwards, methylation pattern of DMRs has to escapes from global epigenetic reprogramming at the pre-implantation embryo (Morgan et al., 2005); (Sasaki and Matsui, 2008).
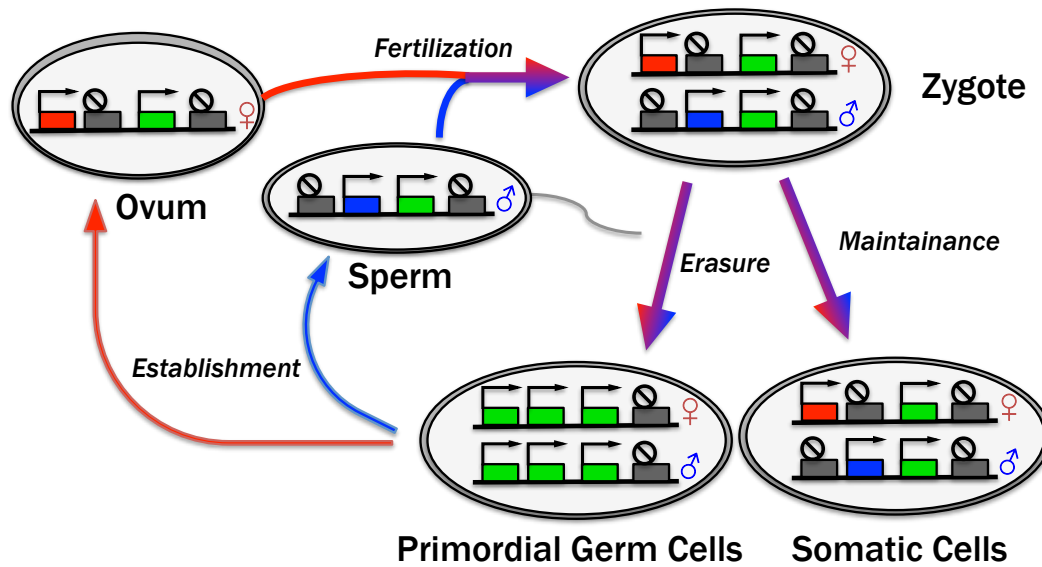
Figure 2: The imprinting cycle

Imprinted marks (DNA methylation) are set during gametogenesis in a sex-specific manner. After fertilization a maternal and a paternal marked genome are present. This zygotic state of imprinted DNA methylation is maintained in all somatic cells, but can be altered. In primordial germ cells the imprinted marks have to be erased to be reestablished depending on the sex of the individual. Biallelic regulated genes never show a significant different expression pattern between their alleles. Maternally (red box), paternally (blue box), biallelic (green box) expressed genes are shown in this figure as well as non-expressed (grey box) genes. In diploid situations (zygote, primordial germ cells, somatic cells) the maternal (♀) and paternal (♂) are shown, whereas in haploid gametes (ovum, sperm) only one of those alleles are shown depending on its origin. (Adapted from (Autuoro et al., 2014))

The next key feature of genomic imprinting is its *cis* acting nature (Barlow, 2011). The imprinting mechanism can only affect genes on the same chromosome, which indicates, that factors or a combination of factors involved in the regulation of genomic imprinting have to interact tightly with just one chromosome. Those factors can regulate gene expression on many different levels from interactions with promoters, enhancers, splicing junctions or polyadenylation sites to cause parental specific differences. Interestingly, more than 80% of imprinted genes in mouse are found clustered (Reik and Walter, 2001); (Barlow and Bartolomei, 2014). This clusters can contain both, paternal and maternal, imprinted genes as well as (long) non-coding RNAs and also non-imprinted genes (Thorvaldsen and Bartolomei, 2007); (Royo and Cavaille, 2008); (Peters and Robson, 2008). 16 clusters with two or more genes showing imprinted expression were described so far (Wan and Bartolomei, 2008), although only 7 have been demonstrated genetically

(Barlow and Bartolomei, 2014) These clusters are controlled by one imprinting control element or ICE. In addition, solo imprinted genes were described (Wang et al., 2011), although the imprinting status of some solo genes has been challenged (Hudson et al., 2011) {Okae, 2012 #589} {Bourc'his, 2008 #619}. Imprinted expressed genes are found on 17 different chromosomes in mouse, although the distribution of imprinted loci is unequal. The murine hotspot is on chromosome 7 where five distinct clusters are found (*Igf2*; *Inpp5f*; *Kcnq1*; *Peg3*; *Pws*) (http:// www.mousebook.org /mousebook-catalogs/ imprinting-resource; 12/8/14).

A common feature of imprinted gene clusters is a differentially DNA methylated region (DMR) known as an imprint control element (ICE) (Wutz et al., 1997). This is a gametic DMR (gDMR) that is established during gametogenesis by the *de novo* DNA methyltransferase DNMT3A together with the non-catalytic DNMT3L (Bourc'his et al., 2001); (Kaneda et al., 2004). DNMT1 is then required to maintain the DNA methylation mark on newly synthesized DNA strains by recognizing hemimethylated DNA and methylating the new strand (Goll and Bestor, 2005). Mutations in DMNT1 disrupt imprinted expression in mouse embryos, which indicates the essential role of this enzyme in the process of maintaining imprinting marks (Li et al., 1993). The unmethylated ICE can initiate imprinted silencing by acting as a promoter for an lncRNA that causes silencing or via binding by the insulator protein CTCF that blocks enhancer access to imprinted genes (Barlow and Bartolomei, 2014). Somatic DMRs (sDMRs) can be later established over the promoters of some imprinted genes, but these play secondary role in silencing {Santoro, 2013 #87}.

The epigenetic modifications at an ICE are established to cause imprinted expression at multiple target loci, which can be up to several hundred kilobases away {Wilkinson, 2007 #621}. There are seven well-characterized imprinted clusters (*Dlk1*; *Gnas*; *Grb10*; *Igf2*; *Igf2r*; *Kcnq1*; *Pws*) in mouse, which contain three to twelve genes and span genomic region of 80 to 3,700kb. The ICE In each of these clusters has been shown by deletion experiments that resulted in the loss of imprinted expression of all genes in

the cluster. Multiple sDMRs can be found in those clusters. In five cases the ICE is maternally methylated (*Gnas*; *Grb10*; *Igf2r*; *Kcnq1*; *Pws*) and in two paternally (*Dlk1*; *Igf2*) (Barlow 2011). In addition to imprinted protein-coding genes, all well characterized imprinted clusters also contain imprinted lncRNAs that may act as regulators (Autuoro et al., 2014).

3.5. Different mechanism of genomic imprinting

3.5.1. *H19/Igf2* cluster

With around 80kb the *H19/Igf2* cluster is one of the smallest imprinted clusters, but with *H19* it contains the first imprinted lncRNA ever described (Bartolomei et al., 1991). Nearly at the same time, imprinted expression was also demonstrated for *Igf2* (DeChiara et al., 1991). This cluster is located on murine chromosome 7 and the lncRNA *H19* is maternally transcribed, while the growth promoter gene *Igf2* shows paternally imprinted expression (Bartolomei and Ferguson-Smith, 2011) (**Figure 3**). The ICE lies between the *H19* lncRNA promoter and protein-coding gene *Igf2* (Tremblay et al., 1997). As an example of the insulator model in imprinted clusters (Thorvaldsen et al., 1998); (Abramowitz and Bartolomei, 2012) it contains four binding sites for the insulator CTCF (Bell and Felsenfeld, 2000); (Kanduri et al., 2000). *H19* and *Igf2* share the same enhancers which are located upstream of *H19*. The maternal ICE is un-methylated and allowing *CTCF* to bind and create an insulator effect (Bell and Felsenfeld, 2000). This binding event creates a physical barrier between the enhancer and the promoter of *Igf2* (Bartolomei, 2009). As a consequence of *CTCF* binding a chromosomal loop is formed (Engel et al., 2008), which influences the spreading of heterochromatin formation (Cuddapah et al., 2009). *Igf2* is silenced and the enhancer interacts with the promoter of *H19* to activate this lncRNA. On the paternal chromosome the ICE is methylated, which inhibits the binding of *CTCF*. Consequently, the enhancer can activate the transcription of *Igf2* (Bartolomei, 2009). No loop is formed between enhancers and *H19* promoter and so causes silencing (Murrell et al., 2004). The *H19/Igf2* cluster is a well-studied example, which shows the essential function of enhancers in controlling imprinted expression.
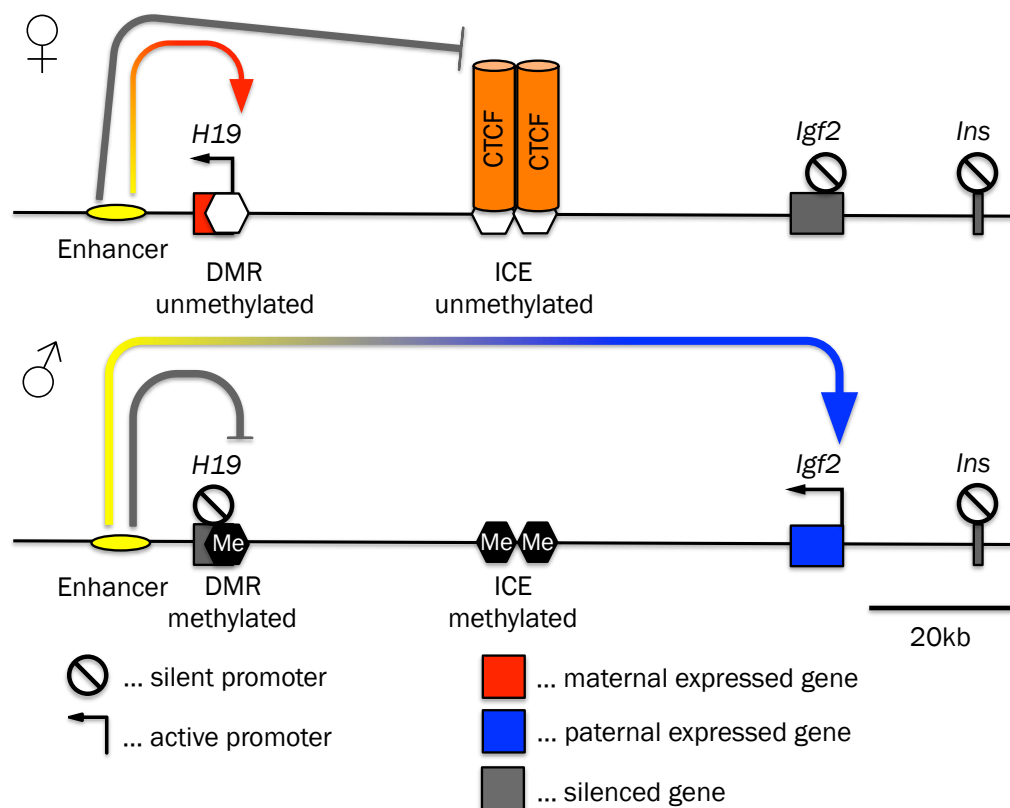
Figure 3: The *H19*/*Igf2* cluster

The ICE of this cluster is located between the protein-coding *Igf2* and lncRNA *H19*. On the maternal allele (♀) the unmethylated ICE is bound by CTCF, which physically blocks enhancer access. The enhancer region cannot form a loop to the promoters of *Igf2*, which stay silenced on this allele. On the paternal allele (♂), the DNA at the ICE is methylated and so the binding of CTCF is inhibited. The enhancer can activate *Igf2* by looping to their promoters. Consequently, the enhancer cannot activate *H19*. In addition, the heterochromatin formation at the ICE spreads to the promoter of *H19* and so reinforces this silencing.

The correct expression pattern of *H19*/*Igf2* plays an important role in embryonic development and deregulation is linked to Beckwith-Wiedermann Syndrome, Silver-Russell Syndrom and cancers (Uribe-Lewis et al., 2011). The imprinted lncRNA *H19* produces a 2.3kb efficiently spliced, noncoding transcript (Milligan et al., 2002), which serves as a precursor for miR675 (microRNA), a repressor of placenta growth (Matouk et al., 2013). This miRNA is acting *in trans* to controlling mice growth (Gabory et al., 2009) and it was demonstrated to have a tumor suppressor function (Hao et al., 1993) as well as oncogenic properties (Chen et al., 2000) and for the loss of imprinted expression of *Igf2* (Kaneda and Feinberg, 2005) underlying the essential importance of these genes.

3.5.2. *Igf2r/Airn* cluster

The murine insulin-like growth factor 2 receptor (*Igf2r*) cluster covers a region on chromosome 17 of around 450kb (**Figure 4**). The growth repressor gene *Igf2r* is maternally expressed and around 1/3 of its gene body is overlapped by an antisense lncRNA transcript called *Airn* (antisense *Igf2r* RNA noncoding) (Lyle et al., 2000). The exclusively paternally expressed, 118kb lncRNA *Airn* is largely un-spliced, nuclear localized and has a short half life (Barlow, 2011). lncRNAs with these specific features similar to *Airn* are defined as macro lncRNAs (Guenzl and Barlow, 2012). The ICE is located in intron 2 of *Igf2r* and contains the promoter of *Airn* in antisense (Wutz et al., 2001). In the pre-implantation embryo (before E4.5) *Igf2r* is expressed biallelic on a very low level, but no *Airn* expression is seen until E6.5. At this time point the upcoming paternal expression of *Airn* represses *Igf2r* on this allele, while maternal *Igf2r* is up-regulated (Lerchner and Barlow, 1997); (Latos et al., 2009). On the maternal allele the ICE is methylated and consequently *Airn* is not expressed (Stöger et al., 1993). The act of transcription of *Airn* via the promoter region of *Igf2r* causes its silencing on the paternal allele by transcriptional interference (Latos et al., 2012). Additionally, neither DNA methylation nor histone modifications are required for silencing *Igf2r* (Latos et al., 2012); (Santoro et al., 2013). *Airn* truncation before it reaches the promoter of *Igf2r* leads to an up-regulation of this gene and results in a biallelic expression and the loss of imprinted expression (Latos et al., 2012). Interestingly, in such a situation the expression of *Igf2r* does not affect the transcription of *Airn.* The movement of the truncated version of *Airn* so close to the promoter of *Igf2r* that an overlap occurs again, results in a normal imprinted expression of the cluster. 97.7% of the un-spliced and 100% of the spliced *Airn* transcript can so be excluded to have a function in silencing *Igf2r*. The transcription of *Airn* through the promoter region of *Igf2r* is required to established imprinted expression in this cluster (Latos et al., 2012). Recently, it was demonstrated that during development the silencing of *Igf2r* via *Airn* transcription requires continuously expression of the RNA and it is not restricted to a developmental window (Santoro et al., 2013).
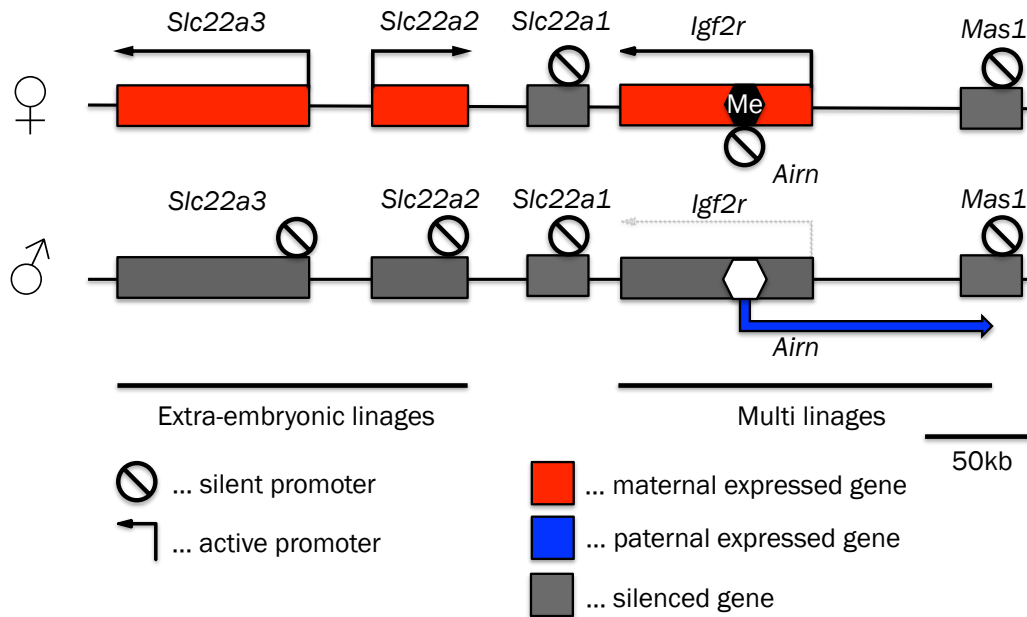
Figure 4: Schematic of the *Igf2r* cluster in mouse

Expression pattern of *Igf2r* cluster from maternal (♀) and paternal (♂) alleles. *Igf2r* and the macro lncRNA *Airn* show multi-linage (ML) imprinted expression. *Slc22a2/a3* show extra-embryonic linages (EXEL) imprinted expression. DNA methylation (Me) on the maternal ICE (hexagon) is required to silence *Airn* on this allele. On the paternal allele *Airn* is transcribed and cause silencing of *Igf2r* by transcription interference. Maternally (red box), paternally (blue arrow) expressed genes and silenced genes (grey box). Black arrows indicate the orientations of the transcripts. Minimal paternal expression of *Igf2r* is indicated by grey, pointed arrow. ⊘ indicates inactive gene.

Only in extra-embryonic tissues the imprinted cluster is spread to neighboring genes, *Slc22a2* and *Slc22a3*, which are consequently transcribed only from the maternal allele (Sleutels et al., 2003). In addition to *Igf2r*, also the imprinted expression of *Slc22a2* and *Slc22a3* is controlled by lncRNA *Airn* (Sleutels et al., 2002). This can be observed for both genes in visceral endoderm and for *Slc22a3* additionally also in placenta, both extra-embryonic linages. *Slc22a2* expression cannot be detected in placenta (Hudson et al., 2011). In placenta it was demonstrated that *Airn* recruits H3K9 histone methyltransferase *G9a* to the promoter of *Slc22a3* and thereby causes silencing (Nagano et al., 2008). Both factors are ubiquitously expressed and so can insufficiently explain this tissue specific expression pattern. Transcription interference with promoters cannot be used to explain, how *Airn* silences these gene, because their promoter regions are located 159kb and 237kb upstream of *Airn* promoter (Sleutels et al., 2003). Consequently, transcription of the lncRNA does not overlap with these promoter regions

(**Figure 4**). The mechanism how, *Airn* is able to silence these genes is unknown so far.
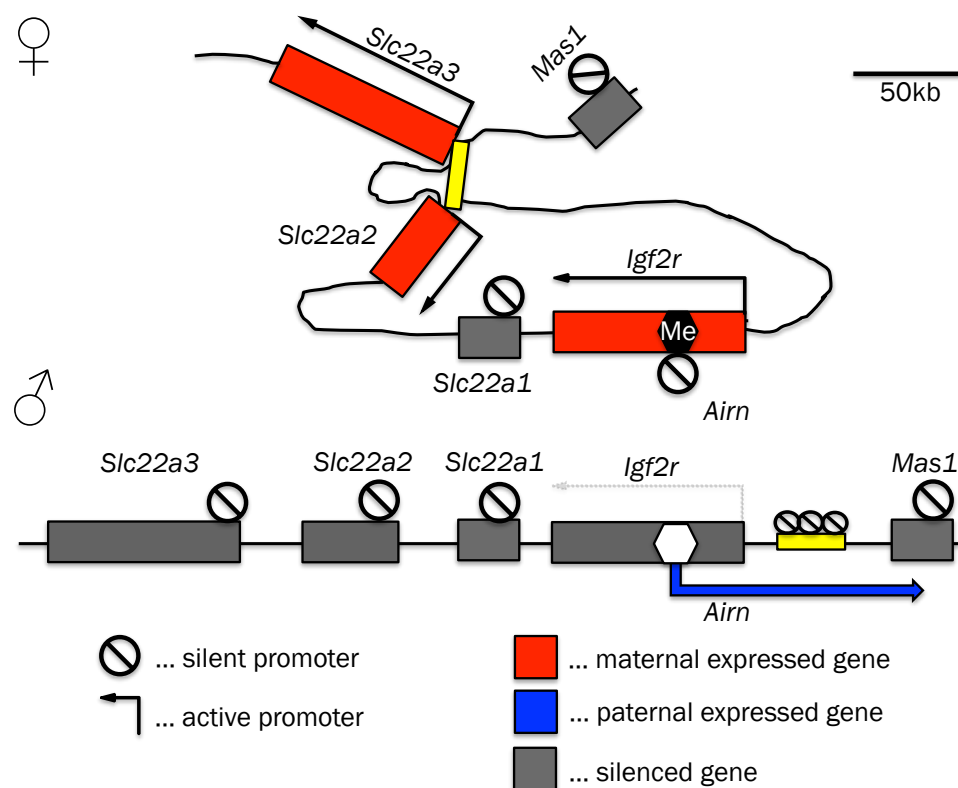
Figure 5: Enhancer interference model in the *Igf2r* cluster

In extra-embryonic tissues the *Igf2r* cluster is extended and the genes *Slc22a2/a3* show imprinted expression. *Airn* silences *Igf2r* via transcription interference with its promoter. On the maternal allele (♀) (top) *Airn* is silenced via DNA methylation on the ICE and therefore a loop between the active enhancer and the promoter regions of *Slc22a2/a3* may be formed. Tissue specific enhancer elements in the *Airn* gene body may be required to activate *Slc22a2/a3* and these elements may be blocked by *Airn* expression on the paternal allele (♂) (bottom). Maternally (red box), paternally (blue arrow), biallelically (green box) expressed genes as well as silenced genes (grey box) are shown. The orientations of the transcripts are indicated by black arrows. Baseline paternal expression of *Igf2r* is indicated by a grey dotted arrow. ⊘ indicates inactive gene.

Based on the unique silencing mechanism of *Igf2r*, an enhancer-interference model was proposed, which links the tissue-specific imprinted expression pattern with silencing of tissue-specific enhancers (Pauler et al., 2012) (**Figure 5**). The concept is, that transcription of *Airn* through the enhancer region might interfere with the ability of these elements to activate target genes and thereby cause silencing. Chromosomal loops may be formed between those enhancer regions and the promoters of *Slc22a2* and *Slc22a3* exclusively on

the maternal allele to activate tissue-specific transcription. The regulation of enhancers is a very complex process, which includes many different factors. Consequently, the activity and strength of enhancers can be regulated in a highly tissue-specific manner (Ong and Corces, 2011b). The silencing of specific enhancers might be an explanation, how the expression levels of *Slc22a2* and *Slc22a3* are controlled by *Airn* lncRNA (**Figure 5**). In addition, the repression of enhancers could be a requirement for the tissue-specific recruitment of an ubiquitously expressed repressive histone modifier (*G9a*) to the promoter of *Slc22a3* in placenta (Nagano et al., 2008)*.*

3.6.   The role of lncRNAs in imprinted clusters

In general, lncRNAs are defined as transcripts of at least 200nt in length, which are not translated into a protein (Guttman and Rinn, 2012). The first lncRNA ever described was *H19*. Before it was characterized as an imprinted lncRNA (Bartolomei et al., 1991), it was thought to be a mRNA with special features: no large open reading frames (ORFs), no translation event and no protein product was detected (Brannan et al., 1990). The next prominent lncRNA described was *Xist*, which is expressed from the inactive X-chromosome and required for X-inactivation (Brown et al., 1991). In the following years new lncRNAs were described like the imprinted lncRNA *Airn* in mouse (Sleutels et al., 2002) and the HOX transcript antisense RNA (*HOTAIR*) in human (Rinn et al., 2007). Another interesting point is the fact that *Xist*, *Airn* and *Kcnq1ot1* (Brown et al., 1991); (Sleutels et al., 2002); (Kanduri et al., 2006) are *cis* acting lncRNAs, while *HOTAIR* acts in *trans* (Rinn et al., 2007). For the miRNA encoded in *H19* (Milligan et al., 2002) a *trans* acting function was proposed to controlling mice growth (Gabory et al., 2009).

In the *Igf2r* cluster the protein-coding gene *Igf2r* is silenced on the paternal chromosome via transcription interference of the macro lncRNA *Airn* (Latos et al., 2012) (**Figure 4**). *Airn* has special features in common with other imprinted lncRNAs, including a low splicing rate, lack of sequence conservation and an

unusually large size (Guenzl and Barlow, 2012), that indicate no function for the transcribed product (Seidl et al., 2006). *Kcnq1ot1*, a macro lncRNA expressed paternally in the *Kcnq1* cluster, shows similar behavior as *Airn*. Also a truncation of *Kcnq1ot1* leads to a loss of imprinted expression in the cluster (Mancini-Dinardo et al., 2006). The same is true for *Nespas* macro lncRNA in the *Gnas* cluster (Williamson et al., 2011).

In these imprinted clusters it was shown, that lncRNAs are required to silence neighboring genes. How lncRNAs cause imprinted silencing of more distant genes remains unclear. Also the mechanism, how tissue-specific extensions of imprinted clusters in extra-embryonic tissues are regulated by lncRNAs is unclear so far (Hudson et al., 2011). One hypothesis that could explain the underlying mechanism is that lncRNAs may interfere with the ability of essential tissue-specific enhancers for imprinted genes to inhibit the up regulation on that allele (Pauler et al., 2012).

3.7. Tissue specificity of the *Igf2r* cluster

In the majority of somatic mouse tissues *Igf2r* and *Airn* are expressed in a parent-of-origin specific manner. This is called multi linage (ML) imprinted expression (**Figure 4**). In contrast, extra-embryonic lineage specific (EXEL) imprinted expression is found in placenta, a short-living extra-embryonic tissue, and visceral yolk sac (VYS), a nutritive membrane surrounding the embryo during development, which is also required for removing toxins at very early stages as well as hematopoiesis (**Figure 6**). *Slc22a2* and *Slc22a3* are examples for EXEL imprinted expression in mouse (Hudson et al., 2011) (**Figure 4**).

Typically placenta is the tissue of choice to test extra-embryonic imprinted expression. However, placenta is a very complex tissue containing multiple cell types, which may lead to biallelic expression in some cell types masking imprinted expression in others (**Figure 6**). In addition, in placenta a high number of maternal cells, like maternal blood vessels and blood, and the

closely interfacing decidua, which is difficult to completely remove, can bias analysis of imprinted expression. Consequently, placenta is not an optimal tissue to study EXEL imprinted expression. A less complex extra-embryonic tissue without maternal contribution would aid investigations of regulation of imprinted expression in the *Igf2r* cluster in extra-embryonic tissues. It was demonstrated that the VYS, fulfills these requirements (Hudson et al., 2011). In addition, if necessary the visceral endoderm (VE) layer, which shows EXEL imprinted expression, can be isolated, although this is usually not necessary for studies of *Slc22a2* and *Slc22a3* as they are only expressed in the VE, and not in the mesoderm layer (Hudson et al., 2011) (**Figure 6**). The VYS is an optimal tissue to answer, how the lncRNA *Airn* causes silencing of genes not overlapped by its transcription in a tissue-specific way. Based on the enhancer interference model, the continuous transcription of the lncRNA *Airn* may interfere on the paternal allele with these enhancers Consequently, these elements stay inactive and their target genes are not transcribed.
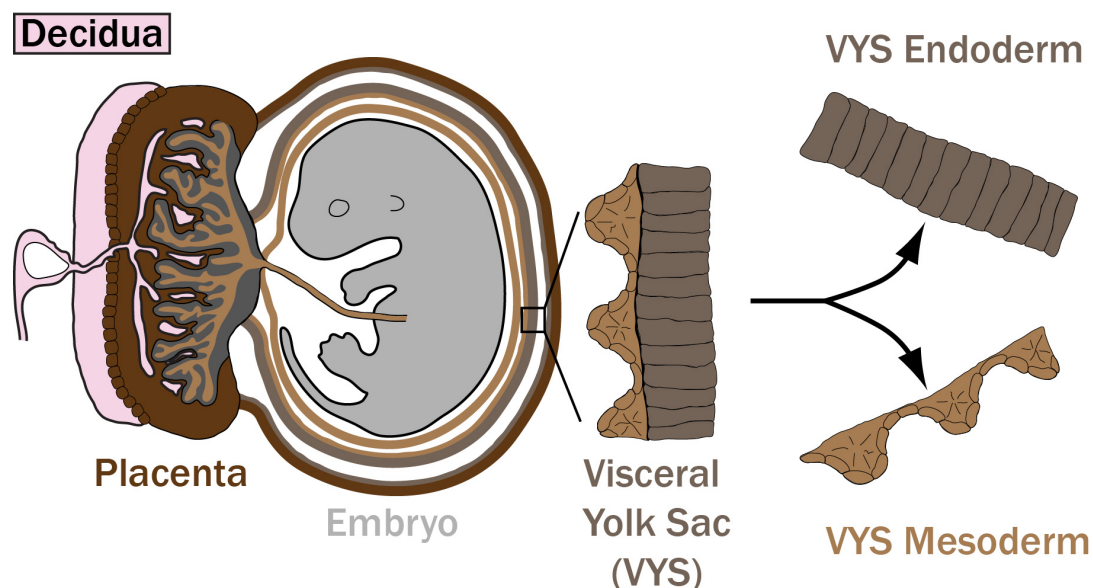
Figure 6: Schematic of a mouse embryo at E12.5

The decidual (pink) is a uterine lining, which forms the maternal part of the placenta. The embryo (grey) is first surrounded by the amnion, then by the visceral yolk sac (VYS) and the most outer membrane is the parental yolk sac. The placenta is infused by maternal cells because of its function to support the embryo with nutrients. The VYS is a less complex tissue (two layers) and has no contact to maternal linages. If necessary, the VYS membrane can be separated into these layers (VYS endoderm (VE); VYS mesoderm (VM)) (Adapted from (Hudson et al., 2011)).

3.8.   The activators of transcription: Enhancers

Enhancers or *cis*-regulatory modules (CRMs) are short genomic regions of 200 bp to 1500 bp in size, which are recognized by activators to up-regulate transcription of target genes (Dickel et al., 2013). This spatiotemporal regulation process of gene expression is required to establish tissue-specific expression profiles that control developmental process (Bulger and Groudine, 2011). Enhancer elements are acting *in cis* and can be located up- or down-stream of their targets. They can be found in introns as well as in gene bodies. Interestingly, in some cases they can also activate target genes at distances of over 1Mb (Maston et al., 2006), like for SOX9 (Smyk et al., 2013). Normally, tissue-specific transcription factors can bind to enhancers to promote the expression of target genes. Multiple binding sites for at least two classes of sequence-specific transcription factors can be found as well as mixtures of transcriptionally activating and repressing domains within the same enhancer region (Dickel et al., 2013). These areas can be seen as operational binding platforms to recruit multiple transcription factors, which recognize a specific sequence motive for binding (transcription factor binding motive). Their binding can lead to the motive recruitment of co-activators or co-repressors and can have an effect on the chromosomal environment (Straub and Becker, 2008). These bound factors determine the activity of the enhancer and contribute to how efficient a target gene is transcribed. The modular nature of enhancers is a feature, which allows multiple enhancers to contribute to the expression level of target genes in an additive manner (Arnone and Davidson, 1997).

An additional hallmark for these elements is that the orientation relative to its target genes do not influence the effect on target (Amano et al., 2009). The distance between enhancers and target promoters may have an effect on its efficiency, because the largest known distance between such regions is 1.6Mb (Smyk et al., 2013). Active enhancers are often found in regions devoid of nucleosomes, where the DNA is more freely accessible. The remaining nucleosomes carry specific active histone marks like H3K4me1 (Heintzman et

al., 2007) and H3K27ac (Creyghton et al., 2010), which can be used to identify new tissue-specific enhancers (Heintzman et al., 2009).

The open chromatin state of an enhancer is required to allow transcription factors to bind their recognition motive. The binding of these activators happens parallel
 to the formation of chromosome loops, which allow the interaction with target transcription start sites. This looping model explains long-range interactions and specific activation of targets, while neighboring genes are not affected (Bulger and Groudine, 1999). Apart from this model, a second mode of action called the linking model was proposed. The concept is, that enhancer-promoter communication is controlled by signals propagated along the intervening DNA. Enhancer facilitators may organize the DNA sequence between promoter and enhancer in a way, that transcription is promoted (Bulger and Groudine, 1999).

To establish such long-range interactions in a specifically direct manner, a mechanism to protect or insulate off-targets and promote target gene interactions has to be established. Results of whole-genome assays suggest that insulator DNA elements, also called chromosomal barrier elements, bound by insulator proteins such as *CTCF* play this role (Filippova et al., 1996). Cohesion proteins are often enriched at *CTCF*-bound insulator sites to increase the stability of these interactions (Parelho et al., 2008). The binding sites are often located in flanking regions of target genes to prevent aberrant activation of neighboring genes by physically blocking enhancers from interacting (Gaszner and Felsenfeld, 2006). The formation of loops is inhibited and this allows enhancer to specifically target genes via long-range interactions and not just genes in their close neighborhood (Kuhn and Geyer, 2003); (Chopra et al., 2009). Enhancers can be controlled on different levels and require the binding of multiple transcription factors and other factors to allow enhancer-promoter interactions meaning enhancers are usually tissue-specific (Ruf et al., 2011). All these features together make it hard to predict potential enhancer targets as well as to link enhancers to target genes.

Novel methods based on Chromosomal Conformation Capturing (3C) (Dekker et al., 2002) allow capturing these long-range interaction and systematic qPCRs or deep-sequencing can be used to visualize interaction patterns for certain genomic regions or on a genome-wide level.

3.9. Chromosomal Conformation Capture (3C)

Chromosome Conformation Capture (3C) based protocols provide medium to high-throughput techniques to analyze overall spatial organization of chromosomes under different conditions. This approach allows analyzing the interaction frequency between any two genomic loci in a nucleus (Dekker et al., 2002).
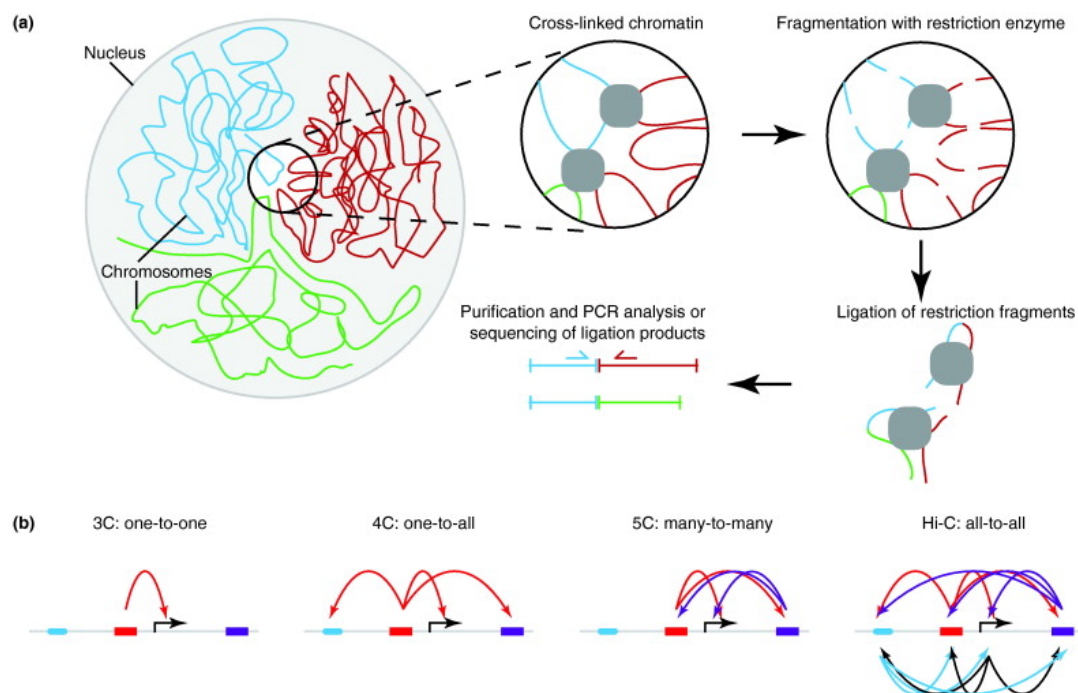
Figure 7 Overview of Chromosome Conformation Capture (3C) based techniques

(A) Summary of the basic concept of 3C including all critical steps (Isolation, cross-linking, digest, ligation, qPCR).
(B) Mode of action for 3C and 3C-based techniques. 3C allows assaying the interaction frequency between any two genomic regions. 4C is used to detect all interactions with one site of interest. With the 5C approach all interactions for multiple regions of interest can be analyzed at once. Hi-C allows a genome-wide picture of interactions.to be obtained (Taken from (Wijchers and de Laat, 2011))

The concept of 3C (**Figure 7A***) is to isolate intact nuclei, which are then treated with formaldehyde to cross-link proteins with other proteins and DNA. Consequently, regions with high level of interactions are more likely cross-linked to each other via protein interactions. For quantification, a restriction digest, followed by ligation reaction is performed leading to the establishment of new, artificial DNA ligation fragments. It is more likely, that cross-linked DNA fragments are ligated than random ligation events, as the protocol is performed in a way to favour intra-molecular ligation. qPCRs are performed to analyze the interaction frequency between two regions of interest.

Novel methods based on the 3C technique were established to obtain chromatin interaction maps at increasing resolution scale (**Figure 7B**). 3C ligation products can be detected by inverse PCR, so called 3C-on-Chip (4C) (Simonis et al., 2006); (Zhao et al., 2006), by multiplexed ligation mediated amplification (5C) (Dostie et al., 2006), or by introducing a biotin mark at the ligation junction to facilitate unbiased purification of ligation junctions (Hi-C) (Lieberman-Aiden et al., 2009) (Belton et al., 2012). Recently, a combination of 3C with deep sequencing approaches allows creating a comprehensive genome-wide interaction map (Rodley et al., 2009); (Sexton et al., 2012). These new high throughput techniques allow studying the chromosomal organization on a genome-wide level at different time points to become a better understanding of the dynamic processes in nuclei.

3.10. Early Mouse Development

Genomic imprinting plays an important role in developmental process of placental mammals and deregulation often results in embryonic lethal phenotypes. After fertilization by a sperm cell, the pronuclei fusion of these two gametes initiates embryogenesis. At the 128-cell stage a structure is formed called the blastocyst, which allows distinguishing between two cell types. At one end, called the embryonic pole, the inner cell mass (ICM) develops. The surrounding layer is called the trophectoderm, the first extra-

embryonic lineage, which is important for the formation of the placenta (Lanner, 2014).

The ICM develops to the epiblast and the primitive endoderm (Fleming et al., 1984), an extra-embronic cell type, which is clearly separated and differentiates into certain linage (Tanaka et al., 1998). By E4.5 three different cell types can be distinguished, the embryonic epiblast and two extra-embryonic lineages, the trophoectoderm and primitive endoderm (Rossant and Tam, 2009). The epiblast cells form all embryonic tissues, while the trophoectoderm only gives rise to the extra-embryonic cell linages (extra-embryonic ectoderm, ectoplacental cone, the primary and secondary giant cells) (Rossant and Tam, 2009). The extra-embryonic visceral and parental endoderm arises from the primitive endoderm gives rise to the visceral and parietal endoderm post-implantation. The implantation into the uterus occurs around E4.5 to E5.0 (Kwon et al., 2008). At E6.5 the embryo starts with the process of gastrulation and during this process the three germ cell layers (ecto-, meso-, endoderm) are formed (Downs and Davies, 1993). Three extra-embryonic membranes surround the embryo at E8.5. While the amnion has no contact points, the visceral as well as the parental yolk sac are in contact with the placenta (Hudson et al., 2011). Between E12.5 (**Figure 6**) and the birth, which normally happens between E19.0 and E21.0, the organs develop further and the embryo increases in size.

