



universität
wien

DISSERTATION

Titel der Dissertation

Dissecting the multiple roles of the *Igf2r* imprint control element

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr. rer.nat.)

Verfasserin:	Martha V. Körner
Matrikel-Nummer:	0007994
Dissertationsgebiet (lt. Studienblatt):	Genetik - Mikrobiologie
Betreuerin:	Prof. Denise P. Barlow, PhD
Wien, im März 2010	

The work presented in this thesis was undertaken in the lab of Prof. Denise P. Barlow from
February 2006 to March 2010 at the
CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences
c/o Department of Microbiology, Immunobiology and Genetics
Max F. Perutz Laboratories
Vienna Biocenter
Dr. Bohrgasse 9/4
1030 Vienna
Austria

Table of contents

TABLE OF CONTENTS	3
A. ZUSAMMENFASSUNG	6
B. ABSTRACT	7
1. INTRODUCTION	8
1.1 GENERAL INTRODUCTION FOR GENOMIC IMPRINTING.....	8
1.1.1 <i>Ploidy and allelic expression</i>	8
1.1.2 <i>Genomic imprinting is an epigenetic process</i>	9
1.2 MECHANISMS REGULATING GENOMIC IMPRINTING	15
1.2.1 <i>Imprinted genes are clustered and controlled by an imprint control element</i>	15
1.2.2 <i>Imprinted clusters contain macro non-coding RNAs</i>	17
1.2.3 <i>Three mechanisms for the regulation of imprinted expression by the ICE</i>	19
1.2.4 <i>Imprinted expression can be restricted to certain tissues or developmental time points</i>	21
1.3 WHICH FUNCTIONS ARE CONTROLLED BY THE ICE AND WHAT CONTROLS THE ICE ITSELF?.....	24
1.3.1 <i>A role for DNA methylation in regulating imprinted expression</i>	24
1.3.2 <i>The ICE controls higher chromosomal organisation in the Igf2 cluster</i>	28
1.3.3 <i>The ICE controls ncRNA expression – does it also control ncRNA biology?</i>	30
1.3.4 <i>Imprinted expression by ncRNAs – transcript or transcription?</i>	31
1.3.5 <i>Histone modifications – epigenetic maintainers at the ICE and imprinted genes</i>	33
1.4 <i>IN VITRO</i> MODELS TO STUDY GENOMIC IMPRINTING	39
1.5 AIM OF THIS STUDY	41
2. RESULTS	43
2.1 THE ROLE OF THE <i>Igf2R</i> ICE ON THE CHROMOSOMAL LEVEL.....	43
2.1.1 <i>Generation of primary mouse embryonic fibroblasts (pMEFs) for DNA FISH</i>	43
2.1.2 <i>Mapping the proximal border of DNA FISH asynchrony in the Igf2r/Airn imprinted gene cluster</i>	47
2.1.3 <i>Do the ICE, the ncRNA Airn or the differentially methylated Igf2r promoter play a role in DNA FISH asynchrony?</i>	50
2.1.4 <i>Can the system detect loss of DNA FISH asynchrony? – Using the Igf2/H19 region as control</i>	53
2.1.5 <i>Modulators of DNA FISH asynchrony: Is histone acetylation involved?</i>	55
2.1.6 <i>3D FISH to analyse chromatin compaction</i>	59
2.1.7 <i>Does DNA FISH asynchrony represent DNA replication asynchrony?</i>	61
2.2 THE ROLE OF THE TANDEM DIRECT REPEATS AND THE CPG ISLAND WITHIN THE <i>Igf2R</i> ICE	63
2.2.1 <i>Elements within the Igf2r ICE</i>	63
2.2.2 <i>Generation of repeat deletion ES cells</i>	65
2.2.3 <i>Does deletion of the Airn tandem direct repeats result in a loss of the differential DNA methylation on the ICE?</i>	76
2.2.4 <i>Does deletion of the Airn tandem direct repeats lead to changes in Airn expression?</i>	80
2.2.5 <i>Does deletion of the Airn tandem direct repeats lead to changes in the low splicing capability of Airn?</i>	85
2.2.6 <i>Does deletion of the Airn tandem direct repeats lead to changes in the nuclear retention of Airn?</i>	88
2.2.7 <i>Can Airn in RΔ800 ES cells still lead to imprinted expression of Igf2r?</i>	90
2.2.8 <i>Generation of CpG island deletion ES cells</i>	99
2.2.9 <i>Does deletion of the Airn CGI lead to an additional gain of DNA methylation on the paternal ICE during in vitro differentiation?</i>	104
2.2.10 <i>Does deletion of the Airn CGI lead to changes in Airn expression?</i>	106
2.2.11 <i>Do CpGΔ cells show imprinted expression of Igf2r?</i>	108

3. DISCUSSION	113
3.1 SUMMARY OF RESULTS	113
3.2 THE ROLE OF THE <i>IGF2R</i> ICE ON THE CHROMOSOMAL LEVEL	113
3.2.1 DNA FISH asynchrony in the imprinted <i>Igf2r</i> cluster extends over 3Mbp	113
3.2.2 DNA FISH asynchrony in the imprinted <i>Igf2r</i> cluster is independent of <i>Airn</i> , the ICE and the sDMR	114
3.2.3 TSA leads to a relaxation of DNA FISH asynchrony in the imprinted <i>Igf2r</i> cluster.....	117
3.2.4 The parental alleles containing the <i>Igf2r</i> cluster do not display differences in chromatin compaction as determined by 3D DNA FISH.....	118
3.2.5 DNA FISH asynchrony in the <i>Igf2r</i> cluster is independent of DNA replication.....	119
3.2.6 DNA FISH asynchrony and imprinted expression – cause, consequence or correlation?.....	123
3.3 THE FUNCTION OF TANDEM DIRECT REPEATS	124
3.3.1 The role of the tandem direct repeats in the <i>Igf2r</i> ICE.....	127
3.3.2 The tandem direct repeats are not needed for the maintenance of the differential DNA methylation on the ICE	128
3.3.3 Deletion of the tandem direct repeats leads to a reduction in <i>Airn</i> length, but upregulation of <i>Airn</i> during <i>in vitro</i> differentiation and the low splicing capability of <i>Airn</i> are unaffected	129
3.3.4 Deletion of the tandem direct repeats does not influence the nuclear localisation of <i>Airn</i>	132
3.3.5 Deletion of the tandem direct repeats leads to a less efficient but not completely abolished imprinted expression of <i>Igf2r</i>	133
3.4 THE ROLE OF THE CPG ISLAND IN THE <i>IGF2R</i> ICE.....	138
3.4.1 Deletion of the CpG island leads to DNA methylation during cell culturing.....	139
3.4.2 Deletion of the CpG island reduces expression of <i>Airn</i>	140
3.4.3 Deletion of the CpG island leads to biallelic expression of <i>Igf2r</i>	142
3.5 THE ROLES OF THE <i>IGF2R</i> ICE	142
4. MATERIALS AND METHODS	144
4.1 METHODS	144
4.1.1 Mini prep of plasmid DNA from bacteria by alkaline lysis.....	144
4.1.2 Midi prep of plasmid and cosmid DNA from bacteria.....	144
4.1.3 Maxi prep of plasmid DNA from bacteria for preparing targeting vectors.....	144
4.1.4 Mini and maxi preps of BAC DNA from bacteria.....	145
4.1.5 Transformation using CaCl ₂ competent bacteria.....	145
4.1.6 Restriction digests.....	145
4.1.7 Gel electrophoresis.....	145
4.1.8 DNA sequencing.....	145
4.1.9 Enzymatic DNA modifications for cloning of DNA.....	145
4.1.10 Gel elution	146
4.1.11 Ligation	146
4.1.12 PCR	146
4.1.13 RT-PCR.....	147
4.1.14 Long PCR/ Proof Reading PCR.....	147
4.1.15 Cloning of pH19P13kb (probe for Southern blotting).....	148
4.1.16 Cloning of pIgf2P6 (for DNA FISH).....	148
4.1.17 Cloning of pNS13DTAmNeoinpBS (targeting vector for Ex12 SNP)	148
4.1.18 Cloning of pRΔ800 (targeting vector for repeat deletion)	149
4.1.19 Cloning of pCpGΔ (targeting vector for CpG island deletion)	150
4.1.20 Preparation of primary MEFs	150
4.1.21 ES cell culture	151
4.1.22 Gene targeting of ES cells by homologous recombination	151
4.1.23 Removal of the selection cassette by transient transfection	152
4.1.24 Subcloning of ES cell clones.....	152
4.1.25 Preparation of genomic DNA	152
4.1.26 Isolation of total RNA.....	152
4.1.27 DNaseI treatment of RNA.....	152
4.1.28 Reverse transcription of RNA.....	153
4.1.29 DNA and RNA blots	153
4.1.30 Real-time qPCR	154
4.1.31 RNA expression tiling array (RETA).....	157

4.1.32 Nuclear/cytoplasmatic RNA extraction	158
4.1.33 Metaphase spreads	159
4.1.34 DNA fluorescence in situ hybridisation (DNA FISH)	159
4.1.35 TSA treatment of cells	162
4.1.36 S phase fractionation	162
4.1.37 Bioinformatics	163
4.1.38 Box plots	164
4.2 MATERIALS	165
4.2.1 Cells	165
4.2.2 Mice	165
4.2.3 Vectors/Plasmids	165
4.2.4 Consumables	165
4.2.5 Kits	168
4.2.6 Solutions	168
4.2.7 Machines	173
4.3 ABBREVIATIONS	173
5. REFERENCES	175
6. CURRICULUM VITAE	189
6.1 PERSONAL DATA	189
6.2 EDUCATION	189
6.3 CONFERENCES	190
6.4 PUBLICATIONS	191
7. ACKNOWLEDGEMENTS	192

A. ZUSAMMENFASSUNG

Genomische Prägung führt zur monoallelischen Expression von Genen von einem elterlichen Allel und involviert epigenetische Mechanismen. Geprägte Gene treten meist in Clustern auf und jeder Cluster hat ein sogenanntes Imprintcontrolelement (ICE), welches die geprägte Expression aller Gene im Cluster reguliert. Der geprägte *Igf2r* Cluster am Mauschromosom 17 umfasst 430kb und enthält drei maternal exprimierte proteincodierende Gene, *Igf2r*, *Slc22a2* und *Slc22a3* und eine paternal exprimierte nichtcodierende RNA (ncRNA), *Airn*, welche die geprägte Expression der proteincodierenden Gene reguliert. Am maternalen Allel ist das ICE DNA methyliert, am paternalen Allel hingegen unmethyliert. Man hat gezeigt, dass eine Deletion des ICE die geprägte Expression aller Gene im *Igf2r* Cluster aufhebt. In dieser Dissertation habe ich analysiert, ob das ICE eine weitere Rolle auf chromosomaler Ebene spielt und ob Elemente innerhalb des ICE verschiedene Rollen im Prägungsprozess spielen. DNA FISH Asynchronie ist ein Merkmal geprägter Cluster, der dahinterliegende Mechanismus ist jedoch noch nicht komplett aufgeklärt. Im ersten Teil meiner Arbeit habe gezeigt, dass eine chromosomale Region von 3Mbp DNA-Fluoreszenz-*in-situ*-Hybridisierungs- (DNA FISH) Asynchronie aufweist. Diese Asynchronie wird jedoch weder vom ICE, noch von *Airn* oder der *Igf2r* Promoterregion kontrolliert. Weiters habe ich gefunden, dass diese DNA FISH Asynchronie weder aufgrund von Unterschieden in der Chromatinkompaktierung noch durch Unterschiede in der DNA Replikation während der S-Phase zustande kommt. Im zweiten Teil meiner Arbeit habe ich den Effekt von zwei Subdeletionen des ICE auf verschiedene Aspekte der genomischen Prägung hin analysiert. Ich habe embryonale Stammzellen (ES Zellen) mit einer paternalen Deletion der tandem direkten Repeats oder der CpG Insel hergestellt und diese ES Zellen mit Hilfe eines *in vitro* Differenzierungssystems analysiert. Ich habe gezeigt, dass die tandem direkten Repeats die Länge von *Airn* und die Fähigkeit von *Airn*, geprägte Expression von *Igf2r* zu verursachen, regulieren, und die CpG Insel von *Airn* benötigt wird, um DNA Methylierung am paternalen Allel zu verhindern. Das zeigt, dass die CpG Insel, die abwärts vom *Airn* Promotor liegt, eine wichtige Rolle in der Transkription von *Airn* spielt.

B. ABSTRACT

Genomic imprinting is a process leading to parent-of-origin dependent monoallelic gene expression and involves epigenetic mechanisms. Imprinted genes mostly occur in clusters and each cluster has an imprint control element (ICE) which regulates imprinted expression of all genes in the respective cluster. The imprinted *Igf2r* cluster on mouse chromosome 17 spans 430kb and contains three maternally expressed protein-coding genes, *Igf2r*, *Slc22a2* and *Slc22a3* and a paternally expressed non-coding RNA (ncRNA), *Airn*, which was shown to induce imprinted expression of the protein-coding genes. The ICE is DNA methylated on the maternal allele but free of methylation on the paternal allele. It was shown that a deletion of the ICE abolishes imprinted expression of all genes in the *Igf2r* cluster. In this thesis I analysed, if the ICE has an additional role on the chromosomal level and if elements within the ICE have distinct roles in the imprinting process. DNA FISH asynchrony is a common feature of imprinted clusters but its mechanism is not fully understood. I showed that a chromosomal region of 3Mbp exhibits DNA FISH asynchrony. I demonstrated that none of the key elements that control or are involved in imprinted expression, the ICE, *Airn*, and the *Igf2r* promoter region, control this DNA FISH asynchrony. I also showed that the observed DNA FISH asynchrony is neither due to differences in chromatin compaction of the parental alleles, nor due to differences in DNA replication during S phase. In the second part of my result section I analysed the effect of two ICE subdeletions on various aspects of genomic imprinting. I generated embryonic stem (ES) cell lines carrying a paternal deletion of the tandem direct repeats or of the *Airn* CpG island, and analysed those ES cell lines using an *in vitro* ES cell differentiation system. I showed, that the tandem direct repeats play a role in regulating the length of *Airn* as well as its capability to induce imprinted expression of *Igf2r*, while the *Airn* CpG island is required to prevent the gain of DNA methylation on the paternal allele. This shows that the CpG island localised downstream of the *Airn* promoter plays an important role in *Airn* transcription.

1. INTRODUCTION

1.1 General introduction for genomic imprinting

1.1.1 Ploidy and allelic expression

Mammals inherit one haploid chromosome set from the father and one from the mother and are thus diploid organisms. Diploidy is generally considered as being beneficial for an organism. If one gene copy gains a deleterious mutation or is deleted, the second copy is still present which can rescue a possible negative effect. This assumes that most genes are expressed biallelically and that the organism can tolerate changes in the gene dose. However, for some genes a reduction to 50% of normal gene expression levels results in a phenotype, a phenomenon termed haploinsufficiency (Wilkie, 1994). Furthermore, three classes of genes exist which show monoallelic gene expression. First, there are X-linked genes in female mammals which show random monoallelic expression (i.e. expression independent on the parental origin of the gene) with the exception of extraembryonic tissues of mice and all tissues of marsupials, where only the maternally inherited gene copy is expressed. Second, there are random monoallelically expressed autosomal genes like olfactory and immune cell receptor genes. Third, there are imprinted genes, where the activity of a gene is dependent on the parent of origin (Zakharova et al., 2009) (see Fig. 1).

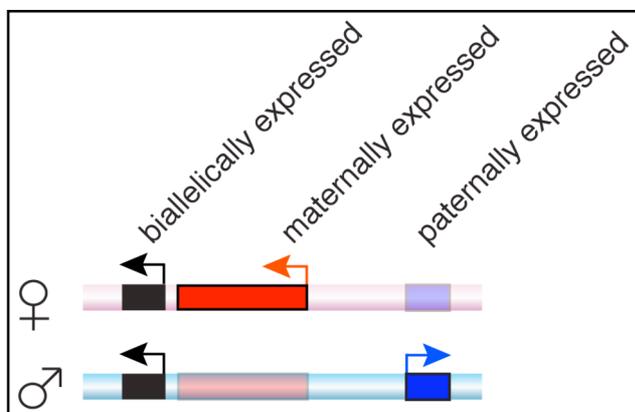


Fig. 1: Most mammalian genes show biallelic expression. In contrast, imprinted genes show expression dependent on the parent of origin, i.e. they are either expressed from the maternal or the paternal allele. Pink bar: maternal chromosome. Light blue bar: paternal chromosome. Arrows indicate transcription. Black boxes: biallelically expressed gene. Solid red box: maternally expressed gene on the expressed allele. Transparent red box: maternally expressed gene on the repressed allele. Solid blue box: paternally expressed gene on the expressed allele. Transparent blue box: paternally expressed gene on the repressed allele.

1.1.2 Genomic imprinting is an epigenetic process

Genomic imprinting describes the phenomenon leading to monoallelic gene expression dependent on the parent of origin. Genomic imprinting has been reported in eutherian mammals, marsupials (but not monotremes), flowering plants and two arthropod families, Coccidae and Sciaridae. These insects use the genomic imprinting process for sex determination, resulting in chromosome elimination or heterochromatinisation of whole chromosomes (Brown and Nur, 1964; Crouse, 1960). However, in mammals, marsupials and plants genomic imprinting leads to monoallelic expression of a subset of autosomal genes (Baroux et al., 2002; Goday and Esteban, 2001; Renfree et al., 2009).

In mammals, many observations from 1970-1980 indicated an effect of the parental source of genetic material on the phenotype. This included the demonstration that X chromosome inactivation occurred in an imprinted fashion in all tissues of marsupials and in mouse extraembryonic tissues (Sharman, 1971; Takagi and Sasaki, 1975). Autosomal parental-specific effects were also demonstrated in mice carrying parental-specific deletions and duplications of chromosome 17 (Johnson, 1975; Lyon and Glenister, 1977). Later in the early 1980s the use of nuclear transfer technology, demonstrated that the haploid maternal and paternal pronucleus of the early mouse zygote contribute differently to embryogenesis. In these experiments, three types of genetically identical embryos were generated containing either one maternal plus one paternal pronucleus (diploid wildtype embryo) or two maternal pronuclei (diploid gynogenetic) or two paternal pronuclei (diploid androgenetic). Only embryos containing both a maternal and a paternal pronucleus were viable, while gynogenetic and androgenetic embryos showed embryonic lethality during early post-implantation development (McGrath and Solter, 1984; Surani et al., 1984). The analysis of a large number of mouse uniparental disomies and duplications arising in mice carrying Robertsonian and reciprocal translocations demonstrated that at least seven chromosomes were involved in these parental-specific developmental effects (Cattanach and Kirk, 1985). The use of inbred mice with genetically identical

parental genomes for these experiments led to the conclusion that parental specific effects arose from an epigenetic mechanism.

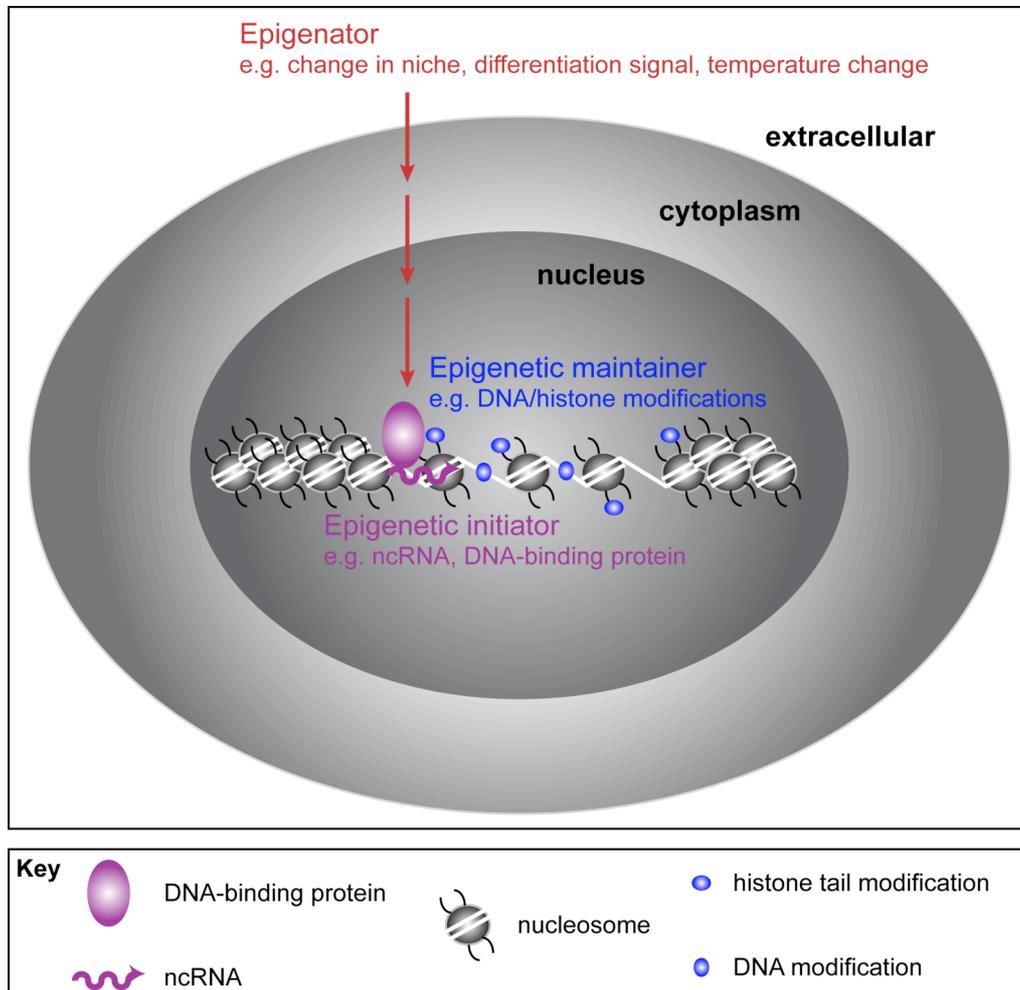


Fig. 2: The epigenetic pathway. The cascade starts with an ‘epigenator’ (red), a signal from outside of the cell, which activates an ‘epigenetic initiator’ (purple). The epigenetic initiator acts locus-specifically and determines the site of the epigenetic state. Then ‘epigenetic maintainers’ (blue) maintain the epigenetic state. See key for details (Figure modified from Berger et al., 2009).

An epigenetic trait is today defined as ‘a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence’ (Berger et al., 2009). The formation of a stable epigenetic state can be separated into three steps: An ‘epigenator’, an ‘epigenetic initiator’ and an ‘epigenetic maintainer’ (see Fig. 2). The epigenator is a signal coming from the outside of the cell, for example a change in the developmental niche, a differentiation signal or a change in the temperature. It triggers the epigenetic initiator which acts locus-specifically on the DNA or the

chromatin. Epigenetic initiators can be for example non-coding RNAs (ncRNAs) or DNA-binding proteins. Finally, epigenetic maintainers fix the epigenetic state by leading to DNA methylation or histone tail modifications (Berger et al., 2009).

After the discovery of parental-specific effects it nearly took another decade until the first imprinted genes were described, namely *Insulin-like growth factor 2 receptor* (*Igf2r*) on mouse chromosome 17 (Barlow et al., 1991), *Insulin-like growth factor 2* (*Igf2*) on mouse chromosome 7 (DeChiara et al., 1991) and *H19*, a non-coding RNA (ncRNA) found in close proximity to *Igf2* (Bartolomei et al., 1991). Up to now approximately 110 imprinted genes are known in mice most of which are conserved in humans (http://www.har.mrc.ac.uk/research/genomic_imprinting/maps.html). Approximately half of them show expression from the maternal allele, the other one from the paternal allele. They are distributed over 15 mouse chromosomes with no apparent bias for a particular chromosomal location (see Fig. 3)

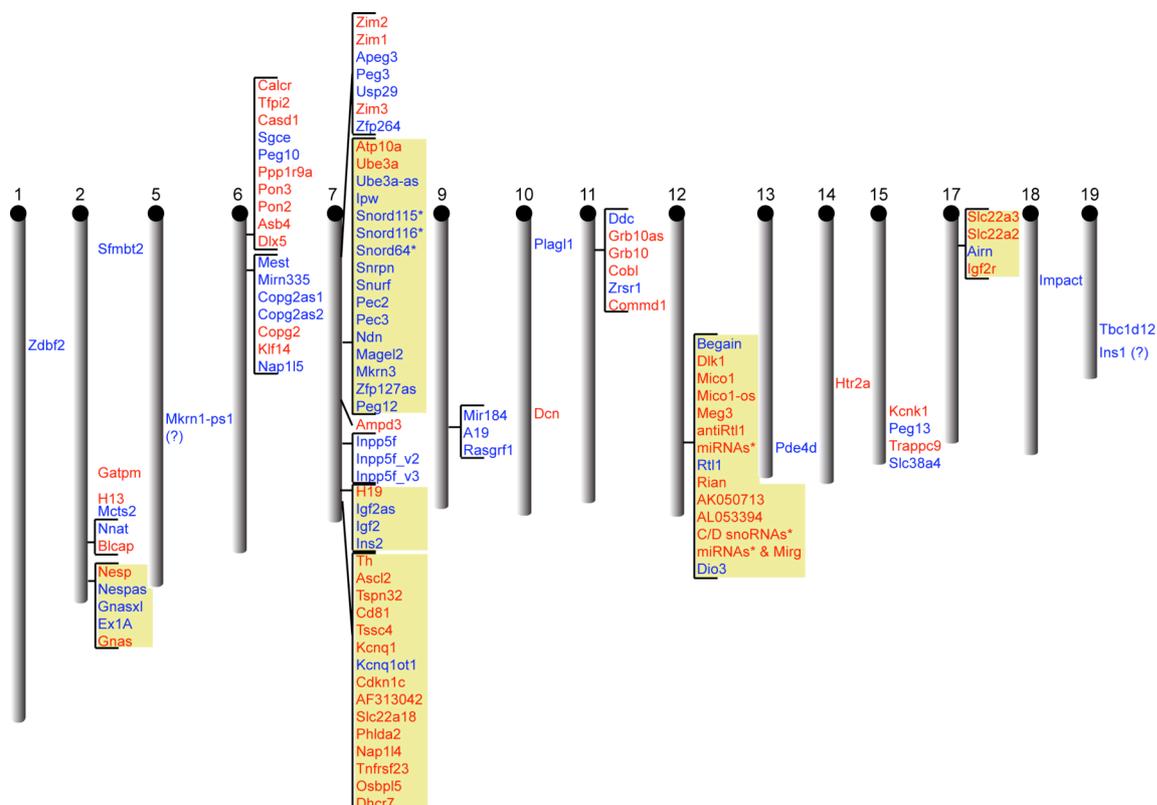


Fig. 3: Imprinted regions in the mouse. Genes in red are maternally expressed, in blue are paternally expressed. Vertical grey bars: chromosomes. *: small nucleolar RNAs and microRNAs. (?): conflicting data. Marked in green are the 6 well-studied imprinted clusters described in detail below (modified from www.har.mrc.ac.uk/research/genomic_imprinting).

During the last two decades many tools have been developed to understand and dissect the genomic imprinting process. Although the epigenators involved in genomic imprinting are hardly known, the epigenetic initiators – non-coding RNAs and DNA-binding proteins like CTCF – are well-established. The way how the epigenetic initiators act is not always understood in detail, however, the epigenetic maintainers are well analysed. Furthermore, also the timing of the epigenetic pathway is well described for imprinted genes. Therefore genomic imprinting is an ideal model system to study genome-wide epigenetic processes. The mechanisms involved in genomic imprinting will be described in detail further below.

1.1.3 Evolution of genomic imprinting

As described above, being diploid provides a big advantage of a second gene copy which can back up mutations occurring on one allele. As genomic imprinting abandons this advantage due to the presence of only one functional allele, several theories have been postulated to explain the evolutionary advantage of genomic imprinting.

One of the earliest proposed explanations for the evolution of genomic imprinting was the prevention of parthenogenesis (Solter, 1988). Most species can reproduce asexually, either as a common mechanism of reproduction or at least after experimental intervention. Nevertheless, mammals are not able to do so. Genomic imprinting could have evolved to prevent parthenogenesis by demanding genetic input from both parents. Sexual reproduction involves recombination between the parental genomes and this leads to two advantages compared to asexual reproduction: First, it allows faster adaptation and second it can remove deleterious mutations from the genome (Engelstadter, 2008). The hypothesis that genomic imprinting arose to prevent parthenogenesis is however not fully sufficient to explain the evolution of genomic imprinting. Maternal-specific gene inactivation alone could lead to the prevention of parthenogenesis, but also paternally inherited alleles are subjected to imprinted gene silencing. However, in support of this theory, that

genomic imprinting evolved to prevent parthenogenesis, bi-maternal embryos were successfully created by constructing oocytes containing one haploid genome from a fully-grown oocyte and one from a non-growing oocyte where the ICEs from the paternally imprinted *Igf2* and the *Dlk1* cluster have been deleted. These modifications led to the expression of the normally maternally silent *Igf2* and *Dlk1* genes, enabling the survival of bi-maternal mice (Kawahara et al., 2007). This experiment indicates that indeed only two maternally silenced genes prevent parthenogenesis. Therefore imprinted gene silencing for at least those two genes could have evolved to prevent parthenogenesis in mammals.

The ovarian-time-bomb hypothesis proposes that imprinting arose to protect females from the danger of trophoblastic diseases (Varmuza and Mann, 1994). One special feature of mammalian development is the strong invasion of the placenta into the uterine tissue of the mother. The first differentiation step in mammalian embryogenesis produces the cells of the inner cell mass which later give rise to the embryo proper and the trophoblast cells which invade into the uterine epithelium of the mother. Since the maternal parent carries several hundred thousand oocytes, this step could potentially pose a threat to the mother if spontaneously activated oocytes lead to parthenogenetic embryos that were able to implant and attach to the maternal blood supply. Interestingly, parthenogenetic embryos produced by nuclear transfer experiments show a strongly reduced trophoblast (Surani et al., 1984). Spontaneously activated oocytes can develop *in situ* and can lead to ovarian tumours, however, they are mostly benign (Varmuza and Mann, 1994). If one assumes that genomic imprinting impairs the development of malign trophoblast diseases by inactivating trophoblast genes in oocytes, it rapidly would be fixed during evolution as it would provide a high selection advantage. An argument against this theory is of course again, that not only maternally inherited genes are subjected to imprinted gene silencing, but also paternally inherited genes.

Up to now the function of 28 imprinted genes has been studied in detail and approximately half of them affect growth of the embryo or extraembryonic tissues (http://www.har.mrc.ac.uk/research/genomic_imprinting/function.html) (see Table 1).

Most maternally expressed growth-regulatory genes repress growth, whereas paternally expressed growth-regulatory genes usually enhance growth. This observation led to the proposal of the ‘parental conflict hypothesis’ (Moore and Haig, 1991). This hypothesis is based on the parent-specific interests of resource flow from the mother to the embryo. Males – especially in polyandrous populations – cannot be sure that they fathered all litters from one female. Therefore they bequeath active growth enhancers to favour the development of their offspring over those sired by other males. Females however provide all the nutrients during intrauterine development. Larger offspring have a higher nutrient and energy consumption and might compete out smaller offspring or weaken the mother thus impairing future pregnancies. Thus the mother expresses growth suppressors, to counteract the paternal interest. Although this theory is applicable to many imprinted genes, there are nevertheless some points which argue against this hypothesis: First, not all imprinted genes regulate growth like *Snrpn* and *Ube3a* which are involved in RNA-editing and ubiquitination (Leff et al., 1992; Rougeulle et al., 1997). Second, not all growth promoting genes are paternally expressed, like the maternally expressed gene *Ascl2* which supports the growth of the placenta (Guillemot et al., 1995). Third, a high competition between different paternal alleles would lead to a rapid evolution of imprinted genes. However, they were found to evolve at a frequency comparable to non-imprinted genes (Hurst and McVean, 1998).

Maternally expressed genes			Paternally expressed genes		
<i>Gnas</i>	-	(1)	<i>Gnasxl</i>	+	(1)
<i>Phlda2</i>	-	(2)	<i>Mest</i>	+	(6)
<i>Ascl2</i>	+	(3)	<i>Peg10</i>	+	(7)
<i>Grb10</i>	-	(4)	<i>Peg3</i>	+	(8)
<i>Igf2r</i>	-	(5)	<i>Igf2</i>	+	(9)
			<i>Kcnq1ot1</i>	+	(10)
			<i>Rasgrf1</i>	+	(11)
			<i>Zac1</i>	+	(12)
			<i>DIK1</i>	+	(13)

Table 1: Imprinted genes affecting embryonic or extraembryonic growth. Genes marked with (+) have a growth promoting effect whereas genes marked with (-) show a growth suppressing effect. References: (1) (Yu et al., 2000) (2) (Frank et al., 2002) (3) (Guillemot et al., 1995) (4) (Charalambous et al., 2003) (5) (Wang et al., 1994) (6) (Lefebvre et al., 1998) (7) (Ono et al., 2006) (8) (Li et al., 1999) (9) (DeChiara et al., 1990) (10) (Mancini-Dinardo et al., 2006) (11) (Itier et al., 1998) (12) (Varrault et al., 2006) (13) (Moon et al., 2002).

Although each of the theories described above give a feasible explanation for the evolution of genomic imprinting at particular loci, it might not be possible to have a unifying theory to explain the origin of genomic imprinting for all imprinted genes. All imprinted genes would be expected to be dose dependent so that imprinted monoallelic expression would show an immediate phenotype that could be selected. However not all genes showing imprinted expression are dosage sensitive. It could be that not all imprinted genes were selected for monoallelic gene expression. It was suggested, that some genes with imprinted expression are innocent bystanders. They show imprinted expression only because they are located close to an imprinted gene selected for monoallelic expression and the mechanism controlling imprinted expression of this gene took control also of other genes adjacent to it (Varmuza and Mann, 1994). Up to now we do not know which genes drove selection and which were carried along - perhaps we could understand this better by knowing how genes are imprinted - to see if there really are innocent bystanders.