3.11. Mouse Chromosomal Engineering

The bacteriophage enzyme Cre recombinase (named as "causes recombination") (Sternberg and Hamilton, 1981) recognizes target DNA sequences and is able to catalyze a site-specific recombination event between two of those sites without any additional cofactors, also in a mammalian system (Sauer and Henderson, 1988). The used sites are called loxP sites and are 34 bp long DNA sequences. Two 13 bp palindromic sequences flank the 8 bp spacer sequence, which gives the orientation to the entire loxP sequence (Nagy, 2000). When two loxP sites are located on the

same chromosome a Cre-mediated recombation *in cis* can happen and results in two possible outcomes (Orban et al., 1992). If the loxP sites are orientated in the same direction, the in-between DNA region is cut out and gets lost. A deletion of the certain genomic area is the outcome. If the loxP sites are oppositely orientated, the intermediate region will be inverted (Nagy, 2000). If the loxP sites are on different homologous chromosomes a recombination *in trans* is also possible, but just between loxP sites with the same orientation, and only during meiotic pairing (Hérault et al., 1998) (**Figure 8**). We used this approach using loxP sites from existing mouse strains in conjunction with a ubiquitously expressed Cre to cause *trans* recombination between homologous chromosomes during male meiotic pairing, a so-called chromosome engineering that should result in a deletion or duplication (**Figure 8**).
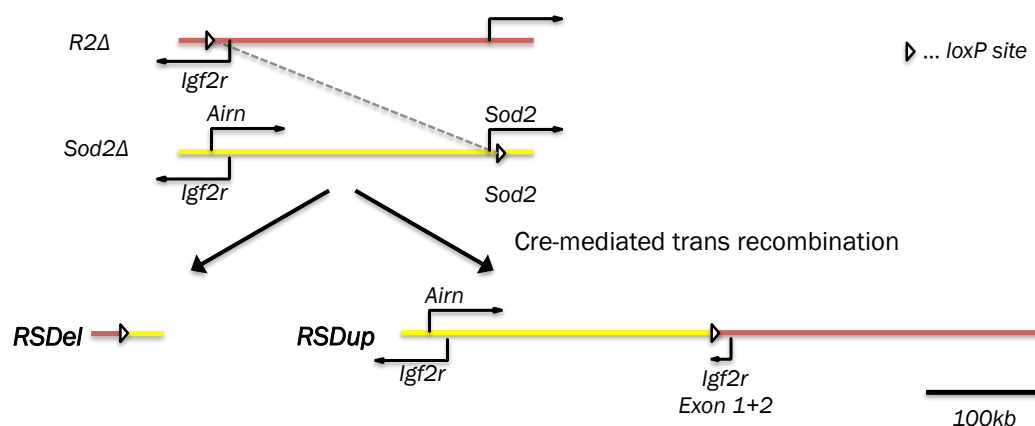
Figure 8: Example for the basic concept of the mouse chromosomal engineering experiment

Two mouse strains (R2Δ; Sod2Δ), with loxP sites at different positions on the same allele, were crossed together. A ubiquitously expressed Cre can mediate a recombination *in trans* resulting in a deletion (RSDel) and a duplication (RSDup).

Deletions in mice are a classical approach to identify and analyze the phenotype of the affected gene or genomic region *in vivo*. In addition, duplications may also give some insight into gene regulation, for example, to study the effect of gene up-regulation, dose-depended phenotypes or the importance of the genomic environment on gene expression levels. Nowadays, new technique like CRISPR/Cas9 can also be used to modify

28

genomic regions (Cong et al., 2013), but few large deletions have yet been reported, and duplications cannot be made using this technique.

3.12. Aim of this project

The practical work for this project was done in the Lab of Denise P. Barlow at the Centre for Molecular Medicine of the Austrian Academy of Science (CeMM) in Vienna, Austria. My work is part of a bigger approach designed by Quanah J. Hudson, a postdoc in the Barlow lab. The aim of my project is to address how in the *Igf2r* cluster the macro lncRNA *Airn* causes tissue specific imprinted silencing of *Slc22a2* and *Slc22a3* in extra-embryonic linages (EXEL).

Before I came to the lab preliminary evidence was obtained supporting an enhancer-interference model where *Airn* expression disrupts extra-embryonic-specific enhancers on the paternal allele to cause imprinted silencing of *Slc22a2/a3* (Pauler et al., 2012). Chromosome Conformation Capture (3C) (Dekker et al., 2002) experiments were performed in MEF (mouse embryonic fibroblasts) cells as well as with mouse tissues. A cross between FVB/NJ and T^{hp} mice were used to analyse relative interaction frequency in E12.5 visceral yolk sac between the promoter regions of *Slc22a2* and *Slc22a3* and the entire gene body of *Airn*. The T^{hp} allele carries a 5Mb deletion, including the *Igf2r* cluster, which allows the gene expression pattern and chromosomal interactions to be assessed in a mono-allelic manner. By comparing the maternal and paternal allele, a higher level of interaction between target areas exclusively on the maternal allele were observed. In addition, active enhancer marks (H3K27ac) were found on the maternal allele in the same region showing increased 3C interactions. An enrichment of the same mark was found by analysing all four possible genotypes of a T^{hp}/AirnT cross by ChIP for the *Airn* truncation alleles (AirnT). All these findings supported the enhancer-interference model of imprinted silencing. Recently, the gene Pde10a was found as novel imprinted gene in placenta (Wang et al., 2011)

and confirmed in the lab by RT-qPCR of placenta samples from different mouse crosses. That's why this gene was included into the 3C analysis.

My Master's project follows up these findings to further test the enhancer interference model. First I compared already prepared 3C samples of E12.5 embryonic and extra-embryonic mouse tissues with a wild type or a T^{hp} allele with a truncated *Airn* allele. The interaction frequencies of promoter regions of target genes (*Slc22a2, Slc22a3, Igf2r, Pde10a*) or control (*Airn*) with the gene body of *Airn* were assayed by Taqman qPCR to identify regions with a gain of interactions. The qPCR results were analysed and plotted by myself.

Second, I was involved in the chromosome engineering experiments, assisting with planning the breeding of the different mouse strains, and designing and conducting genotyping assays. We used loxP sites from existing mice to create a deletion or duplication of the *Airn* gene and the previously identified VYS specific enhancers, to test if the expression levels of *Slc22a2* and/or *Slc22a3* are altered. Using this approach a 28kb deletion/duplication of the first third of the *Airn* gene, and a 270kb deletion/duplication that includes the entire 118kb *Airn* gene were created. I was responsible for designing and performing new Southern blot and PCR assays to screen for these new created deletion and duplication alleles and I further confirmed all new alleles by Sanger sequencing (*Microsynth*). The newly created allele will be subsequently assessed for their phenotype and their effect on imprinted silencing in the *Igf2r* cluster.

4. Results

The transcription of *Airn* lncRNA interferes with the overlapping promoter of *Igf2r* causing silencing of this gene (Latos et al., 2012). The two non-overlapping genes *Slc22a2* and *Slc22a3* show imprinted expression restricted to extra-embryonic tissues and it was demonstrated that also this specific transcription pattern is controlled by *Airn* via an unknown mechanism (Sleutels et al., 2002). The enhancer interference model as a basic concept was proposed to explain the tissue specific regulation mechanism of *Slc22a2/a3* (Pauler et al., 2012). The transcription of *Airn* might interfere with the ability of essential enhancers to loop to their targets (*Slc22a2/a3*) and consequently cause silencing of these non-overlapping genes. The looping between *Airn* gene and *Slc22a2/a3* promoters was by Chromosome Conformation Capture (3C) (Dekker et al., 2002). I performed 3C assays to test, if *Airn* is required to inhibit the loop formation on the paternal allele. In parallel, new mouse strains with deletions and duplications in the *Igf2r* cluster were made by mouse chromosomal engineering to analyze the effects on *Slc22a2* and *Slc22a3* expression.

4.1. Maternal-specific 3C loops and H3K27ac enrichment indicate enhancers within the *Airn* gene

In this section I describe unpublished results, mainly performed by other people in the lab, which provide an important background to my thesis. I have included these results with the permission of the people who conducted the work.

3C experiments (Dekker et al., 2002) were performed in visceral yolk sac (VYS) by Quanah J. Hudson as described in the methods. Embryos of crosses between FVB/NJ and FVB.AK-Del(17)T^{hp} (T^{hp}) mice were used to analyse relative interaction frequency in VYS of E12.5 embryos between the promoter regions of *Slc22a2* and *Slc22a3* and the entire gene body of *Airn* (**Figure 9A**). The T^{hp} deletion allele allowed chromosomal interactions to be assayed in a parental-allele specific manner. The "Hair Pin tail" (*T^{hp})* deletion

is a naturally occurring 5Mb deletion of murine chromosome 17 that contains the entire *Igf2r* cluster (Barlow et al., 1991). The deletion arose on an AKR/J background at the Jackson Laboratory in 1962 (Johnson, 1974), and was backcrossed onto a C3H/HSnJ background (Jax® Mice and Services http://jaxmice.jax.org/strain/001053.html). In the Barlow lab the T[hp] mice were initially backcrossed to BALB/C, but since 1996 have been backcrossed to FVB/NJ (more than 70 generations). The mice show a dominant short tail phenotype and the line can only be maintained paternally due to maternal-specific expression of *Igf2r* (Barlow et al., 1991). The maternal T[hp]-allele is embryonic lethal at around E14.5 on a FVB/NJ background.
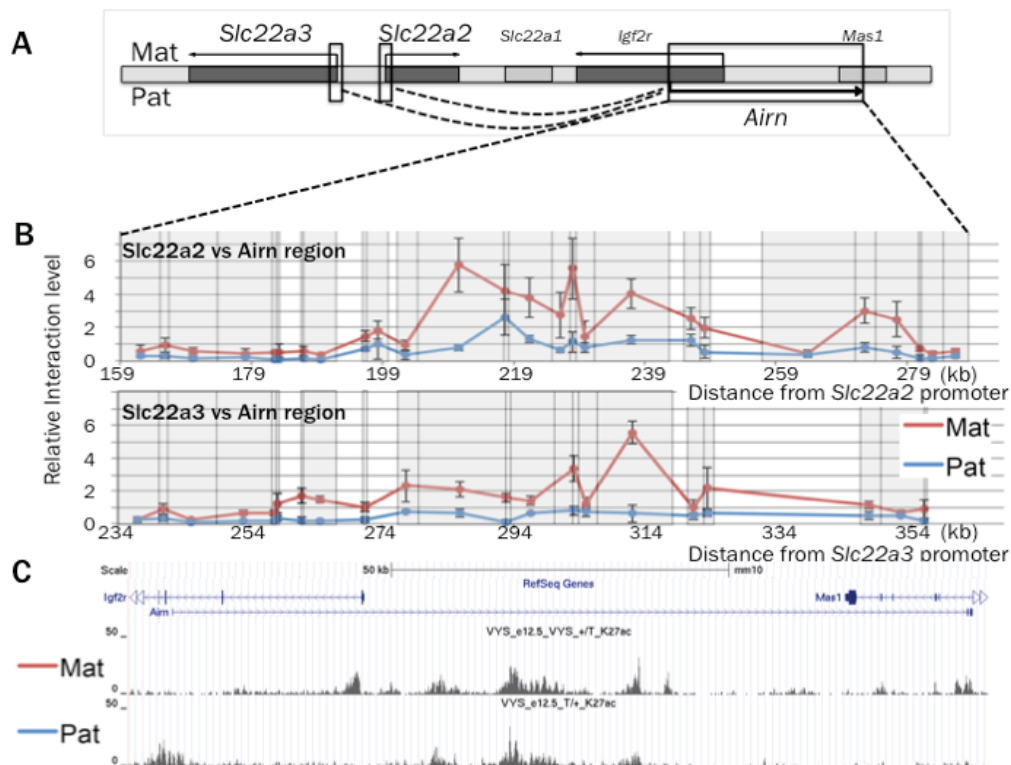
Figure 9: Maternal-specific 3C loops and H3K27ac indicates enhancers in the *Airn* gene body (unpublished data QJ Hudson, with permission)

(A) Schematic of the *Igf2r* cluster indicating the *Slc22a2* and *Slc22a3* promoter regions (small black boxes) that were assayed for 3C interactions with the *Airn* gene (big black box).
(B) In the top blot the interaction frequencies between *Slc22a2* promoter and regions in the *Airn* gene are shown. Below the same comparison is shown for Slc22a3. On the x-axis the distance between bait (*Slc22a2* or *Slc22a3* promoter) and the prey primer (*Airn* gene) is blotted. The relative interaction level on the y-axis was normalized to a control in the *H19* cluster. The grey boxes indicates the EcoRI fragments that were assayed, while the data points are plotted from the middle of each EcoRI fragment. The red line shows the interaction levels only on the maternal (Mat) allele (+/T[hp]), while the blue line (T[hp]/+) shows the paternal (Pat) chromosomal interactions. The interaction frequency of EcoRI digested genomic

fragments was measured by three technical replicates by Taqman qPCR with error bars indicating the standard deviation.

(C) Screenshot of the UCSC genome browser (GRCm38/mm10) showing the entire *Airn* lncRNA transcript according to RefSeq. Below H3K27ac ChiP data from the maternal (+/T^{hp}) and paternal (T^{hp}/+) allele in VYS (E12.5) are shown.

By comparing T^{hp} 3C samples with a maternally or paternally derived allele multiple maternal-specific interactions were detected between the *Slc22a2* promoter and the *Airn* gene in the second 2/3 of the *Airn* gene body (**Figure 9B**). Similarly, maternal-specific interactions with the *Slc22a3* promoter were detected in the same region, although the interaction pattern was not identical with prominent interactions being detected in the *Airn* gene at 215.3, 225.4 and 234.4 kb away from the promoter of *Slc22a2*, whereas a single prominent peak was detected interacting with the *Slc22a3* promoter at 312.0 kb (**Figure 9B**). To detect enhancer activity ChIP-seq was performed on T^{hp} VYS samples using the H3K27ac mark that is associated with active enhancers and promoters (ChIP performed by Quanah Hudson, sequencing library preparation by Philipp Günzl, bioinformatic analysis by Tomasz Kulinski). As expected, H3K27ac was enriched on the maternal *Igf2r* promoter and paternal *Airn* promoter (**Figure 9C**). In addition, maternal enrichment was seen in the second half of the *Airn* gene body correlating with *Slc22a2* and *Slc22a3* promoter 3C interaction peaks. These findings support the predictions of the enhancer-interference model. Furthermore, the same histone mark was tested with all four resulting genotypes of a T^{hp}/AirnT cross (I helped with tissue collection, ChIP and sequencing library preparation by Philipp Bammer, bioinformatic analysis by Florian Pauler) (**Figure 10**). The signal in the wild type (+/+) showed three regions with a high H3K27ac signal in the genomic regions between *Igf2r* and *Mas1*. The paternal truncation of *Airn* (+/AirnT) resulted in an increased enrichment of this active histone mark in the genomic regions mentioned above as well as on the promoter of *Igf2r*. The maternal deleted allele (T^{hp}/+) showed a strongly reduced signal, which was increased in the same regions in the *Airn* truncation (T^{hp}/AirnT) situation. This gain of H3K27ac upon truncation of *Airn* in *Airn* gene regions previously shown to display maternal-specific 3C interactions with the *Slc22a2* and *Slc22a3* promoters supports the enhancer interference model.
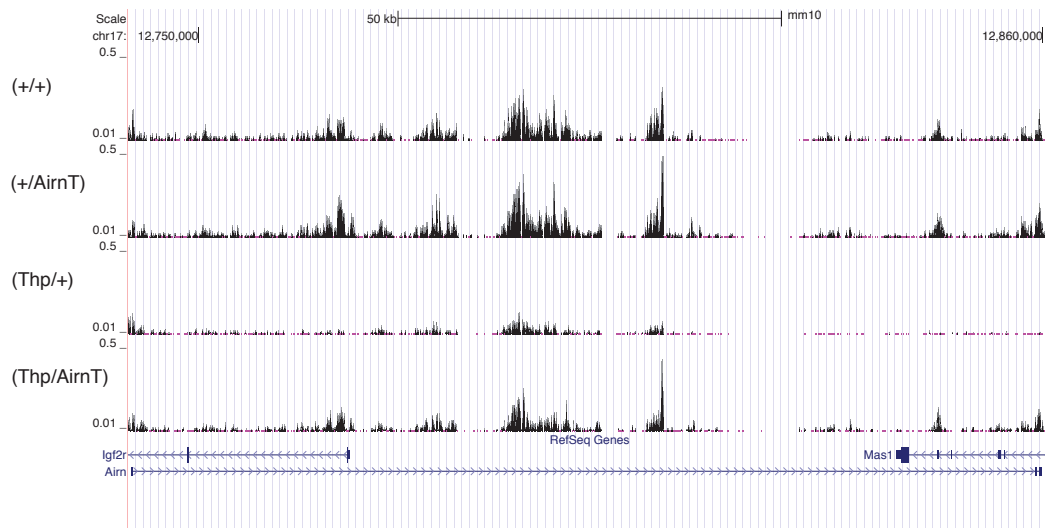
Figure 10: *Airn* reduces enrichment of the enhancer mark H3K27ac on the paternal allele.

UCSC (GRCm38/mm10) screenshot of the Airn lncRNA transcript (118kb). H3K27ac ChiP results are shown for all four resulting genotypes of a T^{hp}/AirnT cross (+/+; +/AirnT; T^{hp}/+; T^{hp}/AirnT).

Recently, we assayed genes within the 5Mb T^{hp} deletion by RT-qPCR for parental-allele specific expression and evidence of silencing by *Airn* in embryo, placenta and visceral yolk sac endoderm (VE) using the T^{hp}, AirnT and R2Δ alleles (Daniel Andergassen master's thesis, Markus Muckenhuber internship report). Using this approach Daniel Andergassen showed that *Pde10a*, a gene that is 4Mb upstream from the TSS of *Airn*, shows imprinted expression in placenta confirming a previous report (Wang et al., 2011). Using a 3C assay designed by Quanah J. Hudson I also examined if interactions between the *Pde10a* promoter and the *Airn* gene could be detected, and if interactions were affected by the truncation of *Airn*.

4.2. Chromosomal loops are controlled by lncRNA *Airn*

One embryonic (trunk) and two extra-embryonic (placenta; visceral yolk sac (VYS)) tissues from E12.5 of T^{hp} x AirnT crosses were used to assess if *Airn* affected 3C looping with the *Slc22a2* and *Slc22a3* promoters. This type of cross gives rise to four different genotypes ((+/+); (+/AirnT); (T^{hp}/+); (T^{hp}/AirnT)). Embryo heads were used for genotyping by Southern blot and individual tissues with the same genotype were pooled. The 3C samples used for this analysis were prepared by Quanah J. Hudson according to the protocol described in the methods section (Dostie and Dekker, 2007), and I performed quality control and 3C qPCR assays. As a negative control a non cross-linked placenta (0% placenta) sample was prepared together with other 3C samples.

During the 3C protocol, a sample was taken from the undigested nuclei (UND) and then after EcoRI digestion (D). Two approaches were taken to assess the quality of the 3C material. First, an aliquot of the UND, D and final 3C samples were separated by gel electrophoresis. The expected pattern of a high molecular weight undigested (UND) DNA, a lower molecular weight smear following EcoRI digestion (D) then an increase in molecular weight following ligation (3C), was seen for all samples except the T^{hp}/AirnT embryo trunk (**Figure 11**). This sample did not show a difference between the UND and D, which indicated an inefficient or no digest.

4.2.1. Quality control of 3C samples
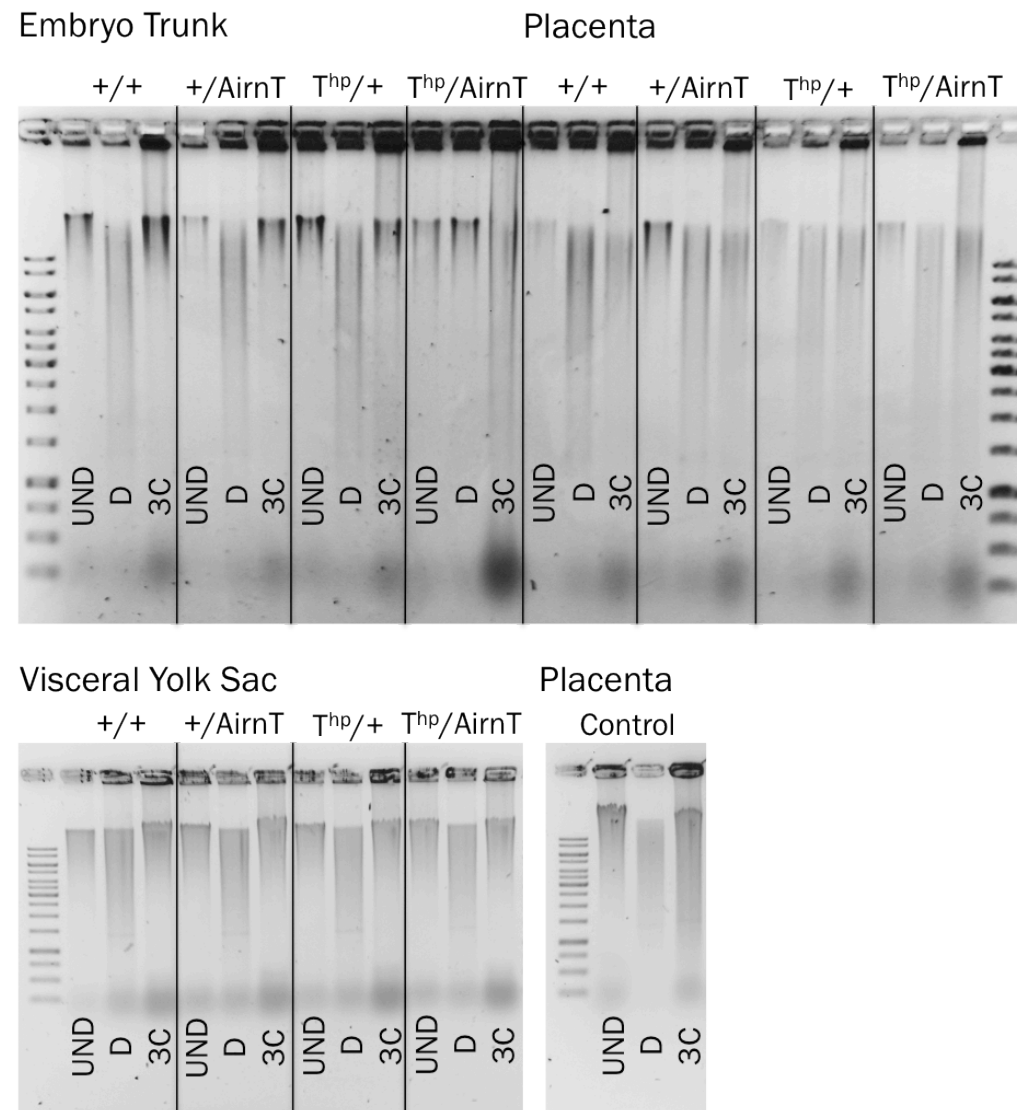
Figure 11: Quality control of 3C samples by gel electrophoresis

Undigested (UND), EcoRI digested (D) and digested and ligated Chromosome Conformation Capture (3C) samples were separated on a 0.8% agarose in 1xTAE gel for 1h at 80V. The gels were stained with EtBr. In three tissues (embryonic trunk; placenta; VYS) four different genotypes were tested. The wild type (+/+) allele, an *Airn* truncation allele with a maternal wild type allele (+/AirnT) and two combinations of a maternal T^{hp} allele with either a wild type paternal allele (T^{hp}/+) or the *Airn* truncation allele (T^{hp}/AirnT) were used for this quality control. In the placenta a control sample was prepared by not formaldehyde cross-linking this sample during the 3C sample preparation. The ladder was GeneRuler^{TM} 100 bp Plus DNA Ladder (Thermo Scientific).

A second test of the 3C material was conducted by calculating the digestion efficiency at selected *EcoRI* sites in the *Airn* gene by qPCR. The digest efficiency was tested using SYBR Green qPCR assays across EcoRI restriction sites comparing UND with D samples. Six primer pairs were used:

Eco_111 F1 and Eco_111_(112) Rev; Eco_112 F1 and Eco_112_(113) Rev; Eco_113 F1 and Eco_113_(114) Rev; Eco_114 F1 and Eco_114_(115) Rev; Eco_116 F1 and Eco_117 R2 (**Table 4** shows genomic location of these assays). All values were then normalized to the Gapdh-3C SYBR Green qPCR assay whose amplicon contained no *Eco*RI sites (Gapdh-3C F and R; **Table 4**). The ratio between normalized D samples to UND samples were calculated and analyzed (**Figure 11**). The ratio between digested to undigested sample was for all samples ranged from 20% up to 75% depending on restriction site. The T^{hp}/AirnT embryo showed a very high ratio, around 100% for all restriction sites, indicating that digestion did not work and confirming the gel result. I therefore excluded this sample from subsequent analysis. The other embryo samples and the VYS samples showed a digestion efficiency of between 25-30% for most assays, whereas placental samples were less efficiently digested with a digestion efficiency of around 35-40% for most assays.
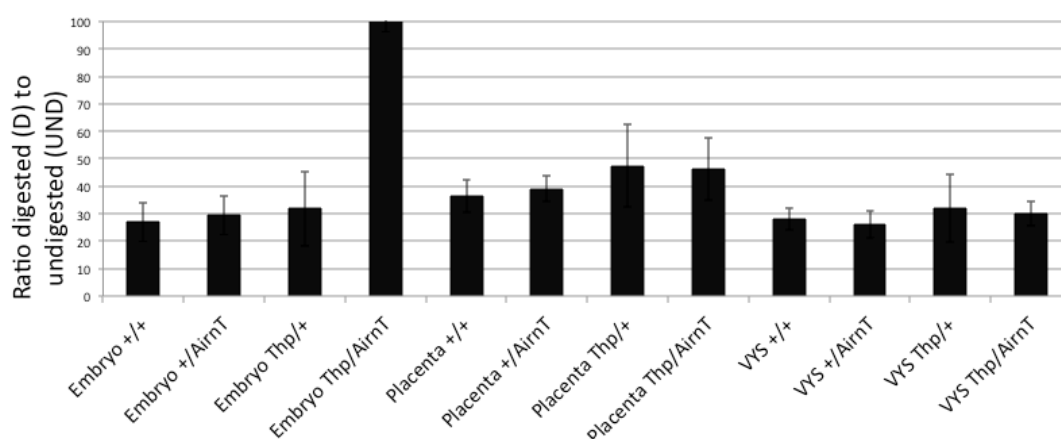
Figure 12: Digestion efficiency of 3C sample assayed by SYBR Green qPCRs

Four genotypes were analyzed in three tissues (embryo trunk, placenta, VYS): wild type (+/+), a maternal wild type allele and paternal *Airn* truncation allele (+/AirnT), a maternal T^{hp} allele with a wild type paternal allele (T^{hp}/+) and a maternal T^{hp} allele with a paternal *Airn* truncation allele (T^{hp}/AirnT). The results of the six primer pairs were averaged and the standard deviation was calculated. The tissues and genotypes are shown on the x-axis. The y-axis shows the digestion efficiency (in %) for each sample and 100% was set as maximum. Each qPCR assay was normalized to a genomic assay at the Gapdh locus and then the digestion efficiency was calculated by dividing the digested sample by the undigested sample and multiplying by 100.

4.2.2. Analysis of 3C samples by Taqman qPCRs

Following the quality control checks, I subjected the 3C samples from embryo trunk, placenta and VYS from T^{hp} x AirnT crosses to Taqman qPCRs 3C assays that were designed by Quanah Hudson. All four genotypes ((+/+); (+/AirnT); (T^{hp}/+); (T^{hp}/AirnT)) were assayed for placenta and VYS, whereas for the embryo only the biallelic samples (+/+ and +/AirnT) were analyzed, while the undigested T^{hp}/AirnT sample and its wild type pair (T^{hp}/+) were not analyzed. All 3C results from the *Igf2r* locus were normalized to the mean of two 3C assays showing interactions with the *H19* ICE, an independent imprinted locus that lies on a different chromosome. The bait primer (Eco004-H19 F1) and a Taqman probe (H19-3C) were used together with two prey primers (Eco013-H19 F3; Eco020-H19 F14) to detect 3C interactions. The standard curves for these two assays were calculated starting from a 1:50 dilution of a PCR amplified fragment for each assay that was derived from a 3C BAC library for *H19* (BAC: RP23-50N22) (dilution series: 1/4, 1/16; 1/64; 1/256). The mean of the two assays for each 3C sample was used to normalize all 3C assays at the *Igf2r* cluster.

For the *Slc22a2*, *Slc22a3*, *Igf2r* and *Airn* 3C assays, a 1:800 dilution of a 3C BAC library (around 800ng/µl) derived from four BACs (RP23-309H20; RP23-367L3; RP23-84H13; RP23-284N16) was used. For the *Pde10a* 3C assays a 1:500 dilution of a second 3C BAC library (842.5 ng/µl) containing five BACs (RP23-309H20; RP23-367L3; RP23-84H13; RP23-284N16; RP23-410M14) was used. In each case the diluted 3C BAC libraries were then used to prepare a dilution series to calculate the standard curves for each 3C assay (dilution series: 1/4; 1/16; 1/64; 1/265). Different sets of Taqman probes, bait and prey primers were used to analyze the interaction frequency between promoter regions of *Slc22a2*, *Slc22a3*, *Igf2r*, *Airn* and *Pde10a* and restriction fragments within the gene body of *Airn* (for details see Materials & Methods). Biallelic (+/+ vs +/*AirnT*) and paternal (T^{hp}/+ vs T^{hp}/*AirnT*) comparisons were made, with a difference considered significant if the error bars did not overlap.

4.2.2.1. *Airn* blocks 3C interactions with the *Slc22a2* promoter

Previous results indicated that there are VYS-specific enhancers within the *Airn* gene that activate maternal expression of *Slc22a2*, implying that *Airn* may silence these enhancers on the paternal allele (**Figure 9B**). Therefore, using 3C material prepared by Quanah Hudson, I conducted 3C assays to test the effect of *Airn* on interactions between the *Slc22a2* promoter and the *Airn* gene body (**Figure 13A**). In the embryo trunk the wild type (+/+) was compared with the *Airn* truncation allele (+/*AirnT*) and showed an overlapping pattern indicating that truncation of *Airn* does not effect interactions with the *Slc22a2* promoter in this tissue (**Figure 13B, top**). Similarly in placenta, where *Slc22a2* is also not expressed, truncation of *Airn* did not effect interactions with the *Slc22a2* promoter in both, the biallelic (+/+ vs +/AirnT) and paternal (T^{hp}/+ vs T^{hp}/AirnT) comparisons (**Figure 13B, middle**). In contrast in VYS, where *Slc22a2* is expressed and shows imprinted expression, multiple regions showed a gain of interactions with the *Slc22a2* promoter upon *Airn* truncation, mostly in the second two thirds of the *Airn* gene body (**Figure 13B, bottom**). These regions were found around 179.7kb, 186.7kb, 208.3kb, 223.6 to 227.3kb, 243.7 to 247.4kb and 270.1kb away from the bait primer in the biallelic situation and between 200.1 to 208.3kb in the paternal plot. For both the placenta and VYS, the overall relative interaction frequency in the paternal (T^{hp}) plot, where only 1 chromosome is present, was reduced to around half the interaction frequency in the biallelic (+) assay, but for each tissue the pattern was similar in both comparisons. In summary, these results indicate that *Airn* blocks loop formation with the *Slc22a2* promoter in VYS.

Figure with panels A and B showing genetic map and interaction frequency plots
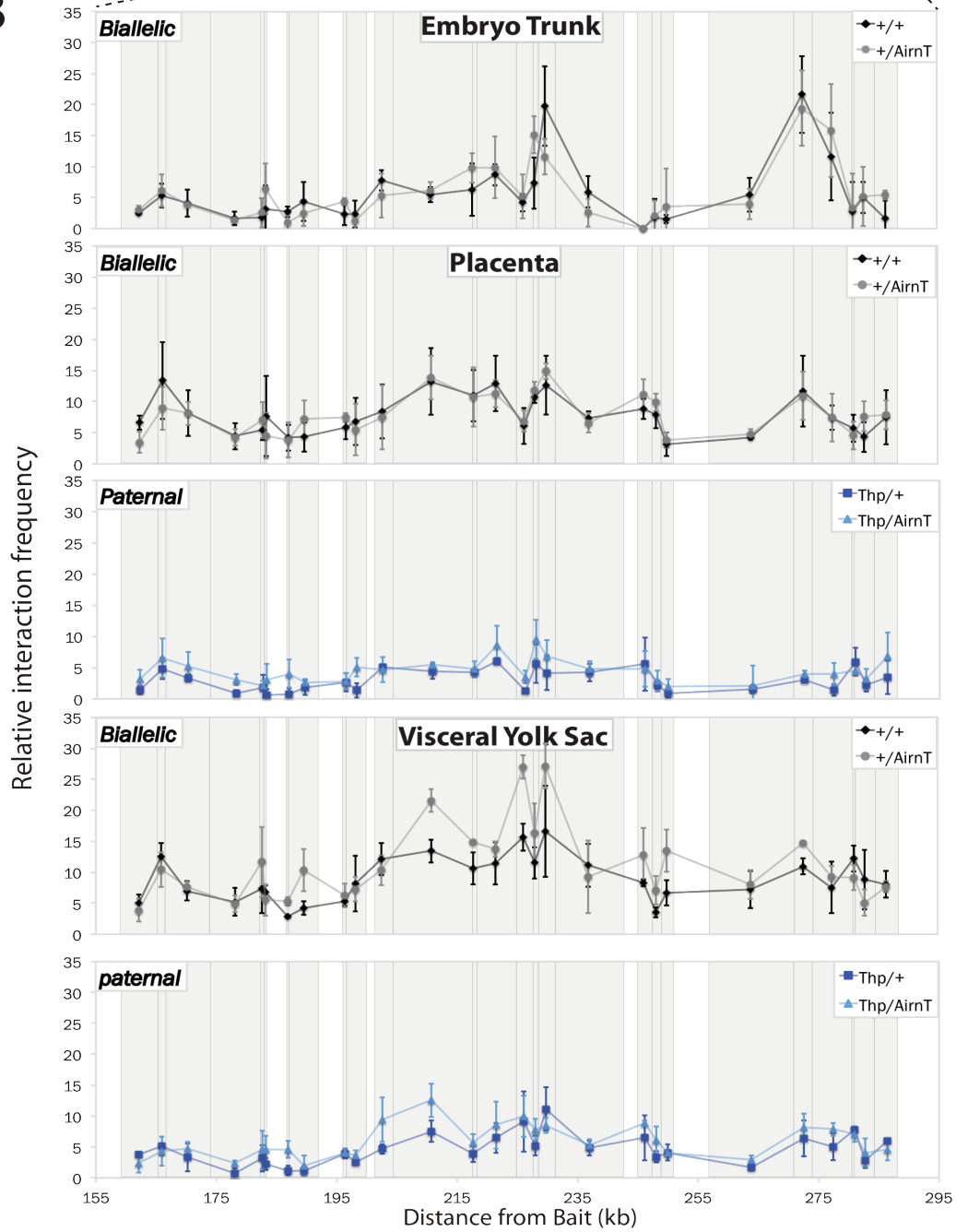
Figure 13: *Airn* **blocks interactions with the** *Slc22a2* **promoter in visceral yolk sac**

(A) Schematic of the *Igf2r* cluster. In this assay the interactions between the promoter region of *Slc22a2* (bait) and regions within the *Airn* gene body (prey) highlighted in black boxes were analyzed.

(B) 3C interaction frequencies between the *Slc22a2* promoter and the 118kb *Airn* gene body in wild type and *AirnT* samples from E12.5 embryos. Top: embryo trunk (biallelic: +/+ vs +/*AirnT*), middle: placenta (biallelic: +/+ vs +/*AirnT* and paternal: T^{hp}/+ vs T^{hp}/AirnT), bottom: visceral yolk sac (VYS) (biallelic: +/+ vs +/*AirnT* and paternal: T^{hp}/+ vs T^{hp}/AirnT). The x-axis shows the distance between the bait fragment (*Slc22a2* promoter) and the prey fragments (*Airn* gene). The relative interaction frequency on the y-axis was normalized to the mean of two *H19* ICE control interactions. Biallelic genotypes are represented in black (+/+) and grey (+/AirnT), paternal genotypes in dark blue (T^{hp}/+) and light blue (T^{hp}/AirnT). Light grey boxes indicate the different sized *Eco*RI fragments. The position and size of these fragments were calculated according to the UCSC genome browser (NCBI37/mm9). Points indicate the mean of three technical replicates, and error bars indicate the standard deviation. Aligned with the 3C results is a screenshot of the UCSC genome browser of the *Airn* region.

4.2.2.2.   *Airn* blocks 3C interactions with the *Slc22a3* promoter in extra-embryonic tissues

For *Slc22a3* previous 3C analyses indicated specific interaction between its promoter and the gene body of *Airn* in both extra-embryonic linages. These interactions were found exclusively on the maternal allele and not on the paternal by Quanah Hudson, indicating that the imprinted lncRNA may play a role in the regulation process. We next tested if interactions between the *Slc22a3* promoter and *Airn* gene body were effected by *Airn* expression (**Figure 14A**). In the embryo the level of interactions was similar between the wild type and *Airn* truncation allele, although a minor increase upon truncation was seen at 257.7kb and 303.0kb from the *Slc22a3* promoter (**Figure 14B, top**). In the biallelic placenta assay, multiple regions (264.7kb; 285.8kb; 303.0kb) showed a gain in interactions in the *Airn* truncation (**Figure 14B, middle**). Although the overall signal of the paternal assay was lower (around 50%), an increase in interactions was again seen at regions 264.7kb and 303.0kb from the *Slc22a3* promoter following *Airn* truncation. Additionally, a gain of interactions was also seen at 241.0 to 245.3kb and 347.7kb, increases that were not detected in the biallelic assay.
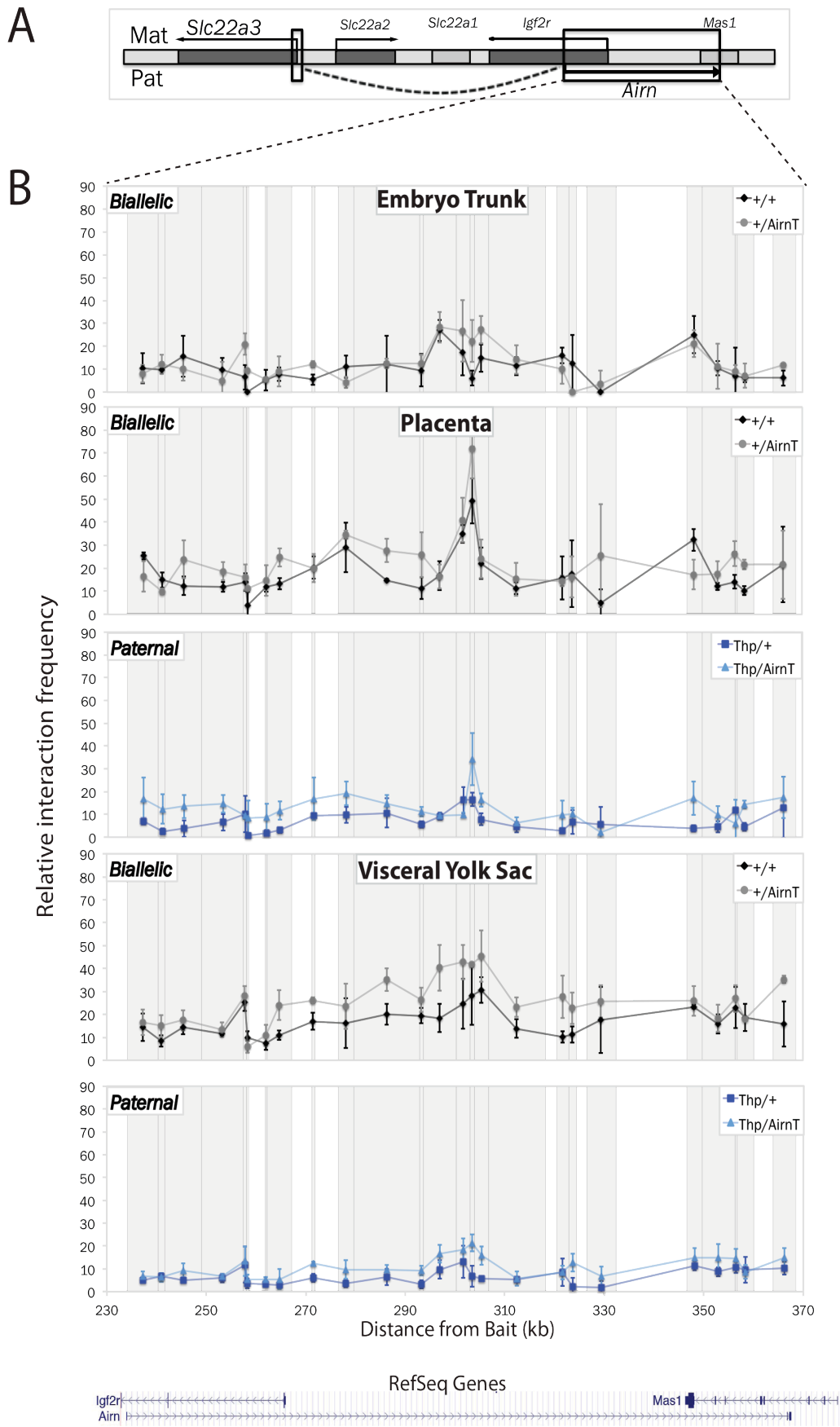
Figure 14: *Slc22a3* promoter and *Airn* gene show a gain of interactions in extra-embryonic tissues

(A) Schematic of the *Igf2r* cluster with the 3C analyzed interaction frequency between *Slc22a3* promoter region and the *Airn* gene, both shown in black boxes.

(B) Taqman qPCR results of an E12.5 T^{hp}/AirnT cross samples to analyze the interaction levels between the bait primer (*Slc22a2* promoter) and the prey primers (*Airn* gene) by 3C. Top: embryo trunk (biallelic: +/+ vs +/AirnT), middle: placenta (biallelic: +/+ vs +/AirnT and paternal: T^{hp}/+ vs T^{hp}/AirnT), bottom: visceral yolk sac (VYS) (biallelic: +/+ vs +/AirnT and paternal: T^{hp}/+ vs T^{hp}/AirnT). The distance between bait and prey primers is blotted on the x-axis. The relative interaction frequency on the y-axis was normalized to the mean of two *H19* ICE interactions. In biallelic comparisons, the black (+/+) sample was compared with the grey (+/AirnT), and in paternal allele comparisons the dark blue (T^{hp}/+) was compared with the light blue (T^{hp}/AirnT). Light grey boxes in the background indicate the different sized EcoRI fragments. The data point was set to the middle of the fragment according to the UCSC coordinates (NCBI37/mm9). Error bars represent the standard deviation of three technical replicates for each sample. Below the 3C results a screenshot of the UCSC genome browser of the *Airn* gene region is aligned with the results.

In the VYS a general increase in interactions was seen following *Airn* truncation in the *Airn* gene body upstream of the *Igf2r* promoter in both the biallelic and paternal allele comparisons. In the biallelic comparison increases were seen in an area between 262.0 to 321.2kb in the presence of *AirnT* (**Figure 14B, bottom**). Similar to placenta, the paternal 3C results showed a lower signal the biallelic samples, but three of the regions identified as increased in the *AirnT* in the biallelic comparison were also seen in the paternal comparison (264.7kb; 304.9kb; 321.2kb). In summary, these results indicate that in extra-embryonic tissues *Airn* blocks interactions on the paternal chromosome between the *Slc22a3* promoter and regions in the *Airn* gene upstream of the *Igf2r* promoter.

4.2.2.3. The *Airn* truncation has no effect on interactions levels with the *Igf2r* promoter

The interaction frequency between regions of the *Airn* gene and *Igf2r* promoter was analyzed. The same embryonic, placental and VYS samples from the T^{hp}/AirnT cross were used as for the *Slc22a2* and *Sla22a3* promoter comparisions. The bait primer was located at the promoter site of *Igf2r* and used in combination with prey primers to an EcoRI site in the *Airn* gene. The analyzed interactions between these genomic regions were pictured as schematic (**Figure 15A**). The embryo trunk showed a high level of interaction at a distance of around 30.3kb away from the *Igf2r* promoter that was not seen in the placenta or VYS (**Figure 15B, top**). However, in the embryo no difference in interactions levels between the wild type (+/+) and the *Airn* truncation allele (+/AirnT) was observed at any point within the *Airn* gene. Similarly, in placenta no difference in interactions levels between the wild type and *Airn* truncation was seen in the biallelic or paternal comparisons (**Figure 15B, middle**). In VYS, the wild type and *AirnT* interaction levels were also very similar, although the biallelic comparison showed a general slightly higher level of interactions on the AirnT allele (**Figure 15B, bottom**). In summary, the *Airn* truncation did not effect interactions with the *Igf2r* promoter in any of the tissues tested.

A

Mat Slc22a3   Slc22a2   Slc22a1   *Igf2r*   *Mas1*
Pat   *Airn*

B

Relative interaction frequency

Biallelic — Embryo Trunk — +/+ , +/AirnT

Biallelic — Placenta — +/+ , +/AirnT

Paternal — Thp/+ , Thp/AirnT

Biallelic — Visceral Yolk Sac — +/+ , +/AirnT

Paternal — Thp/+ , Thp/AirnT

Distance from Bait (kb)

RefSeq Genes

Igf2r   Mas1   Mrgprh
Airn
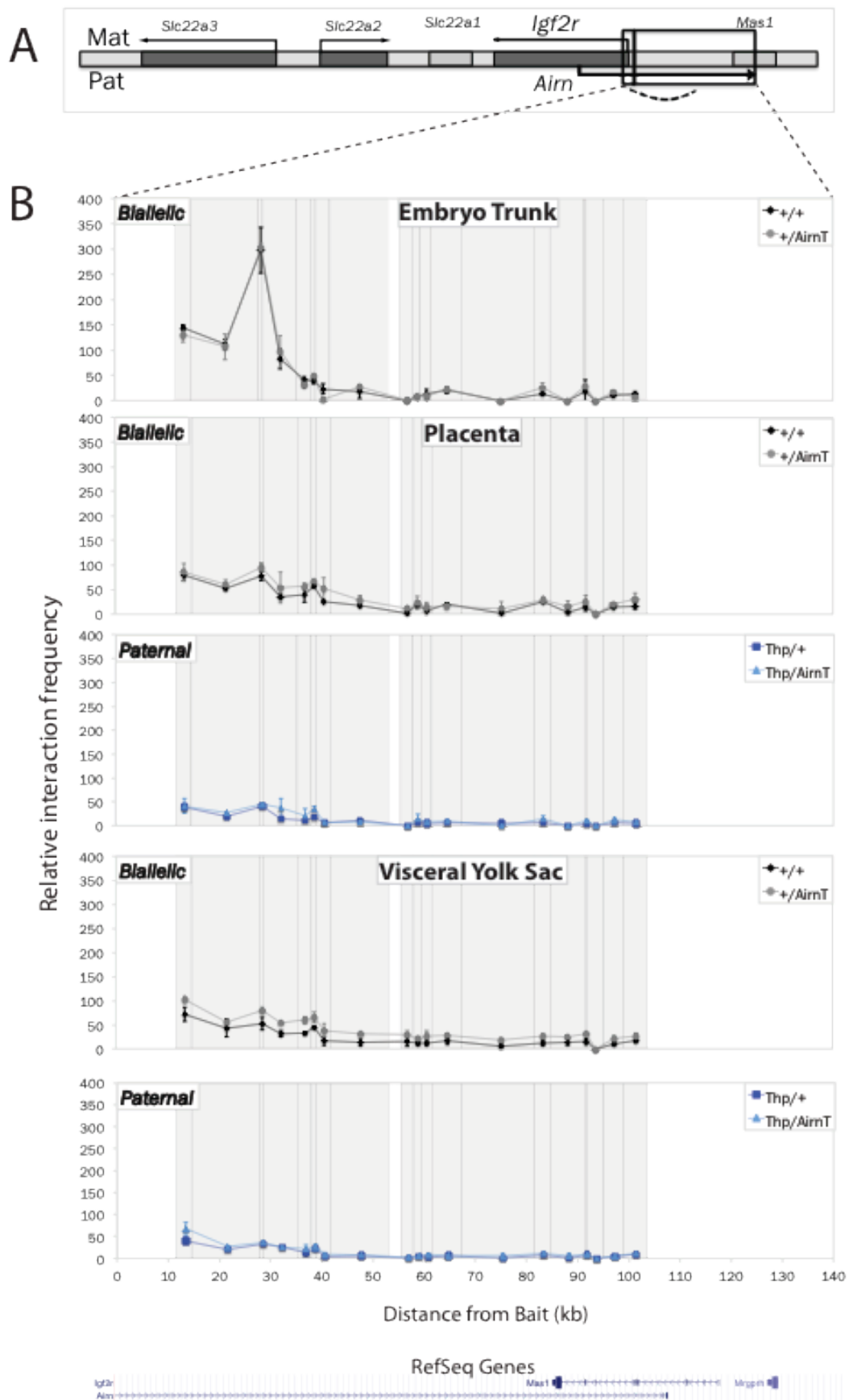
Figure 15: Interactions between the *Igf2r* promoter and the *Airn* gene body are not effected by *Airn* expression

(A) The *Igf2r* cluster, pictured as a schematic, to visualize the analyzed interaction frequency between the promoter region of *Igf2r* (smaller black box) and the *Airn* gene (bigger black box). **(B)** 3C interactions between the *Igf2r* promoter and *Airn* gene body as analyzed by 3C Taqman qPCR assays. Top: embryo trunk (biallelic: +/+ vs +/AirnT), middle: placenta (biallelic: +/+ vs +/AirnT and paternal: T^{hp}/+ vs T^{hp}/AirnT), bottom: visceral yolk sac (VYS) (biallelic: +/+ vs +/AirnT and paternal: T^{hp}/+ vs T^{hp}/AirnT). The distance between bait primer (*Igf2r* promoter) and a set of prey primers (*Airn* gene) was plotted on the x-axis. The y-axis represents the relative interaction frequency normalized to the mean of two *H19* ICE interactions. The data points representing the mean of three technical replicates of each sample and were located in the center of each EcoRI fragment, according to the UCSC genome browser (NCBI37/mm9). Light grey boxes in the background of the blots indicate the size of these fragments. The error bars showed the standard deviation of the three technical replicates. The 3C results are aligned to a screenshot of the UCSC browser, showing the *Airn* gene.

4.2.2.4. Interactions with the *Airn* promoter are largely unaffected by *Airn* expression

The interaction frequencies between the *Airn* promoter and the *Airn* gene body were examined by 3C Taqman qPCR assays using the same embryo trunk, placental and VYS samples from T^{hp}/AirnT as above (**Figure 16A**). In contrast to the previous results described above, which either showed no difference between wild type and *AirnT*, or an increase upon *Airn* truncation, the embryonic samples showed a decrease in interactions following truncation. The signal was decreased between 36.1 to 80.7kb from the *Airn* promoter (**Figure 16B, top**). In placenta, most data points did not show any difference between the biallelic and the paternal comparions (**Figure 16B, middle**). One (T^{hp}/AirnT) data point (67.7kb) in the paternal assay was observed with a higher interaction level compared with the (T^{hp}/+). The following point (74.8kb) showed the opposite picture. Similarly, in VYS there was largely no difference between the wild type and *AirnT* alleles in the biallelic or paternal comparisons (**Figure 16B, bottom**). In both VYS assays one data point (67.7kb) was identified with a higher level of interactions when *Airn* was not present. Similar to other assays, in placenta and VYS, the paternal alleles showed a lower signal level then the biallelic samples.

46
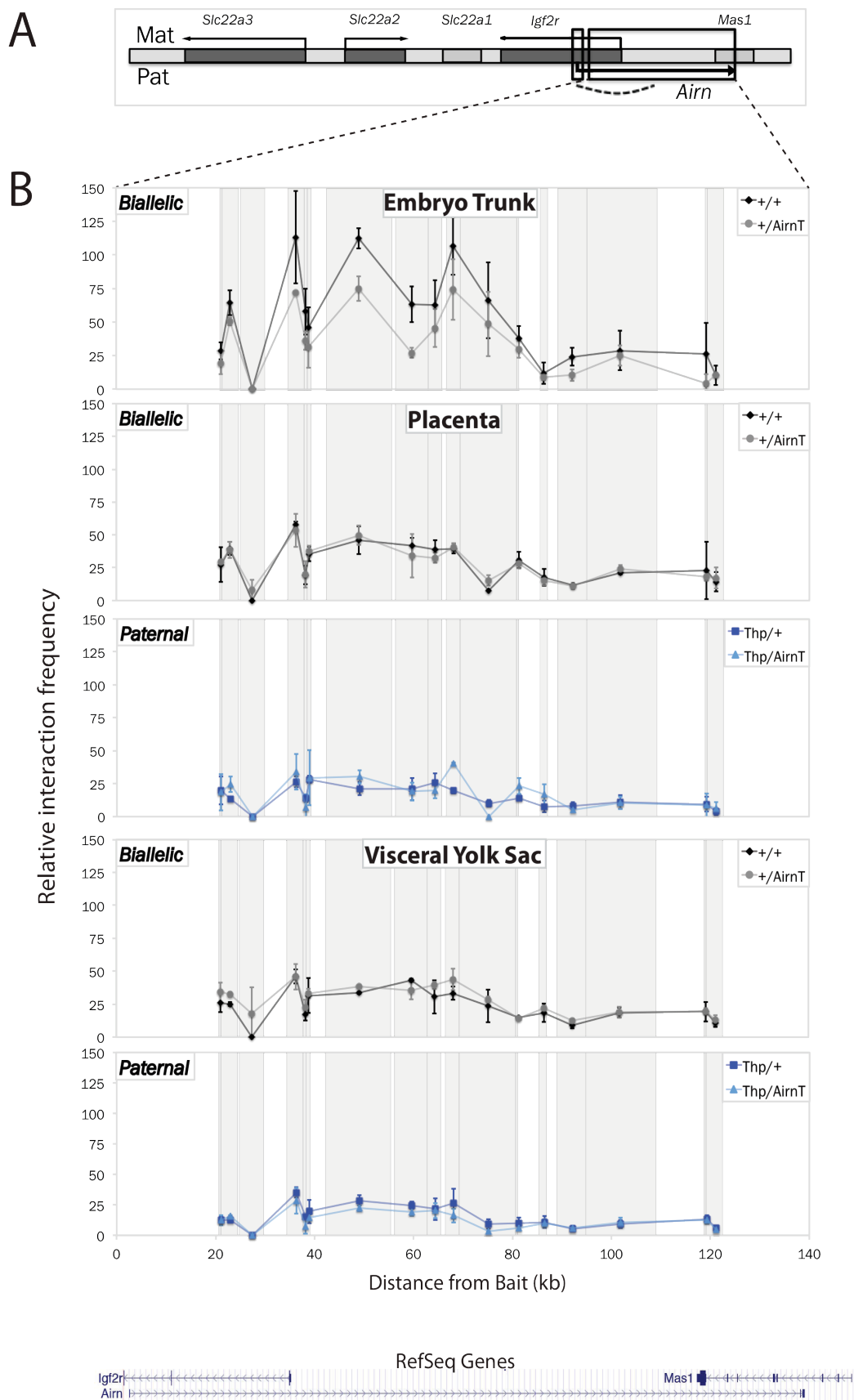
Figure 16: Interactions with the *Airn* promoter are not effected by *Airn* expression in extra-embryonic tissues, but the embryo shows a decrease in interactions at some sites

(A) Schematic of the *Igf2r* cluster in which the interaction between the *Airn* promoter (smaller black box) and *Airn* gene (bigger black box) were shown by a pointed line. The interaction frequency between these regions were assayed

(B) Three tissues of a E12.5 T^{hp} x AirnT cross were used to prepare 3C samples of all four possible genotypes, (+/+) in black; (+/AirnT) in grey; (T^{hp}/+) in dark blue and (T^{hp}/AirnT) in light blue. Top: embryo trunk (biallelic: +/+ vs +/AirnT), middle: placenta (biallelic: +/+ vs +/AirnT and paternal: T^{hp}/+ vs T^{hp}/AirnT), bottom: visceral yolk sac (VYS) (biallelic: +/+ vs +/AirnT and paternal: T^{hp}/+ vs T^{hp}/AirnT). The interaction frequencies, blotted on the y-axis, were assayed by Taqman qPCR and normalized to the mean of two interactions with the *H19* ICE. The x-axis showed the distance between bait (*Airn* promoter) and prey primers (*Airn* gene). A genomic region of 140kb was plotted on this axis. Grey areas in the background of the blots indicate the different sized EcoRI fragments. The mean of three technical replicates is shown with the error bars indicating the standard deviation. The data points were set to the middle of each fragment according to UCSC coordinates (NCBI37/mm9).

4.2.2.5. Interactions between the *Pde10a* promoter and *Airn* gene are too low to analyze

Interactions between the *Pde10a* promoter and the *Airn* gene body were analyzed using 3C Taqman qPCR assays on the same embryonic, placental and VYS samples from T^{hp}/AirnT crosses that were assayed above (**Figure 17A**). In the embryo, a signal was only detected for a limited number of assays, and in most cases not all three replicates showed a signal, resulting in high error bars (**Figure 17B, top**). Similarly, for placenta (**Figure 17B, middle**) and for VYS (**Figure 17B, bottom**) a signal was only detected for a limited number of assays, which showed high error bars. These results indicate that even in placenta where *Pde10a* is expressed the signal is too low to analyze.
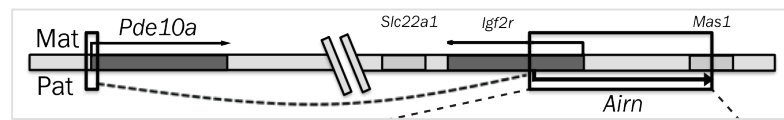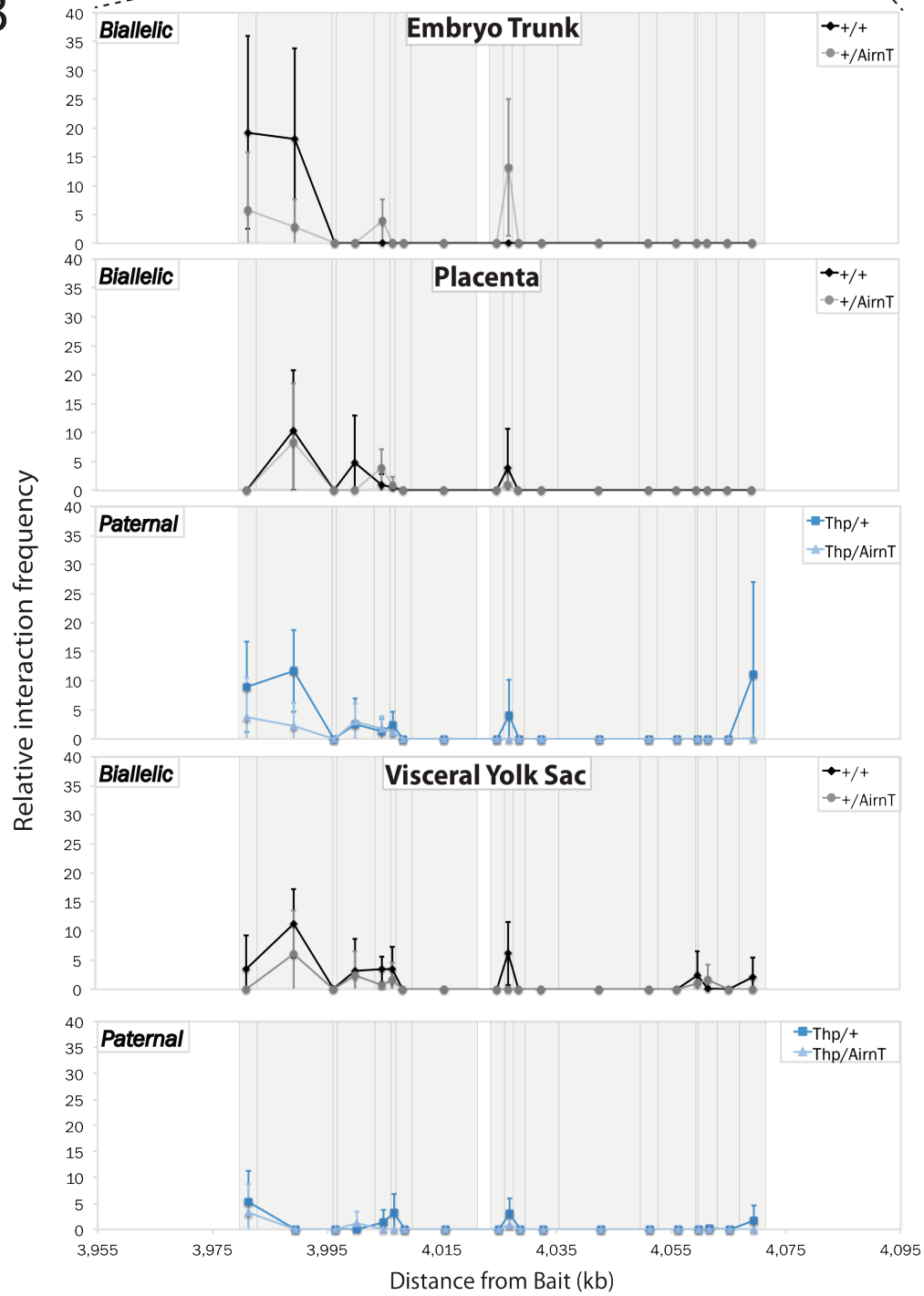
Figure 17: 3C interaction between *Pde10a* promoter and *Airn* could not be validated.

(A) Genes in the *Igf2r* cluster were pictured (*Igf2r, Airn*) together with the 4Mb away genes called *Pde10a*. The 3C assay between the *Pde10a* promoter (smaller black box) and the *Airn* gene body (bigger black box) is indicated in this schematic.

(B) Three tissues of a E12.5 T^{hp} x AirnT cross were used to prepare 3C samples of all four possible genotypes, (+/+) in black; (+/AirnT) in grey; (T^{hp}/+) in dark blue and (T^{hp}/AirnT) in light blue. The biallelic comparison ((+/+) and (+/AirnT)) was done in for three tissues, embryo trunk (top), placenta (middle) and VYS (bottom), while the paternal alleles ((T^{hp}/+) and (T^{hp}/AirnT)) was only done in placenta (middle) and VYS (bottom). The interaction frequencies are plotted on the y-axis, as determined by Taqman qPCR and normalized to the mean of two interactions with the *H19* ICE. The x-axis showed 140kb area of murine chromosome 17, which was scaled to show the distance between *Pde10a* promoter (bait primer) and the in *Airn* gene. The data points were set to the middle of EcoRI fragments in this region according to UCSC genome browser (NCBI37/mm9). Grey areas in the background of the blots indicate the different sized EcoRI fragments. All samples were assayed as technical triplicates and used to calculate the mean and standard deviation.

4.3.　Generation of *Airn* deletion and duplication mice to test the enhancer interference model

The 3C results comparing the *Slc22a2* and *Slc22a3* promoters with the *Airn* gene body indicated that *Airn* expression blocks enhancers for these genes in VYS that lie within the *Airn* gene body (**Figures 13 and 14**). The basic concept was to create a mouse with a deletion or a duplication of the potential enhancer containing DNA region. To do so existing mouse strains were used to perform a chromosome engineering experiment. This technique uses loxP sites from existing mouse strains in conjunction with a ubiquitously expressed Cre to cause *trans* recombination between homologous chromosomes during male meiotic pairing, resulting in a deletion on the one chromosome and a duplication on the other one (Hérault et al., 1998). Constructs used to cause targeted mutations in ES cells by homologous recombination usually include a selection cassette flanked by loxP sites. This cassette is normally removed by transient transfection of Cre recombinase, meaning that mice made from these ES cells contain a single loxP site. We took advantage of this to engineer large deletions and duplications of the *Airn* gene using existing mice with loxP sites, the AirnT, IPΔ and R2Δ mice from our lab (Sleutels et al., 2003; Sleutels et al., 2002; Wutz et al., 2001), and the Sod2**flox/flox** allele obtained from Takahiko Shimizu (Ikegami et al., 2002) together with the ubiquitously expressed Hprt-Cre obtained from Simon Hippenmeyer

(Hippenmeyer et al., 2013; Tang et al., 2002). As described in the following section, by screening the offspring of *AirnT*/IPΔ/Hprt-Cre males we were able to obtain mice with a 28.0kb deletion or 23.5kb duplication of the first third of the *Airn* gene, and by screening the offspring of R2Δ/Sod2/Hprt-Cre males we obtained a 270kb deletion and 266kb duplication including the entire *Airn* gene. My contribution to this project was to design and conduct genotyping assays for these mice, verifying recombination events by Southern blot, PCR and Sanger sequencing (*Microsynth*) as described in the following section.

4.3.1. Validation of loxP site orientations in mouse strains

The Cre-mediated *trans* recombination can only happen, when the required loxP sites are in the same orientation. Therefore, I checked the orientation of all used loxP sites by PCR and followed by Sanger sequencing to confirm the published data. The standard PCR protocol was used to amplify PCR products including the loxP site. Samples were genotyped beforehand by Southern blot to confirm the presence of the mutant allele. The PCR products were subject to Sanger sequencing to identify the genomic region, together with the inserted regions, which cannot be matched to the genome.

The orientation of the loxP site for AirnT was confirmed to be in sense (+) orientation (**Appendix 1**) by forward and reverse sequences using the AirnT_LoxP_F (**Appendix 2**) and AirnT_LoxP_R primers (**Appendix 3**). The full sequence of the AirnT allele was not available complicating sequencing of the loxP site and later verification of the 28kb deletion and 23.5kb duplication allele by Southern blot, PCR and Sanger sequencing. Therefore, I used the primer pair (AirnT_Poly-A_F: ATTCTCCTATGGGGATCCTG; AirnT-Poly-A_R: GAGGGAATAATTGGTT GGGATC) to PCR amplify the entire Poly-A cassette and then to sequence from the outside. The resulted sequence was used to design primer (AirnT_PolyA_R_OUT: GAATATTTCCACGCCAGCCA; AirnT_PolyA_F_OUT: AAGCCCCTTGAGCATCTGAC) for sequencing from the inside out to the genomic sequence. With this strategy the entire Poly-A-cassette of the AirnT allele was sequenced (**Appendix 4**) allowing the sequence of the deletion and duplication alleles to be inferred (including

restriction enzyme sites important for Southern blot assays) and enabling PCR and sequencing primers to be designed in the poly-A cassette for verifying the deletion and duplication alleles.

The sense (+) orientation of the loxP site (**Appendix 5**) in the IPΔ allele was confirmed by sequencing with the IPΔ_loxP_R primer confirming the published orientation (**Appendix 6**) (Sleutels et al., 2003). Sequence from the IPΔ_loxP_F primer stopped at an adenine rich sequence, which is located before the loxP site meaning that the orientation could not be confirmed by sequencing from the other side of the loxP site (**Appendix 7**).

The orientation of the R2Δ loxP site was confirmed to be in the antisense orientation (**Appendix 1**) (-) by forward and reverse sequencing using the R2Δ_loxP_F (**Appendix 8**) and R2Δ_ loxP_R primers (**Appendix 9**).

Genomic DNA from Sod2$^{flox/flox}$ mice was used to confirm the orientation of one of the two inserted loxP sites in this strain. For the loxP1 the primers Sod2_loxP_F_1 and Sod2_loxP_F_2 were used, but no PCR product was amplified. The loxP2 site was shown to be in the antisense (-) orientation by forward and reverse sequencing using the Sod2_loxP_F_2P (**Appendix 10**) and Sod2_loxP_R_2P (**Appendix 11**) primers. Although the orientation of only one of the Sod2$^{flox/flox}$ loxP sites was confirmed, the other loxP site can be assumed to be in the same orientation as after one generation of breeding Southern blotting indicated the floxed allele had been deleted leaving one loxP site.

4.3.2. The mouse breeding plan to create the new strains

The breeding strategies to establish these new strains was made by Quanah J. Hudson (**Figure 18**). The first step was to bring two modified alleles with loxP sites in the same orientation together with a ubiquitously expressed Cre in one male mouse. The X-chromosome linked Hprt-Cre was used as the ubiquitously expressed Cre. The orientation of the loxP sites was confirmed in section 1.3.2. At least two generation of mouse crossing were needed to bring

all three required alleles into one male mouse. These males were then crossed to wild type (FVB/NJ) females and the offspring was screened for recombination events leading to new alleles by Southern blot and PCR.

AIDel = Deletion between AirnT and IPΔ   RSDel = Deletion between R2Δ and Sod2Δ
AIDup = Duplication between AirnT and IPΔ   RSDup = Duplication between R2Δ and Sod2Δ
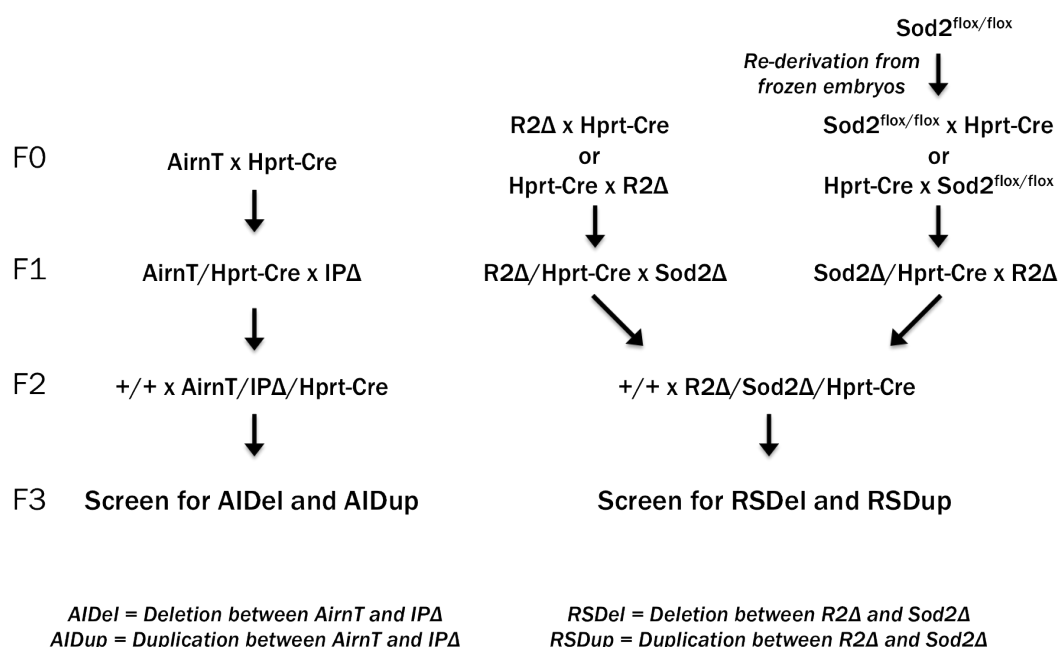
Figure 18: Mouse breeding plan for chromosomal engineering experiment

4.3.3.   Designing genotyping assays for new mouse strains

All four potential new allele were modeled *in silico* to enable Southern blot and PCR genotyping assays to be designed (**Figure 19**). The genomic regions that were modified to model the starting alleles, were downloaded from the UCSC genome browser, version December 2011 (GRCm38/mm10) (http://genome-euro.ucsc.edu/). These genomic sequences were entered into ApE (**A p**lasmid **E**ditor; http://biologylabs.utah.edu/jorgensen/wayned/ape/) and were used as the *in silico* wild type allele. The Sanger sequences of the DNA region surrounding the loxP sites from the starting mouse strains (AirnT, IPΔ, R2Δ, Sod2Δ) were used to define the borders of the deletions (IPΔ, R2Δ, Sod2Δ) and the region, where the poly-A cassette (AirnT) was inserted. In the deletion strains (IPΔ, R2Δ, Sod2Δ), the deleted genomic DNA was replaced by the remaining vector pieces including the loxP site, to construct an *in silico* model of the starting allele. The AirnT allele was re-constructed *in silico* by inserting the sequenced 1.2kb poly-A cassette plus vector and loxP site into the wild type allele. The AirnT and IPΔ alleles were then combined at the loxP

site to construct an *in silico* model of the AirnT/IPΔ deletion (AIDel) and AirnT/IPΔ duplication (AIDup) alleles. A similar step was performed with the R2Δ and Sod2Δ allele to model the R2Δ/Sod2Δ deletion (RSDel) and R2Δ/Sod2Δ duplication (RSDup) alleles. Using these *in silico* models Southern blot assays were designed with a combination of restriction enzyme digests and probes predicted to detect three different sized fragments (wild type, starting strain, deletion or duplication) that could be distinguished on a Southern blot gel. These *in silico* models were then also used to design PCR assays to distinguish wild type and mutant alleles that were also subject to Sanger sequencing to confirm the mutant alleles had the expected sequence at the deletion/duplication boundaries.
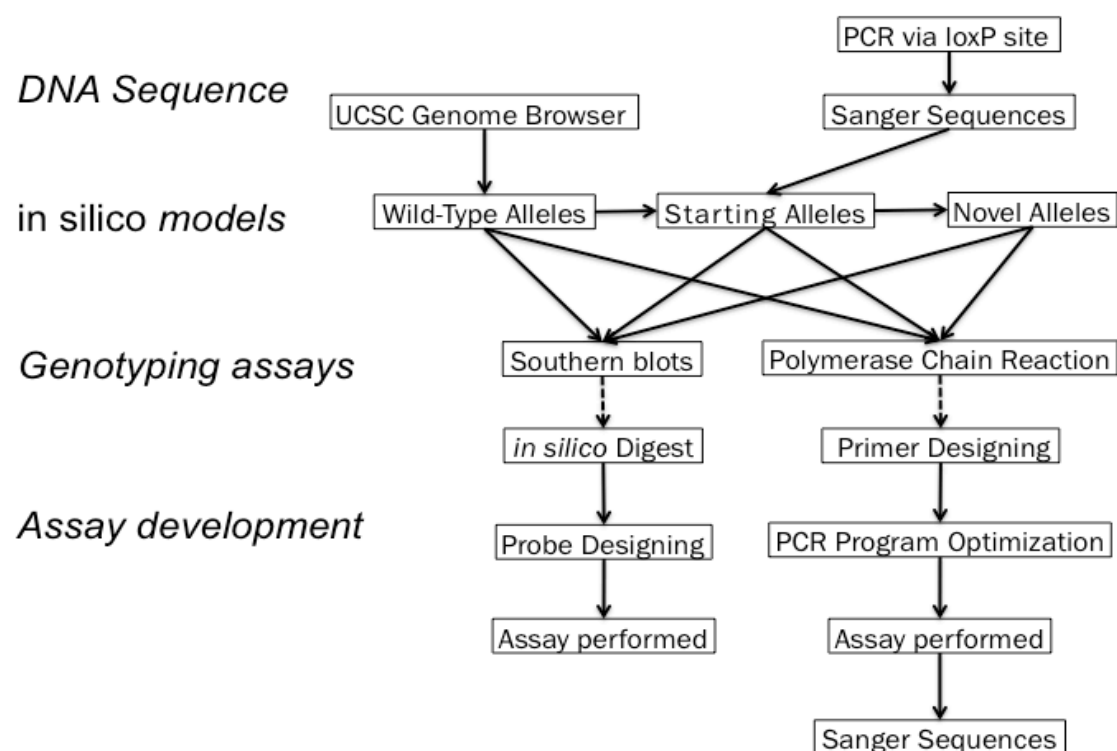
Figure 19: Pipeline to design *in silico* alleles of starting and newly established mouse strains

To establish Southern blot and PCR assays for genotyping the new deletion and duplication alleles the alleles were first constructed *in silico*. The tool ApE was used for this approach (A plasmid Editor; http://biologylabs.utah.edu/jorgensen/wayned/ape/). The wild type alleles were constructed by downloading DNA sequences from the modified genomic regions from the UCSC genome browser (GRCm38/mm10). Predicted deletion and duplication alleles were constructed by combining genomic sequences with construct and loxP site sequences determined by Sanger sequencing (*Microsynth*) around the loxP site. These *in silico* alleles were then used to design new Southern blot and PCR assays for genotyping, which were then tested experimentally.

54

In addition, these *in silico* alleles were also very useful for designing the PCR assays to confirm the Southern blot findings. The presence of restriction enzyme recognition sites has no effect on a PCR reaction, but the presence of the vector pieces have an influence on the predicted length of the amplified product. The remaining vector pieces have a length between 92 and 107 bp including the loxP site. I took advantage of the sequence and size difference between the wild type and mutant to design primers that would amplify different sized fragments distinguishable on a gel. After the designing of the primers was done, the next step was the optimization of the PCR protocol by gradient PCR.

4.3.4. Development of PCR assays

For confirmation of Southern blot results, the *in silico* models were used to design primer pairs for PCR genotyping (**Figure 19**). The basic concept used a common primer in combination with a wild type primer and a deletion/duplication primer to result in different sized PCR fragments. All used primers were designed by primer-blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) with very similar characteristics to find unique primer pairs for the DNA sequence of the *in silico* alleles. In addition, the presence of three primers increases the possibility of dimer primer formation, which can interfere with PCR reactions. Therefore a tool to predict primer-dimer formation was used (http://www.thermoscientificbio.com /webtools/multipleprimer/). The next step was to optimize the standard PCR protocol used in the lab before (for details see 4.2.2.), for the requirements of the new assays. In the PCR reaction mix a higher amount of primers was used (addition third primer) and less water was used to maintain the same total volume. The $MgCl_2$ as well as the Betaine amounts stayed the same. Double the amount of the common primer was used for these PCRs, because these primers were able to interact with both, the wild type and the mutant band primer. In most cases the cycling temperatures and times of PCR programs were unchanged. For assays with bands larger than 500 bp the elongation time was set to 1min. Gradient PCRs were performed to find the optimal annealing temperature for each assay. In

theory the primers should be stable up to 64°C or even higher for specific interactions. Therefore, the gradient PCR was set to analyze annealing temperatures from 53 to 64°C. For the AlDel PCR genotyping assay the optimal temperature was set to 64°C and for AlDup to 62°C. Both assays seemed to work also with lower annealing temperatures without showing unspecific bands. The 266kb duplication (RSDup) genotyping also used an annealing temperature of 64°C. For this assay lower temperatures led to an additional band at around 900 bp. For the 270kb deletion (RSDel) the $MgCl_2$ (25mM) concentration was reduced from 1.5µl to 1.25µl per reaction. The annealing temperature was set to 60°C. Primers were named as *Mus musculus* genotyping primers (MM_typing).

AlDel typing was done with three primers (MM_typing_001; MM_typing_004; MM_typing_007). The common primer MM_typing_001 (antisense -) plus the AlDel specific primer MM_typing_004 (sense +) amplifies a 298 bp AlDel band, while MM_typing_001 plus MM_typing_007 (sense +) amplifies the 557 bp wild type band (**Figure 20A**).

AlDup typing assay was also done with three primers (MM_typing_003_2, MM_typing_005_2, MM_typing_006_2). The 347 bp AlDup band was amplified by the common primer MM_typing_003_2 (antisense -) plus the AlDup primer MM_typing_006_2 (sense +), while the wild type band of 467 bp was amplified by the common primer MM_typing_003_2 (antisense -) plus the wild type primer MM_typing_005_2 (sense +) (**Figure 20B**).

The three primers used for RSDel genotyping were MM_typing_009_2, MM_typing_010 and MM_typing_011_2. The 550 bp RSDel fragment was PCR amplified by the common primer MM_typing_009_2 (antisense -) plus the RSDel primer MM_typing_011_2 (sense +). The 430 bp wild type fragment was amplified by the common primer MM_typing_009_2 (antisense -) plus the wild type primer MM_typing_010 (sense +) (**Figure 20C**).