1.2 Mechanisms regulating genomic imprinting

1.2.1 Imprinted genes are clustered and controlled by an imprint control element

Most imprinted genes are found in clusters containing up to 15 imprinted genes (see Fig. 3, 4). Imprinted clusters contain maternally and paternally expressed genes, but can also contain biallelically expressed genes. Extensive analyses have been done for six of those clusters (*Igf2r*, *Kcnq1*, *Igf2*, *Gnas*, Prader-Willi-/Angelman-Syndrome (*Pws/As*), *Dlk1*) (reviewed by da Rocha et al., 2008; Ideraabdullah et al., 2008; Peters and Williamson, 2008; Regha et al., 2006; Verona et al., 2003). Besides the clustered imprinted genes, several solo-imprinted genes were found, like *Impact* and *Sfmbt2* (see Fig. 3) (Hagiwara et al., 1997; Kuzmin et al., 2008). Future experiments will show if they are really single imprinted genes or if more genes close by are imprinted as well.

The organisation of imprinted genes in clusters allows coordinated imprinted expression of several genes by a single regulatory sequence, called the imprint control element (ICE) or imprint control region (ICR). ICEs are CpG-rich DNA regions and are often identified as CpG islands, genomic regions which display a significantly higher CpG content than the average genome and are usually devoid of DNA methylation (Antequera, 2003). The ICE is genetically defined using deletion experiments in mice or the mapping of minimal naturally occurring deletions in human. In contrast to most CpG islands, the ICE can gain DNA methylation in one specific developmental stage - in gametogenesis. The ICE is modified only on one parental chromosome by a DNA methylation 'imprint', which is acquired during maternal or paternal gametogenesis and is maintained on the same parental chromosome in the diploid embryo (Stoger et al., 1993). Since the other parental chromosome lacks ICE DNA methylation, this region in a diploid cell is also known as a gametic differentially methylated region (gDMR). Deletions of the ICEs have shown that it is the unmethylated ICE which controls the imprinted expression of the whole cluster. Deletion of the methylated ICE does not change imprinted expression in a cluster. Upon deletion of the unmethylated ICE however imprinted genes are no longer expressed in a parental-specific pattern in the six well-studied imprinted clusters mentioned above (Bielinska et al., 2000; Fitzpatrick et al., 2002; Lin et al., 2003; Thorvaldsen et al., 1998; Williamson et al., 2006; Wutz et al., 1997). Note that the term 'imprinted' is used here to refer to the presence of DNA methylation on the ICE and not to gene expression status and that the above-mentioned deletion experiments show that only the unmethylated ICE is active (modified from Koerner et al., 2009).

Four of the six well-studied imprinted clusters in the mouse (*Igf2r*, *Kcnq1*, *Pws/As*, *Gnas*) are maternally imprinted and thus gain their ICE DNA methylation imprint during oogenesis (Coombes et al., 2003; Engemann et al., 2000; Shemer et al., 1997; Stoger et al., 1993). This imprint is maintained only on the maternal chromosome in diploid cells (see Fig. 4 A-D). The remaining two clusters (*Igf2*, *Dlk1*) are paternally imprinted and gain their ICE DNA methylation imprint during spermatogenesis (Takada et al., 2002; Tremblay et al., 1995). This imprint is

maintained only on the paternal chromosome in diploid cells (see Fig. 4 E-F) (modified from Koerner et al., 2009). In addition to gDMRs also so-called somatic DMRs (sDMRs) are sometimes found in imprinted clusters at promoter regions or within imprinted genes but outside of the ICE (see Fig. 4) (Chai et al., 2001; Feil et al., 1994; Hanel and Wevrick, 2001; Kelsey et al., 1999; Lewis et al., 2004b; Stoger et al., 1993; Takada et al., 2002). They are gained later during development at a stage where cells are already diploid. As the appearance of sDMRs is a consequence of the imprinting process, they cannot be used to discriminate the parental origin of an allele at the initiation stage of imprinted expression.

1.2.2 Imprinted clusters contain macro non-coding RNAs

In all six well-studied clusters, at least one gene encodes for a macro ncRNA. The term 'macro' refers to ncRNAs being longer than 200bp and whose function does not depend on the processing to small ncRNAs like siRNAs or miRNAs (Gingeras, 2007). In the *Igf2r*, *Kcnq1* and *Gnas* clusters, the ICE contains a promoter for a macro ncRNA (respectively: *Airn* (108kb), *Kcnq1ot1* (91kb), and *Nespas* (~30kb)) that has an overlap in antisense direction with only one gene in each imprinted cluster (Smilnich et al., 1999; Wroe et al., 2000; Wutz et al., 1997). In the *Pws/As* cluster, the provisionally named *Snrpn-long-transcript* (*Snrpnlt*, also known as *Lncat*) is an unusually long macro ncRNA that may cover 1000kb of genomic sequence (Landers et al., 2004). The *Snrpnlt* ncRNA overlaps in antisense orientation the *Ube3a* gene which is located 720kb downstream of the *Snrpnlt* transcriptional start site. The size of the ICE in this cluster is not precisely determined in the mouse, as the smallest available ICE deletion only shows a partial or mosaic imprinting defect (Bressler et al., 2001). In the paternally-imprinted *Igf2* and *Dlk1* clusters, the ICE is found 5-10kb upstream of the *H19* macro ncRNA (2.2kb) and the provisionally named *Gtl2-long-transcript* (*Gtl2lt*; might span up to 208kb of genomic sequence), respectively. *H19* lacks any known transcriptional overlap with the other genes in the cluster, whereas *Gtl2lt* overlaps *Rtl1* (da Rocha et al., 2008; Pachnis et al., 1984; Seitz et al., 2004; Tierling et al., 2006) (modified from Koerner et al., 2009).

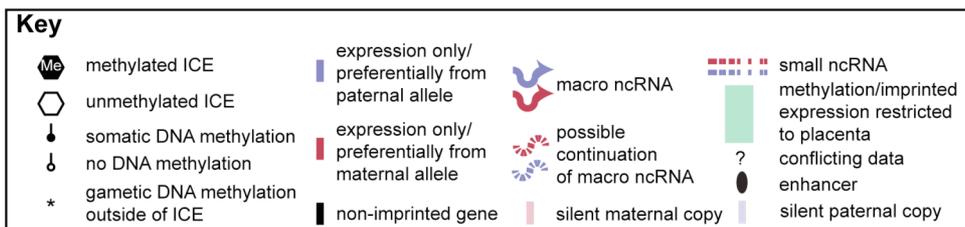
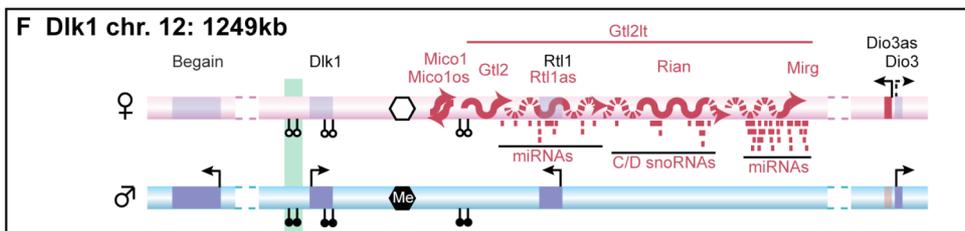
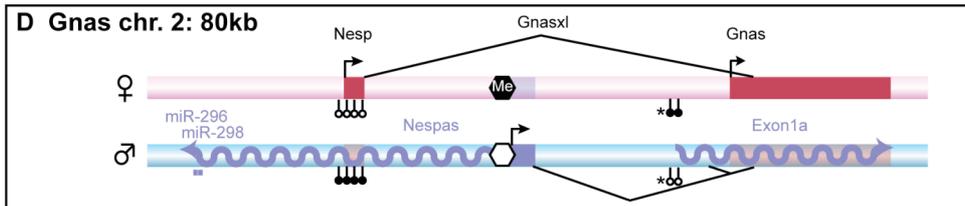
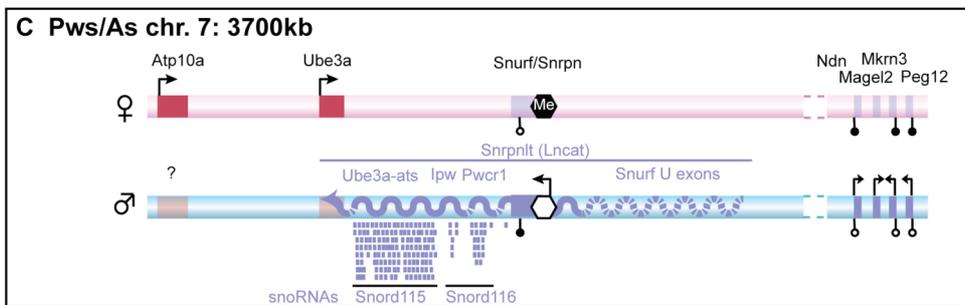
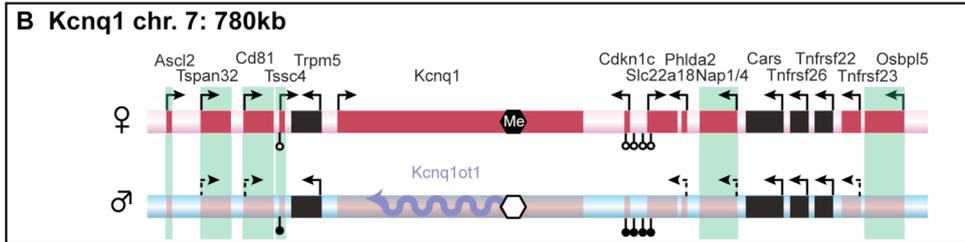
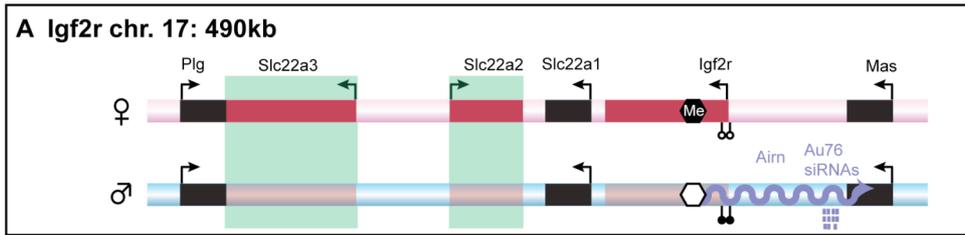


Fig. 4: The genomic organisation of six well-studied mouse imprinted clusters is shown. The maternal chromosome (chr.) is shown as a pink bar, the paternal chromosome as blue bar. Protein-coding genes are shown as boxes. Solid red box: maternally expressed gene on the expressed allele. Transparent red box: maternally expressed gene on the repressed allele. Solid blue box: paternally expressed gene on the expressed allele. Transparent blue box: paternally expressed gene on the repressed allele. Macro ncRNAs are shown as wavy lines, red for maternally expressed, blue for paternally expressed. Arrows indicate transcriptional direction: solid arrows - strong transcription, dashed arrows - weak transcription. Note that many of the indicated genes show tissue- or temporal-restricted gene expression (not indicated). kb: kilobasepairs. See key for further details (modified from Koerner et al., 2009).

Although the organization of these six well-studied imprinted clusters appears to be complex, they generally follow two simple rules: (i) an unmethylated ICE is required for macro ncRNA expression; and (ii) most imprinted mRNA genes are not expressed from the chromosome from which the macro ncRNA is expressed (modified from Koerner et al., 2009).

1.2.3 Three mechanisms for the regulation of imprinted expression by the ICE

Up to now, three different mechanisms have been described how the unmethylated ICE regulates imprinted expression of the genes in its corresponding cluster. In two separate imprinted clusters two paternally-expressed imprinted macro ncRNAs *Airn* and *Kcnq1ot1*, were shown to be necessary for imprinted expression by experiments in which the ncRNA was truncated by the insertion of a polyadenylation cassette into the endogenous locus. Paternal chromosomes that carry a truncated *Airn* or *Kcnq1ot1* ncRNA lose the repression of all protein-coding genes in the imprinted cluster in both embryonic and placental tissues, while maternal alleles were unaffected (Mancini-Dinardo et al., 2006; Shin et al., 2008; Sleutels et al., 2002). These experiments showed that these macro ncRNAs act by repressing multiple flanking genes *in cis* in both embryo and placental tissues (see Fig. 5A). *Nespas* is similar to *Airn* and *Kcnq1ot1* in that it is transcribed from a promoter contained in the unmethylated ICE on the paternal allele and that it has an antisense orientation to the imprinted protein-coding *Nesp* gene. However, it is not yet known if *Nespas* has a similar *cis*-silencing role (modified from Koerner et al., 2009).

In contrast, the maternally-expressed *H19* ncRNA is dispensable for the imprinted

expression of *Igf2* (Schmidt et al., 1999). Instead, a methylation-sensitive insulator contained in the ICE regulates the ability of enhancers that lie downstream of *H19* to interact physically with the upstream *H19*, *Igf2* and *Ins2* promoters. On the unmethylated maternal allele CTCF binds the ICE and restricts the access of enhancers to the *H19* promoter. On the methylated paternal allele CTCF cannot bind and the enhancers interact preferentially with the *Igf2* and *Ins2* promoters, activating their transcription (Bell and Felsenfeld, 2000; Hark et al., 2000) (see Fig. 5B) (modified from Koerner et al., 2009). Interestingly, although *H19* does not have a *cis*-regulatory role in the imprinting process of the *Igf2* cluster it was shown to regulate expression of *Igf2* and at least four other imprinted genes (*Igf2r*, *Cdkn1c*, *Gnas*, *Dlk1* and *Rtl1*) *in trans*. If this function of *H19* is attributable to *H19* itself or to the miRNA encoded within its transcriptional unit is however still unclear (see Fig. 4) (Gabory et al., 2009).

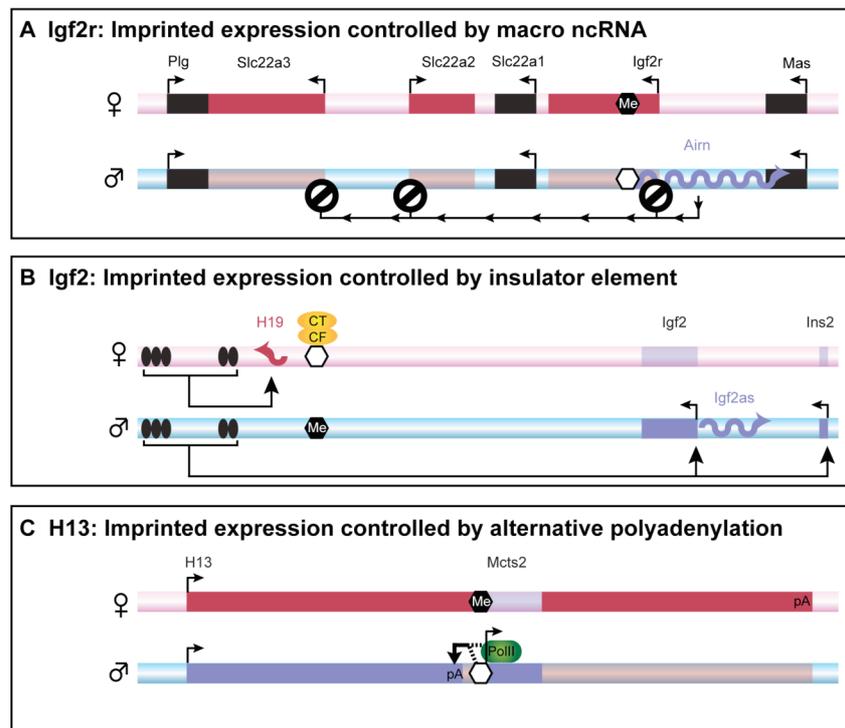


Fig. 5: Three different mechanisms for ICE mediated regulation of imprinted expression. (A) Imprinted expression can be regulated by the expression of a macro ncRNA from a promoter contained in the ICE which leads to silencing of other imprinted genes *in cis*, like in the *Igf2r* cluster. (B) The ICE can act as a methylation-sensitive insulator like in the *Igf2* cluster. (C) The ICE can control differential polyadenylation (pA) like in the *H13* cluster. See text for further details. Key and details as in Fig. 4.

A third mechanism to induce imprinted expression was shown in the newly described *H13* imprinted cluster on mouse chromosome 2 (not shown in Fig. 4). The *H13* (Histocompatibility-13 antigen) gene contains an intronic, maternally-methylated gDMR. The transcription of full-length functional *H13* from the maternal chromosome depends on the methylation of this gDMR. On the paternal allele, the unmethylated gDMR acts as a promoter for the *Mcts2* retrogene, and *Mcts2* expression correlates with the premature polyadenylation of *H13* (Wood et al., 2008). To date, it is not known if *Mcts2* expression or the unmethylated gDMR is required to block production of full-length *H13* transcripts (see Fig. 5C) (modified from Koerner et al., 2009).