For the RSDup genotyping three new primers were used (MM_typing_012, MM_typing_013, MM_typing_014). The 644 bp RSDup band was PCR

amplified by the common primer MM_typing_013 (antisense -) and the RSDup specific primer MM_typing_012 (sense +). The 436 bp wild type control band was PCR amplified by the common primer MM_typing_013 (antisense -) and the wild type primer MM_typing_014 (antisense -) (**Figure 20D**).
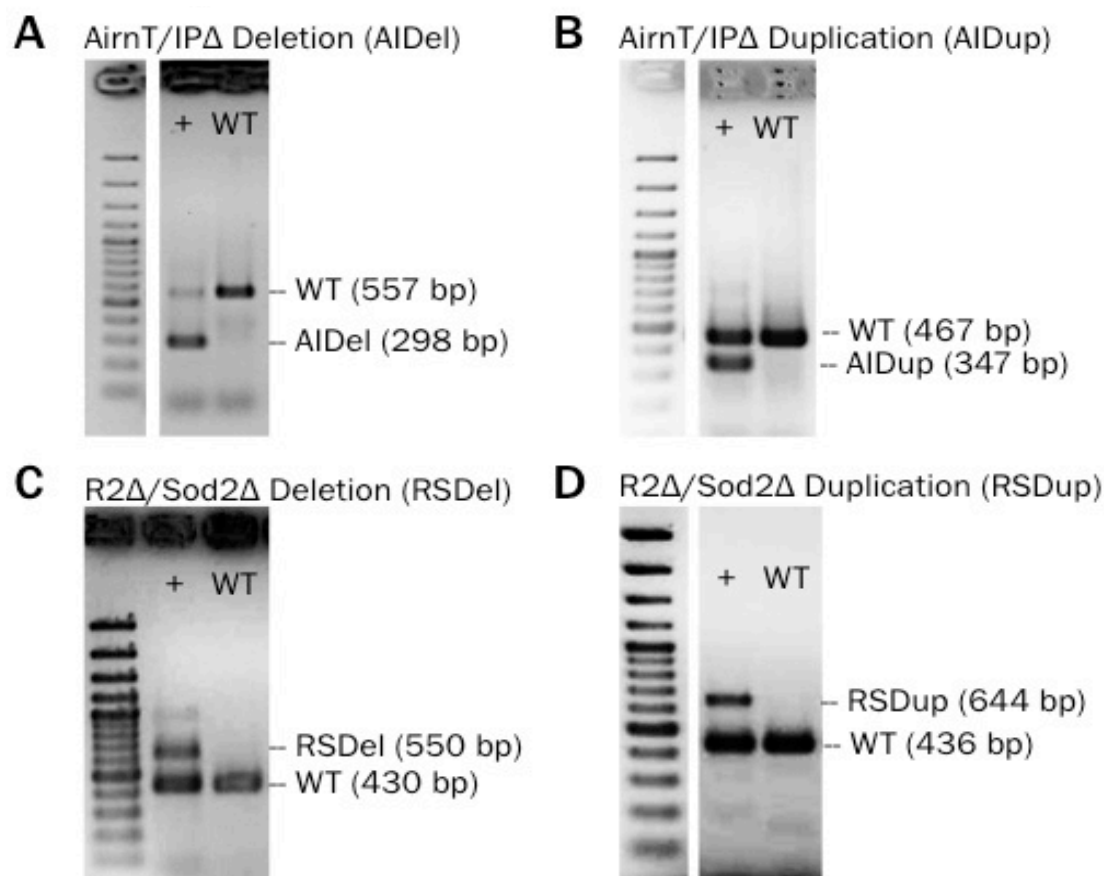
Figure 20: PCR genotyping assays for new mouse strains

(A) The deletion between AirnT and IPΔ (AIDel) is detected by PCR primers MM_typing_001 and MM_typing_004 resulting in a 298 bp band. The 557 bp wild type band is detected by primers MM_typing_001 and MM_typing_007 and is visible in both samples.

(B) The specific band for the duplication between AirnT and IPΔ (AIDup) could be detected at 347 bp with MM_Typing_003_2 and MM_Typing_006_2. In both samples MM_Typing_005_2 and MM_Typing_006_2 resulted in the wild type band at 467 bp.

(C) The RSDel allele showed a specific PCR fragment between MM_Typing_009_2 and MM_Typing_011_2 at 550 bp and the 430 bp control band could be detected in both samples with the primers MM_Typing_009_2 and MM_Typing_010.

(D) The duplication between R2Δ and Sod2Δ (RSDup) was detected by MM_Typing_012 and MM_Typing_013 at 644 bp. Both, the wild type control and the positive sample, showed the control band at 436 bp with the primer pair MM_Typing_013 and MM_Typing_014.

The PCR protocol was used as described in the 'Materials and Methods' part 4.2.2. All samples were diluted (1:100) from genomic DNA sample and were run in 2% agarose stained with ethidium bromide and photographed under UV light. The positive control was beforehand positively genotyped by Southern blot. The wild type control was the genomic DNA of a DR4 mouse. The ladder was GeneRulerTM 100 bp Plus DNA Ladder (Thermo Scientific).

The optimal conditions for the new PCR assays were tested and optimized by gradient PCRs. The same PCR reaction was amplified eight times with different annealing temperatures between 53°C to 64°C (53.0°C; 53.3°C; 54.0°C; 54.8°C; 56.1°C; 57.7°C; 59.6°C; 61.2°C; 62.3°C; 63.2°C; 63.8°C; 64.0°C). The example for RSDup PCR shows that at lower annealing temperature an unspecific band at around 900 bp was detected (**Figure 21**). At a temperature of 63.8°C or higher this unspecific band was gone and only the specific RSDup band (644 bp) and wild type (WT) (436 bp) remained.
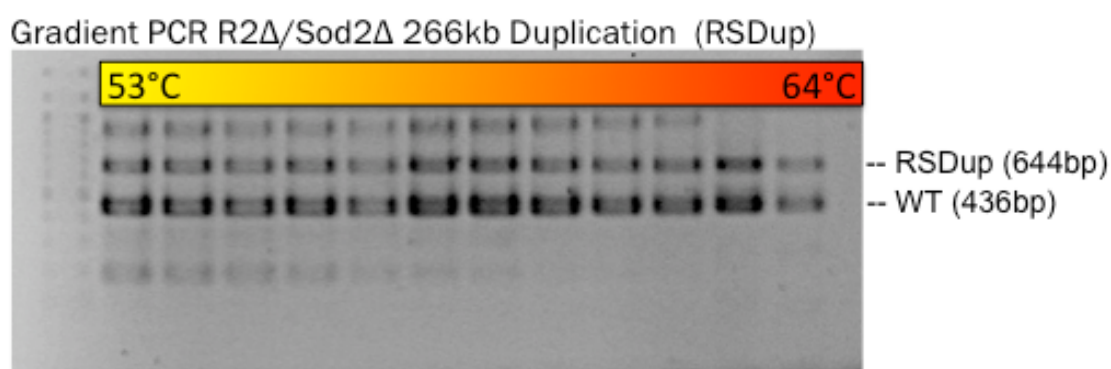
Figure 21: Example of gradient PCR (for RSDup) to optimize PCR protocols

The standard PCR protocol described in the methods was used to prepare this reaction. The DNA template was previously genotyped as RSDup by Southern blot. The standard cycling program was used, only different annealing temperatures between 53°C to 64°C were used to optimize the protocol. The gel electrophoresis running time was 1.5h at 120V for a 400ml of 2% agarose gel and stained with ethidium bromide and photographed under UV light. GeneRulerTM 100 bp Plus DNA Ladder (Thermo Scientific) was used.

4.3.5. Development of Southern blot assays

The *in silico* alleles were used to design Southern blot assays for the newly created alleles (**Figure 19**). The first concept was to test assays already established for the starting strains, or the same restriction enzyme with a new probe. The AIDel Southern blot used the restriction enzyme BglII (A$^\vee$GATCT) together with a new DNA probe (AIDel: chr17:12,772,574-12,773,123; GRCm38/mm10) (**Appendix 12**) located upstream of the IPΔ loxP site to distinguish the three different genotypes (wild type band 9.6kb; IPΔ band 4.8kb; AIDel band 1.0kb). Due to the small size of the AIDel band, the gel electrophorese running time was less than standard with 3h at 120V.

To establish the Southern blot protocol for AIDup all four *in silico* alleles (wild type, two staring strains, duplication) were used for *in silico* digests with different restriction enzymes. The restriction enzyme AanI (PsiI) (TTA$^\vee$TAA) was used together the AIDup probe (chr17:12,746,073-12,746,516; GRCm38/mm10) (**Appendix 13**) to distinguish the different sized fragments after a gel run of 5h (at 120V) (wild type band 7.8kb; AirnT band 3.3kb; AIDup band 4.7kb). The duplication allele contains two binding sites of this probe, one in the wild type allele and a second one in the duplicated allele next to the loxP site. Alternatively, a digest with EcoRI (G$^\vee$AATTC) or PstI (CTGCA$^\vee$G) using the same probe with a gel run of 5h (at 120V), can be used to distinguish between the wild type and AIDup allele (EcoRI: wild type band 6.2kb; AirnT band 2.1kb; AIDup band 2.2kb; or PstI: wild type band 3.8kb; AirnT band 2.4kb; AIDup band 2.4kb). These last 2 assays can only distinguish the wild type and mutant allele, but could be used once the AIDup line is established and no other mutant alleles are present.

The same restriction enzyme (BglII; A$^\vee$GATCT) was used to genotype for the RSDel allele using a new DNA probe located upstream of the R2Δ loxP site that was selected to be repeat free (RSDel: chr17:12,740,089-12,740,420; GRCm38/mm10) (**Appendix 14**). This assay detected 3 clearly separated fragments (wild type band 9.6kb; R2Δ band 6.1kb; RSDel band 2.1kb) The running time on agarose gel was set to 4h (at 120V), which are the standard conditions used for most Southern blot assays in the lab.

For the RSDup Southern blot assay it was possible to use an established assay for genotyping (R2Δ genotyping assay) that also detected the duplicated allele. The restriction enzyme BglII (A$^\vee$GATCT) was used for the digest and the established DNA probe OT KpnI (chr17:12,744,659-12,745,692; GRCm38/mm10) (**Appendix 15**) allowed all three possible alleles to be distinguished by Southern blot (wild type band 9.6kb; R2Δ band 6.1kb; RSDup band 5.5kb). The running time of gel electrophoresis had to be increased from standard conditions to an o/n at 30V followed by 4h at 120V to enable the R2Δ and RSDup bands to be clearly distinguished. The duplicated allele contains two binding sites of this probe, one in the wild type cluster and

a second one next to the loxP site. Consequently, using this assay it is not possible to distinguish between the RSDup allele with a wild type allele (heterozygous) or with a second RSDup allele (homozygous).

Mice with a deletion or duplication allele, which were identified by Southern blot and PCR assay, were used to perform an additional standard PCR with the common primer and the deletion/duplication primer, to amplify only the deletion/duplication band. After a quality control check on an agarose gel, the band was cut out and used for purification. Alternatively, a remaining PCR reaction was purified. The purified PCR fragment was then sent for Sanger sequencing to confirm previous genotyping results by comparing the Sanger sequence to the predicted *in silico* allele.

4.3.6.   AirnT/IPΔ/Hprt-Cre breeding strategy

For this chromosome engineering experiment the AirnT and IPΔ mice were used from the Barlow lab mouse stock, together with the ubiquitously expressed Hprt-Cre. Two males with the X-linked Hprt-Cre and homozygous for MADM7 (6 months old) were received from Simon Hippenmeier (IST Austria) and used to set up these crosses as well as to maintain the strain (Hippenmeyer et al., 2013). The maternal inheritance of the IPΔ allele is embryonic lethal and had to be avoided during breeding to establish the AirnT/IPΔ deletion/duplication.

All crosses were set according to the breeding plan (**Figure 18**). The offspring (F1) of an AirnT x Hprt-Cre (F0) cross were screened for the AirnT allele by Southern blot and for Hprt-Cre by PCR using established assays (http://jaxmice.jax.org/protocolsdb/f?p=116:2:::NO:2:P2_MASTER_PROTOC OL_ID,P2_JRS_CODE:1987,006143). 26 of 43 (60.5%) mice were positive for the AirnT allele and 16 of 43 (37.2%) for the Hprt-Cre allele. 13 of 43 (30.2%) mice (F1) were positive for both alleles. All mice of the F1 generation were genotyped as heterozygous for the MADM7 allele by PCR using an established assay (Hippenmeyer et al., 2010). Double mutant (AirnT/Hprt-Cre) females were set with IPΔ males (**Figure 18**) according to the standard

breeding procedure where two females are set with one male for at least two weeks. In total eight females were used to set up this cross. 11 of 60 (18.3%) mice were positively genotyped for all three modified alleles (AirnT/IPΔ/Hprt-Cre). These mice were genotyped for AirnT and IPΔ alleles by Southern blot and for Hprt-Cre by PCR. In addition, all mice were screened for MADM7 by PCR. 7 of 11 (63.6%) triple mutants were identified as heterozygous MADM7 and in the remaining 4 of 11 (36.4%) mice no MADM7 allele was found. 7 of 11 (63.6%) triple mutants were identified as males (F2). Out of these males, 2 individuals without the MADM7 allele were used to set up the next round of crosses (F3). In addition, two MADM7 heterozygous male were used to set up in total four triple mutant males with two FVB/NJ females each. The males were not removed after two weeks from the breeding cage to speed up the breeding process and to increase the total number of mice for the screens.

Summary of the AirnT/IPΔ/Hprt-Cre screen

	Total Number	Screened Mice by		Mice of Interest		Ratio
		Southern blot	PCR	Male	Female	
AIDel (28.0kb Deletion)	263	152 57.8%	263 100.0%	2	1	1.1%
AIDup (23.5kb Duplication)	263	152 57.8%	263 100.0%	1	1	0.8%

Table 1: Summary of the AirnT/IPΔ/Hprt-Cre screen

263 of 263 (100.0%) mice were genotyped by PCR and 152 of 263 (57.8%) by Southern blot for both alleles. For the 28.0kb deletion 2 males and 1 female were identified. For the second new allele, the 23.5kb duplication, 1 male and 1 female were found. The recombination frequency for the AIDel allele was 1.1% and for the AIDup 0.8%.

The F3 generation was screened for the 28.0kb deletion (AIDel) and 23.5kb duplication (AIDup) between the AirnT and IPΔ loxP sites by newly designed assays protocols described in section 4.3.6 (**Table 1**). 152 of 263 (57.8%) mice were screened for deletion or duplication alleles by Southern blot and 263 of 263 (100.0%). 3 of 263 mice (1.14%), 2 males and 1 female, were positively genotyped for a 28 kb deletion (new F0 of AIDel strain) between the loxP sites of AirnT and IPΔ by Southern blot and PCR. In 2 of 263 mice (0.76%), 1 male and 1 female, the 23.5 kb duplication allele was identified with both assays

(new F0 of AIDup strain). The PCR products of the identified novel alleles were sent for Sanger sequencing to confirm these findings.

4.3.6.1. AirnT/IPΔ/Hprt-Cre genotyping results
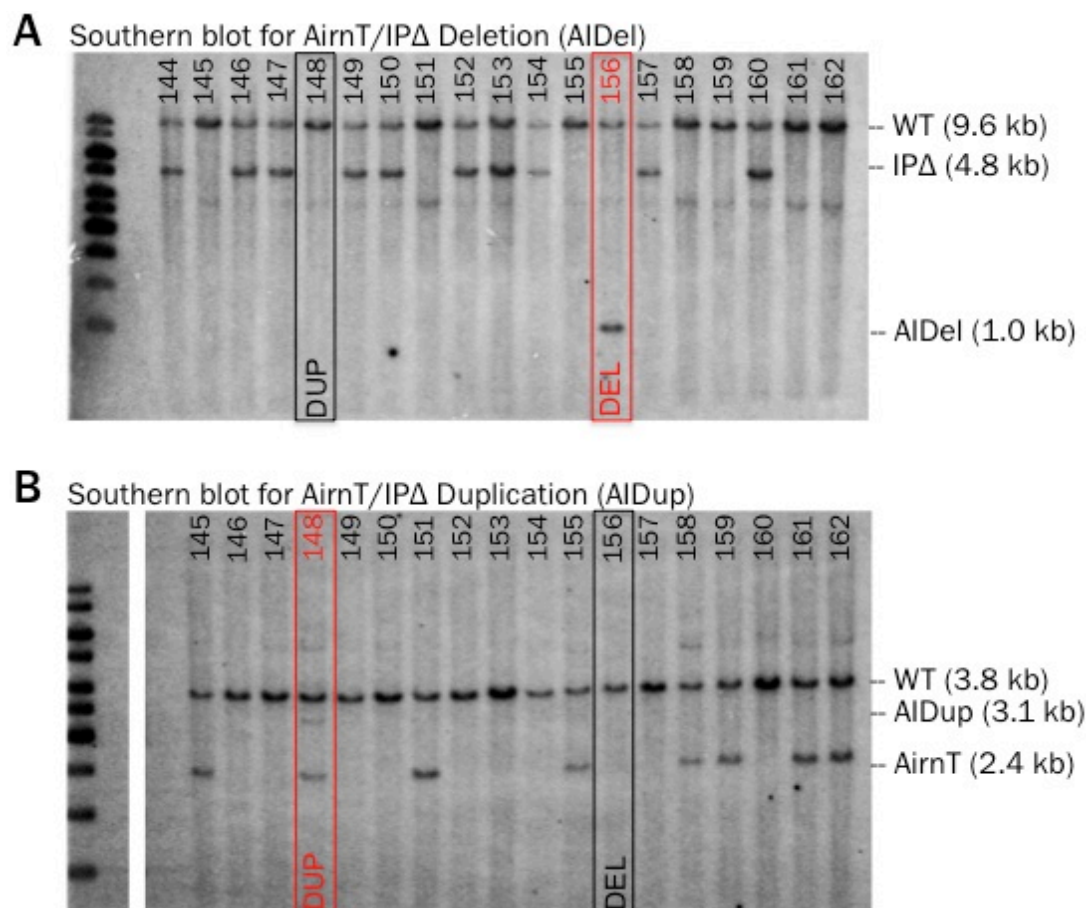
Figure 22: Example of a Southern blot genotype assay for AIDel and AIDup (from founder screen)

(A) BglII digest of genomic DNA was loaded on 0.8% agarose in 1xTBE (400ml) and run for 3h at 120V. The gel was blotted onto a membrane o/n and hybridised with a P32*-labeled AIDel probe. The control wild type band (WT) was present in all samples (19/19; 100.0%) with a fragment size of 9.6 kb. 10 of 19 (52.6%) samples showed an IPΔ band at 4.8 kb. In sample number 156 a deletion band (AIDel) at 1.0 kb was detected. This mutant allele was later confirmed by PCR and Sanger sequencing. Number 148 was identified as AIDup mouse (B), but showed in this assay only the wild type band. Ladder was GeneRuler[TM] 1kb DNA Ladder (Thermo Scientific).

(B) PstI digest of genomic DNA was loaded on 0.8% agarose in 1xTBE (400ml) and run for 5h. The gel was blotted onto a membrane o/n and hybridized with a P32*-labeled AIDup probe. 18 of 18 samples (100.0%) showed the wild type band at 3.8 kb. 7 of 18 (38.9%) of these samples were genotyped for the AirnT allele by showing a band at 2.4 kb. In sample number 148 a weak duplication band of 3.1 kb was detected as well as a second duplication band of 2.4 kb size, that is the same as for the Airn allele. The PCR and Sanger sequencing results confirmed this mouse as an AIDup female. The AIDel allele (number 156) displayed only a wild type band in this assay and no unspecific bands were detected. 30µl of genomic DNA of each sample was used for these assays. Ladder was GeneRuler[TM] 1kb DNA Ladder

(Thermo Scientific). From the FVB/NJ x AirT/IPΔ/Hprt-Cre cross (Tailing date: 17.12.2013) the samples 144 to 162 are shown here.

The newly established Southern blot protocols to screen for AlDel and AlDup alleles protocols described in section 4.3.5 were performed on FVB/NJ x AirnT/IPΔ/Hprt-Cre crosses (**Figure 22**). The wild type band at 9.6kb should be present in all samples, because of the maternal inherited allele. The southern blot for AlDel genotyping resulted in the specific AlDel band at 1.0kb and in addition in an IPΔ specific band at 4.8kb (**Figure 22A**). This assay allowed the AlDel and IPΔ alleles to be distinguished, but the AlDup and AirnT alleles could not be distinguished from wild type. To identify these last two strains the AlDup Southern blot assay had to be performed. The AlDup sample gave a specific band at 3.1kb and the AirnT allele of 2.4kb (**Figure 22B**). The 3.8kb sized wild type band was found in all samples. The AlDel and AlDup assays gave complementary results with samples detected as only wild type in one assay, displaying a mutant allele in the other assay and visa versa.

According to the optimized PCR protocol for each assay (**Figure 20**), all samples of the FVB/NJ x AirnT/IPΔ/Hprt-Cre crosses were genotyped by PCR to confirm Southern blot findings (**Figure 23**). The AlDel assay resulted in a 557 bp wild type band, which was found in all samples, while the specific, 298 bp sized AlDel band was detected in only one sample (**Figure 23A**). These findings were confirmed with both controls. The AlDup PCR assay resulted in a similar picture (**Figure 23B**). The wild type band (467 bp) was found in all samples and only one showed the specific AlDup band (347 bp). In the control samples the expected band pattern was detected.

Samples genotyped for the AlDel allele were sent for Sanger sequencing to confirm the previous findings. For this sequencing the forward primer MM_Typing_004 and the reverse primer MM_Typing_001 were used. The resulting sequences (**Appendix 16**) showed a part of the poly-A-tail of the AirnT allele
(nactcccagtcatagctgtccctcttctcttatggagatccctcgaggtcgacggtatcgataagcttcgag ggaccta) up-stream of the loxP site (**Appendix 1**) followed by a piece of the IPΔ

vector (attaagggttattgaatatgatcggaattcctgcagccc) and finally a piece of the genomic DNA (chr17:12,772,513-12,772,613; GRCm38/mm10) downstream of the IPΔ insertion site. This expected artificial DNA sequence was found for all three mice, genotyped as AIDel.
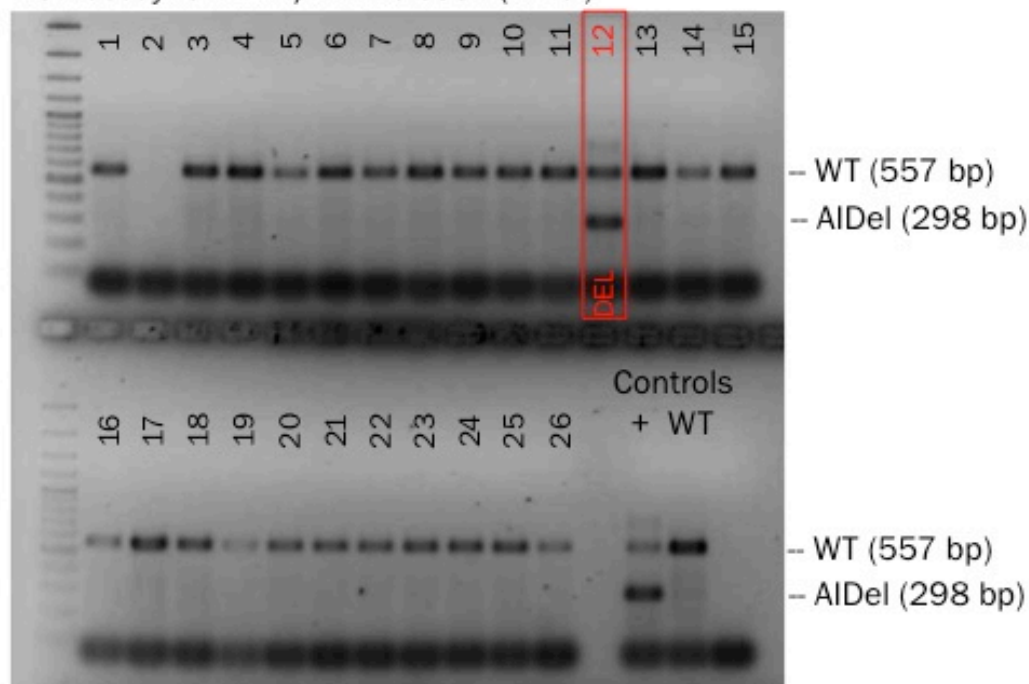
Figure 23: Example of AIDel and AIDup PCR genotyping (from founder screen)

(A) FVB/NJ x AirnT/IPΔ/Hprt-Cre cross (Tailing date: 21.07.2014) (sample 1 -26) genotyped for AIDel allele by PCR. 25 of 26 (96.2%) samples showed the control wild type (WT) band at 557 bp. Sample 2 did not show this band and was repeated. The AIDel band of 298 bp was detected only in sample 12 (1 of 26; 3.8%). No additional, unspecific bands were found in this assay. The positive control (+) showed the double band pattern (298 bp + 557 bp), while the wild type control (WT) only amplified the 557 bp DNA fragment. All samples were diluted (1:100) from genomic DNA sample. The positive and wild type control were beforehand genotyped by Southern blot for the required allele. The ladder was GeneRuler[TM] 100 bp Plus DNA Ladder (Thermo Scientific).
(B) PCR assay to genotype FVB/NJ x AirnT/IPΔ/Hprt-Cre cross (Tailing date: 18.08.2014) (sample 46 -59) for AIDel allele. In 14 of 14 (100.0%) samples the wild type control band at 467 bp was detected. 1 of 14 samples (7.1%) amplified the AIDup band with a size of 347 bp (sample 55) as well as the control band. Both controls, the positive control sample (+) and the

wild type sample (WT), showed the expected band pattern. All samples were diluted (1:100) from genomic DNA sample and were run in 2% agarose stained with ethidium bromide and photographed under UV light. Both control were beforehand positively genotyped by Southern blot. The ladder was GeneRuler[TM] 100 bp Plus DNA Ladder (Thermo Scientific).

This sequencing was done for the AIDup allele with forward MM_Typing_006_2 and reverse MM_Typing_003_2 primers (**Appendix 17**). The genomic DNA upstream of the IPΔ insertion (chr17:12,767,672-12,767,833; GRCm38/mm10) was followed by a short piece of the IPΔ vector (taagcttcgagggaccta) until the loxP site (**Appendix 1**) was detected. After that, the end of the poly-A site of the AirnT insertion was found (tatattaagggttattgaa tatgatcggaattcctgcagcccggg) until again a part of chromosome 17 was detected (chr17:12,744,358-12,744,394; GRCm38/mm10), which normally was located around 23.5kb upstream of the IPΔ insertion site. This indicated that the region between the loxP sites of AirnT and IPΔ was duplicated as expected.

4.3.7. R2Δ/Sod2Δ/Hprt-Cre breeding strategy

This cross was set up with R2Δ mice, from the Barlow lab mouse stock, together with Sod2Δ and Hprt-Cre. Frozen Sod2$^{flox/flox}$/Alb-Cre embryos were obtained from Takahiko Shimizu in Japan and transplanted into a pseudopregnant female by Hans-Christian Theußl at the IMBA/IMP transgenic mouse facility. This resulted in the birth of six mice (two males and four females). These Sod2$^{flox/flox}$ mice were used for the chromosomal engineering experiment together with the offspring of the original two Hprt-Cre males described previously.

The six Sod2$^{flox/flox}$ animals were genotyped for the Sod2$^{flox/flox}$ allele by Southern blot and for Alb-Cre using the generic Cre PCR assay from the Jackson Lab. Four females and one male were identified as homozygous for Sod2$^{flox/flox}$ allele and negative for Alb-Cre. The remaining male was found to be a Sod2$^{+/flox}$ with a copy of Alb-Cre. This male was not used for the R2Δ/Sod2Δ/Hprt-Cre breeding strategy avoiding the need to have unique genotyping assays to distinguish Alb-Cre and Hprt-Cre, and enabling the generic Cre assay to be used for all subsequent Hprt-Cre genotyping. All four

Sod2$^{flox/flox}$ females (F0) were crossed according to the breeding schemata to Hprt-Cre males (F0) and two Hprt-Cre females (F0) were set to be crossed with a Sod2$^{flox/flox}$ male (F0) (**Figure 18**). 21 of 38 (55.3%) mice of the Sod2$^{flox/flox}$ x Hprt-Cre cross were positive for both Sod2Δ by Southern blot and Hprt-Cre by PCR (all these mice were females as Hprt-Cre is X-linked), showing the flox allele had been efficiently removed by Hprt-Cre. In 17 of 38 (44.7%) males of this cross, the Sod2$^{flox/flox}$ was identified by Southern blot and no Hprt-Cre allele (0 of 38 (0.0%)) was detected. 38 of 38 (100.0%) mice from this cross were heterozygous for the MADM7 allele. Therefore, these female double mutants were not used for further breeding as double mutant females wild type at the MADM7 locus were available (see below). In the Hprt-Cre x Sod2$^{flox/flox}$ cross 13 of 18 (72.2%) of all mice showed the Sod2Δ allele with a distribution between the genders of eleven females to two males. For 5 of 18 (27.8%) mice the Southern blot did not show a definitive result. Consequently, these 5 mice were not used for further breeding. The Hprt-Cre allele was found in 10 of 18 (55.6%) and the MADM7 allele in 6 of 18 (33.3%). In total, seven females and two males, or 9 of 18 (50.0%) mice were identified as double mutants (Sod2Δ/Hprt-Cre). Four of these females were crossed to R2Δ males from the Barlow mouse stock to create triple mutants (R2Δ/Sod2Δ/Hprt-Cre) (F2) (**Figure 17**). All of these females were genotyped as wild type at the MADM7 locus. The two males (Sod2Δ/Hprt-Cre), one of which was MADM7 heterozygous while the other had no MADM7 allele, were crossed with R2Δ/Hprt-Cre females.

In parallel, 6 Hprt-Cre females, MADM7 heterozygous, (F0) were crossed to R2Δ males (F0) to generate R2Δ/Hprt-Cre offspring (F1) (**Figure 18**). 24 of 61 (39.3%) of the F1 generation carried the R2Δ allele, which was identified by Southern blot. 27 of 61 (44.3%) were positively genotyped as Hprt-Cre mice by PCR. The overlap of these two results showed, that 11 of 61 (18.0%) mice (F1) were double mutants (R2Δ/Hprt-Cre). The distribution between the genders was 7 males to 4 females. Two of those four females were genotyped as MADM7 heterozygous, where in the remaining two mice no MADM7 allele was detected by PCR. To generate triple male mutants (R2Δ/Sod2Δ/Hprt-Cre) all four R2Δ/Hprt-Cre females (F1) were crossed to

Sod2Δ/Hprt-Cre males with the standard setup of two females per male. The X-linked Hprt-Cre has to be maternally inherited to become triple mutant males in the next generation.

The Sod2Δ/Hprt-Cre females crossed with R2Δ males gave birth to 44 mice (F2) (**Figure 18**). All mice were tested for Hprt-Cre by PCR and 19 of 44 (43.2%) carried this allele. Only these mice were tested for R2Δ by Southern blot. 9 of 19 (47.4%) were identified as R2Δ mice and further tested for the Sod2Δ allele. 6 of these 9 (66.7%, four males and two females), were identified as triple mutants (R2Δ/Sod2Δ/Hprt-Cre), which were 6 of 44 (13.6%) of offspring from this cross. All mice that could be genotyped, 43 of 44 (97.7%), were identified as MADM7 wild type mice by PCR. The R2Δ/Hprt-Cre crossed with Sod2Δ/Hprt-Cre resulted in 33 mice in the F2 generation. The same genotyping approach as for the Sod2Δ/Hprt-Cre x R2Δ cross was used. 25 of 33 (75,8%) mice with Hprt-Cre allele were tested for the R2Δ allele by Southern blot. 15 of 25 (60.0%) mice were found to carry both modified alleles. For these 15 mice, the Sod2Δ Southern blot was performed to find triple mutants. 5 of 15 mice, three males and two females, were identified as triple mutants (R2Δ/Sod2Δ/Hprt-Cre). Therefore, together the Sod2Δ/Hprt-Cre x R2Δ and R2Δ/Hprt-Cre x Sod2Δ/Hprt-Cre crosses resulted in a total number of seven males and four females that were triple mutants. All males were identified as wild types with the MADM7 assay and as a consequence of the location of the Cre allele on the X-chromosome, all 7 males were hemizygous for Hprt-Cre. Four of these males were used to set up crosses with FVB/NJ females according to the standard crossing plan (**Figure 18**).

The next generation (F3) was screened for 270kb deletion (RSDel) and 266kb duplication (RSDup) alleles by Southern blot and PCR (**Table 2**). In total, 72 of 72 (100.0%) mice of this generation were screened by PCR for the RSDel and RSDup allele. 49 of 72 (68.1%) were analyzed by Southern blot. One male (1 of 72; 1.4%) was found to carry the deletion allele (F0 of RSDel) and two males (2 of 72; 2.8%) the duplication allele (F0 of RSDup). Sanger sequencing confirmed all new established alleles.

Summary of the R2Δ/Sod2Δ/Hprt-Cre screen

	Total Number	Screened Mice by		Mice of Interest		Ratio
		Southern blot	PCR	Male	Female	
RSDel (270kb Deletion)	72	49 68.1%	72 100.0%	1	0	1.4%
RSDup (266kb Duplication)	72	49 68.1%	72 100.0%	2	0	2.8%

Table 2: Summary of the R2Δ/Sod2Δ/Hprt-Cre screen

72 of 72 (100.0%) mice were genotyped by PCR and 49 of 72 (68.1%) by Southern blot. For the 270kb deletion (RSDel) one male was identified, while for the 266kb duplication two male mice were found. The recombination frequency for the RSDel allele was 1.4% and for the RSDup 2.8%.

4.3.7.1. R2Δ/Sod2Δ/Hprt-Cre genotyping results

The offspring of the FVB/NJ x R2Δ/Sod2Δ/Hprt-Cre cross was screened by Southern blots for newly created alleles (**Figure 24**). The RSDel and RSDup blots were performed according to the newly developed protocols described in section 4.3.5. The wild type band (9.6kb) was observed in all samples of the RSDel blot (**Figure 24A**). Furthermore, this assay allowed the R2Δ band (6.1kb) and RSDel band at 2.1kb to be distinguished. Samples with only a wild type band in this assay could carry either the Sod2Δ allele or the RSDup, as this assay did not detect these alleles. Similarly, the Southern blot assay for RSDup, detected the RSDup band at 5.5kb as well as the R2Δ band (6.1kb) and the wild type band (9.6kb) identical to the RSDel assay (**Figure 24B**). This assay could not detect the RSDel and Sod2Δ allele so samples lacking a RSDup or R2Δ allele showed only a wild type band.

A Southern blot for R2Δ/Sod2Δ 270kb Deletion (RSDel)

B Southern blot for R2Δ/Sod2Δ 270kb Duplication (RSDup)
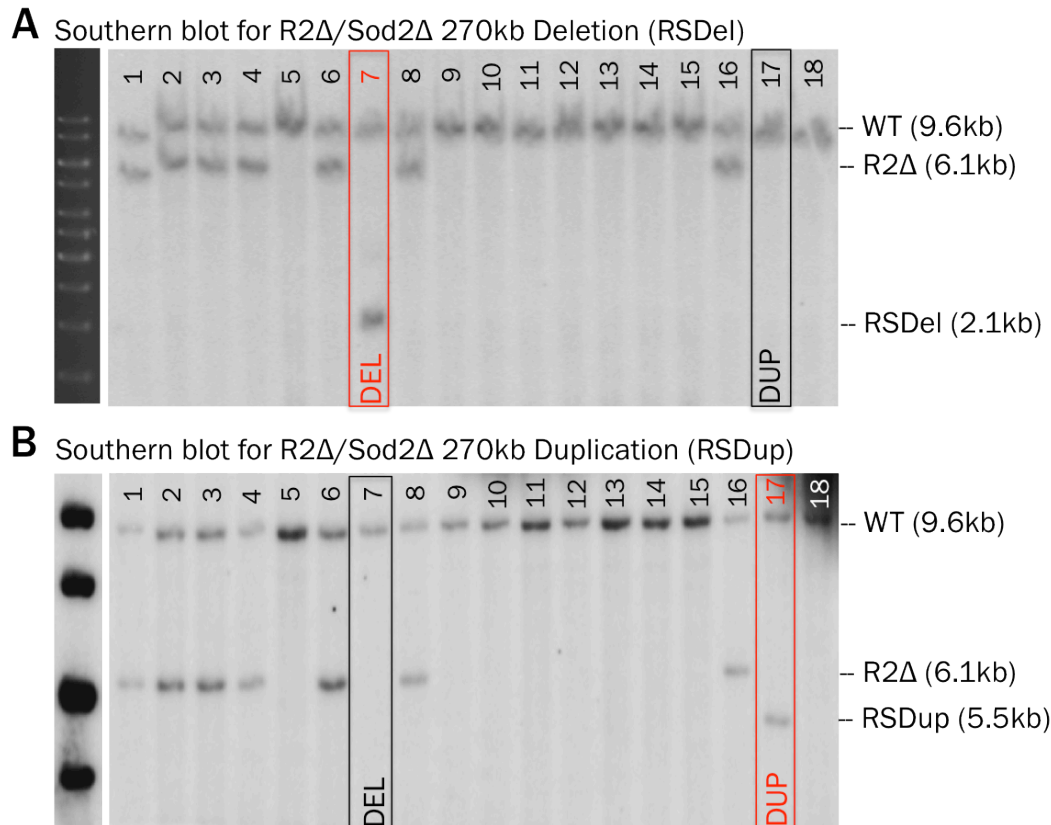
Figure 24: Example of RSDel and RSDup Southern blot assays (from founder screen)

(A) Genomic DNA was digested with BglII and the Southern blot was performed according to the protocol. The RSDel probe was used to detect the DNA fragments. In all samples (19 of 19; 100.0%) the wild type (WT) band was visible with a size of 9.6kb. 7 of 19 (36.8%) samples showed a R2Δ specific band at 6.1kb. 1 of 19 (5.3%) showed the specific band for RSDel (sample 7) at 2.1kb. Sample 17 was identified as RSDup (B) and showed in this assay the expected WT band. The ladder was not stained with this DNA probe, therefore the ladder signal on the gel was used to calculate the size of the resulted bands. No unspecific bands were detected. 30µl of genomic DNA of each sample was used for these assays. Ladder was GeneRulerTM 1kb DNA Ladder (Thermo Scientific).

(B) BglII digest of genomic DNA of 18 samples. The Southern blot was done as described in the protocol with a running time of 3h 16h followed by 5h at 120V. The wild type band (WT) of 9.6kb was visible in all samples (18 of 18; 100.0%), while the R2Δ band (6.1kb) was detected in 7 of 18 samples (36.8%). Sample 17 showed the specific RSDup band at 5.5bk, while in sample 7, the RSDel allele, was detected as wild type with this assay. No unspecific bands were detected. 30µl of genomic DNA of each sample was used for these assays. Ladder was GeneRulerTM 1kb DNA Ladder (Thermo Scientific).

FVB/NJ x R2Δ/Sod2Δ/Hprt-Cre cross (Tailing date: 02.06.14) with samples 1 to 18 are shown here.

All samples of the FVB/NJ x R2Δ/Sod2Δ/Hprt-Cre cross were genotyped by PCR to confirm previous Southern blot results (**Figure 25**). All PCR reactions were performed according to the optimized protocol. The RSDel assay resulted in a wild type band at 430 bp and a specific RSDel band at 550 bp. The positive and wild type control showed both the expected pattern (**Figure**

25A). The RSDup PCRs resulted in a wild type band of 436 bp and a RSDup band of 644 bp (**Figure 25B**). A clear distinction between mutant and wild type allele was possible.

A PCR assay for R2Δ/Sod2Δ 270kb Deletion (RSDel)

B PCR assay for R2Δ/Sod2Δ 270kb Duplication (RSDup)
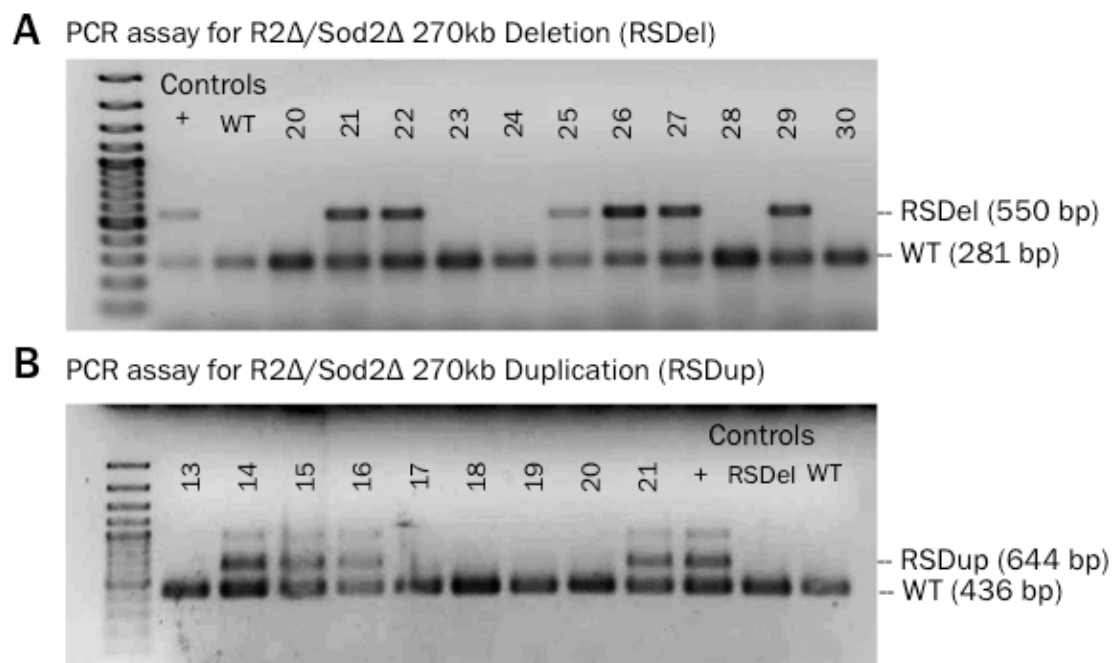
Figure 25: Example of PCR genotyping for the RSDel allele and RSDup allele (from strain maintenance crosses)

(A) RSDel genotyping PCR under standard conditions. DNA samples of mouse embryos were diluted 1:100 in MQH_2O. 11 of 11 (100.0%) samples showed the wild type control band at 281 bp and 6 of 11 (54.5%) were identified as RSDel by the presence of the specific DNA band at 550bp. The two controls were before genotyped as RSDel or wild type mouse and the resulting band pattern confirmed this. The ladder was GeneRulerTM 100 bp Plus DNA Ladder (Thermo Scientific).
(B) RSDup genotyping PCR under standard conditions. Genomic DNA samples were diluted 1:100 in 1xTAE and analyzed. In 9 of 9 (100.0%) cases the wild type (WT) band at 436 bp was detected. 4 of 9 (44.4%) samples showed a specific RSDup band. An additional band with a size between 900 to 1000 bp was detected in samples with the RSDup band. The samples 14, 15, 16 and 21 were confirmed by Southern blot as RSDup mice. All three controls showed the expected band pattern (the positive control sample (+), the RSDel control and the wild type sample (WT)). The wild type control as well as the RSDel control were previously genotyped by Southern blot. The wild type control was the genomic DNA of a DR4 mouse. All samples were run in 2% agarose stained with ethidium bromide and photographed under UV light. The ladder was GeneRulerTM 100 bp Plus DNA Ladder (Thermo Scientific).

The genomic DNA of the one male, which was positively genotyped for RSDel, was used to amplify the specific deletion band, which was then sent for Sanger sequencing. The forward primer MM_Typing_011_2 and the reverse primer MM_Typing_009_2 were used for this approach. The resulting

70

forward 494 bp sequence (**Appendix 18**) identified of a genomic region (377bp) (chr17:12,740,413-12,740,789; GRCm38/mm10) located upstream of the R2Δ deletion site followed by a R2Δ specific vector fragment (tcgaggaattccgatcatattcaataacccttaat) then the loxP site (**Appendix 5**). On the other side of the loxP site the Sod2Δ vector (gtcgccggtatcgatgaattcctgcagccctcga) was found followed up by a short piece of genomic DNA mapping to the Sod2Δ insertion region. With the reverse primer a 490 bp sequence (**Appendix 19**) was identified. The beginning of this sequence could not be linked to the Sod2Δ vector or to an entire loxP site, although one flanking region of this site was found (tatacgaagttat). The entire R2Δ vector site (attaagggttattgaatatgatcggaattcctcga) in addition with a big area of the genomic region (chr17:12,740,367-12,740,789; GRCm38/mm10) was identified in this sequence. The combination of both sequencing results confirmed the finding of Southern blot and PCR assays, and matched the expected sequence of the RSDel deletion.

In parallel, both males positively genotyped for the RSDup allele were processed similarly. The purified duplication PCR band was sent for sequencing with the forward primer MM_Typing_012 and the reverse primer MM_Typing_013. For sample 17 the forward sequencing resulted in a 574 bp piece (**Appendix 20**). The upstream 183 bp long genomic region (chr17:13,010,175-13,010,357; GRCm38/mm10) was mapped in the genome next to the Sod2Δ insertion region and followed by the Sod2Δ specific vector piece (gtcgacaagcttcccgggctcgacgggccccccctcgac). The loxP site (**Appendix 5**) was found right in the middle of this artificial piece of DNA, which ended with a piece of the R2Δ vector (taggtccctcgaanaggttcactagt). The sequence was completed by the downstream sequence of a genomic region (chr17:12,744,359-12,744,650; GRCm38/mm10), which was mapped to the R2Δ site. The reverse primer resulted in a 590 bp fragment (**Appendix 21**), which was bordered by two genomic regions (chr17:12,744,359-12,744,615; chr17:13,010,126-13,010,357; GRCm38/ mm10). This reverse sequence fitted with the forward sequence and together these sequences matched the predicted duplication sequence confirming this mouse as RSDup.

The Sanger sequencing of sample 38, the second sample, which showed specific RSDup bands in both assays, was performed with the same primers. The forward sequence resulted in a 580 bp fragment (**Appendix 22**), while the reverse sequence detected a 593 bp piece (**Appendix 23**). By comparing these two sequences to the results of sample 17 as well as with the *in silico* models, it was confirmed that sample 38 also carried the RSDup allele.

5.   Discussion

Previous 3C experiments in VYS using T^{hp} alleles showed a much higher level of interaction on the maternal allele than the paternal allele between the *Slc22a2* and *Slc22a3* promoters and elements in the second two thirds of the *Airn* gene body. Therefore, I used 3C material generated by Quanah J. Hudson from paternal AirnT (Sleutels et al., 2002) crosses with a T^{hp} female (Johnson, 1974) to determine if expression of *Airn* is responsible for blocking these interactions on the paternal chromosome in extra-embryonic linages (Sleutels et al., 2002). I found that interactions were gained on the paternal allele between the elements in the second two thirds of the Airn and the *Slc22a2* and *Slc22a3* promoters, indicating that indeed *Airn* is responsible for blocking these interactions. These results indicated that essential enhancers for *Slc22a2* and *Slc22a3* lie within the *Airn* gene.

Pde10a, which lies around 4 Mb away from the Airn TSS was reported to show imprinted expression in placenta by RNA-seq of reciprocal F1 crosses (Wang et al., 2011). In his diploma thesis in the Barlow lab, Daniel Andergassen confirmed this result using RT-qPCR of placenta from T^{hp} crosses. Therefore, using a 3C assay designed by Quanah J, Hudson, I also examined if interactions between the *Pde10a* promoter and the *Airn* gene could be detected, and if these interactions were affected by the truncation of *Airn*. Additionally, as controls the interaction frequency of the *Igf2r* promoter and the promoter of the lncRNA *Airn* itself with the *Airn* gene body were assessed following *AirnT* truncation.

To test further if the *Airn* gene contains essential enhancers for *Slc22a2* and *Slc22a3*, I was part of a larger project to generate deletions and duplication of part and the entire *Airn* gene by chromosome engineering, contributing by developing and conducting genotyping assays. By the end of my thesis we had succeeded in creating two deletion alleles (AlDel and RSDel), and two duplication alleles (AlDup and RSDup) of *Airn* that will that will be subsequently tested for their effect on imprinted expression of *Slc22a2* and *Slc22a3*.

5.1.   Chromosomal loops are controlled by *Airn*

The interaction frequencies between the *Slc22a2* promoter and regions in the *Airn* gene were analyzed by Chromosome Conformation Capture (3C) (Dekker et al., 2002). The embryo trunk as embryonic linage served as a control tissue for comparison to the two analyzed extra-embryonic tissues. The overall pattern in the embryonic linage was not influenced by the truncation of the lncRNA *Airn*. The embryo trunk contains mainly precursors to skeletal muscle, skeletal and neural tissue and hence is not a pure cell type and consequently, the interaction patterns of the different cell types are merged. It was reported that *Slc22a2* showed imprinted expression restricted to extra-embryonic linages (Sleutels et al., 2003). The embryo trunk sample where *Slc22a2* is not expressed did not show any difference in interaction levels between wild type and *AirnT*. Similarly in placenta, both the biallelic and the paternal comparisons did not show any effect of *Airn* truncation on interaction levels. This is in agreement with the recently finding that *Slc22a2* is not expressed in placenta (Hudson et al., 2011). Also the overall pattern of both comparisons in placenta was very similar, although in the T^{hp} genotypes the interaction frequency levels was reduced as expected monoallelic comparisons, and was seen for all placental and VYS T^{hp} 3C assays.

In the visceral yolk sac (VYS), *Slc22a2* is expressed and shows maternal imprinted expression (Hudson et al., 2011). Previous expression studies with the AirnT allele resulted in an overexpression of *Slc22a2* when *Airn* was deleted (Sleutels et al., 2002). In both the biallelic and paternal 3C comparisons, multiple regions showed a gain of interaction when *Airn* was truncated. Most of these regions are located in the second third of the *Airn* transcript, the same genomic area where previous analysis in T^{hp} crosses resulted in a higher level of maternal interactions. Also a higher H3K27ac signal on the maternal allele of T^{hp} and a gain of this histone mark in an Airn truncation situation (T^{hp}/AirnT crosses) was found in the same genomic regions by ChiP-Seq. This indicates that in VYS the lncRNA *Airn* is required to inhibit chromosomal loop formation between the *Airn* gene body and the promoter region of *Slc22a2* on the paternal allele. Together with the previous

3C results that indicated these loops were enriched on the maternal allele, and ChiP data showing that enhancer marks were increased following *Airn* truncation, these results provide support for the enhancer interference model of *Airn* mediated silencing (Pauler et al., 2012).

Interactions with the *Slc22a3* promoter in embryo trunk, where the gene is not expressed, were not affected by truncation of *Airn*, similar to what was seen for interactions with the *Slc22a2* promoter. In both placenta and VYS *Slc22a3* shows imprinted expression (Sleutels et al., 2003); (Hudson et al., 2011), and co-related with this I found a gain of interactions between the *Slc22a3* promoter and the second two thirds of the *Airn* gene following *Airn* truncation. This trend was most clear in the VYS samples, especially in the biallelic comparison, while the placenta samples showed more variation and less difference between the genotypes. The interaction pattern between VYS and placenta was not identical indicating that there may exist different enhancers for *Slc22a3* in these two extra-embryonic tissues. In VYS, previous results showed maternal-specific interactions between the *Slc22a3* promoter and the second two thirds of the *Airn* gene (**Figure 9B**), which overlapped with regions that show an increase in enhancer marks following *Airn* truncation. My results showing an increase in interactions between the *Slc22a3* promoter and the same regions of the *Airn* gene, together with these earlier results, also supports the enhancer interference model of *Airn* mediated silencing for *Slc22a2*.

A comparison of the 3C results for *Slc22a2* and *Slc22a3* shows a gain of interaction in the same genomic region upon *Airn* truncation. This indicates that the same common enhancer elements may be responsible for regulating *Slc22a2* and *Slc22a3,* or at least, specific regulatory regions for each gene are found in the same genomic area. However, the interaction patterns are not identical indicating that there may be different enhancers within the *Airn* gene for the two genes. A combination of the 3C data sets with the H3K27ac ChiP data may enable VYS-specific enhancers to be assigned to either *Slc22a2* or *Slc22a3*, although this is complicated by the resolution of the 3C assays and by the fact that the intensity of the 3C signal may not be

quantitative. Therefore, extra evidence may be required to link enhancers with genes. Additionally, it is becoming increasingly clear that genes employ multiple enhancers and that some enhancers may be shared between multiple genes (de Laat and Duboule, 2013).

For the *Igf2r* 3C assays no difference in the interaction signal in any of the five comparisons was found. The *Igf2r* promoter analysis was used as control (**Figure 15B**), because it was demonstrated before that *Igf2r* is silenced by *Airn* via transcription interference (Latos et al., 2012). The imprinted expression pattern for *Igf2r* can be found in all tissues where is co-expressed with *Airn* (Santoro et al., 2013), indicating tissue-specific enhancers do not play a role in imprinted silencing. Following this prediction in all tissues, 3C interactions were assayed between wild type and *AirnT* alleles. The conclusion for both assays in extra-embryonic linages is, that no specific gain of interactions can be observed when *Airn* is truncated. In all paternal comparison ((T^{hp}/+) to (T^{hp}/AirnT)) the maternal copy of the *Igf2r* cluster is not present and as a result of that, an around 50% reduced interaction signal can be detected (Barlow et al., 1991).

In addition, the interaction frequency between the *Airn* gene with its own promoter was assayed (**Figure 16B**). In the embryo trunk samples the truncation of *Airn* (+/AirnT) resulted in a reduction in interactions with its gene body at most points assayed. This result is unexpected, because although *Airn* is truncated, its promoter remains fully active (Sleutels et al., 2002), so interactions with enhancers would not be expected to change. The complexity of the sample could have an influence on the detected signal. However, in placenta and VYS, wild type and *AirnT* alleles show largely similar interaction levels across the *Airn* gene body, indicating that there are no tissue-specific interactions in these linages.

Pde10a is a gene, which is biallelic expressed in adult brain (Fujishige et al., 1999) and recently, it was found by mRNA-Seq as a potential imprinted gene in placenta (Wang et al., 2011). No expression of *Pde10a* was found in VYS (Daniel Andergassen master thesis). More recently in our lab, it was

demonstrated that in placenta *Pde10a* silencing is regulated by *Airn* (Daniel Andergassen masters thesis). However, in placenta or the other tissues, either I detected no interactions or very low levels with high technical variation between the *Pde10a* promoter and the *Airn* gene body (**Figure 17B**). Immunohistochemistry and RNA *in situ* analysis by Rita Casari, a technician in our lab, has shown that *Pde10a* expression in placenta is limited to a sub-population of spongiotrophoblast and trophoblast giant cells (data not shown). Therefore, as these cells form only a small percentage of the total placenta used to generate 3C material, and this may account for why interactions with the *Pde10a* promoter were not detected. To conduct a 3C assay, cells showing imprinted expression of *Pde10a* may need to be enriched in the starting material, or alternatively an alternative method used, for example, DNA FISH on placental sections using probes for *Pde10a* and the *Airn* gene in combination with RNA FISH to identify cells expressing *Pde10a*.

Tissue-specific gene expression is regulated *in cis* by tissue-specific DNA elements known as enhancers being bound by combination of *trans* acting factors, such as transcription factors, leading to the activation of target promoters (Ong and Corces, 2011a); (Dickel et al., 2013). In the *Igf2r* imprinted gene cluster it has been shown that the *Airn* RNA recruits and targets the EHMT2 H3 histone dimethyltransferase to the *Slc22a3* promoter to cause imprinted silencing in placenta (Nagano et al., 2008). However, both *Airn* and EHMT2 show ubiquitous expression and therefore control of tissue-specific silencing of *Slc22a3* remains unexplained. Recently the enhancer interference model was proposed whereby lncRNA transcription could interfere with promoter-enhancer interactions to cause silencing (Pauler et al., 2012). My 3C results comparing wild type with *Airn* truncation alleles, together with previous results examining allele-specific 3C interactions, and ChIP-seq for enhancer marks in these same genotypes, provided support for this hypothesis indicating that *Airn* expression causes imprinted silencing of *Slc22a2* and *Slc22a3* in VYS by blocking enhancers. Additionally, my 3C results indicate that *Slc22a3* is also silenced in placenta by a similar mechanism. Following these results, and to confirm that the *Airn* gene contains essential extra-embryonic specific enhancers for *Slc22a2* and

Slc22a3, a strategy was devised to delete the *Airn* gene as described in the following section.

5.2.   Generation of deletions and duplications of *Airn* to test the enhancer interference model

The mouse chromosomal engineering experiment uses loxP sites from existing mouse strains in combination with a ubiquitously expressed Cre to cause recombination between homologous chromosomes *in trans*. This can only happen during meiotic pairing, and results in a deletion on the one chromosome and an duplication on the other one (Hérault et al., 1998). Three generations of mice are required to bring both loxP sites and the Cre into one male, which can then be bred and the offspring screened for recombination events. We used this approach to generate deletions and duplications of *Airn* to test if the gene contains essential enhancers for *Slc22a2* and *Slc22a3* as predicted by the enhancer interference model (Pauler et al., 2012), and supported by the previously described 3C and ChIP data. Two genomic regions, a 28kb and a 270kb area, were deleted and duplicated using this approach. My contribution to this part of the project was to genotype mice throughout this breeding experiment, and to establish new Southern blot and PCR assays to screen for the new deletion and duplication alleles. Over 300 mice were screened in the third generation to identify founders of the four different alleles. The effect of these deletions and duplications on *Slc22a2* and *Slc22a3* expression will indicate if these regions contain essential enhancers for these genes. Testing the phenotype of these animals is outside the scope of my thesis but in the following section I will discuss the possible phenotypes and their implications for understanding the regulation of imprinted expression in the *Igf2r* imprinted gene cluster.

5.2.1.   Development of genotyping assays for new mouse strains

In silico alleles were used to aid design of genotyping assays for the predicted new alleles resulting from the chromosome engineering experiment. In all used starting strains the loxP site was bordered on both sides with a

remaining piece of the originally plasmid to create the mouse strain. These 18 to 42 bp long DNA pieces are part of a multiple cloning site (MCS) and contained many restriction sites including for commonly used enzymes such as EcoRI (G$^\vee$AATTC) and PstI (CTGCA$^\vee$G), which complicated the design of the Southern blot assays. *In silico* wild type, starting alleles and predicted deletion and duplication alleles were established and used to model different restriction enzyme and probe combinations to determine assays where all possible allele were clearly distinguishable on a Southern blot. These *in silico* alleles were then also used for the design of PCR assays and to align Sanger sequencing of the newly created alleles.

The restriction enzyme EcoRI could not be used for genotyping because a recognition site for this enzyme was found in the remnant MCS close to the loxP resulting in a size difference between the deletion/duplication and the modified starting allele, which was not detectable by Southern blot. BglII (A$^\vee$GATCT) was commonly used in the lab for Southern blot and was not found in the remnant MCS pieces, and was therefore used for three assays (AlDel; RSDel; RSDup). For the AlDup screen the restriction enzyme AanI (PsiI) (TTA$^\vee$TAA) was chosen to give unique bands for the original modified allele and the duplication. However, this enzyme is expensive so it was only used in the original screen for founders. Subsequently, assays using EcoRI (G$^\vee$AATTC) or PstI (CTGCA$^\vee$G) for Southern blot genotyping were used, which do not uniquely distinguish the AlDup allele from other alleles, but can distinguish wild type from modified alleles.

A limitation during the process of designing and testing genotyping assays was that no positive control was available. We expected a low rate of *trans* recombination and consequently a low number of deletion and duplication alleles. It was not possible to optimize the PCRs without a positively genotyped mouse sample, which could only be found by Southern blot. Therefore, a well-designed and functioning Southern blot assay was the initial focus of the screens. The PCR assay was then tested and optimized after samples were successfully genotyped by Southern blot. Later on the optimized PCR genotyping was used in place of Southern blots for

maintenance and for embryo genotyping because of economic and time reasons.

5.2.2. Predicted phenotype of the AIDel (AirnT/IPΔ deletion)

Figure 26: AIDel allele schematic (28.0kb deletion)

Predicting maternal AIDel allele with the 28kb deletion indicated by the loxP site. The approximate position of *Slc22a2/a3* enhancers indicated by 3C and ChIP is shown.

This newly 23.3kb deletion between the loxP site of the AirnT line (Sleutels et al., 2002) and the loxP site of the 4.7kb promoter deletion of *Igf2r* (IPΔ) (Sleutels et al., 2003) was established by chromosome engineering (**Figure 26**). The newly deleted region plus the IPΔ deleted region gives a total 28.0kb deletion. In this allele *Airn* remains truncated to 3kb while *Igf2r* is not expressed due to deletion of its promoter. Due to the deletion of the *Igf2r* promoter maternal inheritance is expected to be perinatal lethal, and therefore males have to be used to maintain this line (Sleutels et al., 2003). The AIDel allele has a mixed background, but is being backcrossed to FVB/NJ.

According to the enhancer-interference hypothesis we expect essential EXEL enhancers for *Slc22a2/a3* to lie in the *Airn* gene (Pauler et al., 2012). Previous 3C and ChIP results and my 3C results indicate that the enhancers are not located in the overlap between *Airn* and *Igf2r*. Therefore, we hypothesis that maternal inheritance of the small deletion including the first third of *Airn* would not affect expression of *Slc22a2/a3* (**Figure 26**). . If this prediction is correct, and no changes in the expression level of target genes can be observed, this deleted region can be excluded from being involved in controlling extra-embryonic specific expression of *Slc22a2/a3*. Paternal inheritance of this deletion should show an increase in *Slc22a2/a3* expression in extra-

embryonic tissues similar to the *AirnT* strain (Sleutels et al., 2002), as the enhancers are not predicted to be affected.

5.2.3.   Predicted phenotype of the AIDup (AirnT/IPΔ duplication)

Figure 27: Schematic of the AIDup allele (23.5kb duplication)
Predicted maternal AIDup allele with the 23.5kb duplicated region indicated by a dashed line. The approximate position of *Slc22a2/a3* enhancers indicated by 3C and ChIP data is shown.

This newly established line created by chromosome engineering contains a genomic duplication between the loxP site of the AirnT line (Sleutels et al., 2002) and the IPΔ line (Sleutels et al., 2003) (**Figure 27**). This 23.5kb duplication between 3kb downstream of the *Airn* TSS, to 26.5kb downstream, results in 52 kb of *Airn* transcription before overlapping the *Igf2r* promoter (instead of 28.5kb of transcription), and duplication of exon 2 of *Igf2r*. Males should be used to maintain this line, because the effect of the exon 2 duplication of *Igf2r* is predicted to result in a frame shift and a non-functional protein. The AIDup allele has a mixed background and is being backcrossed to FVB/NJ.

This small duplication may or may not affect expression of *Slc22a2/a3.* Upon paternal inheritance, the increased distance between the potential enhancers and *Airn* TSS could lead to paternal up-regulation of the two target genes, due to a reduction in the efficiency of *Airn* silencing. Another interesting possibility to test with this allele is if the silencing of *Igf2r* by transcription interference with its promoter is affected by the increased distance between *Igf2r* TSS and *Airn* TSS. Maternal inheritance of the allele could also be used to test, if the distance between *Slc22a2* and *Slc22a3* promoters and the potential enhancers influences the efficiency of these regulatory elements. The target

promoters are 159 kb (*Slc22a2*) and 237 kb (*Slc22a3*) away and additional 23.5 kb will increase the distance to 182.5 kb (+14.8%) and 260.5 kb (+9.0%). It seems unlikely, that these small changes will have an strong effect on the transcription levels, however this needs to be tested experimentally.

5.2.4. Predicted phenotype of the RSDel (R2Δ/Sod2Δ deletion)
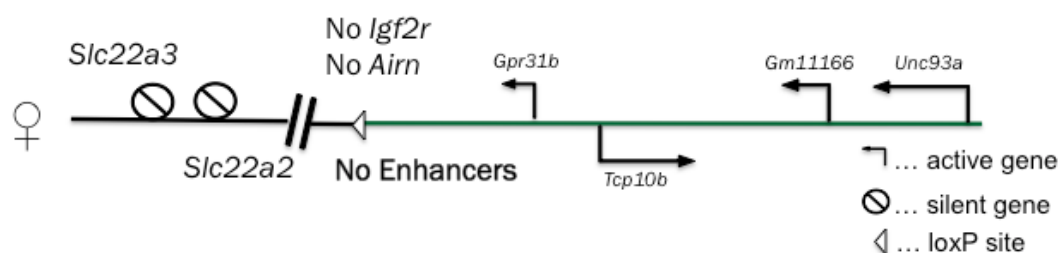
Figure 28: RSDel allele (270kb deletion)

Predicted RSDel allele with the 270kb deletion region indicated by loxP site on the maternal chromosome. The *Slc22a2/a3* enhancers indicated by 3C and ChIP data within the *Airn* gene would be removed.

By chromosome engineering a 270kb deletion was created between the R2Δ and Sod2Δ loxP sites. The newly deleted region is 266 kb in size, which together with the 3.6kb deletion of the *Airn* promoter (R2Δ) (Wutz et al., 2001) and the 0.5kb deletion of exon 3 of *Sod2* (Sod2Δ) (Ikegami et al., 2002) makes a total deletion size of around 270kb (**Figure 28**). *Airn* lncRNA is completely deleted together with all potential enhancers in its gene body. In addition, the first two exons of *Igf2r*, as well as *Mas1*, *Mrgprh*, *Pnldc1*, *Tcp1*, *Mrpl18*, *Acat3*, *Acat2*, *Wtap* and exon 1-3 of *Sod2* are also deleted. The deletion of the *Igf2r* promoter means that similar to the IPΔ strain, this new deletion strain has to be maintaining by paternal matings. The RSDel allele has a mixed background of C57BL/6 and FVB/NJ, but is being backcrossed to FVB/NJ.

The maternal inheritance of the large deletion should lead to a down regulation of *Slc22a2/a3* according to the predictions of the enhancer interference hypothesis (Pauler et al., 2012), and our 3C and ChIP data that indicates there are *Slc22a2/a3* enhancers within the *Airn* gene. The absence

82

of *Igf2r* has no effect on the expression levels of *Slc22a2/a3* (Sleutels et al., 2003). Therefore we expect that other regions of *Airn* outside of the *Igf2r* promoter are required to promote the transcription of *Slc22a2/a3.* We expect to find such regulatory elements in the second two thirds of the *Airn* gene, due to the indications from my 3C results and previous 3C and ChIP results. The inheritance of the RSDel allele on the paternal side should lead to no changes in the silencing of the two target genes, as the absence of *Airn* and the potential enhancers should end up with a situation very similar to a wild type paternal allele where *Airn* would block enhancer activity. In contrast, if a paternal deletion leads to an up-regulation of *Slc22a2/a3 to* maternal wild type levels, this would argue against the presence of tissue specific enhancers in the *Airn* gene, and rather support a tissue specific regulatory function of the *Airn* lncRNA transcript to recruit silencing via EHMT2. A third possibility is an intermediate phenotype with a faint up-regulation of *Slc22a2/a3* compared to a wild type situation, which would indicate additional regulatory elements outside of the *Airn* gene.

Apart from *Igf2r,* where a maternal deletion is perinatal lethal (Sleutels et al., 2003), other genes in the RSDel deletion, namely *Mas1*, *Mrgprh*, *Pnldc1*, *Tcp1*, *Mrpl18*, *Acat3*, *Acat2, Wtap* and exon1-2 of *Sod2,* were not reported to have a heterozygous phenotype in embryogenesis. The gene *Wtap* is part of the deleted region and can lead to a lethal phenotype in very early embryonic development of C57BL/6 in a homozygous knock out situation (Fukusumi et al., 2008). As a consequence of that, a cross between RSDel and T[hp] mice would lead to this full knock out situation. Consequently, only wild type mice will be used to set crosses with RSDel to keep embryos alive until E12.5.

5.2.5. Predicted phenotype of the RSDup (R2Δ/Sod2Δ duplication)
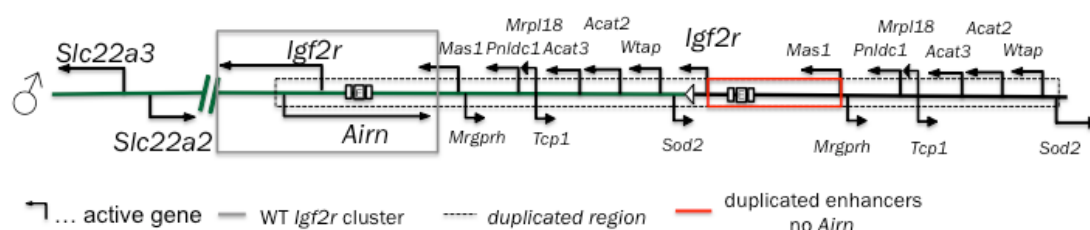
Figure 29: Schematic showing the RSDup allele (266kb duplication)

Predicted paternal RSDel allele with the 266kb duplicated region indicated by the dashed line (includes 532kb). The wild type version of the *Igf2r* cluster (grey box), with a fully functional *Igf2r, Airn* gene and the potential *Slc22a2/a3* enhancer regions, is around 266kb away from the duplicated cluster, in which only the first two exons of *Igf2r* and no *Airn* can be found. Therefore, the duplicated *Airn* gene body (red box) with the enhancer regions should not be overlapped by *Airn* transcription.

This newly established RSDup line created by chromosome engineering contains a 266kb duplication between the R2Δ (Wutz et al., 2001) and Sod2Δ (Ikegami et al., 2002) loxP sites, 3kb to 269kb downstream of the *Airn* TSS (**Figure 29**). This results in a normal wild type allele centromeric of the *Sod2* loxP site, including wild type *Airn* and *Igf2r*, but with a *Sod2* gene that contains only the first two exons. A second duplicated region telemetric of the *Sod2* loxP site has an *Igf2r* gene containing only the first two exons, but no *Airn* should be expressed as the first 3kb of *Airn* plus the promoter are missing at the duplication boundary. Telomeric of *Igf2r* are duplicated copies of *Mas1*, *Mrgprh*, *Pnldc1*, *Tcp1*, *Mrpl18*, *Acat3*, *Acat2*, *Wtap* and a second intact copy of *Sod2* (**Figure 29**). A wild type version of *Igf2r* is present in this allele so both sexes should be able to be used to maintain the strain. Additionally, in contrast to RSDel, RSDup should be able to be crossed to T[hp] as 2 intact copies of Wtap are present, allowing allele-specific analysis at E12.5. RSDup has a mixed background and is being backcrossed to FVB/NJ.

The RSDup allele will result in duplication of the putative *Slc22a2/a3* enhancers 266kb downstream where no overlap by *Airn* transcription occurs. If *Airn* transcription is required for enhancer blocking, the uncovered enhancers may result in an up-regulation of *Slc22a2/a3* upon paternal inheritance, and perhaps also upon maternal inheritance. However, it may also be that at this increased distance and changed genomic context, these

84

enhancers may no longer activate *Slc22a2/a3.* Also, given that the *Sod2* promoter and first two exons remain, it could be that transcription continues over the *Airn* gene resulting in the continued blocking of these enhancers by this artificial transcript. In both scenarios it is also possible that the enhancers are less efficient in activating *Slc22a2/a3* than the enhancers in their wild type position, resulting in weak activation. Given all these feasible possibilities it is difficult to predict the phenotype of the RSDup allele.

5.3.   Future perspective

Three out of four stains (AIDel, RSDel, RSDup) were maintained via a male carrying the newly established allele, and are being used to set up crosses to collect embryonic material for analysis. The AIDup strain has to be maintained before further crosses can be set. AIDel, AIDup and RSDup males and females of the next generation will be set with T^{hp} mice, while RSDel mice will be crossed to FVB/NJ wild type mice. Embryonic and extra-embryonic samples at E12.5 will be collected, genotyped and used for RT-qPCR to analyze the expression levels of *Slc22a2* and *Slc22a3* in these mice. Additionally, the effect on *Igf2r* and *Airn* will be examined along with biallelic controls that are expressed in VE, *Tcp1, Phf10* and *Psmb1*. *Tcp1* is a biallelic gene within the Thp region, while *Phf10* and *Psmb1* are biallelic genes that lie outside of the Thp region. *Tcp1* was shown to be biallelic and unaffected by *Airn* (Daniel Andergassen, Diploma thesis). In previous RT-qPCR assays it was shown, that these gene are not affected by the T^{hp} deletion or a loss of *Airn* (Markus Muckenhuber internship report).

The most powerful allele to test the enhancer interference model is a maternal RSDel allele. The RT-qPCR for *Slc22a2* and *Slc22a3* will be the key experiment and also the most relevant to plan addition experiments with this mouse strain. If the expected down-regulation of the two target genes can be observed, *Pde10a* will also be assayed with the same method in placenta. Upon RSDel maternal deletion active histone marks at the TSS of *Slc22a2* and *Slc22a3* may be lost and replaced by repressive chromatin to keep these

genes silenced. A ChIP on the RSDel allele could be performed to test this. Whether or not the RSDup allele shows a phenotype, one possible experiment would be to perform ChIP-Seq for H3K27ac with the RSDup allele to test if the duplicated enhancers are active. The sequencing cannot distinguish between the identical duplicated regions, but the signals of both will be summed up allowing the activity of the duplicated enhancers to be inferred. Additionally, 3C analyses could be performed for different crosses to determine their effect on interactions in the *Igf2r* cluster.

The successful detection of such tissue specific enhancers in the *Airn* gene will cause more interest in identifying their exact DNA sequence. Enhancers are classified as regulatory elements with a sequence length up to 1,500 bp (Dickel et al., 2013). The combination of my and previous 3C results with ChIP data for active histone marks identify regions that should contain *Slc22a2* or *Slc22a3* enhancers. Precise definition of enhancer sequences may require extra data and motif analysis. Often multiple enhancer elements in the same genomic region can work together to boost the transcription of a target gene (Arnone and Davidson, 1997). This could also be true for the *Airn* region and has to be investigated in the future. CRISPR/Cas9 or similar techniques could be performed to generate mice with smaller deletions of these regions to analyze which of these regions are essential for activating *Slc22a2/a3* and which are supporting elements.

6.   Methods

6.1.   Chromatin Conformation Capture (3C)

The Chromatin Conformation Capture (3C) samples analyzed in this thesis were prepared by Quanah J. Hudson. I performed the quality control and 3C analysis on these samples. E12.5 embryonic (trunk) and two extra-embryonic (placenta; visceral yolk sac (VYS)) tissues from T^{hp} x AirnT crosses were used for these analyses. This type of cross can give rise to four different genotypes ((+/+); (+/AirnT); (T^{hp}/+); (T^{hp}/AirnT)). Embryo heads were used for genotyping by Southern blot. Individual tissues with the same genotype were pooled and used to prepare 3C samples according to the protocol (Dostie and Dekker, 2007). The 3C samples were analyzed by TaqMan qPCR to evaluate the relative interaction levels.

6.1.1.   Quantitative polymerase chain reaction (qPCR)

Quantitative polymerase chain reaction (qPCR) combines cycles of PCR (see 4.2.2) (denaturation, annealing, elongation) with a fluorescent dye reporter system that recognizes the amount of PCR product, together with a qPCR machine and computer that measures and records fluorescence every cycle. I performed two different types of qPCR assays: SYBR Green and TaqMan. The SYBR Green technique uses a non-specific fluorescent dye, which interacts with any double-stranded DNA to measure the amount of PCR product. The TaqMan assay uses PCR primers in combination with a sequence-specific DNA probe, which is modified on one end with a fluorescent reporter and on the other end with a quencher. If these two components are in close proximity, the quencher absorbs the signal of the reporter. During the elongation step the probe is digested by the exonuclease function of a specific polymerase. Thereby the reporter is released from close proximity to the quencher resulting in a florescent signal that can be detected.

Mastermix

SYBR Green	Final Conc.			Cycling	Conditions	
2x MESA GREEN qPCR MasterMix Plus	1x	12.50	µl	95°C	5'	
Primer mix (Forward+Reverse)	0.05µM	2.50	µl	95°C	15"	
Water for embryo transfer	-	5.00	µl	60°C	1'	40x
DNA template	-	5.00	µl	95°C	15"	
				60°C	20"	
				95°C	15"	

TaqMan	Final Conc.			Cycling	Conditions	
2x qPCR MasterMix Plus	1x	12.50	µl	50°C	2'	
Forward primer	0.9µM	0.23	µl	95°C	10'	
Reverse primer	0.9µM	0.23	µl	95°C	15"	40x - 50x
TaqMan probe	0.2µM	0.05	µl	60°C	1'	
Water for embryo transfer	-	7.00	µl			
DNA template	-	5.00	µl			

Table 3: Mastermixes and cycling conditions for both types of qPCR assays

To quantify the resulting signals the 7900HT Fast Real Time PCR machine with Sequence Detection System Software, Version 2.3 (Applied Biosystems) was used. The qPCR reaction took place in a 96-well plate and each tube contained a total reaction volume of 25µl (20µl mastermix + 5µl DNA template). The mastermix was prepared following an established recipe (**Table 3**) and two different cycling protocols for the two different qPCR systems were performed (**Table 3**). Each sample was loaded three times to give technical replicates according to the pipetting plan (**Figure 30**). This loading plan was similar for SYBR Green and TaqMan assays.

	1	2	3	4	5	6	7	8	9	10	11	12	
A	Igf2r_Pde10a BAC library 1:4	Igf2r_Pde10a BAC library 1:4	Igf2r_Pde10a BAC library 1:16	Igf2r_Pde10a BAC library 1:16	Igf2r_Pde10a BAC library 1:64	Igf2r_Pde10a BAC library 1:64	Igf2r_Pde10a BAC library 1:256	Igf2r_Pde10a BAC library 1:256	✕	✕	✕	✕	**Eco_054-a2_F2 vs Eco_096_F3 (Slc22a2_3C_Probe)**
B	Placenta 0% 3C	Placenta 0% 3C	Placenta 0% 3C	Embryo +/+ 3C	Embryo +/+ 3C	Embryo +/+ 3C	Embryo +/AirT 3C	Embryo +/AirT 3C	Embryo +/AirT 3C	Embryo Thp/+ 3C	Embryo Thp/+ 3C	Embryo Thp/+ 3C	
C	Placenta +/+ 3C	Placenta +/+ 3C	Placenta +/+ 3C	Placenta +/AirT 3C	Placenta +/AirT 3C	Placenta +/AirT 3C	Placenta Thp/+ 3C	Placenta Thp/+ 3C	Placenta Thp/+ 3C	Placenta Thp/AirT 3C	Placenta Thp/AirT 3C	Placenta Thp/AirT 3C	
D	VYS +/+ 3C	VYS +/+ 3C	VYS +/+ 3C	VYS +/AirT 3C	VYS +/AirT 3C	VYS +/AirT 3C	VYS Thp/+ 3C	VYS Thp/+ 3C	VYS Thp/+ 3C	VYS Thp/AirT 3C	VYS Thp/AirT 3C	VYS Thp/AirT 3C	
E	Igf2r_Pde10a BAC library 1:4	Igf2r_Pde10a BAC library 1:4	Igf2r_Pde10a BAC library 1:16	Igf2r_Pde10a BAC library 1:16	Igf2r_Pde10a BAC library 1:64	Igf2r_Pde10a BAC library 1:64	Igf2r_Pde10a BAC library 1:256	Igf2r_Pde10a BAC library 1:256	✕	✕	✕	✕	**Eco_054-a2_F2 vs Eco_097_F1 (Slc22a2_3C_Probe)**
F	Placenta 0% 3C	Placenta 0% 3C	Placenta 0% 3C	Embryo +/+ 3C	Embryo +/+ 3C	Embryo +/+ 3C	Embryo +/AirT 3C	Embryo +/AirT 3C	Embryo +/AirT 3C	Embryo Thp/+ 3C	Embryo Thp/+ 3C	Embryo Thp/+ 3C	
G	Placenta +/+ 3C	Placenta +/+ 3C	Placenta +/+ 3C	Placenta +/AirT 3C	Placenta +/AirT 3C	Placenta +/AirT 3C	Placenta Thp/+ 3C	Placenta Thp/+ 3C	Placenta Thp/+ 3C	Placenta Thp/AirT 3C	Placenta Thp/AirT 3C	Placenta Thp/AirT 3C	
H	VYS +/+ 3C	VYS +/+ 3C	VYS +/+ 3C	VYS +/AirT 3C	VYS +/AirT 3C	VYS +/AirT 3C	VYS Thp/+ 3C	VYS Thp/+ 3C	VYS Thp/+ 3C	VYS Thp/AirT 3C	VYS Thp/AirT 3C	VYS Thp/AirT 3C	

Figure 30: Example of 3C TaqMan qPCR loading plan

6.1.2. Quality control of 3C samples

According to the 3C protocol a quality control of the 3C was performed by gel electrophoresis. Aliquots of undigested (UND) control samples and digested (D) control samples of each 3C sample were prepared. The concentrations were measured by *Nanodrop ®-1000 spectrophotometer* and adjusted to around 200ng/µl. 0.5µg of each UND, D and 3C was loaded onto 0.8% agarose gel in 1xTAE and ran for 3hrs at 80V (**Figure 11**).