1.2.4 Imprinted expression can be restricted to certain tissues or developmental time points

Genomic imprinting consists of a cycle of events that begins when the ICE DNA methylation imprint is established on one parental allele during gametogenesis. After fertilization, when the embryo is diploid, the ICE methylation imprint is maintained on the same parental allele through the action of the DNA methyltransferase DNMT1 (Li et al., 1993). All tissues in the embryo and the adult with the exception of late germ cells maintain the gDMR and it can be used to faithfully identify the parental origin of chromosomes in somatic cells. Despite the universal presence of the DNA methylation imprint, not all imprinted genes in a cluster show ubiquitous imprinted expression, but their imprinted expression might be restricted to certain tissues or certain developmental time points. Imprinted expression can be maintained by developmentally or tissue-specifically regulated factors or lost in the absence of such factors. Temporal- and tissue-specific imprinted expression could be achieved by regulating ncRNA expression (for the *Igf2r* and *Kcnq1* clusters) or regulating insulator formation (for the *Igf2* cluster). To complete the genomic imprinting life cycle, the ICE methylation imprint is erased during early germ-cell development, to allow the parental gametes to acquire the correct DNA methylation imprint ready for the next generation (see Fig. 6) (Barlow and Bartolomei, 2007) (modified from

Koerner et al., 2009).

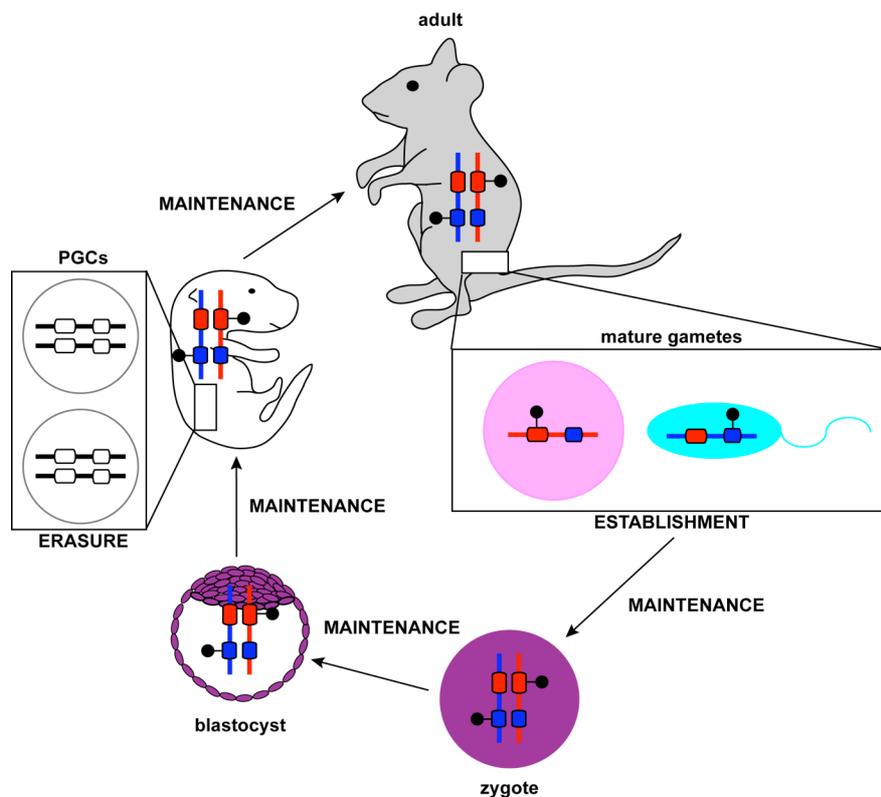


Fig. 6: Lifecycle of the imprint. The DNA methylation imprint (black lollipop) is established on the ICE (blue and red rectangle) during gamete formation that occurs shortly after birth. In oocytes (pink filled circle), methylation is acquired on maternally methylated ICEs (red rectangle) whereas in sperm cells (blue filled circle with tail) methylation occurs on paternally methylated ICEs (blue rectangle). Imprints are maintained on the same chromosome when the embryo becomes diploid after fertilization (purple filled circle) and during embryonic development. Finally, imprints are erased in the primordial germ cells (PGCs; white circles) of the developing embryo. Shown is a chromosome with two ICEs, one is maternally methylated, the other one paternally methylated. Red line: maternal chromosome. Blue line: paternal chromosome.

Tissue-specific variation in imprinted expression has been well studied in only a few imprinted clusters. In the *Igf2r* cluster, the ICE shows maternal-specific methylation in all diploid cells. However, imprinted expression of the genes in the cluster does not take place in all tissues and during whole development (see Fig. 4A). Undifferentiated ES cells and preimplantation embryos do not show imprinted expression at all. At this stage, *Igf2r* is biallelically expressed whereas *Airn* expression is not detectable (Latos et al., 2009; Lerchner and Barlow, 1997; Szabo and Mann, 1995; Terranova et al., 2008). After implantation or during differentiation of ES cells *Airn* starts to be expressed and leads to imprinted expression of *Igf2r* (Latos et al., 2009; Lerchner and Barlow, 1997; Szabo and Mann, 1995; Terranova

et al., 2008). As soon as imprinted expression of *Igf2r* is established, *Igf2r* and *Airn* show ubiquitous imprinted expression in post-implantation embryos, extraembryonic and adult tissues (Latos et al., 2009; Terranova et al., 2008). The only exception are post-mitotic neurons which lack *Airn* expression but show biallelic expression of *Igf2r* (Yamasaki et al., 2005). Imprinted expression of *Slc22a2* and *Slc22a3* is restricted to the placenta. *Slc22a2* shows imprinted expression at all stages examined (Zwart et al., 2001). *Slc22a3* shows imprinted expression 11.5 days post coitum (dpc), at 15.5dpc however, it shows biallelic expression and expression levels decrease at later stages (Verhaagh et al., 2001; Zwart et al., 2001). Both genes display biallelic expression in some embryonic and adult tissues. *Slc22a2* shows biallelic expression in embryonic and adult kidney (Mooslehner and Allen, 1999; Seidl, 2006; Zwart et al., 2001), whereas *Slc22a3* is biallelically expressed in adult heart (Zwart et al., 2001).

Also in the *Kcnq1* cluster the ICE shows differential DNA methylation of the ICE in all tissues (see Fig. 4B). However only two genes, *Kcnq1ot1* and *Cdkn1c* show ubiquitous imprinted expression in embryonic, extraembryonic and adult tissues (Mancini-DiNardo et al., 2003; Smilnich et al., 1999; Umlauf et al., 2004). Imprinted expression of *Cd81*, *Slc22a18*, *Phlda2* and *Kcnq1* is temporally regulated. *Kcnq1* shows imprinted expression in the placenta at all stages examined as well as in the early embryo; from 14.5dpc however it is biallelically expressed in embryonic tissues (Caspary et al., 1998; Gould and Pfeifer, 1998; Mancini-DiNardo et al., 2003; Umlauf et al., 2004). *Cd81* shows imprinted expression in both, embryo and placenta at 8.5dpc, later however it is biallelically expressed in both lineages (Caspary et al., 1998; Umlauf et al., 2004). *Slc22a18* shows imprinted expression only during embryonic development, in the embryo proper as well as the placenta, but in the adult expression is mostly biallelic (Dao et al., 1998). Similarly, *Phlda2* which shows a strong maternal bias in embryonic and extraembryonic tissues, shows a relaxation of imprinted expression in adult spleen and kidney (Qian et al., 1997). Imprinted expression of the other genes in the cluster is regulated in a tissue-specific manner. *Tnfrsf23* shows a strong maternal bias in extraembryonic but only a weak maternal bias in embryonic tissues (Clark et al., 2002). Imprinted expression of five genes in

this cluster, *Osbp15*, *Nap114*, *Tssc4*, *Tspan32* and *Ascl2*, is restricted to extraembryonic tissues only (Engemann et al., 2000; Umlauf et al., 2004).

Generally, extraembryonic tissues provide a good example of tissue-specific variation in imprinted expression as the majority of imprinted genes in the mouse only show imprinted expression in the placenta. In many imprinted clusters, a small number of centrally positioned genes show 'ubiquitous' imprinted expression (i.e., in embryo, placenta and adult), whereas additional genes in the cluster that lie further away from the ICE have imprinted expression only in the placenta (see Fig. 4). As experiments that involve either ICE deletion or ncRNA truncation (described above) show that imprinted expression in the embryo and placenta is controlled by the same elements, there are two possible explanations for this phenomenon: either the ICE or the ncRNA act differently in these two tissues to repress genes; or the placenta allows spreading of the basic mechanism that operates in embryonic tissue (Hudson et al., 2010; Miri and Varmuza, 2009; Wagschal and Feil, 2006) (modified from Koerner et al., 2009).

1.3 Which functions are controlled by the ICE and what controls the ICE itself?

1.3.1 A role for DNA methylation in regulating imprinted expression

As described above, deletion experiments showed that only the unmethylated ICE is active in inducing silencing of flanking protein-coding genes either by activating a ncRNA promoter, by forming an insulator or by regulating alternative polyadenylation. Thus the ICE can be viewed as a *cis*-acting repressor and DNA methylation as a modification to repress this repressor. Analysis of ICE methylation can offer insights into how this epigenetic modification is attracted to specific sequences and how it is used to inhibit ncRNA transcription and insulator function. In the maternal germline, the DNA methyltransferase-like protein DNMT3L, in concert with the DNA methyltransferase DNMT3A, are crucial players in the establishment of

ICE germline DNA methylation (Bourc'his and Proudhon, 2008). The subsequent maintenance of ICE methylation requires the DNMT1 family of DNA methyltransferases (Hirasawa et al., 2008). Additional proteins, such as the Krüppel-associated box zinc finger protein ZFP57, are also required for acquiring ICE methylation in the *Pws/As* cluster and for maintaining ICE methylation in the *Dlk1* and three other imprinted clusters not shown in Fig. 4, but play no role in the *Igf2* and *Igf2r* clusters (Li et al., 2008). Although the exact mechanism by which ZFP57 acts is unknown, this finding raises the possibility that each ICE requires different additional factors for the acquisition and maintenance of germline DNA methylation (modified from Koerner et al., 2009).

Exactly how *de novo* DNA methylation enzymes recognize ICE sequences is unclear. As described above, epigenetic initiators acting locus-specifically are needed to recruit DNA methylation and histone modifications which are not locus-specific but rather could follow gene expression (Berger et al., 2009). The identity of those initiators however is still not clear. Furthermore, it is still unknown if the DNA methylation is actively targeted to the ICE or if DNA methylation is the default state and the unmethylated ICE has to be protected from it. For the *Igf2r* ICE specific sequences have been identified by the injection of DNA fragments into the male or female pronucleus: an allele discrimination sequence (ADS) and a *de novo* methylation sequence (DNS) which are necessary for the establishment but not for the maintenance of the differential DNA methylation of the *Igf2r* ICE (Birger et al., 1999). Interestingly it was shown by a band-shift assay using nuclear extracts from normal, parthenogenetic and androgenetic ES cells that an unknown protein binds to the DNS in normal and androgenetic, but not in parthenogenetic ES cells (Birger et al., 1999). Therefore it was concluded, that the protein is of paternal origin and it binds to the DNS to protect it from acquiring DNA methylation (Birger et al., 1999).

If DNA methylation however would be directly targeted to the ICE, this could be achieved by a specific sequence, a specific secondary structure or another specific feature like transcription. No conserved sequence can be found between ICEs of different imprinted clusters. Many ICEs however contain a run of tandem direct

repeats that have been suggested to form a secondary structure that induces DNA methylation (Neumann et al., 1995). Tandem direct repeats are significantly enriched in imprinted genes compared to non-imprinted control genes (Hutter et al., 2006). A comparison of the tandem direct repeats of *Snurf/Snrpn*, *Kcnq1* and *Igf2r* in mouse, human and other primates revealed, that although the tandem direct repeats in those three imprinted clusters have a unique sequence, they all have a similar size, a high CpG content, an ordered arrangement of CpG dinucleotides and similar predicted secondary structures (Paoloni-Giacobino et al., 2007). Therefore structural features of the DNA induced by the tandem direct repeats could serve as the imprinting signal. Using crystallography it was shown, that DNMT3A and DNMT3L interact directly and form a tetrameric complex with two active sites, which are separated from each other by the size of about one DNA helical turn. Further analysis indicated, that DNMT3A methylates DNA in a periodic pattern of 8 to 10 basepairs (Jia et al., 2007). Thus, the tandem nature of the repeats could support DNA methylation by DNMT3A.

Interestingly, the transcription of overlapping imprinted protein-coding genes in oocytes, is needed to acquire methylation imprints in the *Gnas* cluster (Chotalia et al., 2009). This work shows that truncation of the *Nesp* mRNA transcript (Fig. 4D) by insertion of a polyadenylation cassette that abolishes transcription through the ICE, impairs acquisition of the ICE methylation mark in oocytes. This might represent a common theme for oocyte-specific DNA methylation imprints, given that an overlapping mRNA gene was shown to be transcribed through five other maternal gDMRs in oocytes, but not in sperm. However, further work is still needed to determine exactly how the methylation machinery in oocytes targets ICE sequences (modified from Koerner et al., 2009).

As mentioned above, in addition to the gametic differential DNA methylation some imprinted genes gain a somatic differentially methylated region (sDMR) during development. One of those sDMRs is found on the *Igf2r* promoter (see Fig. 4A). In undifferentiated ES cells and preimplantation embryos when *Igf2r* shows biallelic expression, both alleles are unmethylated. However, upon imprinted expression of

Igf2r, the paternal promoter gains DNA methylation, whereas the maternal, highly active promoter does not (Latos et al., 2009; Lerchner and Barlow, 1997; Stoger et al., 1993; Szabo and Mann, 1995; Terranova et al., 2008). Interestingly, the two other imprinted mRNA genes in this cluster, *Slc22a2* and *Slc22a3* that show imprinted expression only in placenta do not gain an sDMR on its promoter in this tissue (Zwart et al., 2001). In the *Kcnq1* cluster, three mRNA genes, *Cdkn1c*, *Slc22a18* and *Tssc4* have sDMRs (see Fig. 4B). Whereas the first two show this sDMR in all embryonic, extraembryonic and adult tissues examined, *Tssc4* only shows it in extraembryonic tissues (Lewis et al., 2004b). The presence of an sDMR also reflects the expression status, as *Cdkn1c* and *Slc22a18* show imprinted expression in both embryonic and extraembryonic tissues, whereas imprinted expression of *Tssc4* is restricted to the extraembryonic lineage (Lewis et al., 2004b; Umlauf et al., 2004).

The function of DNA methylation on DMRs was analysed using targeted deletions of DNMT1 in mice. These knock-out mice have basically no or very little DNA methylation and die during embryonic development at approximately 8.5-9.5dpc (Li et al., 1993). In the *Igf2r* cluster, absence of DNA methylation leads to twofold upregulation of *Airn* expression, as *Airn* now can be expressed from both parental alleles. *Igf2r* expression is reduced to basal levels, as *Airn* now represses it on both parental alleles (Li et al., 1993; Seidl et al., 2006). This experiment demonstrates, that *Airn* expression is controlled directly by DNA methylation, whereas *Igf2r* expression is in turn repressed by *Airn* in the absence of DNA methylation. The gDMR therefore has a direct function, whereas the sDMR seems to be rather a consequence than a cause of the silencing of the paternal *Igf2r* allele by *Airn*. In the *Kcnq1* locus, disruption of DNMT1 leads to loss of methylation on the ICE and thus biallelic expression of *Kcnq1ot1*. Surprisingly, also *Kcnq1*, *Slc22a18* and *Cdkn1c* showed biallelic expression in embryonic tissues, although with decreased expression levels from the maternal allele which indicates a direct role of DNA methylation in the silencing of those genes (Lewis et al., 2004b). An early study however found biallelic expression of *Cdkn1c*, but not of *Kcnq1* in embryos deficient for DNMT1 (Caspary et al., 1998). But even more surprising was the effect of a

deficiency of DNMT1 in extraembryonic tissues: Whereas *Kcnq1ot1*, *Cdkn1c* and *Slc22a18* showed biallelic expression, *Kcnq1*, *Cd81*, *Ascl2* and *Tssc4* maintained imprinted expression (Caspary et al., 1998; Lewis et al., 2004b), indicating that in extraembryonic tissues imprinted expression of the latter four genes is independent of DNA methylation. However, it is clearly dependent on the ICE, as deletion of the ICE in a *Dnmt1*^{-/-} background led to biallelic expression of *Kcnq1*, *Cd81*, *Ascl2* and *Tssc4* (Lewis et al., 2004b). Therefore, another mechanism than DNA methylation must operate in extraembryonic tissues to control imprinted expression. In agreement with this theory it was shown, that differential histone modifications mark imprinted genes in extraembryonic tissues in the *Kcnq1* cluster and they themselves are regulated by the ICE (Lewis et al., 2004b).

1.3.2 The ICE controls higher chromosomal organisation in the *Igf2* cluster

For the imprinted *Igf2* cluster it has been shown, that the ICE controls higher chromosomal organisation which in turn plays a crucial role in regulating imprinted expression. Depending on its methylation status it interacts with different sequences on the maternal and the paternal allele to control imprinted expression by regulating enhancer access (Engel et al., 2008; Murrell et al., 2004).

Furthermore it was shown, that the ICE of the *Igf2* cluster also controls DNA FISH asynchrony. The phenomenon of imprinted genes showing DNA FISH asynchrony was discovered in the early 1990s. Unsynchronised cells are subjected to DNA fluorescence *in situ* hybridisation (DNA FISH) (Selig et al., 1992). The hybridisation of sequence specific fluorescently labelled probes like cosmids or BACs in the course of the DNA FISH method results in three different hybridisation patterns in cells which are in S phase: Two single spots (SS), if only one sister chromatid for each allele can be detected, two double spots (DD) if both sister chromatids for both alleles can be detected, or a single and a double spot (SD) if for one allele one sister chromatid and for the other both sister chromatids can be detected. For most genomic regions, mainly SS and DD patterns are present, an SD pattern is only

found in a minority of cells (below 20%). However, imprinted regions were shown to be enriched for the SD pattern (30 to 40%) (Kitsberg et al., 1993). Subsequently, a whole range of imprinted clusters was analysed for the presence of this phenomenon and it was consistently found to be present in different clusters like the *Igf2r*, *Igf2*, *Pws/As* and *Kcnq1* cluster in mouse and human (Bickmore and Carothers, 1995; Gribnau et al., 2003; Kagotani et al., 2002; Kitsberg et al., 1993; Simon et al., 1999; Smrzka et al., 1995). For the paternally imprinted *Igf2* cluster a genomic 13kb deletion comprising *H19*, the ICE and a downstream region, was shown to lead to a loss of DNA FISH asynchrony upon maternal transmission of the deletion that removed the unmethylated active ICE (Greally et al., 1998; Gribnau et al., 2003; Leighton et al., 1995).

Also a differential subnuclear localisation was shown to be associated with the *Igf2* gene cluster. In fetal liver cells, the paternal allele was found at the periphery, the maternal allele was found rather in the nuclear center. In ES cells, the parental alleles showed the opposite orientation with the maternal allele found close to the nuclear periphery. A different subnuclear localisation was therefore present in both analysed cell types, however, a correlation between the allele first showing a double spot in the DNA FISH assay and the nuclear localisation could not be drawn, the localisation rather seems to depend on the cell type (Gribnau et al., 2003). Interestingly, also the differential subnuclear localisation was disturbed upon maternal transmission of the 13kb deletion described above, because in fetal liver cells both alleles were now localised at the periphery (Gribnau et al., 2003).

In summary, the ICE of the *Igf2* cluster does control three features of higher chromosomal organisation: (i) three dimensional chromosomal conformation, (ii) DNA FISH asynchrony and (iii) subnuclear localisation of the parental alleles. It is not known yet, if all ICEs control those features, as it was shown for the maternally imprinted *Snrpn* cluster in humans, that a 5-30kb big deletion comprising the ICE on the paternal allele containing the unmethylated active ICE does not abolish DNA FISH asynchrony (Gunaratne et al., 1995).

1.3.3 The ICE controls ncRNA expression – does it also control ncRNA biology?

As shown by the ICE deletion experiments, the ICE clearly controls imprinted expression of the ncRNAs found in imprinted clusters, and as the function of the ICE is controlled by DNA methylation, ncRNAs are controlled by DNA methylation as well (Bressler et al., 2001; Fitzpatrick et al., 2002; Lin et al., 2003; Thorvaldsen et al., 1998; Williamson et al., 2006; Wutz et al., 2001). In contrast to the majority of mammalian mRNA-encoding genes that are intron rich, many imprinted ncRNAs are unspliced or spliced with a low intron/exon ratio. Notably, the export of the ncRNA to the cytoplasm correlates with splicing. The *H19* ncRNA is fully spliced and exported to the cytoplasm (Brannan et al., 1990; Pachnis et al., 1984), whereas the *Kcnq1ot1* ncRNA is unspliced and retained in the nucleus (Pandey et al., 2008). The *Airn* ncRNA produces unspliced to spliced transcripts at a ratio of 95:5, and only the spliced transcripts are exported to the cytoplasm (Seidl et al., 2006). In the *Gnas* cluster, both *Exon1A* and *Nespas* ncRNAs are also found as spliced and unspliced forms, but their cellular localization is unknown (Holmes et al., 2003; Li et al., 2000; Liu et al., 2000; Williamson et al., 2006). Both *Airn* and *Kcnq1ot1* were shown by RNA Fluorescence In Situ Hybridization (RNA FISH) to form RNA ‘clouds’ at their site of transcription (Braidotti et al., 2004; Nagano et al., 2008; Terranova et al., 2008). We do not yet know whether these ncRNA ‘clouds’ explain their ability to repress flanking genes or whether this ‘cloud-like’ appearance is a consequence of a lack of splicing, as mRNA genes mutated to inhibit splicing also show nuclear retention and intranuclear RNA focus formation reminiscent of clouds (Custodio et al., 1999; Ryu and Mertz, 1989) (modified from Koerner et al., 2009).

Even though not all imprinted macro ncRNAs have been studied in sufficient detail, at least some have been shown to have unusual transcriptional features that are not generally associated with mammalian mRNA genes, such as reduced splicing potential or low intron/exon ratio, nuclear retention and accumulation at the site of transcription. Investigations into the control of these unusual transcriptional features, and into their role in the functional properties of imprinted ncRNAs, are only just

beginning. Of course it would be tempting to speculate, that the ICE controls those properties. Three ncRNAs, *Airn*, *Kcnq1ot1* and *Nespas* originate from a promoter within the ICE. Therefore at least in those cases an obvious genetic element localised within the ICE that might account for unusual transcriptional features is the promoter. It was recently shown in fission yeast that splicing regulation is promoter-driven (Moldon et al., 2008). It was therefore surprising to find that the endogenous *Airn* promoter does not control the low splicing capability of the *Airn* ncRNA as replacement of the endogenous *Airn* promoter with a *Pgk* promoter did not change the low splicing ability of *Airn* (Stricker et al., 2008). Other elements within the ICE but outside of the promoter could also influence the biology of those ncRNAs which are transcribed through the whole – as *Snrpnlt* – or parts of the ICE – as *Airn*, *Kcnq1ot1* and *Nespas*. Such an element could be for example tandem direct repeats as they might lead to secondary structures within the ncRNA which then in turn could directly or after binding by specific proteins modify ncRNA features (modified from Koerner et al., 2009).

1.3.4 Imprinted expression by ncRNAs – transcript or transcription?

In both placental and embryonic tissues, the repression of multiple genes in the *Igf2r* and the *Kcnq1* clusters on the paternal chromosome depends on the *Airn* and *Kcnq1ot1* macro ncRNAs (Mancini-Dinardo et al., 2006; Shin et al., 2008; Sleutels et al., 2002). However, the mechanism by which these ncRNAs induce repression is unknown. One significant open question is whether the ncRNA itself or the act of its transcription is required for silencing (Koerner et al., 2009; Pauler et al., 2007). If it is the ncRNA product which leads to imprinted expression of the mRNA genes, it could be either by an RNAi pathway or by RNA-directed targeting.

In the RNAi pathway, double stranded RNAs (dsRNAs) complementary to the target gene lead to repression by inducing degradation or translational inhibition of the target (Mattick and Makunin, 2006). As *Airn* and *Kcnq1ot1* are both overlapping one imprinted gene in their respective cluster, RNAi was an early proposed mechanism

for imprinted gene silencing. It was suggested, that the RNAi machinery would first silence the overlapping transcript and the silencing mechanism then would spread to the other imprinted genes in the cluster (Rougeulle and Heard, 2002). At least for the *Igf2r* cluster this model however was excluded, as mice deficient for *Igf2r* expression still were able to show imprinted expression of *Slc22a2* and *Slc22a3* (Sleutels et al., 2003). Alternatively, local regions of sequence homology shared between the ncRNA and all imprinted mRNA genes, or local inverted regions forming short hairpins, would allow an RNAi-mediated silencing mechanism (see Fig. 7A). However, the regulation of imprinted expression takes place *in cis* and RNAi is generally a *trans*-acting mechanism. Although a *cis*-acting RNAi mechanism has been discovered in *Schizosaccharomyces pombe*, homologues for the involved *cis*-restricted RITS (RNA-induced transcriptional silencing) complex have not yet been identified in mammals (Wassenegger, 2005).

The RNA-directed targeting model follows the generally accepted model for *Xist* ncRNA mediated silencing involved in X chromosome inactivation. On the future inactive X chromosome *Xist* is expressed, spreads and localises to the whole X chromosome *in cis*. This leads to the recruitment of several histone modifying proteins and histone variants as well as DNA methylation, which results in gene silencing (Wutz and Gribnau, 2007). Imprinted ncRNAs could function in an analogous way (see Fig. 7B).

In contrast to RNA-based models, it was proposed previously that the *Airn* ncRNA might silence *Igf2r* because of transcriptional interference, and that *Airn* may act solely by transcription *per se* (Pauler et al., 2007). According to this model, ncRNA transcription either interferes directly with transcriptional initiation or with the activity of essential *cis*-regulatory elements (see Fig. 7C).

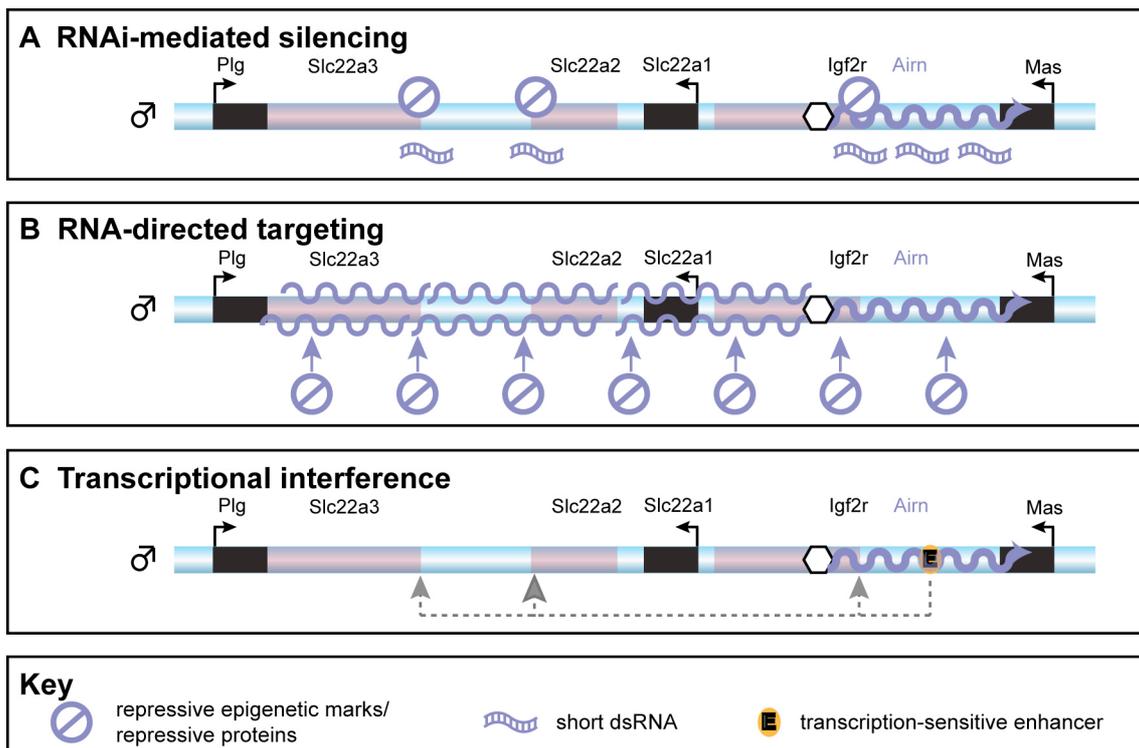


Fig. 7: Models for ncRNA mediated silencing. Only the paternal allele is shown. (A) RNAi-mediated silencing: dsRNAs are generated from the ncRNA *Airn*, either due to transcriptional overlap with *Igf2r* or by the formation of regional double stranded regions within *Airn* itself. These dsRNAs trigger silencing by recruiting the RNAi machinery. (B) RNA-directed targeting: The *Airn* transcript spreads along the whole region and leads to the recruitment of repressive proteins or repressive epigenetic marks. (C) Transcriptional interference: The act of transcription of the ncRNA interferes directly with the transcription of *Igf2r* or with a transcription-sensitive enhancer (orange ellipse with E) crucial for the expression of the imprinted mRNA genes. All other elements as in Fig. 4.

1.3.5 Histone modifications – epigenetic maintainers at the ICE and imprinted genes

Above I described how DNA methylation directly regulates ICE activity but does not directly silence imprinted protein-coding genes. This section focuses on current progress in understanding the potential roles played by histone modifications in restricting macro ncRNA expression to one parental allele and imprinted protein-coding genes to the other parental allele. Recent publications have shown that the ICE carries allele-specific histone modifications that are specific to the DNA methylated or the unmethylated allele (see Fig. 8). Genome-wide sequencing, as well as oligo tiling arrays, have been used to show that the DNA methylated ICE is marked by focal repressive histone modifications of the type found in constitutive

centromeric and telomeric heterochromatin such as H3K9me3 (trimethylation of lysine 9 on histone 3), H4K20me3 (trimethylation of lysine 20 on histone 4) and HP1 (heterochromatin protein 1) (Mikkelsen et al., 2007; Regha et al., 2007). In the *Igf2r* and *Kcnq1* clusters, the repressive H3K27me3 (trimethylation of lysine 27 on histone 3) mark is present on the methylated ICE only in undifferentiated ES cells. In the *Igf2r* cluster, H3K27me3 is absent from embryonic fibroblasts, but in the *Kcnq1* cluster, it is present in both embryo and placenta (Latos et al., 2009; Lewis et al., 2006; Lewis et al., 2004b; Umlauf et al., 2004). The unmethylated ICE lacks repressive modifications but carries active histone modifications, such as H3K4me and H3/H4 acetylation. The presence of active and repressive histone modifications on the same DNA sequence that modify different parental chromosomes, can be used to identify an ICE. The usefulness of this approach was demonstrated in a genome-wide study of diploid ES cells that identified short regions carrying both repressive H3K9me3 and active H3K4me3 on the ICE of the six imprinted clusters in Fig. 4 (Mikkelsen et al., 2007) (modified from Koerner et al., 2009).

The histone modification profiles established so far show that repressive marks are associated with the DNA methylated ICE, whereas active marks are associated with the unmethylated ICE. Although it has proven to be relatively simple to assign a function to DNA methylation in regulating ICE activity, a general function for histone modifications has not yet been identified. Repressive H3K9me3 modifications are regulated by three known histone methyltransferases, SUV39H1, SUV39H2 and ESET, whereas the repressive H4K20me3 is regulated by SUV4-20H1 and SUV4-20H2 (Allis et al., 2007a; Allis et al., 2007b; Kouzarides, 2007). The repressive H3K9me3 mark is maintained and even enhanced on the ICE in embryonic fibroblasts lacking SUV39H1 and SUV39H2, whereas the repressive H4K20me3 is reduced in embryonic cells lacking SUV4-20H1 and SUV4-20H2, without removing either the repressive H3K9me3 or DNA methylation (Pannetier et al., 2008; Regha et al., 2007). The ESET methyltransferase was found to bind to the *Igf2r* ICE as well as the *Igf2r* promoter and is a good candidate for the histone methyltransferase setting the H3K9me3 mark. However, its role could not be tested directly because ESET deficient cells are not viable at any developmental stage (Dodge et al., 2004; Regha

et al., 2007). Thus suitable genetic systems are not yet available to test the role of the repressive H3K9me3 in regulating ICE activity (modified from Koerner et al., 2009).

In contrast to the lack of a defined role for histone modifying enzymes regulating ICE activity, several reports describe a role for these enzymes in regulating placental but not embryonic, imprinted expression. The Polycomb group protein EED, which is required for repressive H3K27me3 modifications, has been shown to repress the paternal allele of 4 out of 18 tested imprinted genes in embryos at 7.5dpc that mainly consist of extraembryonic tissue at this stage (Mager et al., 2003). The affected genes were contained in three different imprinted clusters, in which the majority of genes maintained correct imprinted expression. This indicates that EED does not play a general role in regulating imprinted expression, but is attracted to specific genes. The G9A histone lysine methyltransferase that dimerises with G9A-like protein (GLP) to induce repressive H3K9me2 (dimethylation of lysine 9 on histone 3) modifications is necessary for the paternal repression of some genes in the *Kcnq1* and *Igf2r* clusters in the placenta, but not in the embryo (Nagano et al., 2008; Wagschal et al., 2008). As mentioned above, in embryonic tissue repressive heterochromatin (H3K9me3, H4K20me3, HP1), but not repressive H3K27me3, modifies the DNA methylated ICE in a focal manner. Not all of these repressive modifications have been mapped throughout imprinted clusters in placenta, but repressive H3K9me2/3 and H3K27me3 marks were found at the promoters of monoallelically silenced mRNA genes in the *Igf2r* cluster (Nagano et al., 2008). Tissue-specifically silenced genes however showed widespread H3K27me3 in embryonic tissues (Pauler et al., 2009). Both, H3K27me3 and H3K9me2/3 were found to be widespread in the placenta on the chromosome carrying the silenced mRNA genes in the *Kcnq1* cluster (Umlauf et al., 2004). In one study in placenta (Nagano et al., 2008), both active and repressive histone modifications were found on genes that showed placental-specific imprinted expression. Although this might indicate the existence of 'bivalent' domains (Bernstein et al., 2006), care should be taken in interpreting these results owing to the risk of maternal tissue contamination in placenta samples (modified from Koerner et al., 2009).