6.1.3. Digestion efficiency check of 3C samples

Digestion efficiency of 3C samples was checked for five *EcoRI* restriction sites within the *Airn* gene (**Figure 12**). For this approach the starting concentration of all samples was diluted from 200ng/µl to 100ng/µl. SYBR Green assays were designed across these restriction sites using standard conditions (**Table 3**). Undigested and digested controls were taken during the 3C protocol. All samples were analyzed three times per qPCR assay as technical replicates and averaged afterwards to calculate the mean. These values were normalized to the Gapdh-3C results. The normalized digested samples (D) were divided by the undigested samples (UND) to calculate the percentage of digestion efficiency. All five control SYBR Green assays (Eco_111 F1; Eco_112 F1; Eco_113 F1; Eco_114 F1; Eco_116 F1) were analyzed identical (**Table 4**). The mean and standard deviation of the 5 assays was then calculated for each tissue (**Figure 12**).

3C Quality Control SYBR Green primers

Gapdh-3C F	ACAGTCCATGCCATCACTGCC	chr6:125,162,685-125,162,705
Gapdh-3C R	GCCTGCTTCACCACCTTCTTG	chr6:125,162,356-125,162,376
Eco_111 F1	TTGGTGGACTAAAAGTACCTAGA GAA	chr17:12,799,100-12,799,125
Eco_111_(112) Rev	CTGTGCGGCCCTGTTCTCCC	chr17:12,799,507-12,799,526
Eco_112 F1	ATTTCTTGGCAATGGTTCAGTGA	chr17:12,799,711-12,799,733
Eco_112_(113) Rev	GCTCATCCCCAGCAAGCCCA	chr17:12,800,020-12,800,039
Eco_113 F1	CTTTGCAAATGTTCTCAACTGG	chr17:12,806,429-12,806,450
Eco_113_(114) Rev	GCCATCGCTCCCGCACCATT	chr17:12,806,535-12,806,554
Eco_114 F1	TCCCACAATAGCAAGACTTAAAA	chr17:12,809,132-12,809,154
Eco_114_(115) Rev	TCAGGGAATGCAAGCCGCAACT	chr17:12,809,393-12,809,414
Eco_116 F1	GGATTTCTAAGGAGTCCACAACC	chr17:12,812,886-12,812,908
Eco_117 R2	CAGACTGAGCCTCCTAGATGACA A	chr17:12,813,040-12,813,063

Table 4: 3C Quality Control SYBR Green primers

Names, DNA sequences of primers and genomic co-ordinates of the amplicons (GRC38/mm10) for 3C digestion efficiency check

6.1.4. BAC 3C library preparation

Four or five different BACs were used to prepare BAC 3C libraries to calculate standard curves for each 3C qPCR assay performed at the *Igf2r* locus, while one BAC (RP23-50N22) was used to prepare libraries for the *H19/Igf2* control locus (**Figure 31**). The *Igf2r, Airn, Slc22a2*, and *Slc22a3* genes were covered by four BACs, *Igf2r* BAC I (RP23-309H20; 261,145 bp), *Igf2r* BAC II (RP23-367L3; 226,917 bp), *Igf2r* BAC III (RP23-84H13; 206,247 bp), *Igf2r* BAC IV (RP23-284N16; 192,477 bp), and the around 4Mb away *Pde10a* BAC V (RP23-410M14; 189,710 bp) covered the entire *Pde10a* gene body. The BAC 3C library containing the four *Igf2r* BACs and the library for *H19/Igf2* cluster were prepared previously by Quanah J. Hudson and used as a standard for *Airn*, *Igf2r*, *Slc22a2*, *Slc22a3* and *H19* 3C qPCRs. Additionally, I prepared a second library containing these four BACs plus the *Pde10a* BAC using the following protocol.
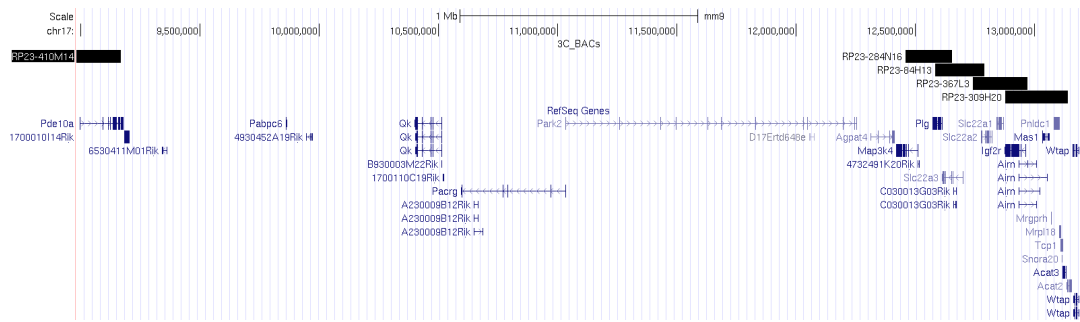
Figure 31: UCSC genome browser (GRC38/mm10) screenshot of the *Igf2r* cluster including *Pde10a* and showing the position of the BACs used to prepare 3C libraries

E. coli (DH5α) glycerol stocks of all five BACs were streaked out on *circlegrow agar* plates containing chloramphenicol (1:1000 from 25mg/ml CAM stock). After incubation at 37°C o/n, a single colony was picked from each plate and used to inoculate 4ml of LB+CAM. After 6hrs at 37°C, the liquid culture was added to 100ml LB+CAM and incubated at 37°C o/n. Isolation of BAC DNA with *Epicentre MACMAX^TM^ DNA purification kit* was done as described in the user manual. A BAC dilution series (1:2; 1:4; 1:8) was loaded onto a 0.8% agarose in 1xTAE gel (18hrs; 65V) together with λ mix marker (*λ Mix 19 (0.5µg/µl); Thermo Scientific*) (**Figure 32**). The BAC bands were compared with bands of the ladder to calculate the amount of BAC DNA. All five BAC samples were pooled in equal molar amounts to give a total amount of 20ng of DNA. Phenol:Chloroform: Isoamylalcohol extraction was performed two times. The DNA was precipitated (0.1Vol NaAcO; 2.5Vol 96% EtOH; 1µl glycerol) at -20°C o/n. The precipitate was centrifuged (16,100xg; 15min; 4°C), washed with 200µl 70% EtOH and centrifuged again (16,100xg; 7min; 4°C). The DNA pellet was air-dried for 5min and resuspended in 165µl MQ H$_2$O. The BACs were then digested with EcoRI (160µl BAC DNA; 20µl 10x Buffer; 20µl EcoRI (10U/µl)) and incubated at 37°C o/n. 5µl of the undigested (UD) BACs was saved as a quality control. On the next day an EtOH precipitation and 70% EtOH wash was performed. After resuspending the pellet in 145µl MQ H$_2$O, the digested BAC DNA was ligated (140µl EcoRI digested BAC DNA mix; 40µl 5x DNA Ligase Reaction Buffer; 20µl T4 DNA ligase) and incubated at 16°C o/n to create the 3C BAC library. 5µl of the digested (D) BACs was saved as a quality control. A visual quality control was done by running UND, D and 3C samples on a 0.8% agarose gel in 1xTAE

(60min; 100V). *Nanodrop ®-1000 spectrophotometer* was used to measure the final *Igf2r/Pde10a* BAC library concentration (842.50 ng/µl). A dilution series of the ligated BAC 3C library was tested by qPCR to find a dilution range that was similar to 3C interaction levels from the tissues. 1:500 was then set as the starting concentration for the dilution series to calculate the 3C standard curve for the five BACs containing library. The originally dilution series of the four *Igf2r* BACs containing library was used as a 1:800 dilution. In both cases the final working concentration of the 3C BAC library was approximately 1ng/ul.
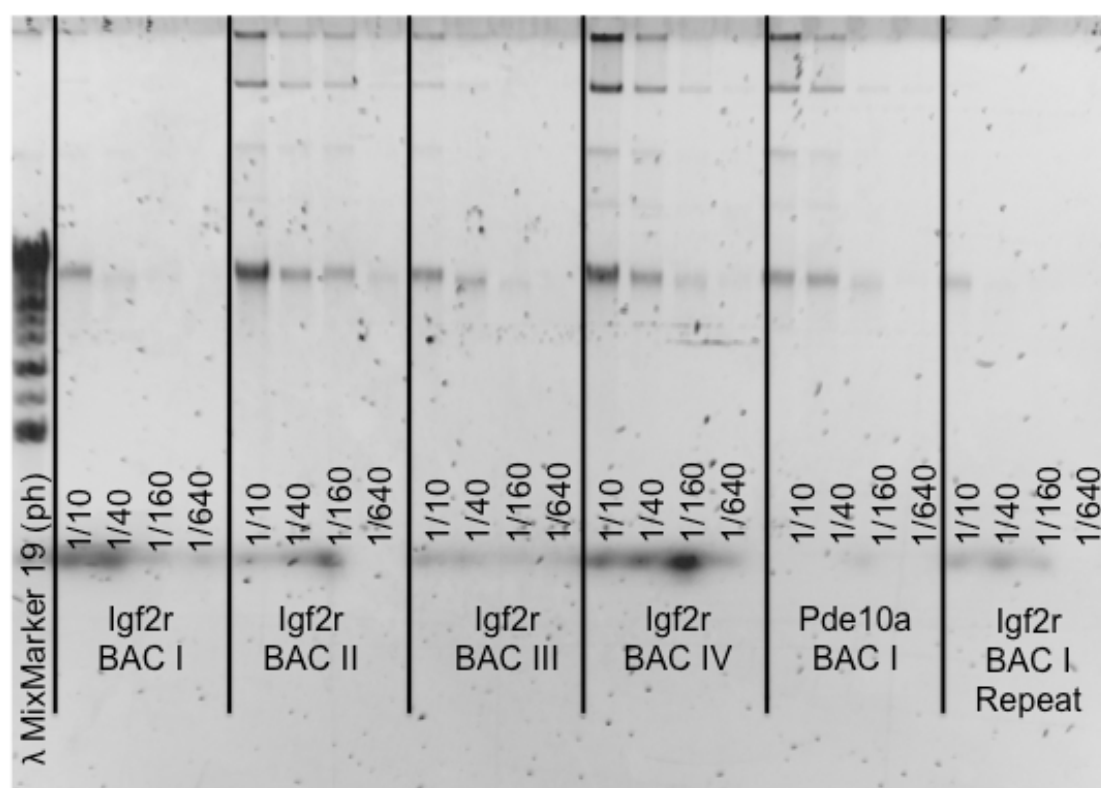
Figure 32: Quality control of isolated BACs after EcoRI treatment

A dilution series of each isolated BAC (1/10; 1/40; 1/160; 1/640) was prepared with MQ H_2O. Separation was done by gel electrophoresis on 0.8% agarose in 1xTAE for 18hrs at 65V. All samples were run in 0.8% agarose stained with ethidium bromide and photographed under UV light. 20µl of the Lambda Mix Marker 19 (Thermo Scientific) ladder was loaded.

6.1.5.   Analysis of 3C samples by qPCR

Chromosome interactions can be detected from 3C material using Taqman qPCR with a bait primer in combination with a Taqman probe located on a "bait" restriction fragment near the cutting site, and a prey primer located near

92

the cutting site of a "prey" fragment. Relative 3C interaction levels are calculated by dividing the interactions at the locus of interest by interactions at an independent locus. I used two interactions in the *H19/Igf2* cluster located on a different chromosome to normalize 3C interactions in the *Igf2r* cluster. Interaction levels between the ICE in the *H19/Igf2* cluster and nearby regions was analyzed by Taqman qPCR using a Taqman probe (**Table 5**) located on an EcoRI fragment containing the ICE. The bait primers were also located in the *H19* ICE (**Table 6**) and two different prey primers on two different EcoRI fragments (**Table 7**). A bait primer and Taqman probe located in one of the promoter regions of *Airn*, *Igf2r*, *Slc22a2*, *Slc22a3* or *Pde10a* were used together with a certain set of prey primers, covering the entire gene body of the lncRNA *Airn*, to analyze the interaction levels between those regions. Different sets of bait-prey-primer pairs were used to analyze the interaction frequency between a certain promoter and the *Airn* gene body (**Table 8**). 3C samples were diluted to final working concentration of 100ng/µl before being subject to TaqMan qPCRs. For each primer pair a 1:500 starting dilution of a BAC 3C library (1:4; 1:16; 1:64; 1:256) was used to calculate a standard curve as described above.

Taqman probes for 3C analysis

Gene	TaqMan Probe	Sequence	
H19	H19-3C	AATGGCAATGCTGTGGGTCACCCA	chr7:142,581,953-142,581,976
Airn	DMR2-3C	TGCTTTGCCAACAGGGACAACTTCAA	chr17:12,746,521-12,746,546
Igf2r	DMR1-3C	AGCTTGAGGCTGGGCTGATATACAGGATG	chr17:12,773,380-12,773,408
Pde10a	Pde10a-3C	CGGATGATTTTCATGTTTTGATTTGCAGTG	chr17:8,799,887-8,799,916
Slc22a2	Slc22a2-3C	AACCATGCCGACCGTGGATGATATTC	chr17:12,584,277-12,584,302
Slc22a3	Slc22a3-3C	AGACTCTCCGGAAAACAGCGGTGCTT	chr17:12,507,941-12,507,966

Table 5: Taqman probes in target promoters to analyze 3C interactions by qPCR

The raw data was processed in Microsoft Excel 2011 to calculate interaction frequencies. The standard curve for each primer pair was used to convert exponential CT values to the linear scale. The mean and the standard deviation of three technical replicates were used to calculate 3C interactions

and technical variation. The 3C results of the *Igf2r* cluster were normalized to mean of two interactions with the *H19* ICE as described above. The relative interaction levels for all four genotypes at all prey primer binding sites were plotted per tissue. The wild type paternal allele was compared with a paternal allele with a truncated Airn for both the wild type maternal allele ((+/+) vs (+/AirnT)) and a maternal T^{hp} allele ((T^{hp}/+) vs (T^{hp}/AirnT)).

qPCR primers (bait) for 3C analysis

Gene	Bait primer	Sequence	
H19	Eco_004-H19_DMR_F1	ACCAGCCTAGAAAATGCATGTGT	chr7:142,581,885-142,581,907
	Eco_096 F3	AACTCCACCCCAGCATGAAC	chr17:12,746,496-12,746,515
	Eco_096 F4	CCACAGCTCAGCCTCATCTCT	chr17:12,746,455-12,746,475
Airn	Eco_096 F5	TCAGATGTGGGCTCCTCCTAA	chr17:12,746,419-12,746,439
	Eco_096 F6	CCTCATCTCTCTCAGATCCTGAAAT	chr17:12,746,466-12,746,490
	Eco_096 F46	CTCAGCCTCATCTCTCTCAGATC	chr17:12,746,461-12,746,483
Igf2r	Eco_104-DMR1_F1	CATCACACATGGTGGTGCAT	chr17:12,773,317-12,773,336
Pde10a	Pde10a Bait	CTGCAGGGATGGCAGGAAC	chr17:8,799,948-8,799,967
Slc22a2	Eco_054-a2 F2	GCAGCATTTGCAACCCTGTA	chr17:12,584,227-12,584,246
Slc22a3	Eco_034-a3_F3	TCCCAAAGTTCAAAAGGAGAAGTC	chr17:12,507,886-12,507,909
	Eco_034-a3_F4	GCCCACAGGGCTAAAAGATTT	chr17:12,507,859-12,507,879

Table 6: Bait primers detect 3C interaction frequency by Taqman qPCRs

qPCR primers (prey) for 3C analysis

Gene	Prey primer	Sequence	
H19	Eco013-H19 F3	GGGACAGTCCCGAAGAATGAG	chr7:142,611,988-142,612,008
	Eco020-H19 F14	CAGCAAGCGGCTGGCTAA	chr7:142,645,277-142,645,295
	Eco_096 F3	AACTCCACCCCAGCATGAAC	chr17:12,746,496-12,746,515
	Eco_097 F1	GGAGTCAATGCCTTCTTCTGAATT	chr17:12,747,861-12,747,884
	Eco_098 F1	ACCAAAAGGTCACGAATAAGGA	chr17:12,755,215-12,755,236
	Eco_099 R1	CTTCAGTGTCATGTGGCATTTT	chr17:12,755,340-12,755,361
	Eco_100 F1	CATGGTGAGATTAGATGCCTTT	chr17:12,764,265-12,764,286
Airn	Eco_101 F1	CGTTTAAAAGAATAGACATCCTGATT	chr17:12,764,705-12,764,730
	Eco_102 F1	CCAGCTCCAAGGACTCTGAGA	chr17:12,767,977-12,767,997
	Eco_103 F1	TTTGGGCAGCTAATTGGTCTAT	chr17:12,768,360-12,768,381
	Eco_104 R1	GGTAGTCCAGGCTCTTTAGTCC	chr17:12,768,517-12,768,538
	Eco_106 R1	TGGAGCTCAGACATTTATACCG	chr17:12,777,465-12,777,486
	Eco_107 F1	CTGTGAGTTCTGTATGGGGACT	chr17:12,781,296-12,781,317

Eco_108 F1	CTGAGTTCCAGTTTCAAACCAC	chr17:12,781,884-12,781,905
Eco_109 F1	GCCAGGCAGGTAGCTTTAATAC	chr17:12,782,702-12,782,723
Eco_110 R1	TCAAAACACCTCTTTTGGGAAT	chr17:12,782,783-12,782,804
Eco_111 F1	TTGGTGGACTAAAAGTACCTAGAGAA	chr17:12,799,100-12,799,125
Eco_112 F1	ATTTCTTGGCAATGGTTCAGTGA	chr17:12,799,711-12,799,733
Eco_113 F1	CTTTGCAAATGTTCTCAACTGG	chr17:12,806,429-12,806,450
Eco_114 F1	TCCCACAATAGCAAGACTTAAAA	chr17:12,809,132-12,809,154
Eco_115 F1	TGCATATTAACTGCATGGCATA	chr17:12,810,120-12,810,141
Eco_116 F1	GGATTTCTAAGGAGTCCACAACC	chr17:12,812,886-12,812,908
Eco_117 R2	CAGACTGAGCCTCCTAGATGACAA	chr17:12,813,040-12,813,063
Eco_118 F1	GTTAAAACTGTTGGAGAAATCAGGAAT	chr17:12,824,785-12,824,811
Eco_119 F1	AATGACCATCTTACCAAAAGCAATC	chr17:12,825,367-12,825,392
Eco_120 F1	TTACACCAATCAGAATGGCTAAGATC	chr17:12,826,617-12,826,643
Eco_121 F1	ACAGCCAGTAGCTGGACAGGAA	chr17:12,829,049-12,829,070
Eco_122 F1	CGCCCACAGTAGTAACAAAACC	chr17:12,830,629-12,830,650
Eco_123 F1	CACTGCAGAAACTACCCAACAA	chr17:12,832,576-12,832,597
Eco_124 F2	GACCCTAGTTTTAATCTATTGTTCACTTTG	chr17:12,838,523-12,838,552
Eco_125 F1	TTCCCAGGAATTGTGCTAGGTA	chr17:12,852,701-12,852,722
Eco_126 F1	ACACAGAATGGGTGGTAGTCCT	chr17:12,855,854-12,855,875
Eco_127 F1	AGTTCTAAGTGTTGGGCAAAGG	chr17:12,862,468-12,862,489
Eco_128 F1	GACCATATCTGAAACCAGACCA	chr17:12,862,870-12,862,891
Eco_129 F1	TATATCACCTCAGATGTATCTTCATGC	chr17:12,866,253-12,866,279
Eco_130 F1	TGACTTTTCTCACTGGTGTGGT	chr17:12,870,054-12,870,075
Eco_131 R1	CAAAGCTGGTGGGACACTATTT	chr17:12,870,180-12,870,201

Table 7: List of all used prey primers

Prey primer names and sequences near EcoRI junctions in the *Airn* gene body that were used in conjunction with the bait primers and Taqman probes to assay 3C interactions.

Airn		*Igf2r*		*Pde10a*	
Bait	Prey	Bait	Prey	Bait	Prey
Eco_096 F46	Eco_101 F1	Eco_104-DMR1_F1	Eco_110 R1	Pde10a	Eco_110 R1
Eco_096 F3	Eco_102 F1	Eco_104-DMR1_F1	Eco_111 F1	Pde10a	Eco_111 F1
Eco_096 F3	Eco_104 R1	Eco_104-DMR1_F1	Eco_112 F2	Pde10a	Eco_112 F2
Eco_096 F46	Eco_107 F1	Eco_104-DMR1_F1	Eco_113 F1	Pde10a	Eco_113 F1
Eco_096 F46	Eco_108 F1	Eco_104-DMR1_F1	Eco_114 F1	Pde10a	Eco_114 F1
Eco_096 F3	Eco_109 F1	Eco_104-DMR1_F1	Eco_115 F1	Pde10a	Eco_115 F1
Eco_096 F3	Eco_111 F1	Eco_104-DMR1_F1	Eco_116 F1	Pde10a	Eco_116 F1
Eco_096 F6	Eco_113 F1	Eco_104-DMR1_F1	Eco_117 R2	Pde10a	Eco_117 R2
Eco_096 F46	Eco_114 F1	Eco_104-DMR1_F1	Eco_121 F1	Pde10a	Eco_121 F1
Eco_096 F3	Eco_116 F1	Eco_104-DMR1_F1	Eco_122 F1	Pde10a	Eco_122 F1
Eco_096 F5	Eco_117 R2	Eco_104-DMR1_F1	Eco_123 F1	Pde10a	Eco_123 F1

Eco_096 F3	Eco_118 F1	Eco_104-DMR1_F1	Eco_124 F2	Pde10a	Eco_124 F2
Eco_096 F3	Eco_120 F1	Eco_104-DMR1_F1	Eco_125 F1	Pde10a	Eco_125 F1
Eco_096 F3	Eco_122 F1	Eco_104-DMR1_F1	Eco_126 F1	Pde10a	Eco_126 F1
Eco_096 F4	Eco_124 F2	Eco_104-DMR1_F1	Eco_127 F1	Pde10a	Eco_127 F1
Eco_096 F3	Eco_125 F1	Eco_104-DMR1_F1	Eco_128 F1	Pde10a	Eco_128 F1
Eco_096 F46	Eco_128 F1	Eco_104-DMR1_F1	Eco_129_F1	Pde10a	Eco_129_F1
Eco_096 F5	Eco_129 F1	Eco_104-DMR1_F1	Eco_130 F1	Pde10a	Eco_130 F1
		Eco_104-DMR1_F1	Eco_131_R1	Pde10a	Eco_131_R1

Slc22a2		*Slc22a3*	
Bait	Prey	Bait	Prey
Eco_054-a2_F2	Eco_096 F3	Eco_034-a3_F3	Eco_096 F3
Eco_054-a2_F2	Eco_097 F1	Eco_034-a3_F3	Eco_097 F1
Eco_054-a2_F2	Eco_098 F1	Eco_034-a3_F3	Eco_098 F1
Eco_054-a2_F2	Eco_099 R1	Eco_034-a3_F3	Eco_099 R1
Eco_054-a2_F2	Eco_100 F1	Eco_034-a3_F4	Eco_100 F1
Eco_054-a2_F2	Eco_101 F1	Eco_034-a3_F3	Eco_101 F1
Eco_054-a2_F2	Eco_103 F1	Eco_034-a3_F3	Eco_103 F1
Eco_054-a2_F2	Eco_104 R1	Eco_034-a3_F3	Eco_104 R1
Eco_054-a2_F2	Eco_106 R1	Eco_034-a3_F3	Eco_106 R1
Eco_054-a2_F2	Eco_107 F1	Eco_034-a3_F3	Eco_110 R1
Eco_054-a2_F2	Eco_110 R1	Eco_034-a3_F4	Eco_111 F1
Eco_054-a2_F2	Eco_111 F1	Eco_034-a3_F4	Eco_112 F1
Eco_054-a2_F2	Eco_112 F1	Eco_034-a3_F4	Eco_113 F1
Eco_054-a2_F2	Eco_113 F1	Eco_034-a3_F3	Eco_114 F1
Eco_054-a2_F2	Eco_114 F1	Eco_034-a3_F3	Eco_115 F1
Eco_054-a2_F2	Eco_115 F1	Eco_034-a3_F3	Eco_116 F1
Eco_054-a2_F2	Eco_116 F1	Eco_034-a3_F4	Eco_117 R2
Eco_054-a2_F2	Eco_117 R2	Eco_034-a3_F3	Eco_121 F1
Eco_054-a2_F2	Eco_121 F1	Eco_034-a3_F4	Eco_122 F1
Eco_054-a2_F2	Eco_122 F1	Eco_034-a3_F3	Eco_124 F2
Eco_054-a2_F2	Eco_123 F1	Eco_034-a3_F3	Eco_126 F1
Eco_054-a2_F2	Eco_125 F1	Eco_034-a3_F3	Eco_127 F1
Eco_054-a2_F2	Eco_126 F1	Eco_034-a3_F3	Eco_128 F1
Eco_054-a2_F2	Eco_127 F1	Eco_034-a3_F3	Eco_129_F1
Eco_054-a2_F2	Eco_128 F1	Eco_034-a3_F3	Eco_131_R1
Eco_054-a2_F2	Eco_129 F1		
Eco_054-a2_F2	Eco_130 F1		

Table 8: Summary of used bait and prey primer pair combinations

6.2. Chromosome Engineering Experiment

6.2.1. Breeding scheme for mouse crosses

The chromosome engineering method uses loxP sites in the same orientation from existing mouse strains in conjunction with a ubiquitously expressed Cre to cause *trans* recombination between homologous chromosomes during male meiotic pairing, resulting in a deletion or duplication (Hérault et al., 1998). This requires two generations of breeding in order to obtain triple mutant males with the two loxP sites *in trans* together with the Cre. We used the ubiquitously expressed X-linked *Hprt-Cre* ((Tang et al., 2002), obtained from Simon Hippenmeyer), the *Airn* promoter deletion allele (R2Δ), the *Airn* truncation allele (AirnT) and *Igf2r* promoter deletion allele (IPΔ) from our lab (Sleutels et al., 2003; Sleutels et al., 2002 ; Wutz et al., 2001) and Sod2$^{flox/flox}$ ((Ikegami et al., 2002), obtained from Takahiko Shimizu). I determined the orientation of the loxP sites from the R2Δ, AirnT, IPΔ and Sod2$^{flox/flox}$ alleles by PCR (see 1.1.1.2) across the loxP site followed by Sanger sequencing (**Table 9**).

Primer pairs to check loxP orientation

Forward	Sequence	Reverse	Sequence
AirnT-F	GCCCCTTTGTAGGTGCCTCTGC	AirnT-R	ACTCCCAGTCATAGCTGTCCCTCT
IPD-F	GACGCTCAGTTCTCCTTGACAAATTCT	IPD-R	TGGCTAGGGCTGGGGATATAAAGTT
R2D-F	CCTGTTTCCAGCCCAGGAGGC	R2D-R	CTGCTTGCCCCTTTGTAGGTGCC
Sod2 LoxP1 F	CGAGGGGCATCTAGTGGAGAAG	Sod2 LoxP1 R	TTACGTCCAGCCAAGCT
Sod2 LoxP2 F	GGGTGGTGGAGAACCCAAA	Sod2 LoxP2.2 R	GAAGAAAGTCACCTCCACACACAG

Table 9: PCR primer pairs to amplify the loxP-containing region

The purified product was sent for sequencing (by Microsynth) according to the user manual to check the orientation of the loxP site.

We bred together the AirnT and IPΔ alleles with Hprt-Cre to engineer a 28kb deletion and a 23.5kb duplication between the loxP sites (+ orientation), and bred together the R2Δ and Sod2$^{flox/flox}$ with Hprt-Cre to engineer a 270kb deletion and a 266kb duplication between the loxP sites (- orientation). Quanah J. Hudson made the breeding schema (**Figure 33**) and all mice used in

this approach were genotyped using previously established Southern blot assays (see below).

AIDel = Deletion between AirnT and IPΔ
AIDup = Duplication between AirnT and IPΔ

RSDel = Deletion between R2Δ and Sod2Δ
RSDup = Duplication between R2Δ and Sod2Δ
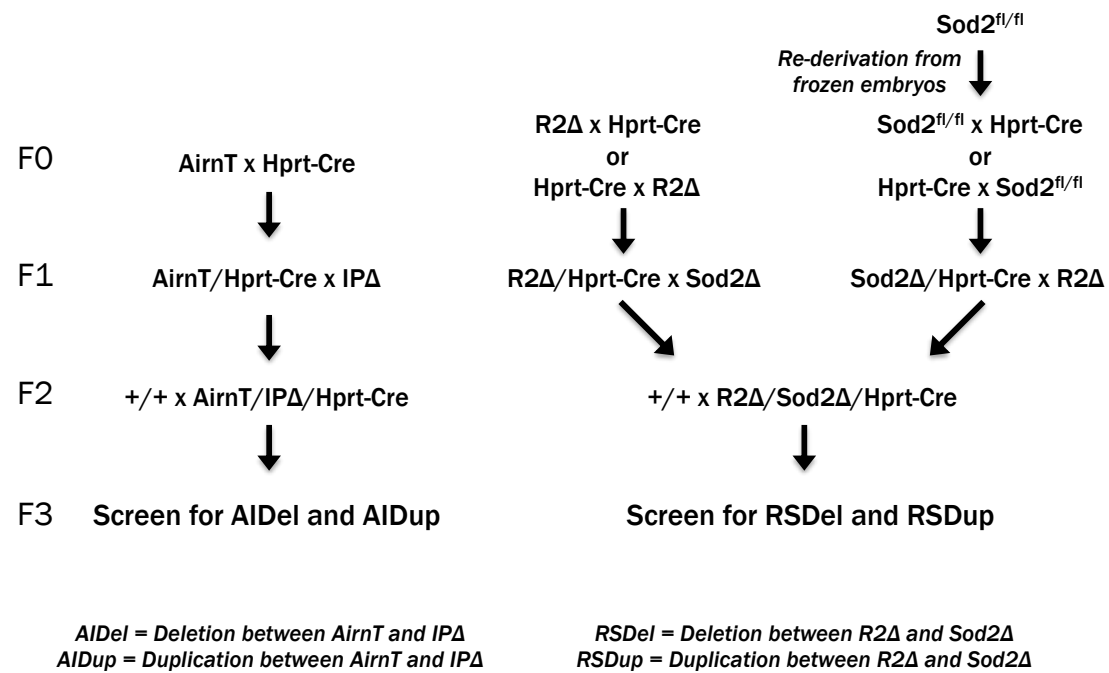
Figure 33: The crossing schema of the chromosomal engineering experiment

To generate the 28kb deletion/23.5kb duplication, in the F0 generation AirnT females were crossed with Hprt-Cre males (**Figure 33**). Female offspring (F1), positively genotyped for both mutant alleles, were crossed with an IPΔ male as maternal inheritance of an *Igf2r* promoter deletion is embryonic lethal. Triple mutant male offspring (F2) were crossed with FVB/NJ (+/+) females and the next generation (F3) was screened by newly designed Southern Blot and PCR assays to identify deletions and duplications between AirnT and IPΔ (AirnT/IPΔ-Deletion= AIDel; AirnT/IPΔ-Duplication= AIDup). An *in silico* model of the genomic changes shows that in the deletion no *Igf2r* as well as no functional *Airn* is present, while in the duplication *Igf2r* carries a second exon 2. (**Figure 34A**).
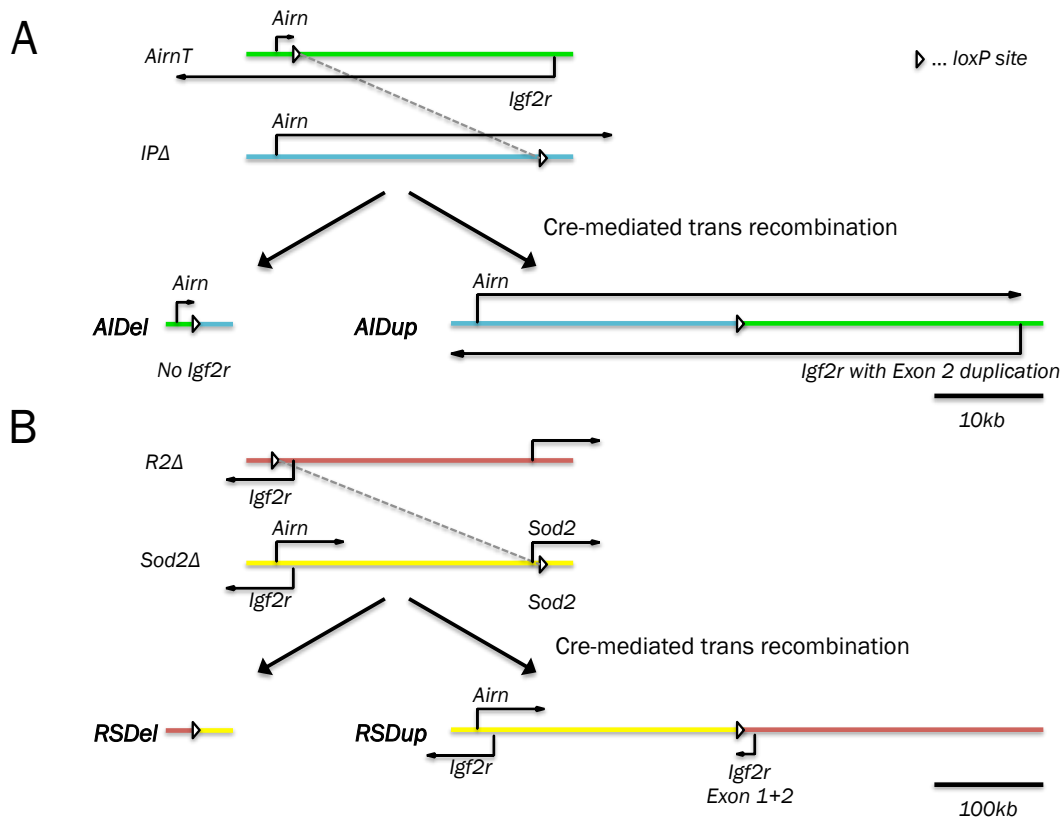
Figure 34: Schematic of the resulting new alleles of the mouse chromosomal experiment

(A) AirnT and IPΔ allele carry both a loxP site (white triangle). Cre-mediated *trans* recombination can lead to the two new mutant alleles, the AlDel or the AlDup allele. The shown genomic region is about 30kb in size.

(B) Similarly, for R2Δ and Sod2Δ alleles in the presence of the ubiquitously expressed Cre recombination event can result in RSDel or RSDup allele. The region between the loxP sites is around 270kb.

In parallel to generate the 270kb deletion/ 266kb duplication, first Sod2$^{flox/flox}$ frozen embryos were transplanted into pseudopregnant C57BL/6 females by Hans-Christian Theußl (IMBA/IMP transgenic facility) to reestablish this strain (**Figure 33**). Sod2$^{flox/flox}$ mice (F0) were crossed with Hprt-Cre mice and the offspring (F1) were screened for Sod2Δ (Hprt-Cre should cause germline deletion of Sod2$^{flox/flox}$) as well as Hprt-Cre mutant alleles by Southern Blot or PCR assays. Double mutant females were then crossed with R2Δ males. In addition, Hprt-Cre mice were crossed with R2Δ mice (Hprt-Cre x R2Δ or R2Δ x Hprt-Cre). Double mutant, female offspring (F1) were then crossed with Sod2Δ males. Triple mutant males from both crosses (F2) were then paired with FVB/NJ (+/+) females. The F3 generation was then screened for deletions or duplications between R2Δ and Sod2Δ by novel Southern Blot and

PCR assays (R2Δ/Sod2Δ-Deletion= RSDel; R2Δ/Sod2Δ-Duplication= RSDup). In RSDel *Airn* is deleted together with the *Igf2r* promoter and its first 2 exons, while the RSDup contains a wildtype *Igf2r* cluster as well as a duplicated allele including *Igf2r* exon 1+2 a and the *Airn* gene minus its promoter (**Figure 34B**).

6.2.2. Polymerase chain reaction (PCR)

The polymerase chain reaction is a biochemical technique that relies on thermal cycling to amplify a certain piece or pieces of DNA millions of times. The classical PCR mixture contains the heat-stable *Taq* polymerase (Chien et al., 1976) in an optimized buffer and at least one pair of DNA primers, short DNA fragments, which are complementary to the ends of the target DNA. A mixture of dNTPs, DNA nucleotides (adenine (A), cytosine (C), guanine (G), thymine (T)) containing triphosphate groups, for polymerization and $MgCl_2$ to activate the enzymatic activity of the polymerase are also in the mix. To improve the amplification of the PCR reaction betaine, a neutral chemical compound with a positively charged cationic functional group to reduce the formation of secondary structures of the DNA, can be added.

The typical cycling steps for a standard PCR starts with the *Initialization step*: During this phase double stranded template DNA is heated to 94-96°C for 1 to 9 min to separate or denature the DNA strands and to activate the Hot-start polymerase. This is done once before the cycling starts. *Denaturation step*: This is the first cycling step and is normally performed at 94-96°C for 20-30" to melt the DNA strand. Each of the 35 cycles started with this step. *Annealing step*: During this next phase the primers bound specifically each the sense and anti-sense strand of the template DNA. This step can be performed between 50°C and 65°C for 20-40". *Elongation step*: The cycle is finished when the DNA between the primer pair is amplified by the *Taq* polymerase at 72°C. dNTPs are part of the reaction mix and are used by the *Taq* polymerase to synthesize a new strand with a polymerization speed of approximately 1,000 bp per minute at 72°C. The annealing temperature and the length of the denaturation step and elongation step can be adapted for each PCR, as can

the number of PCR cycles. The standard protocol for PCRs done in the Barlow lab was adjusted and optimized for these new genotyping assays (**Table 10**).