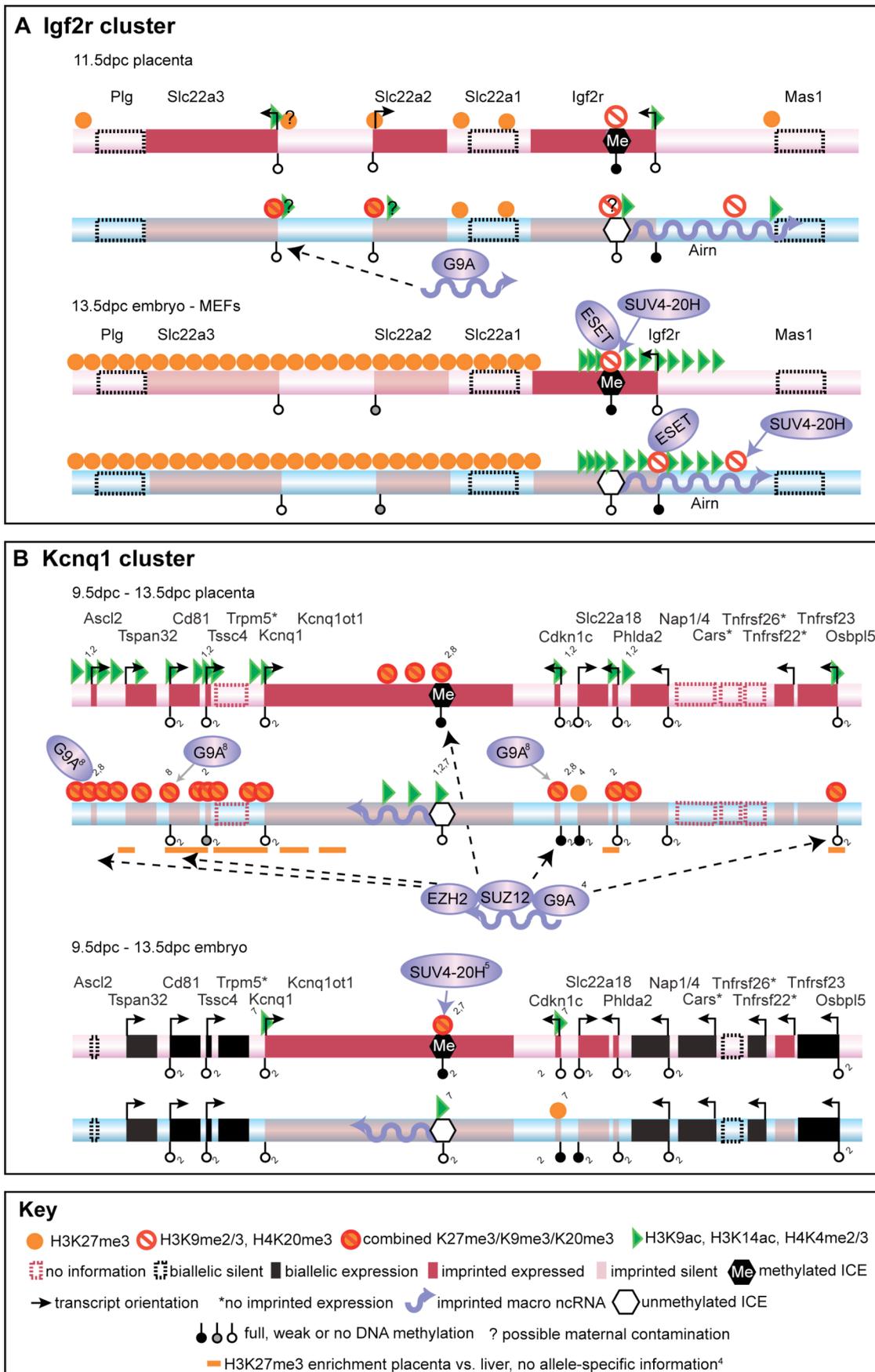


Fig. 8: Allele-specific histone modifications at the *Igf2r* and *Kcnq1* clusters. Active and repressive histone modifications on the maternal or paternal allele are shown for the *Igf2r* (A) and *Kcnq1* (B) clusters for embryo and placenta. For simplicity, histone modifications are combined into three groups: repressive constitutive heterochromatin (H3K9me2/3, H4K20me3), repressive facultative heterochromatin (H3K27me3) and active euchromatin (H3K9ac, H3K4me2/3). In both clusters, only the indicated positions were assayed in the placenta, with the exception of one ChIP-chip mapping of H3K27me3 (orange lines). (A) In the *Igf2r* cluster, an unbiased continuous ChIP-chip mapping was performed in the embryo, with the indicated positions showing enrichment. (B) In the *Kcnq1* cluster, all positions shown in the placenta were also assayed in the embryo, and only indicated positions show allele-specific enrichment. Blue shaded ellipses indicate histone-modifying enzymes, and the solid blue arrows extending from the ellipses indicate the involvement of the respective enzyme in the setting or maintaining of a histone mark. (A, B) Both *Kcnq1ot1* and *Airn* ncRNAs bind to histone modifying enzymes. (B) Dotted arrows indicate binding of *Kcnq1ot1* ncRNA to chromatin. Abbreviations: H3: Histone 3, H4: Histone 4, K: Lysine, me3: trimethylation, me2: dimethylation, ac: acetylation. For symbols see key. Details as in Fig. 4. Gene expression marked by an asterisk is inferred from the presence of active histone modifications (Mikkelsen et al., 2007). References: *Igf2r* placenta: (Nagano et al., 2008), *Igf2r* embryo: (Regha et al., 2007). *Kcnq1* placenta: all histone modification without a numbered reference: (Umlauf et al., 2004). Others: ¹(Green et al., 2007), ²(Lewis et al., 2004b), ³(Mikkelsen et al., 2007), ⁴(Pandey et al., 2008), ⁵(Pannetier et al., 2008), ⁶(Regha et al., 2007), ⁷(Umlauf et al., 2004), ⁸(Wagschal et al., 2008) (Figure modified from Koerner et al., 2009).

In summary, the analysis of histone modifications shows that the same active and repressive histone modifications that correlate with expressed and silent genes, also modify imprinted genes in an allele-specific manner. Further work is needed to determine which modifications reflect the cause or consequence of imprinted expression. While there is currently no indication that histone modifications cooperate with DNA methylation to restrict macro ncRNA expression to one parental allele, there is emerging data that in the placenta, histone modifications may play a role together with a macro ncRNA in repressing imprinted mRNA genes *in cis* (modified from Koerner et al., 2009).

Two recent studies indicate that the *Airn* and *Kcnq1ot1* ncRNAs themselves are directly involved in silencing genes in the placenta. *Kcnq1ot1* was found to localize physically to several silent genes on the paternal allele that lay hundreds of kilobase pairs away from the *Kcnq1ot1* promoter (Pandey et al., 2008). This finding is supported by RNA/DNA FISH, which showed partial overlap between the *Kcnq1ot1* RNA and the flanking imprinted genes in the *Kcnq1* cluster in the trophoblast cells of early embryos that contributes to the placenta (Terranova et al., 2008). Furthermore, *Kcnq1ot1* also directly interacts with the Polycomb group proteins EZH2 and SUZ12, which are necessary for establishing the repressive H3K27me3

mark, and with G9A, which is involved in setting the repressive H3K9me2 mark (Pandey et al., 2008). Together, this indicates that in the placenta, the *Kcnq1ot1* ncRNA localizes to chromatin and targets histone methyltransferases to the whole imprinted cluster (see Fig. 8). Notably, embryos that are deficient for G9A and the Polycomb proteins EZH2 and RNF2 show a loss of paternal repression of some of the placental-specific imprinted genes in the *Kcnq1* cluster (Terranova et al., 2008; Wagschal et al., 2008). Similarly in the placenta, the *Airn* ncRNA in the *Igf2r* cluster lies in close proximity to the silent *Slc22a3* promoter and was shown to bind G9A (see Fig. 8). In addition, G9A null embryos show a loss of placental imprinted expression of *Slc22a3* but maintain imprinted *Igf2r* expression (Nagano et al., 2008). An RNA FISH study of the *Airn* and *Kcnq1ot1* ncRNAs in TS cells and in preimplantation trophectoderm cells has also shown that both ncRNAs are located in nuclear domains that are characterized by a high density of repressive H3K27me3 and a lack of active histone modifications and RNA polymerase II (Terranova et al., 2008). In summary, the evidence so far indicates that the *Airn* and *Kcnq1ot1* ncRNAs induce imprinted expression by an RNA-directed targeting mechanism in the placenta that only effects genes showing placental-specific imprinted expression. According to the proposed model, the ncRNA expressed from the unmethylated ICE is maintained at the site of transcription and associates with chromatin *in cis*. The ncRNA could localize throughout the cluster or to specific gene promoters by looping and subsequently attracts specific histone modifications that repress transcription of multiple genes lying at some distance from the ncRNA gene itself (modified from Koerner et al., 2009).

As described above, recent work suggests that in placenta the ncRNA transcript itself mediates gene silencing. However, imprinted ncRNAs may exert different effects in the mouse embryo compared with the placenta, as only a subset of the genes in the imprinted gene clusters show imprinted expression in embryonic and adult somatic tissue. In the *Igf2r* cluster, the *Airn* ncRNA represses *Igf2r* in embryos, but represses *Igf2r*, *Slc22a2* and *Slc22a3* in the placenta. In the *Kcnq1* cluster, the *Kcnq1ot1* ncRNA silences *Kcnq1*, *Cdkn1c*, *Slc22a18* and *Phlda2* in the embryo, but an additional six genes in the placenta. A further indication of differences between

the embryo and the placenta is that G9A and EED, required respectively for repressive H3K9me2 and H3K27me3 modifications and for imprinted expression of some placental genes, appear to play no role in the imprinted expression in the *Igf2r* and *Kcnq1* clusters in embryonic tissues (Mager et al., 2003; Nagano et al., 2008; Wagschal and Feil, 2006). Therefore, different silencing mechanisms may be used in embryonic and extraembryonic tissues. Several lines of evidence support a model in which *Airn* silences *Igf2r* by transcriptional interference in embryonic tissue. First, *Airn* has a low half-life of approximately 90 minutes, which argues against a function for the ncRNA in targeting repressive chromatin, as this would require it to be stable at least for one cell cycle (Seidl et al., 2006). Second, *Airn* does not induce widespread repressive chromatin in embryos (Regha et al., 2007). Third, the ability of *Airn* to silence *Igf2r* is dependent on promoter strength, a feature associated with transcriptional interference (Shearwin et al., 2005; Stricker et al., 2008). The *Igf2r* and *Kcnq1* clusters differ in that the *Kcnq1ot1* ncRNA represses multiple genes in the embryo, and as it is contained entirely within *Kcnq1* (Pandey et al., 2008), it does not overlap a promoter. However, it is possible to propose a transcriptional interference model of silencing for this ncRNA by postulating the existence of crucial *cis*-regulatory elements like enhancers overlapped by *Kcnq1ot1*. Although there is less evidence to support a transcriptional interference model for *Kcnq1ot1*, the lack of widespread repressive chromatin marks on genes in this cluster that show imprinted expression in the embryo (Pandey et al., 2008; Umlauf et al., 2004), as well as the absence of a role for G9A and EED (Mager et al., 2003; Wagschal et al., 2008), does indicate that RNA-mediated targeting does not operate in embryonic tissues (modified from Koerner et al., 2009).

1.4 *In vitro* models to study genomic imprinting

An excellent tool to study mechanisms that regulate imprinted expression are *in vitro* models. It was shown recently for the *Igf2r* cluster that embryonic stem (ES) cell differentiation mimics the onset of imprinted expression and the gain of epigenetic modifications seen in the developing embryo (Latos et al., 2009) (see Fig. 9). This

work establishes the utility of ES cells to study imprinted expression typical of embryonic tissue. However, since imprinted expression of the *Slc22a2* and *Slc22a3* genes in the *Igf2r* cluster is restricted to extraembryonic tissues, they could not be analysed in the ES cell differentiation system. Trophoblast stem (TS) cells are an obvious ES cell analogue for the study of genes that show imprinted expression only in placental tissues. However, differentiated TS cells appear to be an unsuitable model for the later stages of placental development, as the expression patterns and epigenetic modifications detected *in vivo* are not recapitulated *in vitro* (Lewis et al., 2006).

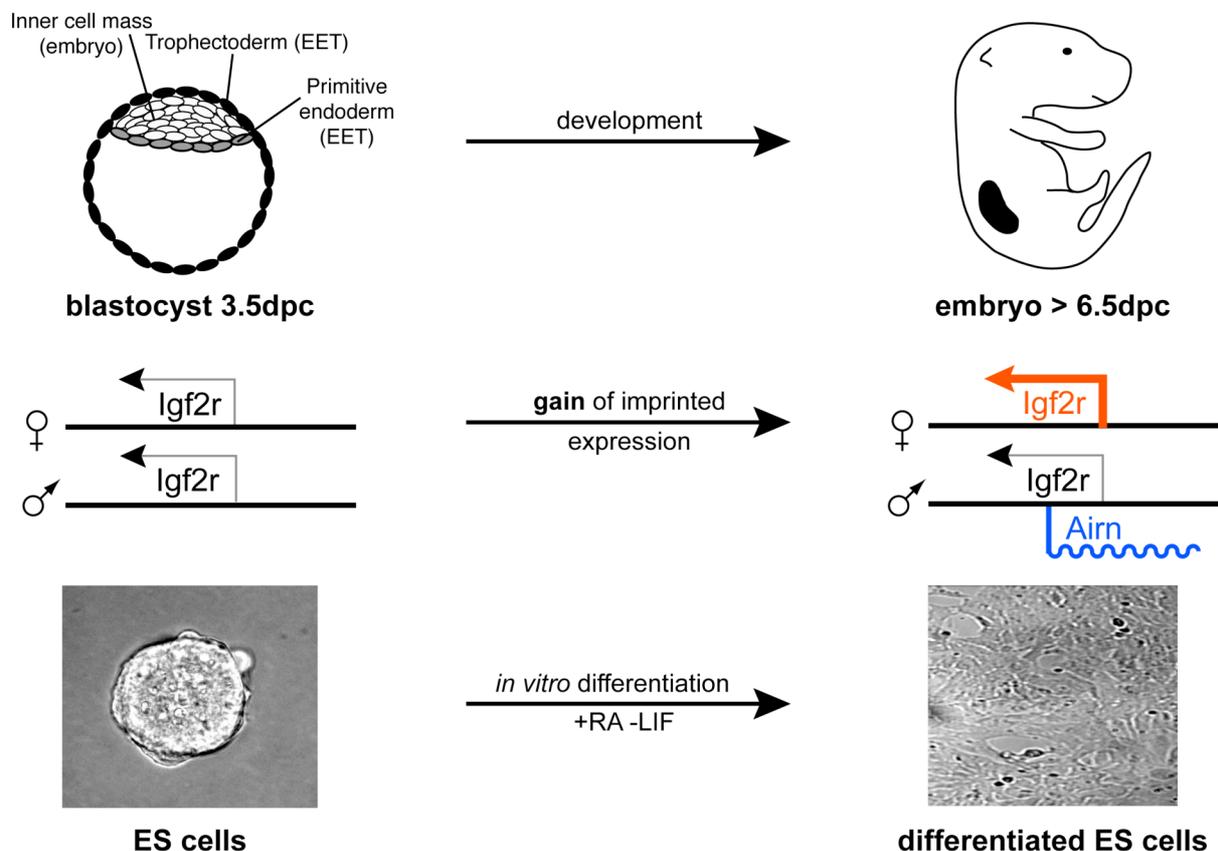


Fig. 9: *In vitro* differentiation system to study imprinted expression of the *Igf2r* cluster. Preimplantation embryos show low level but biallelic expression of *Igf2r* and no expression of *Airn*. After implantation *Airn* is expressed and prevents the upregulation of *Igf2r* from the paternal allele leading to imprinted expression of *Igf2r*. This gain of imprinted expression can be recapitulated during ES cell differentiation. At 3.5dpc cells from the inner cell mass of the blastocyst can be taken into culture and grown as ES cells, showing biallelic *Igf2r* and no *Airn* expression. During differentiation by the addition of retinoic acid (RA) and LIF withdrawal *Airn* starts to be expressed leading to imprinted expression of *Igf2r* (Figure modified from Latos et al., 2009).

1.5 Aim of this study

In my thesis I have investigated the role of the ICE in the mouse *Igf2r* imprinted cluster in regulating information at two levels. First, I analysed the role of the ICE at the chromosomal level in regulating the phenomenon of DNA FISH asynchrony. Second, I examined the role of the tandem direct repeats and the CpG island within the ICE in influencing the maintenance of DNA methylation at the gDMR and in controlling *Airn* ncRNA expression and function. Furthermore, I set up a model system that can be used in the future to test if the tandem direct repeats in the ICE play a role in the establishment of the differential DNA methylation of the gDMR *in vivo* during gamete formation.

For the first part I generated a whole range of primary mouse embryonic fibroblasts (pMEFs) with targeted manipulations of the *Igf2r* cluster on one or both parental alleles using mice which have been generated previously. I used those pMEFs for DNA FISH and set up S phase fractionation to analyse allelic DNA replication.

By DNA FISH I mapped the proximal border of DNA FISH asynchrony in the *Igf2r* cluster and could therefore show that DNA FISH asynchrony in the *Igf2r* cluster extends over 3Mbp of genomic sequence. I furthermore showed, that the ICE does not regulate this DNA FISH asynchrony and neither does *Airn* or the sDMR. For the *Igf2* cluster I could confirm, using a 3.8kb deletion, that in this cluster the maternal unmethylated ICE controls DNA FISH asynchrony. Furthermore I was able to show that treatment with Trichostatin A, a general histone deacetylase inhibitor, leads to a relaxation of the DNA FISH asynchrony observed in the *Igf2r* cluster. By using three-dimensional DNA FISH I found that chromatin compaction is not different between the maternal and the paternal allele containing the *Igf2r* cluster. Lastly, I present data that indicates, that DNA FISH asynchrony between the maternal and paternal *Igf2r* allele is not due to a difference in DNA replication during S phase.

For the second part of the thesis I participated in the generation of ES cells carrying a SNP in *Igf2r*, providing a valuable tool for *in vitro* experiments analysing the

regulation of imprinted expression. I furthermore created ES cell lines with a targeted deletion of the tandem direct repeats within the ICE of the *Igf2r* cluster. This deletion was made in inbred wildtype ES cells, inbred ES cells with the SNP in *Igf2r* and in an intraspecies ES cell line. Furthermore I improved a qPCR assay for detection of allelic *Igf2r* expression levels making use of the SNP in *Igf2r*. In addition, in the inbred ES cells with the SNP in *Igf2r* I also generated a targeted deletion of the CpG island localised within the ICE.

The inbred ES cells carrying the tandem direct repeat deletion allele were analysed using the *in vitro* ES cell model system described with respect to possible changes in DNA methylation or *Airn* ncRNA features. I could show that deletion of the repeats does not change the maintenance of the gDMR nor the gain of the sDMR in these cells before and after differentiation. With respect to RNA biology, deletion of the repeats reduces the length of *Airn*, but does neither change steady-state levels of the known *Airn* splice variants nor nuclear localisation of *Airn*. Deletion of the repeats however seems to lower the efficiency of *Airn* to prevent upregulation of the paternal *Igf2r* allele during ES cell differentiation. In the future, the intraspecies ES cell line carrying a repeat deletion allele can be used for generating a knock-out mouse to answer the question if the repeats play a role in the establishment of the gDMR or in regulating imprinted expression of *Slc22a2* and *Slc22a3*.

In ES cells, the deletion of the CpG island from the paternal allele leads to a gain of DNA methylation *in cis*, indicating, that one function of the CpG island is to keep the paternal ICE unmethylated. Furthermore, expression of *Airn* is dramatically decreased and *Igf2r* shows biallelic expression upon deletion of the *Airn* CpG island.

2. RESULTS

2.1 The role of the *Igf2r* ICE on the chromosomal level

2.1.1 Generation of primary mouse embryonic fibroblasts (pMEFs) for DNA FISH

In the *Igf2r/Airn* region, there are two differentially methylated regions (DMRs) present, the gametic DMR (gDMR) in the imprint control element (ICE) containing the *Airn* promoter and a somatic DMR (sDMR) on the *Igf2r* promoter. Mice carrying a deletion of those elements have been created previously, namely R2 Δ (region 2 deletion which is the deletion genetically defining the ICE) (Wutz et al., 2001) and IP Δ (*Igf2r* promoter deletion) (Sleutels et al., 2003). Paternal inheritance of the R2 Δ allele leads to complete loss of imprinted expression in the whole cluster, whereas inheritance of the IP Δ allele only abrogates expression of *Igf2r*. A truncation of the ncRNA *Airn* from 108kb to 3kb by insertion of a polyadenylation signal also has been generated previously in mice and has been named AirT (*Airn* truncation) (Sleutels et al., 2002). A paternally inherited AirT allele shows imprinted expression of the truncated *Airn*, but imprinted expression of *Igf2r*, *Slc22a2* and *Slc22a3* is lost. In this section I analysed, if the ICE, the sDMR or the ncRNA *Airn* have an effect on DNA FISH asynchrony. Since for the DNA FISH assay diploid cells are necessary, the use of established cell lines, which are mostly poly- and/or aneuploid is not possible (Todaro and Green, 1963). However, primary mouse embryonic fibroblasts (pMEFs) are mainly diploid. Therefore, I generated pMEFs lacking the ICE, the sDMR or full-length *Airn* either on the paternal, the maternal or both parental alleles.

For compaction analysis and allelic DNA replication studies I specifically wanted to investigate one parental allele. Therefore I also generated pMEFs carrying the naturally occurring 6Mbp large Thp deletion, which comprises the whole *Igf2r/Airn* cluster, either on the maternal or the paternal allele (Johnson, 1974; Johnson, 1975).

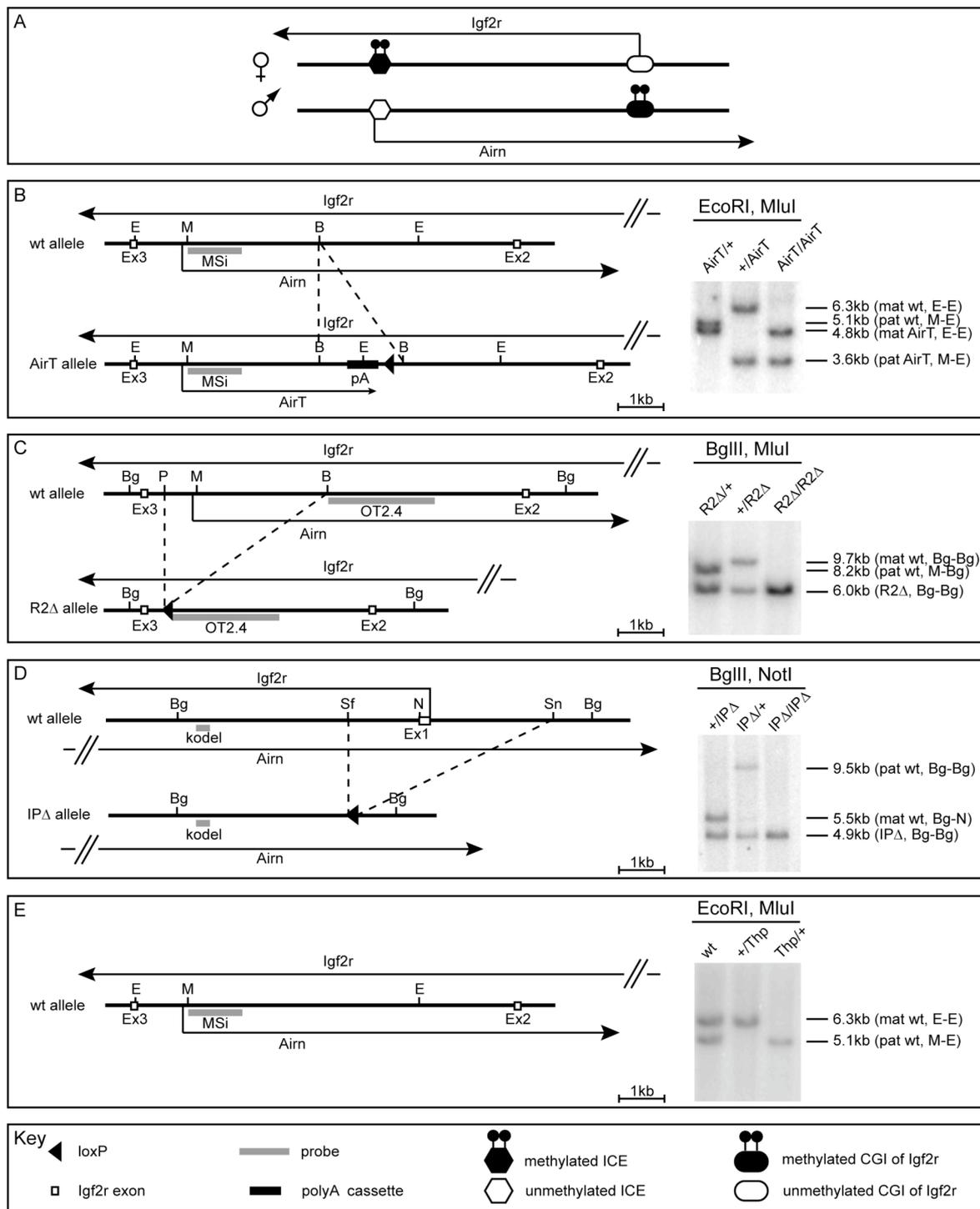


Fig. 10: Genotyping of embryos carrying a mutant allele in the *Igf2r/Airn* region by DNA blotting. To discriminate between a maternal and a paternal origin of the mutant allele, methylation-sensitive restriction enzymes were used in combination with a second restriction enzyme. Due to the allelic DNA methylation at the gDMR and sDMR the parental origin of the allele can be detected. (A) Schematic overview of the genomic regions carrying the *Igf2r* and *Airn* promoters. (B)-(E) On the left schematic drawings of the mutant and the corresponding wildtype (wt) allele are shown. For all genotypes the maternal allele is indicated first. On the right representative DNA blots for genotyping are shown. Probes used for DNA blotting are indicated as grey bars, used restriction enzymes are shown above the DNA blot picture. Shown are only restriction sites relevant for the generation or

screening of the mutant alleles. (B) The AirT allele was generated by the introduction of a rabbit β -globin polyadenylation cassette into the BamHI site 3kb downstream of the *Airn* transcriptional start (Sleutels et al., 2002). (C) The R2 Δ allele was generated by a deletion of the PacI-BamHI region surrounding the *Airn* promoter and genetically identified the ICE (Wutz et al., 2001). (D) The IP Δ allele was generated by deleting the SfuI-SnaBI fragment comprising the *Igf2r* promoter (Sleutels et al., 2003). Note that at 13.5dpc methylation of the paternal CpG island (CGI) comprising the *Igf2r* promoter is not always 100% complete (variation between littermates, unpublished). Therefore a certain percentage of paternal alleles still can be cut in an IP Δ /+ embryo, thus resulting in an additional 5.5kb fragment. (E) The Thp allele is a naturally occurring 6Mbp big deletion comprising the whole *Igf2r/Airn* imprinted cluster. Therefore only the schematic drawing of the wt allele is shown. Restriction enzymes: E: EcoRI, M: MluI, B: BamHI, Bg: BglIII, P: PacI, Sf: SfuI, N: NotI, Sn: SnaBI. Note that MluI and NotI are methylation-sensitive enzymes. Arrows indicate transcription in (A) and transcriptional orientation in (B)-(E).

To determine the parental origin of the mutant allele, methylation-sensitive restriction enzymes were used which only can cut if the underlying DNA sequence is unmethylated. In combination with methylation-insensitive restriction enzymes specific DNA fragments are generated due to the gDMR or the sDMR, which allow to clearly distinguish the parental origin of the mutant and wildtype allele after crossing heterozygous parents (see Fig. 10A). Representative examples of DNA blots are shown in Fig. 10B-E. Using this approach, I was able to generate not only wildtype (+/+) pMEFs, but also pMEFs carrying an AirT allele, an R2 Δ allele, an IP Δ allele or a Thp allele, either on the maternal allele (AirT/+, R2 Δ /+, IP Δ /+, Thp/+; note that the maternal allele is written on the left), the paternal allele (+/AirT, +/R2 Δ , +/IP Δ , +/Thp), or on both parental alleles (AirT/AirT, R2 Δ /R2 Δ , IP Δ /IP Δ). Before using the cells for experiments, their genotype was verified again using DNA harvested from pMEFs. In addition, their karyotype was analysed by metaphase spreads, ensuring that at least 80% of all cells still had 40 chromosomes (data not shown).

DNA FISH asynchrony has been well studied in the *Igf2/H19* cluster which loses DNA FISH asynchrony upon maternal transmission of the phil allele, where a neomycin resistance gene replaced 13kb of genomic DNA comprising the *H19* gene and a 9.9kb long upstream region including the ICE (Greally et al., 1998; Gribnau et al., 2003; Leighton et al., 1995). I received pMEFs containing this 13kb deletion upon maternal and paternal transmission as a gift from Karl Pfeifer (Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland, USA). Furthermore, I received cells from Marisa Bartolomei (University of

Pennsylvania, Philadelphia, USA) containing a 3.8kb deletion either on the paternal or the maternal allele which removes the ICE but leaves the *H19* gene intact (Thorvaldsen et al., 2002). To confirm the correct genotype I analysed genomic DNA from those cells again by methylation-sensitive restriction digest and DNA blotting (see Fig. 11). For both deletions, a maternal (*phil*+/+, 3.8kb Δ /+) and a paternal (+/*phil*, +/3.8kb Δ) transmission were confirmed.

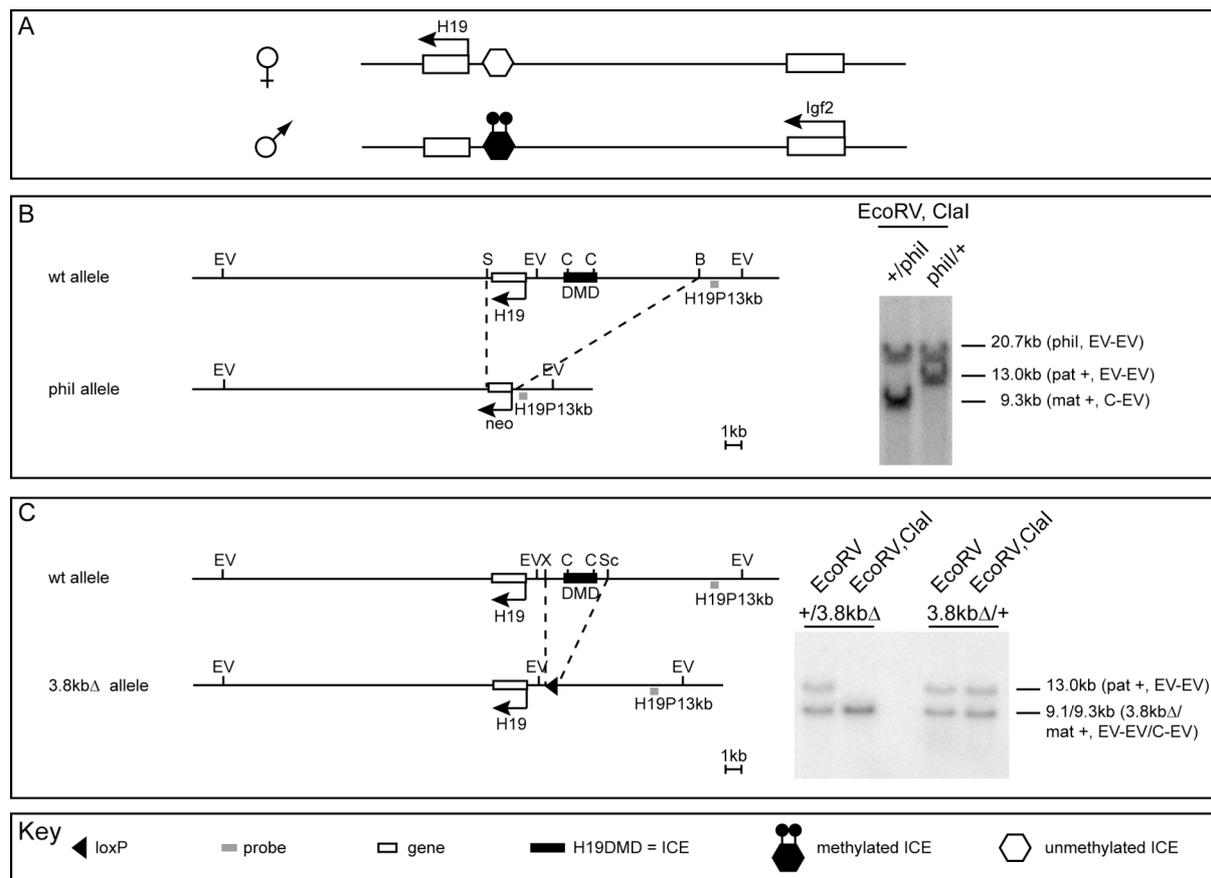


Fig. 11: Genotyping of embryos carrying a mutant allele in the *Igf2/H19* region. (A) Schematic overview of the genomic regions carrying the *Igf2* and *H19* genes as well as the H19DMD, the ICE for this region. (B)-(C) As in Fig. 10. (B) The *phil* allele was generated by replacing the Sall-BamHI region containing the *H19* gene and a 9.9kb large upstream region including the ICE with a neomycin resistance gene (*neo*) (Leighton et al., 1995). (C) The 3.8kb Δ allele was generated by deleting the XbaI-SacI region surrounding the ICE (Thorvaldsen et al., 2002). Please note that in the +/3.8kb Δ , the 9.1kb band corresponding to the 3.8kb deletion allele and the 9.3kb band corresponding to the maternal wt allele could not be separated by standard gel electrophoresis. However, loss of the 13.9kb band in the EcoRI/ClaI digest compared to the EcoRI digest alone is indicative for the correct genotype. Restriction enzymes: EV: EcoRV, S: Sall, C: ClaI, B: BamHI, X: XbaI, Sc: SacI. Note that ClaI is a methylation sensitive enzyme and only cuts the unmethylated allele. Arrows indicate transcription in (A) and transcriptional orientation in (B)-(C).