AlDel

PCR Mix

H_2O	8.9	µl
5x Buffer	5.0	µl
Betaine (5M)	4.0	µl
$MgCl_2$ (25mM)	1.5	µl
dNTPs (10mM)	0.5	µl
MM_Typing_001 (10uM)	2.0	µl
MM_Typing_004 (10uM)	1.0	µl
MM_Typing_007 (10uM)	1.0	µl
Taq Polymerase	0.1	µl
DNA Template	1.0	µl
Total	**25.0**	µl

PCR Program

95°C	5'	
95°C	30"	
64°C	30"	35x
72°C	1'	
72°C	5'	
10°C	∞	

AlDup

PCR Mix

H_2O	8.9	µl
5x Buffer	5.0	µl
Betaine (5M)	4.0	µl
$MgCl_2$ (25mM)	1.5	µl
dNTPs (10mM)	0.5	µl
MM_Typing_003_2 (10uM)	1.0	µl
MM_Typing_005_2 (10uM)	1.0	µl
MM_Typing_006_2 (10uM)	2.0	µl
Taq Polymerase	0.1	µl
DNA Template	1.0	µl
Total	**25.0**	µl

PCR Program

95°C	5'	
95°C	30"	
62°C	30"	35x
72°C	1'	
72°C	5'	
10°C	∞	

RSDel

PCR Mix

H2O	8.9	µl
5x Buffer	5	µl
Betaine (5M)	4	µl
MgCl2 (25mM)	1.5	µl
dNTPs (10mM)	0.5	µl
MM_Typing_009_2 (10uM)	1	µl
MM_Typing_010 (10uM)	1	µl
MM_Typing_011_2 (10uM)	2	µl
Taq Polymerase	0.1	µl
Template	1	µl
Total	**25**	µl

RSDup

PCR Mix

H2O	9.9	µl
5x Buffer	5	µl
Betaine (5M)	4	µl
MgCl2 (25mM)	1.5	µl
dNTPs (10mM)	0.5	µl
MM_Typing_012 (10uM)	1	µl
MM_Typing_013 (10uM)	1	µl
MM_Typing_014 (10uM)	1	µl
Taq Polymerase	0.1	µl
Template	1	µl
Total	**25**	µl

PCR Program				PCR Program		
95°C	5'			95°C	5'	
95°C	30"			95°C	30"	
64°C	30"	35x		64°C	30"	35x
72°C	1'			72°C	1'	
72°C	5'			72°C	5'	
10°C	∞			10°C	∞	

Table 10: PCR protocol for all four new established mouse lines

For the AIDel, AIDup, RSDel and RSDup strain the optimized containing PCR master mixes and PCR cycling programs were summarized.

Based on this protocol new PCR assays were designed to have an alternative to Southern blot for genotyping the AIDel, AIDup, RSDel and RSDup alleles (**Table 10**). Genomic DNA was diluted 1:100 in MQ H_2O and was used as template for these PCRs. The standard reaction volume was 25µl (20µl mastermix and 5µl DNA template). The optimal condition for each assay was tested by gradient PCRs with melting temperatures between 53-64°C. Different sets of *Mus musculus* typing primers (MM_Typing) were used to identify mutant alleles. The common primers can be used together with one of two specific primers to result in different sized fragments depending on the genomic situation (**Table 11**).

AIDel

Primers	Sequence	Sense	Function
MM_Typing_001	TGGCTAGGGCTGGGGATATA AAGTT	-	Common primer
MM_Typing_004	ACTCCCAGTCATAGCTGTCC CTCT	+	AIDel primer
MM_Typing_007	GGGTATGTTGTGCTTTTATTG TCCC	+	WT primer
Fragments			
WT	557 bp	(MM_Typing_001 + MM_Typing_007)	
AIDel	298 bp	(MM_Typing_001 + MM_Typing_004)	

AIDup

Primers	Sequence	Sense	Function
MM_Typing_003_ 2	ACAGCTCAGCAGTTAGGAGT GC	-	Common primer
MM_Typing_005_ 2	TCCCAGTTCCTATCAATCGCT GA	+	WT primer
MM_Typing_006_ 2	ATTGTGGCCCAGGAGAGGTG	+	AIDup primer

Fragments

WT	467 bp	(MM_Typing_005_2 + MM_Typing_006_2)
AlDel	347 bp	(MM_Typing_003_2 + MM_Typing_006_2)

RSDel

Primers	Sequence	Sense	Function
MM_Typing_009_2	TCGGGGCTAGTGAGATGGC	-	Common primer
MM_Typing_010	GGGTGGTGGAGAACCCAAAGG	+	WT primer
MM_Typing_011_2	TCCAGGAGTGATCTGGCTGA	+	RSDel primer

Fragments

WT	430 bp	(MM_Typing_009_2 + MM_Typing_010)
RSDel	550 bp	(MM_Typing_009_2 + MM_Typing_011_2)

RSDup

Primers	Sequence	Sense	Function
MM_Typing_012	TGGCTAGGGCTGGGGATATAAAGTT	+	RSDup primer
MM_Typing_013	ACTCCCAGTCATAGCTGTCCCTCT	-	Common primer
MM_Typing_014	GGGTATGTTGTGCTTTTATTGTCCC	+	WT primer

Fragments

WT	436 bp	(MM_Typing_013 + MM_Typing_014)
RSDup	644 bp	(MM_Typing_012 + MM_Typing_013)

Table 11:PCR primers to genotype AlDel, AlDup, RSDel and RSDup

Names, sequences, orientation and function of PCR primers used to genotype new established mouse lines. In addition, the used primer combinations as well as the size of the resulting PCR fragment are shown.

6.2.3. PCR genotyping for Hprt-Cre and MADM7

Established PCR assays were used to genotype mice for the Hprt-Cre and MADM7 allele (**Table 12**). The protocol for genotyping the Hprt-Cre was downloaded from the webpage of the Jackson Laboratory (http://jaxmice.jax.org/protocolsdb/f?p=116:2:::NO:2:P2_MASTER_PROTOC OL_ID,P2_JRS_CODE:1987,006143). The original protocol for genotyping mice with the MADM7 allele was obtained from Simon Hippenmeyer and described first in 2010 (Hippenmeyer et al., 2010) (**Table 13**).

Hprt-Cre

PCR Mix

H2O	6.9	µl
5x Buffer	5	µl
Betaine (5M)	4	µl
MgCl2 (25mM)	1.5	µl
dNTPs (10mM)	0.5	µl
oIMR1084 (10µM)	2	µl
oIMR1085 (10µM)	2	µl
oIMR7338 (10µM)	1	µl
oIMR7339 (10µM)	1	µl
Taq Polymerase	0.1	µl
Template	1	µl
Total	25	µl

PCR Program

94°C	3'	
94°C	30"	
51.7°C	30"	35x
72°C	30"	
72°C	2'	
10°C	∞	

MADM7

PCR Mix

H2O	9.9	µl
5x Buffer	5	µl
Betaine (5M)	4	µl
MgCl2 (25mM)	1.5	µl
dNTPs (10mM)	0.5	µl
SH177 (10µM)	2	µl
SH193 (10µM)	2	µl
SH194 (10µM)	1	µl
Taq Polymerase	0.1	µl
Template	1	µl
Total	25	µl

PCR Program

95°C	5'	
95°C	30"	
58°C	30"	35x
72°C	50"	
72°C	10'	
10°C	∞	

Table 12: PCR protocol for Hprt-Cre and MADM7 genotyping

The optimized containing PCR master mixes and PCR cycling programs were summarized for the genotyping assay of these alleles.

Hprt-Cre

Primers	Sequence	Function
oIMR1084	GCGGTCTGGCAGTAAAAACTATC	Cre
oIMR1085	GTGAAACAGCATTGCTGTCACTT	Cre
oIMR7338	CTAGGCCACAGAATTGAAAGATCT	Wild type
oIMR7339	GTAGGTGGAAATTCTAGCATCATCC	Wild type

Fragments

WT	423 bp	(oIMR7338 + oIMR7339)
Cre	100 bp	(oIMR1084 + oIMR1085)

MADM7

Primers	Sequence
SH177	TCAATGGGCGGGGGTCGTT
SH193	GGCTGAACTGGAGGATTGTCTAGTTTTCC
SH194	GCTCTTGCTCTCATTTCTCAGTTCAAACCGCGTATC

Fragments	
WT	470 bp
MADM7	570 bp

Table 13: PCR primers to genotype for Hprt-Cre and MADM7

Names, sequences, orientation and function of PCR primers used to genotype for these alleles were shown.

6.2.4.   Cloning of Southern blot probes in E. coli (DH5α)

PCR (**Table 14**) was done with specific primer pairs designed to amplify novel Southern blot probes (**Table 15**) using a 1:100 diluted BAC (RP23-309H20). The PCR product was cleaned up with *WIZARD® SV Gel and PCR Clean-Up System* according to the user manual and the concentration was measured by a *Nanodrop ®-1000 spectrophotometer*. The purified PCR fragment was sent for Sanger sequencing (*Microsynth*) to confirm the PCR product. The concentration of the probe was used to calculate the amount of needed DNA to end up with a 1:3 ration between insert (Southern blot probe) and vector (pGEM®-T Easy Vector Systems; Promega). Following the standard protocol, a ligation (T4 DNA ligase (3U/µl); Promega) was done at 4°C o/n. Dialysis (*Membrane Filter, 0.05µm VMWP; Merck Millipore*) was performed with the ligation mix to reduce the salt concentration. The mix was electroporated into electro-competent bacteria (*E. coli* (DH5α)) and incubated in a thermomixer (300rpm; 60min; 37°C). The amount of bacteria was split into two and separately plated on circlegrow agar plates (+ 1:1000 AMP; 1:625 X-Gal; 1:200 IPTG) for incubation at 37°C o/n.

PCR Mix

H$_2$O	21.8	μl
5x Buffer	10.0	μl
Betaine (5M)	8.0	μl
MgCl$_2$ (25mM)	3.0	μl
dNTPs (10mM)	1.0	μl
Forward Primer	2.0	μl
cxReverse Primer	2.0	μl
Taq Polymerase	0.2	μl
DNA Template	2.0	μl
Total	**50.0**	μl

PCR Program

94°C	5'	
94°C	30"	
60°C	30"	35x
72°C	1'	
72°C	5'	
10°C	∞	

Table 14: PCR protocol for amplifying new Southern blot probes

Southern Blot Probes

Probe Name	Forward Primer	Reverse Primer	Size
AlDel	GGTGCTAAAGATGGAAGCCAGGG	GGCTGGCAGATGAGACCCAA	550 bp
AlDup	TTCAATGAGCAGCCTCAGCC	AGTTCATGCTGGGGTGGAGT	444 bp
Kodel	CTCCAGATTGGCAGTTCACACAGTG	ACGTTTTGGGGTTGCTGGGT	325 bp
Msi	CGCGTTAGAGGATTCCGCAAA	AATCCTCCCCTTGTGCAGTTTGC	1013 bp
RSDel	AGTTTCACTCCAATTCTCACC	TCCTCTCCAAACAGGTGACTC	332 bp
OT KnpI / RSDup	GGTACCTACAACTTCAAATTGGGC	GGTACCCTTTGGCCATGAAG	1034 bp
Sod2	CCTTGCGGTGGGACGTAGGC	GCCATGCAGGGAGCAGGACC	666 bp

Table 15: Southern blot probes for new mouse strains

List of new designed Southern blot probes for new established mouse lines (AlDel, AlDup, RSDel and RSDup) as well as probes for starting strains (AirnT, IPΔ, R2Δ and Sod2). The forward and reverse primers of a probe amplify a DNA fragment of the stated size.

White colonies were picked and grown shaking (200rpm) in 3ml LB+AMP (1:1000) at 37°C o/n. Miniprep (*QIAgen Spin Miniprep Kit*) was done with 1.5ml of the o/n cultures according to the user manual. The re-suspended DNA was digested (DNA single enzyme digest; EcoRI) and the quality was

checked on 2% agarose gel in 1xTAE (60min; 80V). One sample was chosen and checked by Sanger sequencing (*Microsynth*) for sequence identity with the predicted probe. The remaining second half of the o/n culture was used to inoculate 100ml LB+AMP (1:1000) and grown o/n (200rpm; 37°C). 500μl of the overnight culture was used to prepare a glycerol stock of this bacterial strain. A midiprep was conducted to isolate the plasmid from the remaining culture using *QIAfilter Plasmid Midi Kit (Qiagen)* according to the user manual. The concentration was measured by *Nanodrop ®-1000 spectrophotometer*. 30μg of plasmid DNA was digested by EcoRI incubated at 37°C o/n. The insert was separated from the plasmid on a 2% agarose gel in 1x TAE (90min; 80V), and cut out and cleaned up with *WIZARD® SV Gel and PCR Clean-Up System* according to the user manual. The concentration of the new probe was measured (*Nanodrop ®-1000 spectrophotometer*) and adjusted to a final concentration of around 20ng/μl. 20ng of the probe DNA (1 μl) was used as a template to synthesis a Southern blot probe. For strains maintained in the lab (AirnT, IPΔ and R2Δ), and Sod2$^{flox/flox}$ / Sod2Δ the Southern blot probes were previously prepared in the same way and were available in the lab (**Table 15**) (Prepared by Justyna Konecka, Ruth Klement and Daniela Mayer).

6.2.5. Genomic DNA isolation

Around 1 cm of the tail of a three weeks old mouse was cut off and placed into a 1.5ml tube. 700μl of DNA lysis buffer was added and incubated at 55°C o/n. 300μl of saturated NaCl solution were added to the tube before shaking (20sec). After the formation of bubbles the tube was centrifuged (16,100xg; 10min; RT) and supernatant of the DNA lysis solution was transferred into new 1.5ml tubes containing 600μl (0.6vol) isopropanol. The genomic DNA was precipitated by inverting the tube a couple of times. The tube was centrifuged (16,100xg; 7min; RT) and the supernatant discarded by carefully pouring out. To wash the DNA pellet, 800μl of EtOH (70%) was added and the tube was shaken vigorously (20sec). The tube was kept at RT for at least 15min and shaken every 5min to wash the DNA pellet. Afterwards the tube was centrifuged (16,100xg; 7min; RT), the supernatant was poured out and centrifuged again (16,100xg; 1min; RT) to collect remaining EtOH at the

bottom. The liquid was pipetted off and the pellet was air-dried for 3-5min. 120µl of TE buffer was used to dissolve the genomic DNA by incubating at 55°C (o/n).

6.2.6. Digest of genomic DNA

Mouse Strain	Digest			Agarose Gel			
	Restriction Enzyme	Buffer	Probe	Size	Running Time	Tension	Rows
AirnT	EcoRI	EcoRI	Msi	24 cm	5h	5 V/cm	2
AlDel	BglII	O(range)	AlDel	24 cm	3h	5 V/cm	2
AlDup	AanI	Tango 1x	AlDup	24 cm	5h	5 V/cm	2
IPΔ	BglII	O(range)	KodelOT	24 cm	4h	5 V/cm	2
R2Δ	BglII	O(range)	KnpI	24 cm	4h	5 V/cm	2
RSDel	BglII	O(range)	RSDelOT	24 cm	3h	5 V/cm	2
RSDup	BglII	O(range)	KnpI	24 cm	o/n + 4h	1 + 5 V/cm	1
Sod2Δ	EcoRV	R(ed)	Sod2	24 cm	o/n + 4h	1 + 5 V/cm	1
Thp	EcoRI + MluI	R(ed)	MSi	24 cm	5h	5 V/cm	2

Table 16: Southern blot setups for different mouse strains

The list of mouse strains contains established and new created strain. Restriction enzymes (Fermentas) were always used with the recommended buffer (Fermentas). DNA probe was cut out and purified from a vector. 400ml of a 0.8% agarose in 1xTBE was used for these blots.

The type of restriction enzyme used as well as the gel running time differed between Southern blot assays (**Table 16**) and was optimized for each assay. The genomic DNA was incubated at 55°C (30min) to ensure that it was completely dissolved before usage. 30µl of the pre-heated DNA was transferred into a new tube and 10µl of Single Enzyme Digest mix or Double Enzyme Digest mix were added. The mixture was incubated at 37°C (o/n or a minimum time of 6hrs). To the 40µl DNA digest, 4µl red loading dye was added, vortexed, and spun down. The entire sample was loaded onto a 0.8% agarose in 1x TBE gel and the running time and voltage was set according to the Southern Blot assay (**Table 16**). For all assays a 1kb ladder was used (GeneRuler™ 1kb DNA Ladder; Thermo Scientific). The gel was placed for 60min in an EtBr bath to stain the digested DNA and then photographed. After

this quality check the gel was incubated 2 x 30min in a denaturation solution on a shaker. During this time, the bands of the ladder were used to calculate the optimal size of the membrane (Amersham Hybond™-XL, GE Healthcare). The Southern Blot sandwich was assembled (**Figure 35**) and the blotting was done onto a membrane o/n at RT.
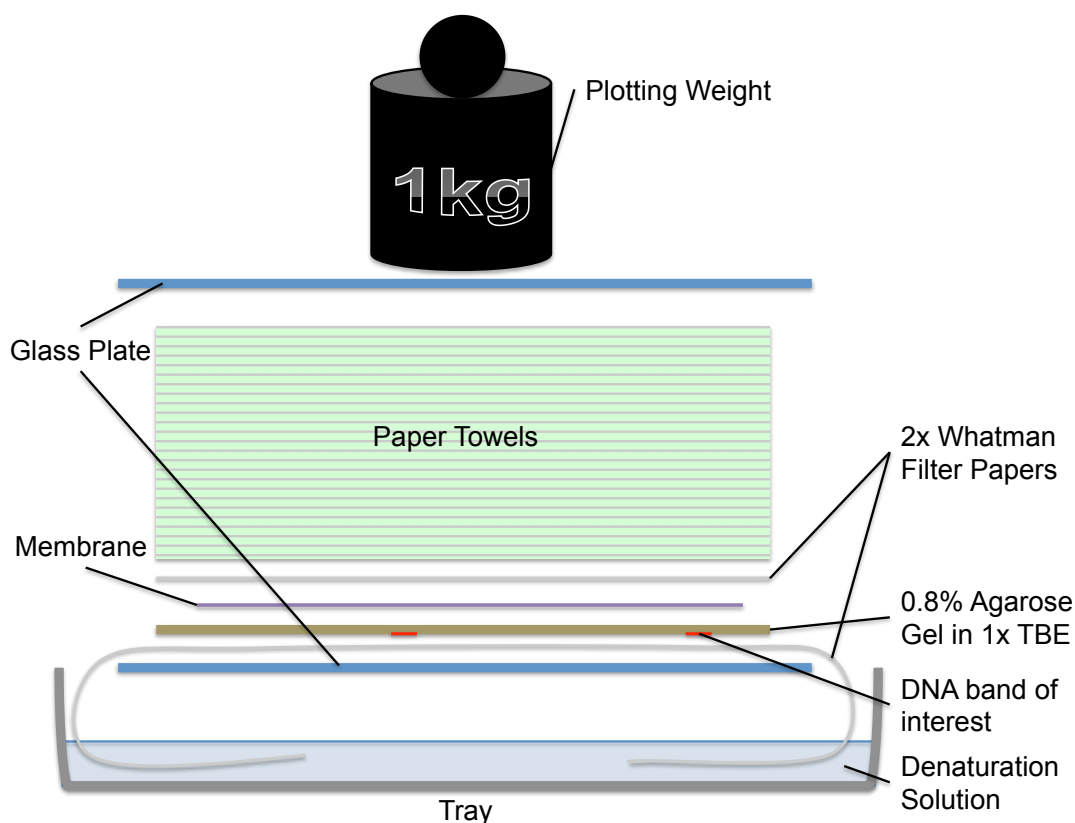

Figure 35: Assembly of the Southern blot sandwich

The bottom filter papers were pre-wet in denaturation solution and than placed on the bottom glass plate, allowing the filter papers to contact the denaturation solution below. The gel was carefully removed out of the running tray, flipped by 180° and than placed with the upper side on the filter paper. The membrane was placed on top of the gel in a way that the entire sample area was covered. On top of the membrane were placed pre-wet filter papers, a stack of dry paper towels, a heavy glass plate and a blotting weight. The blotting was done at RT o/n.

6.2.7. Labeling of DNA probe

20ng of a purified DNA probe was put into a screw top tube and filled up with MQ H_2O to a final volume of 14µl. The Cold Mix (20µl LS; 6µl CTG-Mix; 1µl Klenow fragment) was freshly prepared. In the Radioactive Lab the DNA fragment was denatured by heating (5min; 100°C) and then cooled down on

ice (2min). In parallel 2μl α-^{32}P-dATP (+ 1μl per seven days after delivering) was added to the Cold Mix, vortexed and spun down. This Hot Mix was then added to the probe on ice, vortexed and spun down. The mix was incubated at RT (6–18hrs.). The probe was cleaned with a Sephadex G50 spin column. The column was constructed by pushing aquarium wool into a 1ml syringe, filling the syringe with Sephadex G50 beads in TE buffer, and then spinning down the syringe in a 15ml Falcon tube (1450xg; 3min). The excess TE buffer was spun out and Sephadex G50 column was compressed. In the radioactive lab 60μl of TE buffer was added to the ^{32}P labeled probe. The entire amount was loaded into the Sephadex G50 column and then spun down in a 15ml Falcon (1450xg; 3min) to remove unincorporated radioactive nucleotides. 1μl of the cleaned probe was added to 99μl TE buffer to measure the counts on a scintillation counter (Tri-Carb 2810 TR, PerkinElmer) as a quality control. The probe was stored in a Plexiglas box for a short time at RT or o/n at 4°C.

6.2.8. Hybridization

The Southern Blot sandwich was dissembled and the properly labeled membrane was shaken in 200ml 20mM Na_2HPO_4 (2min, RT) to neutralize it. Afterwards the membrane was placed into a pre-heated hybridization tube (65°C), 20ml Church Buffer was added and the membrane pre-hybridized in the hybridization oven (65°C) for 30min to 3hrs. The cleaned DNA probe was denatured by heating (5min; 100°C) and afterwards cooled down on ice (2min). The Church Buffer for pre-hybridization of the membrane was discarded and 20ml fresh Church Buffer was added. The denatured probe was then added to the hybridization tube and incubated rotating in the hybridization oven (65°C) for 18hrs. On the following day, the probe was removed and 30ml pre-warmed (65°C) Church Wash Buffer was added to the membrane. The tube was rotated by hand for a couple of times and then the buffer was removed. 30ml fresh pre-warmed (65°C) Church Wash Buffer was then added to wash the membrane in the hybridization oven for 30min (65°C). The buffer was changed once and the blot was washed again for 30min. After that, the liquid was discarded, the membrane removed from the tube, wrapped in tissues and dried by air. The membrane was then sealed in foil

and placed into a film cassette with the DNA side up. A BAS-MS2025 Imaging Plate (Fujifilm) was added on the top of the membrane and exposed o/n or up to three days in an X-ray cassettes (*Dr. Goos-Suprema*).

6.2.9. Scanning

The imaging plate was scanned (Typhoon 8600 Variable Mode Image scanner) and linear adjustments of brightness and contrast were done in Photoshop to optimize the image.

6.2.10. Stripping Blot

Southern Blot membranes can be re-used by stripping and hybridized again. A stained membrane was incubated shaking at RT for 2 x 30min in 200ml 40mM NaOH, 2 x 5min in 200ml MQ H_2O and 2 x 10min in 200ml 20mM Na_2HPO_4. After this treatment the membrane was directly placed into hybridization tube and pre-hybridized in 20ml pre-heated Church Buffer (65°C) for at least 30min. The buffer was discarded and 20ml fresh preheated Church Buffer (65°C) was added. The purified P^{32} labeled denatured probe was added to the stripped blot and the standard Southern blot protocol was followed.

7. Materials

7.1. Mouse strains

The "Hair Pin tail" (T^{hp}) deletion is a naturally occurring 5Mb deletion of murine chromosome 17 that contains the entire *Igf2r* cluster (Barlow et al., 1991). The deletion arose on an AKR/J background at the Jackson Laboratory in 1962 (Johnson, 1974), and was backcrossed onto a C3H/HSnJ background (Jax® Mice and Services http://jaxmice.jax.org/strain/001053.html). In the Barlow lab the T^{hp} mouse was first crossed to BALB/C, but since 1996 has been backcrossed to FVB/NJ meaning this mouse has an FVB/NJ background. The T^{hp} mouse strain has to be maintained paternally. A maternal inheritance causes an embryonic lethal phenotype between E15.5 to E17.5.

In the AirnT strain a 1.2kb long poly-A-cassette was inserted around 3kb downstream of the *Airn* promoter start site, which resulted in a truncation of the lncRNA *Airn* (Sleutels et al., 2002). The 118kb long RNA was reduced to around 3kb. The allele was constructed by homologous recombination in 129 ES cells than backcrossed onto a FVB/NJ background. The silencing effect of *Airn* on *Igf2r* is lost. Consequently *Igf2r* is expressed biallelic and the offspring is 15% smaller (Sleutels et al., 2002). Males and females can be used for maintenance.

The IPΔ allele carries a 4.7kb deletion including the promoter region of *Igf2r* (Sleutels et al., 2003). The allele was constructed by homologous recombination in 129 ES cells onto a FVB/NJ background for more than 50 generations. This allele is lethal on maternal inheritance, but shows no phenotype if maintained paternally.

R2Δ mice carry a 3.6kb deletion including the *Igf2r*-ICE, which contains the promoter for *Airn.* The allele was constructed by homologous recombination in 129 ES cells and backcrossed onto a FVB/NJ background. This results in a biallelic expression of *Igf2r* and a reduction in body size of around 20% (Wutz et al., 2001). Maintaining can be done maternally or paternally.

The Sod2**flox/flox** mice carry a floxed allele with loxP sites flanking exon3 of *Sod2* on chromosome 17. The allele was constructed by homologous recombination in ES#298 cells and transplanted into C57BL/6CrSlc females. The maintenance was done with C57BL/6 background. In Sod2Δ exon 3 is floxed out by Cre and this leaded to a 0.5kb deletion (Ikegami et al., 2002). Maintenance can be done maternally or paternally.

FVB/NJ is an inbred mouse strain used to backcross to all mouse strains in this approach. The AirnT and IPΔ mouse strains were maintained in the Barlow lab since 2002 and 2003. R2Δ mouse strain was maintained in the lab since 2001. Sod2**flox/flox** mice we obtained from Takahiko Shimizu (Department of Advanced Aging Medicine, Chiba University, Japan) as frozen embryos and transplanted into pseudopregnant C57BL/6 females by Hans-Christian Theußl

(IMBA/IMP transgenic facility, Vienna, Austria) to reestablish this strain. Mice with the Hprt-Cre allele and MADM7 allele (Hippenmeyer et al., 2013) were provided by Simon Hippenmeyer (IST, Klosterneuburg, Austria).

7.2. Solutions

5x TEN 9

> 250mM Tris pH 9.0
> 100 mM EDTA pH 8.0
> 200mM NaCl

Ampicillin stock

> 1g ampicillin (50mg/ml)
> 50ml H_2O
> 50ml EtOH (pure)

Circlegrow Agar

> 16g circlegrow broth
> 6g agar
> 400ml H_2O

Church buffer

> 20mM Na_2HPO_4
> 7% SDS
> 1mM EDTA

Church wash buffer

> 20mM Na_2HPO_4
> 1% SDS

Cold mix

> 20µl LS
> 6µl CTG mix
> 1µl Klenow fragment

CTG-mix

> 100µM dTTP
> 100µM dCTP
> 100µM dGTP
> BSA (2mg/ml)

*Denaturation
solution*

0.5M NaOH
1.5M NaCl

DNA single digest

4µl MQ H2O
4µl 10xBuffer
2µl Restriction Enzyme 10U/µl
30µl DNA

DNA double digest

2µl MQ H2O
4µl 10xBuffer
2µl Restriction Enzyme I 10U/µl
2µl Restriction Enzyme II 10U/µl
30µl DNA

DNA lysis buffer

1x TEN 9
1% SDS
Proteinase K (0.5mgs/ml)

IPTG

1g IPTG
41.7ml H_2O

LB broth medium

10g LB-medium
400ml MQ H_2O

LS

25ml 1M Hepes pH 6.6
1ml OL (hexamer primer)
25ml TM

OL

50U pd(N)$_6$ random hexamer
primer
1.6ml 1xTE buffer pH 8.0

10x red loading dye	
	0.125g 0.5% xylenol orange
	7.5ml 30% glycerol
	17.5ml 1x TAE

Sephadex G50	
	10g Sephadex G50
	200ml 1xTE buffer pH 8.0

TAE buffer (50x)	
	242g Tris
	57.1ml acetic acid
	100ml 0.5M EDTA pH 8.0
	with MQ H_2O up to 1000ml

TBE Bbuffer (10x)	
	108g Tris
	55g Boric Acid
	40ml 0.5M EDTA pH 8.0
	with MQ H_2O up to 1000ml

TE Buffer	
	10ml 1M Tris pH 8.0
	2ml 0.5M EDTA
	with MQ H_2O up to 1000ml

TM	
	25ml 1M Tris ph 8.0
	10ml 25mM $MgCl_2$
	0.35ml β-Mercaptoethanol

X-Gal	
	50mg X-Gal
	1ml N,N-dimethylformamid

7.3. Chemicals and consumables

100μM dCTP	Bioron
100μM dGTP	Bioron
100μM dTTP	Bioron
10mM dNTPs	Thermo Scientific
2-Propanol	Merck
25mM Magnesium chloride (MgCl2)	Promega
2x MESA GREEN qPCR MasterMix Plus	Eurogentec
2x qPCR MasterMix Plus	Eurogentec
50U pd(N)$_6$ Random Hexamer Primer	Ambion
Acetic Acid 100%	Roth
Agar for molecular biology	AppliChem
Amersham Hybond^{TM}-XL membrane	GE Healthcare
Betaine	Sigma
Boric Acid	AppliChem
Bovine serum albumin (BSA) (2mg/ml)	Ambion
Circlegrow (Powder)	MP Biomedicals
Disodium hydrogen phosphate (Na$_2$HPO$_{4)}$	Merck
Ethanol (EtOH) 96%	Merck
Ethylenediaminetetraacetic acid (EDTA)	AppliChem
Etidium Bromide Solution 1%	AppliChem
GeneRulerTM 100 bp Plus DNA Ladder	Thermo Scientific
GeneRulerTM 1kb DNA Ladder	Thermo Scientific
Glycerol anhydrous for molecular biology	AppliChem
Hepes Buffer	AppliChem
Lambda Mix Marker,19	Thermo Scientific
LB-medium	MP Biomedicals
Membrane Filter, 0.05μm VMWP	Promega
MQ H$_2$O	Merck Millipore
N,N-dimethylformamid	Sigma
Phenol stable : Chloroform : Isoamylalcohol 25:24:1	AppliChem
Sephadex G50 Medium	Amersham Bioscience
Sodium Chlorid (NaCl) pure	AppliChem
Sodium dodecyl sulfate (SDS)	AppliChem
Sodium hydroxide (NaOH)	Merck
Tris for molecular biology	AppliChem
Water for embryo transfer	Sigma
Whatman Filter Papers	Whatman
X-Gal	Roth
α-^{32}P-dATP	Hartmann
β-Mercaptoethanol	Sigma

7.4. Enzymes and buffers

10x Buffer Orange	Fermentas
10x Buffer Tango 1x	Fermentas
5x Colorless GoTaq Flexi Buffer	Promega
5x DNA Ligase Reaction Buffer	Promega
5x Green GoTaq Flexi Buffer	Promega
AanI (10U/µl)	Fermentas
BglII (10U/µl)	Fermentas
Buffer Red	Fermentas
EcoRI (10U/µl)	Fermentas
EcoRI buffer	Fermentas
EcoRV (10U/µl)	Fermentas
GoTaq Polymerase (5U/µl)	Promega
Klenow fragment (2U/µl)	Promega
MluI (10U/µl)	Fermentas
pGEM®-T Easy Vector Systems	Promega
Proteinase K (0.5mgs/ml)	AppliChem
T4 DNA ligase (3U/µl)	Promega

7.5. Appendix

Listed DNA sequences of used or sequenced pieces. Big letters representing genomic DNA areas (mouse), while small letters stand for non-genomic vector pieces. Bold DNA sequences are loxP sites.

Appendix 1: Sense orientation of loxP site (+)

> **ataacttcgtatagcatacattatacgaagttat**

Appendix 2: Forward primer to control loxP site on AirnT allele by AirnT_LoxP_F

> cgataagcttcgagggaccta**ataacttcgtatagcatacattatacgaagttat**attaaggggttattgaatatgatcgga attcctgcagcccgggGGATCCCAACCAATTATTCCCTCAACAAGAGCCCTGCAATTTTTC ATTCCAGTGGTATTGGAACTGGAGCCAACACCTCTCCTGGGCCACAATATGACAA CCTACAGTCCCTTGAAAACATAAGGGTGGATGGGGAGACACCTACTAGTCCTTCC TCAAATACTGTTGTGTTGATTTTGTTAGCCCAGGAGTACAGGCAGAGGCA

Appendix 3: Reverse primer to control loxP site on AirnT allele by AirnT_LoxP_R

> GTATTTGAGGAAGGACTAGTAGGTGTCTCCCCATCCACCCTTATGTTTTCAAGGGA CTGTAGGTTGTCATATTGTGGCCCAGGAGAGGTGTTGGCTCCAGTTCCAATACCA CTGGAATGAAAAATTGCAGGGCTCTTGTTGAGGGAATAATTGGTTGGGATCCcccgg gctgcaggaattccgatcatattcaataaccccttaat**ataacttcgtataatgtatgctatacgaagttat**taggtccctcg aagcttatcgataccgtcgacctcgagggatctccataagagaagagggacagc

Appendix 4: AirnT poly-A cassette sequence

> agaacttcagggtgagtttgggggacccttgattgttctttnttttttcgctattgtaaaattcatgttnnatggaggggggcaaagttt tcagggtgttgtttagaatgggaagatgtcccttgtatcaccatggaccctcatgataaattttgtttctttcactttctactctgttga

caaccattgtctcctcttattttcttttcattttctgtaactttttcgttaaactttagcttgcatttgtaacgaatttttaaattcacttttgtt
tatttgtcagattgtaagtactttctctaatcacttttttttcaaggcaatcagggtatattatattgtacttcagcacagttttagag
aacaattgttataattaaatgataaggtagaatatttctgcatataaattctggctggcgtggaaatattcttattggtagaaac
aactacaccctggtcatcatcctgcctttctctttatggttacaatgatatacactgtttgagatgaggataaaatactctgagtc
caaaccgggcccctctgctaaccatgttcatgccttcttctctttcctacagctcctgggcaacgtgctggttgttgtgctgtctc
atcattttggcaaagaattcactcctcaggtgcaggctgcctatcagaaggtggtggctggtgtggccaatgccctggctca
caaataccactgagatcttttccctctgccaaaaattatggggacatcatgaagcccttgagcatctgacttctggctaat
aaaggaaatttattttcattgcaatagtgtgttggaatttttgtgtctctcactcggaaggacatatgggagggcaaatcattta
aaacatcagaatgagtatttggttagagtttggcaacatatgcccatatgctggctgccatgaacaaaggttggctataaa
gaggtcatcagtatatgaaacagcccctgctgtccattccttattccatagaaaagccttgacttgaggttagatttttttata
ttttgttttgtgttatttttttctttaacatccctaaaattttccttacatgtttactagccagattttcctcctctcctgactactcccagt
catagctgtccctcatctcttatggagatccctcgaggtcgacggtatcgataagcttcgagggacctaataacttcgtatag
catacattatncgaagttatattaagggttattgaatatgatcggaattcctgcagcccggg

Appendix 5: Anti-sense orientation of loxP site (-)

ataacttcgtataatgtatgctatacgaagttat

Appendix 6: Forward primer to control loxP site on IPΔ allele by IPΔ_loxP_F

CTAAACTGGGTGGGATTAAAGTGCCGCCTTTAATCCTGGCACTTGGGAGCACACA
GGCAGACAGATTTATGGGTTCAAGGCCAGTCTGGTCTACAGAGTGAGTTCAAGAA
TAGCCAGGGTGACACAGAGCAACCCTATCTCCCCTCCTTCCCCTCACCAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

Appendix 7: Reverse primer to control loxP site on IPΔ allele by IPΔ_loxP_R

GGCCCTGGCTTCCATCTTTAGCACCACAGACACAAATGTGTCTCTCTGTGTGTGTC
TCTGTGTGTATAAGCTTCAATGAAAAGCTagggctgcaggaattccgatcatattcaataacccttaat
ataacttcgtataatgtatgctatacgaagttattaggtccctcgaagcttaTCGAATTTATGTCTGTTTACC
ACATACCTTACGGTGTCCTTGAAGAAGGCACCAGattcCCCTGAAACTGAAGTTACA
GATGGCTGTGAGGCTCCACGTGGATGCTGGGCATTGAACATGGATCCTGTGCGA
AAGCAGCAAGCACTCCTAACTGCTGAGCTGTTAACTGTCTAGTCTCCTAGAGTTAA
GTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGGGGGGG

Appendix 8: Forward primer to control loxP site on R2Δ allele by R2Δ_loxP_F

GCTATTTCTGTAAGGCCAGGCCACAGCAGCTGCCCCAGACCCCTCCCATGAGCCT
GCTTCCCCCAGACTCCTCCCATGAGCCTACTTCCTTAATCTATAGAATCTATCTATA
CAGAAGACTCACTTGAACCTCACAAAAGAATCAATGTAGTCATAAGTAAAGATTTAT
TGAACACATGGGATGGAGTTAtcgaggaattccgatcatattcaataacccttaat**ataacttcgtataatgt**
atgctatacgaagttattaggtccctcgaagaggttcactagtGATCCCAACCAATTATTCCCTCAACA
AGAGCCCTGCAATTTTTCATTCCAGTGGTATTGGAACTGGAGCCAACACCTCTCCT
GGGCCACAATATGACAACCTACAGTCCCTTGAAAACATAAGGGTGGATGGGGAGA
CACCTACTAGTCCTTCCTCAAATACTGTTGTGTTGATTTTGTTAGCCCAGGAGTAC
AGGCAGAGGCACCTACAAAGG

Appendix 9: Reverse primer to control loxP site on R2Δ allele by R2Δ_loxP_R

GTATTTGAGGAAGGACTAGTAGGTGTCTCCCCATCCACCCTTATGTTTTCAAGGGA
CTGTAGGTTGTCATATTGTGGCCCAGGAGAGGTGTTGGCTCCAGTTCCAATACCA
CTGGAATGAAAAATTGCAGGGCTCTTGTTGAGGGAATAATTGGTTGGGATCactagtg
aacctcttcgagggaccta**ataacttcgtatagcatacattatacgaagttat**attaagggttattgaatatgatcggaatt
cctcgaTAACTCCATCCCATGTGTTCAATAAATCTTTACTTATGACTACATTGATTCTT
TTGTGAGGTTCAAGTGAGTCTTCTGTATAGATAGATTCTATAGATTAAGGAAGTAG
GCTCATGGGAGGAGTCTGGGGGAAGCAGGCTCATGGGAGGGGTCTGGGGCAGC
TGCTGTGGCCTGGCCTTACAGAAATAGCACAGAGGTCACCAGGCTATTAACCACC
CCTGCCCCCAGCCTCCTGGGTGGAAAAACAGGGAGGCTGGGGGCAGGGGTGGT
TAATAGCCTGGTGACCTCTGT

Appendix 10: Forward primer to control second loxP site in Sod2^{flox/flox} (Sod2_loxP_F_2P)

GTAGGAATGATTGATGTTTTTTCTTTTTTTAAGATTTATTTATTTATTATATGTNCGTT
ATTATATGCAACTACACTGTANCTGTCTTCAGACACACCAGGAGAGGGSGTCAGAT
CTCATTACGGGNGGTTGTGAGCCACCATGTGTTGCTGGtcgacctcgac**ataacttcgtataa
tgtatgctatacgaagttat**gtcgccggtatcgatgaattcctgcagccctcgaGCTGAGCCATCTCACTAG
CCCCGATTGTTGGTTTTTTCCCCCTCTCTTTCTTAACTTTTAAAATTTTCTTTTACAA
GAGTTTTACTGTGTAGTTCTTGGAAATCAGTATGTAGACAGTCTTTATAAATTTGTT
TATTTTTTTGAGGCAGTTTATTATATAATTTTTATTATGTGTATATATATATCTGTG
TGNGGAGGTGACTTTCTTAA

Appendix 11: Reverse primer to control second loxP site in Sod2^{flox/flox} (Sod2_loxP_R_2P)

(TAAACTGCCTCAAAAAAATAAACAAATTTATAAAGACTGTCTACATACTGNTTTCCA
AGAACTACACAGTAAAACTCTTGTAAAAGAAAATTTTAAAAGTTAAGAAAGAGAGG
GGGAAAAAACCAACAATCGGGGCTAGTGAGATGGCTCAGCtcgagggctgcaggaattcat
cgataccggcgac**ataacttcgtatagcatacattatacgaagttat**gtcgaggtcGACCAGCAACACATG
GTGGCTCACANCCACCCGTAATGAGATCTGACGCCCTCTCCTGGTGTGTCTGAAG
ACAGCTACAGTGTANTTGCATATAATAACGTACATATAATAAATAAATAAATCTTAA
AAAAANAAAAAACATCAATCATTCCTACTGTCCACCAAATATGAAGCTGTAGCAGC
TTATACCAACCTTTGGGTTCTCCACCACCCCAAAAAGGGTGATNTNGNTGCTACAG
CTTATATTTGGTGGACGNAGGAATGATNGANGTTTTTTCTTTTTTTAAGATTTATTTA
TTTATTATNTGTACGTTATTATATGC)

Appendix 12: AlDel_probe (chr17:12,772,574-12,773,123; GRCm38/mm10)

GGTGCTAAAGATGGAAGCCAGGGCCTTGGCTCTTAAGGCAAGAGCTTTACTAAAC
ACTAACTTTATATCCCCAGCCCTAGCCAGAAATATTTTGATCTATATCTGTCTTTCA
CAATGTTCTAGACAAGGCATGCTATGACTTATATTGGCATCCTTACCAAGAAGTTC
ATCTTAACTTAAATCCGCATAGTAAAGCATGATTGACAAAAACTGCACAGGAGGAT
TCCAAACGCTCTGCAAACCTCCAGCACTCTCCCTCCCCGTGAGACACTGCCCTAG
AGCACTGGATCATGCTATAGCCCATGGGTCGAATCAGGTCACACAGCTATTTTGTA
AATATTTTTACTGGGATACAGCCATCCCCATTGCTTATGTGTCGGACTGGCTGATT
TTGCCCACAAGAGCAGAGATAATAGAAACCATGGCCTGAAGAGCCTAAAGTATCT
ACATTTAACCATGACCCTCACCAAGAAAACTCACCCAACACCTCTTTTAGGTTCCC
AGTTCCCAGGAGTGCTGCCTTTAAACCAGCAGTTTGGGTCTCATCTGC

Appendix 13: AlDup_probe (chr17:12,746,073-12,746,516; GRCm38/mm10)

TTCAATGAGCAGCCTCAGCCACTCAAGATGGAGTCTGAGTTCTTCCCTCACTTCTT
ACAATGCCCACCAAGATGTCCTAGAGCATACTTACCTTATACTAAAAACAATATTTG
TATTTTGCTAATTAAAACAATGAATCCTGTAGAGACTCCAGAGTCACTACAGGACC
CATTTTCACTTCTAATCACTTAAAAGTTTAGGTCAGATGGTTTATGGAGCACAAACT
ACATGGAGGGAGCCATGCAGTCAGTAAGTAAATCTTGGTTCTGTACATCCAGTGT
CCCTTCTTCATTCCGTCAGGTAGATACAGTTACACACGCTCTGATAGGCTCCATAA
TGGCTCTGATCAGATGTGGGCTCCTCCTAAAGACTCGTGTGCTGCCCACAGCTCA
GCCTCATCTCTCTCAGATCCTGAAATGGGCCAACTCCACCCCAGCATGAACT

Appendix 14: RSDel_probe (chr17:12,740,089-12,740,420; GRCm38/mm10)

AGTTTCACTCCAATTCTCACCATCTTTAAGACACATTTAGTTGGTGGTTCTGGGTG
GTAGTGACGTGATGTGAATATAAATTGGGGGGAGGGGGAAATTGACCATTTAGTC
CTCCAGGATCAGAATTCCTGTAACAAGAACACCAGAAGGGGAAAATGTTCTTAACA
AAGAGCTTGAAGAGGGCCCCACTCACCAGGGTCTTCCCACACAGAAATGTAATGC
TAGTCTGGATTCTGTGCTGTGAATCTGAAGGCTGGCAGCCCATAGTGGTGTTGAA
TTCCAGGAGTGATCTGGCTGAACTTCTCAAAAGTGAGTCACCTGTTTGGAGAGGA

Appendix 15: RSDup probe (chr17:12,744,659-12,745,692; GRCm38/mm10)

GGTACCTACAACTTCAAATTGGGCTTATACATCCATGGGGCAAAGAGCCTCCTCC
AAACTTTACAACTGTCATTTCTGTTCCGATGAGCACAAAACCTCAAGACTGAGATA
CCACAAAGAGAGTCAAGGACAGTCACCCCTTCTTTAATCTCAACACTCCTCTCATA
GAAACACTACCATTGGCAAATTCCAGCCATCTCTGATCCAGGGCTGGGGAAGGAC

TCCTGGATTTGTCTGAGGGCTGCCAAATAACTTATACTAGGGTAGGCATTCAACTT
AAGACACACCAACCATGAACATATAAATCTGTTAGCAGGACCCCTGGCTGGGACA
GCTGGGGAGACCAACTAAATTCAGTACCAGAAACAAATTGTCCCCATTTGCCTTGC
CTTGGGAACTTGGGACTTATGTGCACATCAGGACAACACTCACATGATTGCTACTC
CAGAGGATACTCTTTCTCACCCTGGACTAAGATGTTCCAGGGGTGGGAAGTCTAC
AGACATCTTATCTGTCCATAACAGTCAGAGAAAGAGAAACTTCAAGCTATTCCCTG
TCCACCTGGGAAAGACCTAGCAGGGCCATCAGCTACGAGCACACCAGTGACAGG
GCTCCAACAGGTATAACACAGACAACTGACCAGTTAGACTAGGAAAGACCTGACC
CCAGTATAGCAAAGTATGGTCACCTTACAAAATCCAATACACGCCCTGGCTACAGC
GATCCTCCAGAAGAGTGGGCACTAGATGTAACAGCTGCTGGAAGCGGAGCATCT
GTGCCTTCCTAGGGAAGACTGCTGGCTCACAGGACCATCCAGCCCGGCAGGGGA
AAGGCTAAACGATTATTAGGAGAGAGAGAGTAGAAAACACTTGAGATACCTGGAC
TATTTGGGGACCATGTGGCAGGCCAAATGTTCACTTTTACAGCTTTGCTGTTTGGC
CCTCCTTTTCAGCTGTCTAACTAAGTTCCTCCAACAAAGGCTCCAAGTTTTTATTGC
CAAGACAGTCACAAACTCTTCATGGCCAAAGGGTACC

Appendix 16: Forward Sequence of AlDel loxP site by MM_Typing_004

nactcccagtcatagctgtccctcttctcttatggagatccctcgaggtcgacggtatcgataagcttcgagggaccta**ata
acttcgtatagcatacattatacgaagttatat**taagggttattgaatatgatcggaattcctgcagcccTAGCTTTT
CATTGAAGCTTATACACACAGAGACACACACAGAGAGACACATTTGTGTCTGTGGT
GCTAAAGATGGAAGCCAGGGCCTTGGCTCTAAGGCA

Appendix 17: Reverse complimentary reverse sequence of AlDup loxP site by MM_Typing_003_2

CAGCTCAGCAGTTAGGAGTGCTTGCTGCTTTCGCACAGGATCCATGTTCAATGCC
CAGCATCCACGTGGAGCCTCACAGCCATCTGTAACTTCAGTTTCAGGGGAATCTG
GTGCCTTCTTCAAGGACACCGTAAGGTATGTGGTAAACAGACATAAATTCGAtaagct
tcgagggacctaataacttcgtatagcatacattatacgaagttatattaagggttattgaatatgatcggaattcctgcagcc
cgggGGATCCCAACCAATTATTCCCTCAACAAGAGCCCTGC

Appendix 18: Forward Sequence of RSDel loxP site by MM_Typing_011_2

GGAGAGGATAAGAATGAACACACAATGAACAAACATAAAAGCACGTAGGCTCTTTT
CTCACAGATAAGCACTAGGCAAGCCAGGAATGTTAAACACGCAGGGTTGTCTCTT
CAAAGGGAGAGTAGATAACCTGTTTCCAGCCCAGGAGGCTGGGGGCAGGGGTGG
TTAATAGCCTGGTGACCTCTGTGCTATTTCTGTAAGGCCAGGCCACAGCAGCTGC
CCCAGACCCCTCCCATGAGCCTGCTTCCCCCAGACTCCTCCCATGAGCCTACTTC
CTTAATCTATAGAATCTATCTATACAGAAGACTCACTTGAACCTCACAAAGAATCA
ATGTAGTCATAAGTAAAGATTTATTGAACACATGGGATGGAGTTAtcgaggaattccgatc
atattcaataacccttaatataacttcgtataatgtatgctatacgaagttatgtcgccggtatcgatgaattcctgcagccctc
gaGCTGAGCCATCTCA

Appendix 19: Reverse Sequence of RSDel loxP site by MM_Typing_009_2

cttacnnatagcatacattatacgaagttatattaagggttattgaatatgatcggaattcctcgaTAACTCCATCCC
ATGTGTTCAATAAATCTTTACTTATGACTACATTGATTCTTTTGTGAGGTTCAAGTG
AGTCTTCTGTATAGATAGATTCTATAGATTAAGGAAGTAGGCTCATGGGAGGAGTC
TGGGGGAAGCAGGCTCATGGGAGGGGTCTGGGGCAGCTGCTGTGGCCTGGCCT
TACAGAAATAGCACAGAGGTCACCAGGCTATTAACCACCCCTGCCCCCAGCCTCC
TGGGCTGGAAACAGGTTATCTACTCTCCCTTTGAAGAGACAACCCTGCGTGTTTAA
CATTCCTGGCTTGCCTAGTGCTTATCTGTGAGAAAAGAGCCTACGTGCTTTTATGT
TTGTTCATTGTGTGTTCATTCTTATCCTCTCCAAACAGGTGACTCACTTTTGAGAAG
TTCAGCCAGATCACCTCCTGGA

Appendix 20: Forward Sequence of RSDup loxP site sample 17 by MM_Typing_012

AAGAGAGGGAAATATTACCACATTCTGGAAATTTTACTTGCAATAAGCAAATCACA
CAAATCTTAAATACGGGAAGAGACTGGGATTTTTGGAAATTGCAGATCTGGGGAG
GATGTGGTAATAGTGAAGCAGGGGAATAGCCCAGTTGGGAAAGCATTTCCTTTAA
GGTGACATCTGGCATCTgtcgacaagcttcccgggctcgacgggcccccccctcgacataacttcgtataatgt

atgctatacgaagttattaggtccctcgaanaggttcactagtGATCCCAACCAATTATTCCCTCAACAA
GAGCCCTGCAATTTTTCATTCCAGTGGTATTGGAACTGGAGCCAACACCTCTCCTG
GGCCACAATATGACAACCTACAGTCCCTTGAAAACATAAGGGTGGATGGGGAGAC
ACCTACTAGTCCTTCCTCAAATACTGTTGTGTTGATTTTGTTAGCCCAGGAGTACA
GGCAGAGGCACCTACAAAGGGGCAAGCAGTTGAGTCCCATACAGGAAGGGGGTA
ACATAAGGATGACATAATTTGGGAGTTGCCCAGGAAAAAACCCT

Appendix 21: Reverse Sequence of RSDup loxP site sample 17 by MM_Typing_013

TNTCCTTATGTTACCCCCTTCCTGTATGGGACTCAACTGCTTGCCCCTTTGTAGGT
GCCTCTGCCTGTACTCCTGGGCTAACAAAATCAACACAACAGTATTTGAGGAAGG
ACTAGTAGGTGTCTCCCCATCCACCCTTATGTTTTCAAGGGACTGTAGGTTGTCAT
ATTGTGGCCCAGGAGAGGTGTTGGCTCCAGTTCCAATACCACTGGAATGAAAAAT
TGCAGGGCTCTTGTTGAGGGAATAATTGGTTGGGATCactagtgaacctcttcgagggaccta
ataacttcgtatagcatacattatacgaagttatgtcgaggggggggcccgtcgagcccgggaagcttgtcgacAGATG
CCAGATGTCACCTTAAAGGAAATGCTTTCCCAACTGGGCTATTCCCCTGCTTCACT
ATTACCACATCCTCCCCAGATCTGCAATTTCCAAAAATCCCAGTCTCTTCCCGTATT
TAAGATTTGTGTGATTTGCTTATTGCAAGTAAAATTTCCAGAATGTGGTAATATTTC
CCTCTCTTGGGTTGGCCACTTACACAATTATGTTGAAATACTATACTTCTCCACTAG

Appendix 22: Forward Sequence of RSDup loxP site sample 38 by MM_Typing_012

AACCCAAGAGAGGGAAATATTACCACATTCTGGAAATTTTACTTGCAATAAGCAAA
TCACACAAATCTTAAATACGGGAAGAGACTGGGATTTTTGGAAATTGCAGATCTGG
GGAGGATGTGGTAATAGTGAAGCAGGGGAATAGCCCAGTTGGGAAAGCATTTCCT
TTAAGGTGACATCTGGCATCTgtcgacaagcttcccgggctcgacgggccccccctcgacataacttcgta
taatgtatgctatacgaagttattaggtccctcgaagaggttcactagtGATCCCAACCAATTATTCCCTCAA
CAAGAGCCCTGCAATTTTTCATTCCAGTGGTATTGGAACTGGAGCCAACACCTCTC
CTGGGCCACAATATGACAACCTACAGTCCCTTGAAAACATAAGGGTGGATGGGGA
GACACCTACTAGTCCTTCCTCAAATACTGTTGTGTTGATTTTGTTAGCCCAGGAGT
ACAGGCAGAGGCACCTACAAAGGGGCAAGCAGTTGAGTCCCATACAGGAAGGGG
GTAACATAAGGATGACATAATTTGGGAGTTGCCCAGGAAAAAACCCTC

Appendix 23: Reverse Sequence of RSDup loxP site sample 38 by MM_Typing_013

GGTNTCCTTATGTTACCCCCTTCCTGTATGGGACTCAACTGCTTGCCCCTTTGTAG
GTGCCTCTGCCTGTACTCCTGGGCTAACAAAATCAACACAACAGTATTTGAGGAA
GGACTAGTAGGTGTCTCCCCATCCACCCTTATGTTTTCAAGGGACTGTAGGTTGTC
ATATTGTGGCCCAGGAGAGGTGTTGGCTCCAGTTCCAATACCACTGGAATGAAAA
ATTGCAGGGCTCTTGTTGAGGGAATAATTGGTTGGGATCactagtgaacctcttcgagggacc
taataacttcgtatagcatacattatacgaagttatgtcgaggggggggcccgtcgagcccgggaagcttgtcgacAGAT
GCCAGATGTCACCTTAAAGGAAATGCTTTCCCAACTGGGCTATTCCCCTGCTTCAC
TATTACCACATCCTCCCCAGATCTGCAATTTCCAAAAATCCCAGTCTCTTCCCGTAT
TTAAGATTTGTGTGATTTGCTTATTGCAAGTAAAATTTCCAGAATGTGGTAATATTT
CCCTCTCTTGGGTTGGCCACTTACACAATTATGTTGAAATACTATACTTCTCCACTA
GT

List of figures

List of tables

Bibliography

I have tried to make all holders of the Copyright identified and obtained their consent to the use of the images in this work. Should a copyright infringement become known, I request to send me a message.

Ich habe mich bemüht, sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

1. Abramowitz, L.K., and Bartolomei, M.S. (2012). Genomic imprinting: recognition and marking of imprinted loci. Current opinion in genetics & development *22*, 72-78.

2. Amano, T., Sagai, T., Tanabe, H., Mizushina, Y., Nakazawa, H., and Shiroishi, T. (2009). Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. Developmental Cell *16*, 47-57.

3. Angelman, H. (1965). 'Puppet' Children A Report on Three Cases. Developmental Medicine & Child Neurology *7*, 681-688.

4. Arnone, M.I., and Davidson, E.H. (1997). The hardwiring of development: organization and function of genomic regulatory systems. Development (Cambridge, England) *124*, 1851-1864.

5. Autuoro, J.M., Pirnie, S.P., and Carmichael, G.G. (2014). Long Noncoding RNAs in Imprinting and X Chromosome Inactivation. Biomolecules *4*, 76-100.

6. Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell research *21*, 381-395.

7. Barlow, D.P. (2011). Genomic Imprinting: A Mammalian Epigenetic Discovery Model. Annual Review of Genetics *45*, 379-403.

8. Barlow, D.P., and Bartolomei, M.S. (2014). Genomic imprinting in mammals. Cold Spring Harbor Perspectives in Biology *6*.

9. Barlow, D.P., Stöger, R., Herrmann, B.G., Saito, K., and Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature *349*, 84-87.

10. Bartholdi, D., Krajewska-Walasek, M., Ounap, K., Gaspar, H., Chrzanowska, K.H., Ilyana, H., Kayserili, H., Lurie, I.W., Schinzel, A., and Baumer, A. (2009). Epigenetic mutations of the imprinted IGF2-H19 domain in Silver-Russell syndrome (SRS): results from a large cohort of patients with SRS and SRS-like phenotypes. Journal of Medical Genetics *46*, 192-197.

11. Bartolomei, M.S. (2009). Genomic imprinting: employing and avoiding epigenetic processes. Genes & development *23*, 2124-2133.

12. Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian genomic imprinting. Cold Spring Harbor Perspectives in Biology *3*.

13. Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153-155.

14. Barton, S.C., Surani, M.A., and Norris, M.L. (1984). Role of paternal and maternal genomes in mouse development. Nature *311*, 374-376.

15. Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482-485.

16. Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi-C: a comprehensive technique to capture the conformation of genomes. Methods *58*, 268-276.

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

18. Brannan, C.I., Dees, E.C., Ingram, R.S., and Tilghman, S.M. (1990). The product of the H19 gene may function as an RNA. Molecular and cellular biology *10*, 28-36.

19. Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38-44.

20. Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes & development *13*, 2465-2477.

21. Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327-339.

22. Cassidy, S.B., Schwartz, S., Miller, J.L., and Driscoll, D.J. (2012). Prader-Willi syndrome. Genetics in Medicine: Official Journal of the American College of Medical Genetics *14*, 10-26.

23. Cattanach, B.M., and Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. Nature *315*, 496-498.

24. Chen, C.L., Ip, S.M., Cheng, D., Wong, L.C., and Ngan, H.Y. (2000). Loss of imprinting of the IGF-II and H19 genes in epithelial ovarian cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *6*, 474-479.

25. Cheng, X. (2014). Structural and Functional Coordination of DNA and Histone Methylation. Cold Spring Harbor Perspectives in Biology *6*.

26. Chien, A., Edgar, D.B., and Trela, J.M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. Journal of bacteriology *127*, 1550-1557.

27. Chopra, V.S., Cande, J., Hong, J.-W., and Levine, M. (2009). Stalled Hox promoters as chromosomal boundaries. Genes & development *23*, 1505-1509.

28. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science (New York, NY) *339*, 819-823.

29. Cooper, D.W., VandeBerg, J.L., Sharman, G.B., and Poole, W.E. (1971). Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. Nature: New Biology *230*, 155-157.

30. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America *107*, 21931-21936.

31. Crouse, H.V., Brown, A., and Mumford, B.C. (1971). --chromosome inheritance and the porblem of chromosome "imprinting" in Sciara (Sciaridae, Diptera). Chromosoma *34*, 324-339 328.

32. Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.-Y., Cui, K., and Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome research *19*, 24-32.

33. Davis, T.L., Yang, G.J., McCarrey, J.R., and Bartolomei, M.S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Human molecular genetics *9*, 2885-2894.

34. de Laat, W., and Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. Nature *502*, 499-506.

35. DeChiara, T.M., Robertson, E.J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. Cell *64*, 849-859.

36. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science (New York, NY) *295*, 1306-1311.

37. Dickel, D.E., Visel, A., and Pennacchio, L.A. (2013). Functional anatomy of distant-acting mammalian enhancers. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences *368*, 20120359.

38. Dostie, J., and Dekker, J. (2007). Mapping networks of physical interactions between genomic elements using 5C technology. Nature protocols *2*, 988-1002.

39. Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C.*, et al.* (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome research *16*, 1299-1309.

40. Downs, K.M., and Davies, T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. Development (Cambridge, England) *118*, 1255-1266.

41. Ehrlich, M., and Wang, R.Y. (1981). 5-Methylcytosine in eukaryotic DNA. Science (New York, NY) *212*, 1350-1357.

42. Engel, N., Raval, A.K., Thorvaldsen, J.L., and Bartolomei, S.M. (2008). Three-dimensional conformation at the H19/Igf2 locus supports a model of enhancer tracking. Human molecular genetics *17*, 3021-3029.

43. Ferguson-Smith, A.C. (2011). Genomic imprinting: the emergence of an epigenetic paradigm. Nature reviews Genetics *12*, 565-575.

44. Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Molecular and cellular biology *16*, 2802-2813.

45. Fleming, T.P., Warren, P.D., Chisholm, J.C., and Johnson, M.H. (1984). Trophectodermal processes regulate the expression of totipotency within the inner cell mass of the mouse expanding blastocyst. Journal of embryology and experimental morphology *84*, 63-90.

46. Fujishige, K., Kotera, J., Michibata, H., Yuasa, K., Takebayashi, S., Okumura, K., and Omori, K. (1999). Cloning and characterization of a novel human phosphodiesterase that hydrolyzes both cAMP and cGMP (PDE10A). The Journal of biological chemistry *274*, 18438-18445.

47. Fukusumi, Y., Naruse, C., and Asano, M. (2008). Wtap is required for differentiation of endoderm and mesoderm in the mouse embryo. Developmental dynamics : an official publication of the American Association of Anatomists *237*, 618-629.

48. Gabory, A., Ripoche, M.-A., Le Digarcher, A., Watrin, F., Ziyyat, A., Forné, T., Jammes, H., Ainscough, J.F.X., Surani, M.A., Journot, L*., et al.* (2009). H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. Development (Cambridge, England) *136*, 3413-3421.

49. Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nature reviews Genetics *7*, 703-713.

50. Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annual review of biochemistry *74*, 481-514.

51. Guenzl, P.M., and Barlow, D.P. (2012). Macro lncRNAs: a new layer of cis-regulatory information in the mammalian genome. RNA biology *9*, 731-741.

52. Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature *482*, 339-346.

53. Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E., and Tycko, B. (1993). Tumour-suppressor activity of H19 RNA. Nature *365*, 764-767.

54. Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W*., et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

55. Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A*., et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics *39*, 311-318.

56. Hérault, Y., Rassoulzadegan, M., Cuzin, F., and Duboule, D. (1998). Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). Nature genetics *20*, 381-384.

57. Hippenmeyer, S., Johnson, R.L., and Luo, L. (2013). Mosaic analysis with double markers reveals cell-type-specific paternal growth dominance. Cell reports *3*, 960-967.

58. Hippenmeyer, S., Youn, Y.H., Moon, H.M., Miyamichi, K., Zong, H., Wynshaw-Boris, A., and Luo, L. (2010). Genetic mosaic dissection of Lis1 and Ndel1 in neuronal migration. Neuron *68*, 695-709.

59. Hudson, Q.J., Seidl, C.I.M., Kulinski, T.M., Huang, R., Warczok, K.E., Bittner, R., Bartolomei, M.S., and Barlow, D.P. (2011). Extra-embryonic-specific imprinted expression is restricted to defined lineages in the post-implantation embryo. Developmental biology *353*, 420-431.

60. Ideraabdullah, F.Y., Vigneau, S., and Bartolomei, M.S. (2008). Genomic Imprinting Mechanisms in Mammals. Mutation research *647*, 77-85.

61. Ikegami, T., Suzuki, Y., Shimizu, T., Isono, K., Koseki, H., and Shirasawa, T. (2002). Model mice for tissue-specific deletion of the manganese superoxide dismutase (MnSOD) gene. Biochemical and biophysical research communications *296*, 729-736.

62. Isles, A.R., and Holland, A.J. (2005). Imprinted genes and mother-offspring interactions. Early Human Development *81*, 73-77.

63. Johnson, D.R. (1974). Further observations on the haipin-tail (Thp) mutation in the mouse. Genetical research *24*, 207-213.

64. Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics *13*, 484-492.

65. Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C.F., Wolffe, A., Ohlsson, R., and Lobanenkov, V.V. (2000). Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. Current biology : CB *10*, 853-856.

66. Kanduri, C., Thakur, N., and Pandey, R.R. (2006). The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing. The EMBO journal *25*, 2096-2106.

67. Kaneda, A., and Feinberg, A.P. (2005). Loss of imprinting of IGF2: a common epigenetic modifier of intestinal tumor risk. Cancer research *65*, 11236-11240.

68. Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900-903.

69. Kaufman, M.H. (1983). Early Mammalian Development: Parthenogenetic Studies (CUP Archive).

70. Khosla, S., Mendiratta, G., and Brahmachari, V. (2006). Genomic imprinting in the mealybugs. Cytogenetic and Genome Research *113*, 41-52.

71. Kinoshita, T., Ikeda, Y., and Ishikawa, R. (2008). Genomic imprinting: a balance between antagonistic roles of parental chromosomes. Seminars in Cell & Developmental Biology *19*, 574-579.

72. Kuhn, E.J., and Geyer, P.K. (2003). Genomic insulators: connecting properties to mechanism. Current Opinion in Cell Biology *15*, 259-265.

73. Kwon, G.S., Viotti, M., and Hadjantonakis, A.-K. (2008). The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. Developmental Cell *15*, 509-520.

74. Lanner, F. (2014). Lineage specification in the early mouse embryo. Experimental cell research *321*, 32-39.

75. Latos, P.A., Pauler, F.M., Koerner, M.V., Şenergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E*., et al.* (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. Science (New York, NY) *338*, 1469-1472.

76. Latos, P.A., Stricker, S.H., Steenpass, L., Pauler, F.M., Huang, R., Senergin, B.H., Regha, K., Koerner, M.V., Warczok, K.E., Unger, C*., et al.* (2009). An in vitro ES cell

imprinting model shows that imprinted expression of the Igf2r gene arises from an allele-specific expression bias. Development (Cambridge, England) *136*, 437-448.

77. Lerchner, W., and Barlow, D.P. (1997). Paternal repression of the imprinted mouse Igf2r locus occurs during implantation and is stable in all tissues of the post-implantation mouse embryo. Mechanisms of development *61*, 141-149.

78. Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.

79. Li, J.Y., Lees-Murdock, D.J., Xu, G.L., and Walsh, C.P. (2004). Timing of establishment of paternal methylation imprints in the mouse. Genomics *84*, 952-960.

80. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

81. Lyle, R., Watanabe, D., te Vruchte, D., Lerchner, W., Smrzka, O.W., Wutz, A., Schageman, J., Hahner, L., Davies, C., and Barlow, D.P. (2000). The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. Nature genetics *25*, 19-21.

82. Mancini-Dinardo, D., Steele, S.J.S., Levorse, J.M., Ingram, R.S., and Tilghman, S.M. (2006). Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. Genes & development *20*, 1268-1282.

83. Marmorstein, R., and Zhou, M.-M. (2014). Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. Cold Spring Harbor Perspectives in Biology *6*.

84. Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. Annual Review of Genomics and Human Genetics *7*, 29-59.

85. Matouk, I., Raveh, E., Ohana, P., Lail, R.A., Gershtain, E., Gilon, M., De Groot, N., Czerniak, A., and Hochberg, A. (2013). The increasing complexity of the oncofetal h19 gene locus: functional dissection and therapeutic intervention. International Journal of Molecular Sciences *14*, 4298-4316.

86. McGrath, J., and Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell *37*, 179-183.

87. Milligan, L., Forne, T., Antoine, E., Weber, M., Hemonnot, B., Dandolo, L., Brunel, C., and Cathala, G. (2002). Turnover of primary transcripts is a major step in the regulation of mouse H19 gene expression. EMBO reports *3*, 774-779.

88. Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. Trends in genetics: TIG *7*, 45-49.

89. Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. Human molecular genetics *14 Spec No 1*, R47-58.

90. Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nature genetics *36*, 889-893.

91. Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

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

93. Ong, C.-T., and Corces, V.G. (2011a). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nature reviews Genetics *12*, 283-293.

94. Ong, C.T., and Corces, V.G. (2011b). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nature reviews Genetics *12*, 283-293.

95. Orban, P.C., Chui, D., and Marth, J.D. (1992). Tissue- and site-specific DNA recombination in transgenic mice. Proceedings of the National Academy of Sciences of the United States of America *89*, 6861-6865.

96. Otto, S.P., and Goldstein, D.B. (1992). Recombination and the evolution of diploidy. Genetics *131*, 745-751.

97. Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T.*, et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.

98. Pauler, F.M., Barlow, D.P., and Hudson, Q.J. (2012). Mechanisms of long range silencing by imprinted macro non-coding RNAs. Current opinion in genetics & development *22*, 283-289.

99. Perrot, V., Richerd, S., and Valéro, M. (1991). Transition from haploidy to diploidy. Nature *351*, 315-317.

100. Peters, J., and Robson, J.E. (2008). Imprinted noncoding RNAs. Mammalian genome : official journal of the International Mammalian Genome Society *19*, 493-502.

101. Reik, W., and Lewis, A. (2005). Co-evolution of X-chromosome inactivation and imprinting in mammals. Nature reviews Genetics *6*, 403-410.

102. Reik, W., and Walter, J. (2001). Genomic imprinting: parental influence on the genome. Nature reviews Genetics *2*, 21-32.

103. Renfree, M.B., Hore, T.A., Shaw, G., Graves, J.A.M., and Pask, A.J. (2009). Evolution of genomic imprinting: insights from marsupials and monotremes. Annual Review of Genomics and Human Genetics *10*, 241-262.

104. Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

105. Rippe, K., Schrader, A., Riede, P., Strohner, R., Lehmann, E., and Langst, G. (2007). DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. Proceedings of the National Academy of Sciences of the United States of America *104*, 15635-15640.

106. Rodley, C.D., Bertels, F., Jones, B., and O'Sullivan, J.M. (2009). Global identification of yeast chromosome interactions using Genome conformation capture. Fungal genetics and biology : FG & B *46*, 879-886.

107. Rossant, J., and Tam, P.P.L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development (Cambridge, England) *136*, 701-713.

108. Royo, H., and Cavaille, J. (2008). Non-coding RNAs in imprinted gene clusters. Biology of the cell / under the auspices of the European Cell Biology Organization *100*, 149-166.

109. Ruf, S., Symmons, O., Uslu, V.V., Dolle, D., Hot, C., Ettwiller, L., and Spitz, F. (2011). Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. Nature genetics *43*, 379-386.

110. Santoro, F., Mayer, D., Klement, R.M., Warczok, K.E., Stukalov, A., Barlow, D.P., and Pauler, F.M. (2013). Imprinted Igf2r silencing depends on continuous Airn lncRNA expression and is not restricted to a developmental window. Development (Cambridge, England) *140*, 1184-1195.

111. Sasaki, H., and Matsui, Y. (2008). Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nature reviews Genetics *9*, 129-140.

112. Sauer, B., and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proceedings of the National Academy of Sciences of the United States of America *85*, 5166-5170.

113. Seidl, C.I.M., Stricker, S.H., and Barlow, D.P. (2006). The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. The EMBO journal *25*, 3565-3575.

114. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

115. Shuman, C., Beckwith, J.B., Smith, A.C., and Weksberg, R. (1993). Beckwith-Wiedemann Syndrome. In GeneReviews(®), R.A. Pagon, M.P. Adam, H.H. Ardinger, T.D. Bird, C.R. Dolan, C.-T. Fong, R.J. Smith, and K. Stephens, eds. (Seattle (WA): University of Washington, Seattle).

116. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics *38*, 1348-1354.

117. Sleutels, F., Tjon, G., Ludwig, T., and Barlow, D.P. (2003). Imprinted silencing of Slc22a2 and Slc22a3 does not need transcriptional overlap between Igf2r and Air. The EMBO journal *22*, 3696-3704.

118. Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature *415*, 810-813.

119. Smyk, M., Szafranski, P., Startek, M., Gambin, A., and Stankiewicz, P. (2013). Chromosome conformation capture-on-chip analysis of long-range cis-interactions of the SOX9 promoter. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *21*, 781-788.

120. Solter, D. (1988). Differential imprinting and expression of maternal and paternal genomes. Annual Review of Genetics *22*, 127-146.

121. Sternberg, N., and Hamilton, D. (1981). Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. Journal of Molecular Biology *150*, 467-486.

122. Stöger, R., Kubicka, P., Liu, C.G., Kafri, T., Razin, A., Cedar, H., and Barlow, D.P. (1993). Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. Cell *73*, 61-71.

123. Straub, T., and Becker, P.B. (2008). DNA sequence and the organization of chromosomal domains. Current opinion in genetics & development *18*, 175-180.

124. Stringer, J.M., Pask, A.J., Shaw, G., and Renfree, M.B. (2014). Post-natal imprinting: evidence from marsupials. Heredity *113*, 145-155.

125. Surani, M.A., Barton, S.C., and Norris, M.L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature *308*, 548-550.

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

127. Tang, S.-H.E., Silva, F.J., Tsark, W.M.K., and Mann, J.R. (2002). A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvImJ. Genesis (New York, NY: 2000) *32*, 199-202.

128. Thorvaldsen, J.L., and Bartolomei, M.S. (2007). SnapShot: imprinted gene clusters. Cell *130*, 958.

129. Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes & development *12*, 3693-3702.

130. Tremblay, K.D., Duran, K.L., and Bartolomei, M.S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. Molecular and cellular biology *17*, 4322-4329.

131. Uribe-Lewis, S., Woodfine, K., Stojic, L., and Murrell, A. (2011). Molecular mechanisms of genomic imprinting and clinical implications for cancer. Expert Reviews in Molecular Medicine *13*, e2.

132. Varmuza, S., and Mann, M. (1994). Genomic imprinting--defusing the ovarian time bomb. Trends in genetics: TIG *10*, 118-123.

133. Wan, L.-B., and Bartolomei, M.S. (2008). Regulation of imprinting in clusters: noncoding RNAs versus insulators. Advances in Genetics *61*, 207-223.

134. Wang, X., Soloway, P.D., and Clark, A.G. (2011). A survey for novel imprinted genes in the mouse placenta by mRNA-seq. Genetics *189*, 109-122.

135. Waters, A.J., Bilinski, P., Eichten, S.R., Vaughn, M.W., Ross-Ibarra, J., Gehring, M., and Springer, N.M. (2013). Comprehensive analysis of imprinted genes in maize reveals allelic variation for imprinting and limited conservation with other species. Proceedings of the National Academy of Sciences of the United States of America *110*, 19639-19644.

136. Wijchers, P.J., and de Laat, W. (2011). Genome organization influences partner selection for chromosomal rearrangements. Trends in genetics : TIG *27*, 63-71.

137. Williamson, C.M., Ball, S.T., Dawson, C., Mehta, S., Beechey, C.V., Fray, M., Teboul, L., Dear, T.N., Kelsey, G., and Peters, J. (2011). Uncoupling antisense-mediated silencing and DNA methylation in the imprinted Gnas cluster. PLoS genetics *7*, e1001347.

138. Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature *389*, 745-749.

139. Wutz, A., Theussl, H.C., Dausman, J., Jaenisch, R., Barlow, D.P., and Wagner, E.F. (2001). Non-imprinted Igf2r expression decreases growth and rescues the Tme mutation in mice. Development *128*, 1881-1887.

140. Yagi, S., Hirabayashi, K., Sato, S., Li, W., Takahashi, Y., Hirakawa, T., Wu, G., Hattori, N., Hattori, N., Ohgane, J*., et al.* (2008). DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) in mouse promoter regions demonstrating tissue-specific gene expression. Genome research *18*, 1969-1978.

141. Yotova, I.Y., Vlatkovic, I.M., Pauler, F.M., Warczok, K.E., Ambros, P.F., Oshimura, M., Theussl, H.-C., Gessler, M., Wagner, E.F., and Barlow, D.P. (2008). Identification of the human homolog of the imprinted mouse Air non-coding RNA. Genomics *92*, 464-473.

142. Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U*., et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nature genetics *38*, 1341-1347.

Curriculum Vitae

Markus Muckenhuber

Education

11/2011 – 12/2014   **Graduate studies** (Master's program) of *Molecular Biology* at the University of Vienna, Vienna, Austria

10/2007 – 11/2011   **Undergraduate studies** (Bachelor's program) of *Biology* at the University of Vienna, Vienna, Austria

Research Experience

09/2013 – 12/2014   Master's project on '*The impact of Airn lncRNA expression on enhancer-promoter interactions*', group of Denise P. Barlow, CeMM (Research Center for Molecular Medicine), Vienna, Austria

05/2013 – 09/2013   Research project on '*The role of post-transcriptional regulation in establishing specific miR expression pattern during embryonic development in C. elegans*', group of Luisa Cochella, IMP (Research Institute of Molecular Pathology), Vienna, Austria

03/2013 – 05/2013   Research project on '*Analysis of differences in imprinted expression in Igf2r cluster of murine embryonic and extra-embryonic tissues by qPCR*', group of Denise P. Barlow, CeMM (Research Center for Molecular Medicine), Vienna, Austria

06/2011   Bachelor Thesis '*Optimisation of the electrotransfer of serum via SDS-PAGE and western blot onto a PVDF-membrane*', group of Univ. Prof. Dr. Kurt Herkner at the Department of Pediatrics and Adolescent Medicine, General Hospital, Vienna, Austria

Scholarship

10/2012 – 09/2013   **Performance Scholarship** for graduate studies (Master) of *Molecular Biology* at the University of Vienna, Vienna, Austria

Acknowledgements

Before I start with the people, who supported me during my masters' thesis in the lab, I want to name the wonderful people, who started to do so much earlier. Thanks a lot to my mother and father for making it possible for me to start with this 'exotic' studies and also for supporting me during the last years in many ways. Special thanks to my little sister for not finishing with her studies before me.

Denise Barlow gave me the opportunity to do the practical work for this thesis in her lab at CeMM and for this I have to thank her a lot. The inspiring scientific environment in the lab and the entire CeMM as well as the enthusiasm of its members supported me during the last year and helped me a lot to prepare my way for the near future. I learned what it means to work as a scientist and how important it is to be well-organised and structured in the lab. Especially Denise's support and her experience helped me a lot to prepare myself your PhD selections and finally to find a great position for the next years. In general, the scientific atmosphere at CeMM and the community of CeMMies create an amazing institute to do science on high levels and it was a pleasure for me to be part of this.

The next person I want to thank a lot is my supervisor Quanah Hudson. He managed to provide me the right amount of guidance, when this was required, and also a lot of freedom to develop my own ideas. This was the right mixture for me to make my first steps on a way to become an independent scientist. He was an amazing help by correcting this thesis and I know it was sometimes an extensive amount of work to do this. With his highly developed writing skills, he taught me a lot about scientific writing and I'm very thankful about it.

A big thanks to our (former) technician Justyna for teaching me Southern blots, the Barlow-Lab-organisation, (of course) how to make a proper list and three Polish words. Also our shared breaks and the tours to the coffee pirates

were necessary to make the last year to such a great one for me. It was a pleasure.

The next two ladies, Sabine and Sarah, did an amazing job during the last year and because of their work the project was a success. Thanks for your accurate way of working. With your help I learned a lot about breeding plans of mice and how to set up a useful system to keep the overview. Also our after-mouse meetings were a great time, which I will miss.

Although we were working on totally different stories, the time with Christoph, especially the Halloween party, I enjoyed it a lot. Also the remaining (and former) members of the Barlow lab; Florian, Philipp B, Philipp G, Alexandra and Ruth; are great characters, always very helpful and every one of these people contributed in a very unique way to my last year. Thanks.

Last but not least I want to mention a very special person, who was a big help for me at the very beginning and she is an amazing person with a huge heart: Rita, all the best for your future and stay strong.

There are many other people; old friends, new friends, colleges; not mentioned above who helped me during the last years continuing this way and finally to reach this goal. Thank you very much.