2.1.2 Mapping the proximal border of DNA FISH asynchrony in the *Igf2r/Airn* imprinted gene cluster

In my previous diploma work I showed that the whole imprinted *Igf2r/Airn* cluster shows DNA FISH asynchrony (Koerner, 2006). By using DNA FISH with 7 cosmid and one plasmid probe I showed that in wildtype pMEFs the distal border of the asynchrony maps between *Sod2* and *Tcte2* and at the proximal side it extends at least until *Map3k4* (Koerner, 2006). To map the proximal border of this DNA FISH asynchrony I used 6 more BAC probes for the DNA FISH analysis. For each probe, at least 100 BrdU-positive nuclei were analysed for their hybridisation pattern and the percentage of cells showing one of the three possible FISH-signals, single single spots (SS), single double spots (SD) or double double spots (DD), was scored. An SD pattern between 25% to 40% is considered to be indicative of DNA FISH asynchrony, whereas an SD pattern between 10% to 20% is considered to be indicative of DNA FISH synchrony (Gribnau et al., 2003). In my diploma work I found asynchronous regions to show an SD signal between 26.2% to 41.9% and synchronous regions to show an SD signal between 6.7% to 15.8% (Koerner, 2006). Therefore I consider regions showing in more than 25% an SD signal as asynchronous and regions with less than 17% as synchronous. For each DNA FISH experiment two biological replicates were performed (see Table 2, Fig. 12). An SD pattern in more than 25% of the BrdU-positive nuclei and therefore DNA FISH asynchrony was found for the 4 BACs mapping closest to the imprinted *Igf2r/Airn* region, namely RP24-99O8, RP24-281L7, RP24-362H17 and RP24-222A12. In contrast, the two BACs located further proximally, namely RP24-173F10 and RP24-223O4 showed an SD pattern in less than 17% of the BrdU-positive nuclei and therefore displayed DNA FISH synchrony. Therefore, I could map the proximal border of the DNA FISH asynchrony to a 307kb long genomic region between the putative ncRNA *Papbc3* and *Qk*. In combination with the data obtained previously (Koerner, 2006), I therefore was able to show that the region exhibiting DNA FISH asynchrony extends over 3Mbp, whereas the region showing imprinted expression in pMEFs is just 167kb, as only *Igf2r* and *Airn* show imprinted expression in those cells,

and the size of the whole imprinted cluster, including *Slc22a2* and *Slc22a3*, which show imprinted expression in extraembryonic tissues only, is 430kb.

In addition to the percentage of cells showing an SD pattern also the percentages of cells showing an SS and a DD pattern were calculated (see Fig. 12D and Table 2). Interestingly, the percentages are not the same for different probes, either indicating differences in replication timing where a high percentage of SS would indicate late replication of the first allele and a high percentage of SS + SD would indicate late replication of the second allele. Alternatively, differences in sister chromatid cohesion or chromatin accessibility along the region examined could also be the reason for this phenomenon.

probe	cell type	SS	SD	DD	cells scored	SS%	SD%	DD%
RP24-223O4	+/+ BR1	68	13	22	103	66.0	12.6	21.4
	+/+ BR2	63	15	26	104	60.6	14.4	25.0
	+/+ mean				207	63.3	13.5	23.2
RP24-173F10	+/+ BR1	57	14	33	104	54.8	13.5	31.7
	+/+ BR2	61	13	29	103	59.2	12.6	28.2
	+/+ mean				207	57.0	13.0	29.9
RP24-222A12	+/+ BR1	58	37	22	117	49.6	31.6	18.8
	+/+ BR2	49	34	25	108	45.4	31.5	23.1
	+/+ mean				225	47.5	31.6	21.0
RP24-362H17	+/+ BR1	31	32	39	102	30.4	31.4	38.2
	+/+ BR2	44	32	39	115	38.3	27.8	33.9
	+/+ mean				217	34.3	29.6	36.1
RP24-281L7	+/+ BR1	46	32	23	101	45.5	31.7	22.8
	+/+ BR2	48	39	18	105	45.7	37.1	17.1
	+/+ mean				206	45.6	34.4	20.0
RP24-9908	+/+ BR1	57	30	14	101	56.4	29.7	13.9
	+/+ BR2	54	31	18	103	52.4	30.1	17.5
	+/+ mean				204	54.4	29.9	15.7

Table 2: DNA FISH asynchrony in wildtype pMEFs. For each used probe two biological replicates (BR) were analysed. BrdU-positive nuclei were scored for their FISH signals – two single spots (SS), a single and a double spot (SD) or two double spots (DD) and their percentages were calculated. At least 100 cells were scored per biological replicate. The rows in grey give the total number of cells scored in both biological replicates and the mean percentage of the respective FISH signals of the two biological replicates.

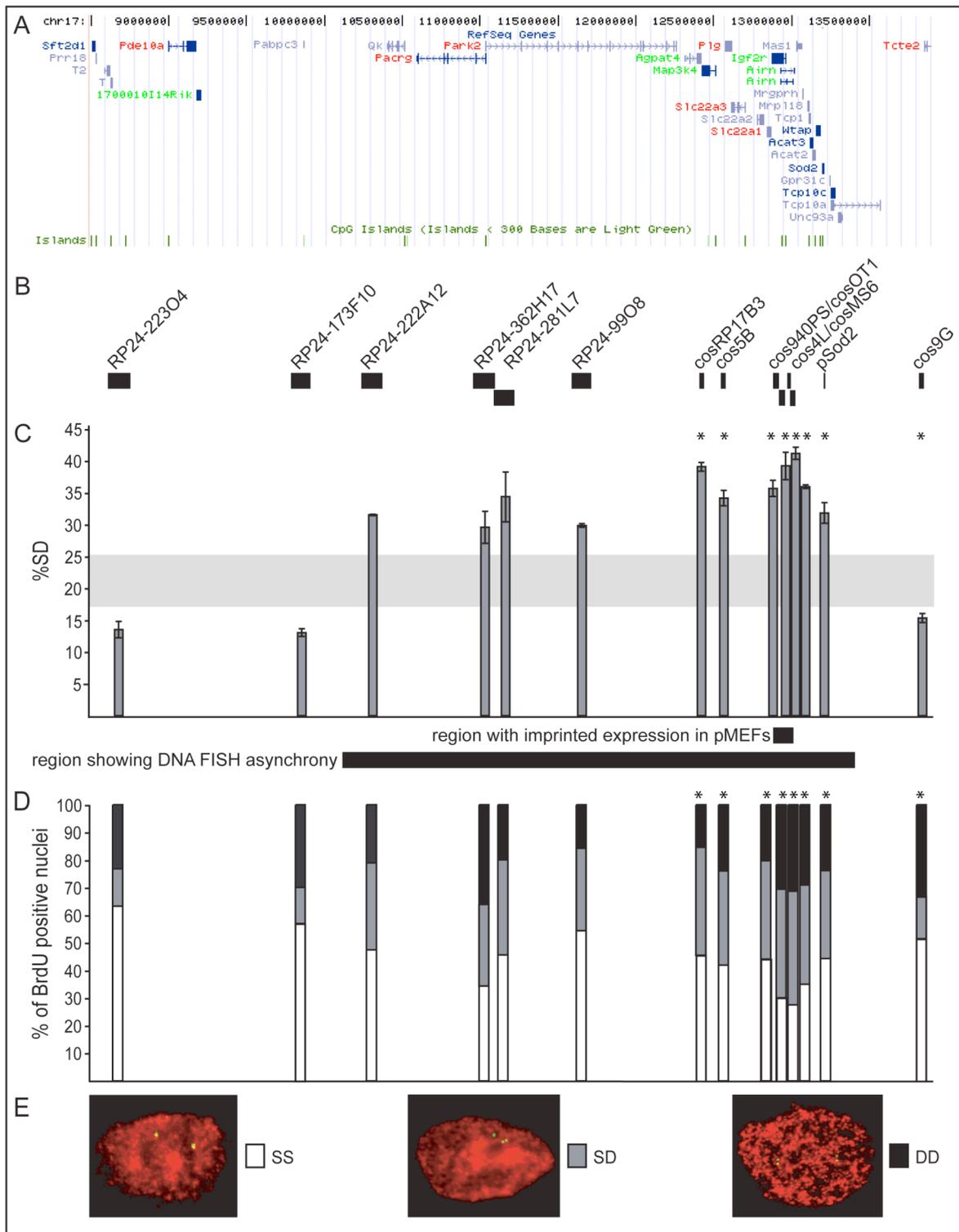


Fig. 12: DNA FISH asynchrony in wildtype pMEFs. (A) 5.5Mbp large genomic region of chr. 17 containing the imprinted *Igf2r/Airn* gene cluster taken from the UCSC genome browser. On the top the bp positions of chr. 17 are shown. Below RefSeq genes are shown in blue. Names in green indicate expressed, names in red not expressed, names in blue not informative genes in a mouse embryonic fibroblast cell line as determined by RNA chip (Regha et al., 2007). CpG islands are shown as green bars. (B) Location and names of BACs, cosmids and plasmids used as probes for DNA-FISH are indicated relative to the map above. (C) Percentage of cells (wildtype primary MEFs)

showing an SD pattern during DNA-FISH using the probe above. Each bar represents the mean of two biological replicates, the error bars mark the standard deviation. Bars extending above the grey line mark DNA FISH asynchrony, bars below SD synchrony. The short black horizontal bar indicates the region which shows imprinted expression in MEFs - it extends over 167kb (n.b. that *Slc22a2* and *Slc22a3* are not expressed in MEFs). The long black horizontal bar shows the region of DNA FISH asynchrony - it extends over 3Mbp. (D) Percentage of cells (wildtype primary MEFs) showing an SS (white), SD (grey) and DD (black) pattern as determined by DNA-FISH using the probe above. Each bar represents the mean of two biological replicates. (E) Representative pictures for each of the FISH signal patterns are shown below (red – BrdU immunofluorescence, green – sequence-specific probe). Note that bars marked with * represent data which have been shown already in (Koerner, 2006).

2.1.3 Do the ICE, the ncRNA *Airn* or the differentially methylated *Igf2r* promoter play a role in DNA FISH asynchrony?

During my previous work I already analysed cells carrying one or two AirT or R2Δ alleles in one biological replicate with probes indicated in Fig. 13 and DNA FISH to see, if those alleles have an effect on the DNA FISH asynchrony observed for wildtype cells (Koerner, 2006). In the current work I analysed a second biological replicate for the probes already presented in (Koerner, 2006). In addition, I performed two biological replicates for one more probe (cos940PS) on +/-AirT cells and two more cell types (AirT/AirT, R2Δ/R2Δ) for probe cos9G. Furthermore, I also analysed two biological replicates for all 3 genotypes with an IPΔ allele (IPΔ/+, +/-IPΔ, IPΔ/IPΔ) and probe cos940PS for a possible effect of the somatic DNA methylation on the *Igf2r* promoter on DNA FISH asynchrony. Again, at least 100 BrdU-positive nuclei per experiment were analysed with respect to their DNA FISH signals.

An SD pattern in more than 25% of the BrdU-positive nuclei and therefore DNA FISH asynchrony was found for all used probes (cosRP18B3, cos5B, cos940PS, cosOT1) except cos9G and all used cell types. The values obtained are comparable to those obtained for wildtype cells. +/-AirT and AirT/AirT cells still show an SD value in more than 25% of the examined cells and therefore DNA FISH asynchrony for all probes except cos9G despite the lack of imprinted expression of *Igf2r*. Therefore, full-length *Airn* expression is not needed for DNA FISH asynchrony. +/-R2Δ and R2Δ/R2Δ cells do not express any *Airn* at all and therefore also lack imprinted expression of *Igf2r*. However, all probes except cos9G showed DNA FISH asynchrony. Therefore, *Airn* is

not needed at all for DNA FISH asynchrony. Also cells lacking the methylated ICE ($R2\Delta/+$) exhibit an SD value in more than 25% of the examined cells for all probes except *cos9G* and therefore show DNA FISH asynchrony. This result shows that also the imprint mark is dispensable for DNA FISH asynchrony. $+/IP\Delta$ cells lack the methylated *Igf2r* promoter, $IP\Delta/+$ cells lack the unmethylated *Igf2r* promoter and therefore only display a low level of *Igf2r* expression from the paternal allele. $IP\Delta/IP\Delta$ cells do not have any *Igf2r* promoter and are therefore completely free of *Igf2r* expression. However, all three cell types show an SD value in 33.3% to 39.1% percent of cells and therefore DNA FISH asynchrony. Therefore neither expression of *Igf2r* nor the methylated or the unmethylated promoter of *Igf2r* is needed for DNA FISH asynchrony. In summary, neither full-length expression of *Airn* (as in $+/AirT$ and $AirT/AirT$ cells) nor the methylated ICE (as it is absent in $R2\Delta/+$ cells) nor the unmethylated ICE (as it is absent in $+/R2\Delta$ cells) nor the maternal or the paternal *Igf2r* promoter are needed for DNA FISH asynchrony (see Fig. 13, Table 3).

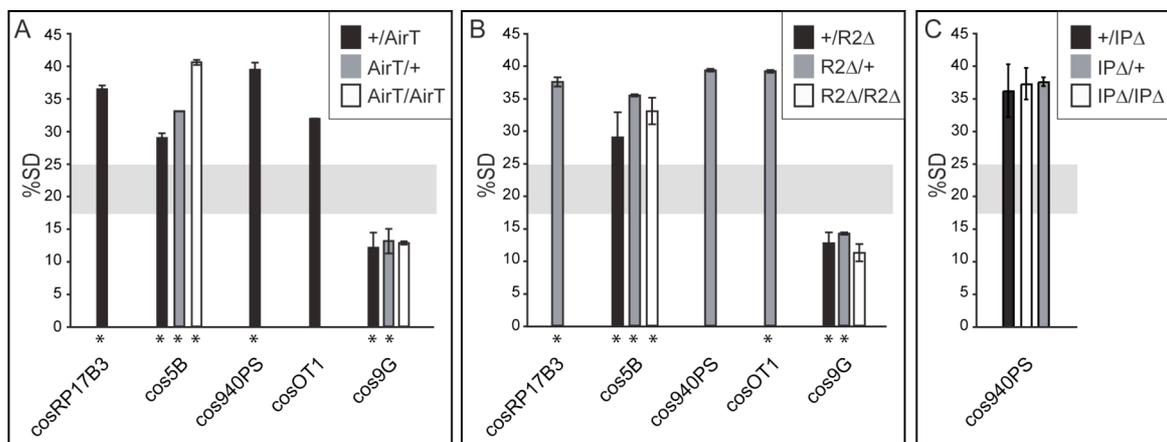


Fig. 13: DNA FISH asynchrony in pMEFs carrying an AirT (A), $R2\Delta$ (B) or $IP\Delta$ (C) allele on one or both parental chromosomes. For each used probe two biological replicates (BR) were analysed. Details as in Fig. 12C. Please note that for bars marked with * one biological replicate was already shown in (Koerner, 2006).

probe	cell type	SS	SD	DD	cells scored	SS%	SD%	DD%
cosRP17B3	+/AirT BR1	48	41	25	114	42.1	36.0	21.9
	+/AirT BR2	45	39	22	106	42.5	36.8	20.8
	+/AirT mean				220	42.3	36.4	21.3
	R2D/+ BR1	48	38	14	100	48.0	38.0	14.0
	R2D/+ BR2	51	37	12	100	51.0	37.0	12.0
	R2D/+ mean				200	49.5	37.5	13.0
cos5B	+/AirT BR1	38	29	35	102	37.3	28.4	34.3
	+/AirT BR2	37	30	35	102	36.3	29.4	34.3
	+/AirT mean				204	36.8	28.9	34.3
	AirT/+ BR1	46	34	23	103	44.7	33.0	22.3
	AirT/+ BR2	42	33	25	100	42.0	33.0	25.0
	AirT/+ mean				203	43.3	33.0	23.7
	AirT/AirT BR1	29	41	32	102	28.4	40.2	31.4
	AirT/AirT BR2	30	42	31	103	29.1	40.8	30.1
	AirT/AirT mean				205	28.8	40.5	30.7
	+/R2D BR1	28	27	48	103	27.2	26.2	46.6
	+/R2D BR2	29	32	40	101	28.7	31.7	39.6
	+/R2D mean				204	27.9	28.9	43.1
	R2D/+ BR1	28	38	41	107	26.2	35.5	38.3
	R2D/+ BR2	26	36	40	102	25.5	35.3	39.2
	R2D/+ mean				209	25.8	35.4	38.8
	R2D/R2D BR1	40	35	27	102	39.2	34.3	26.5
R2D/R2D BR2	42	33	30	105	40.0	31.4	28.6	
R2D/R2D mean				207	39.6	32.9	27.5	
cos940PS	+/AirT BR1	47	55	35	137	34.3	40.1	25.5
	+/AirT BR2	41	44	29	114	36.0	38.6	25.4
	+/AirT mean				251	35.1	39.4	25.5
	R2D/+ BR1	40	45	30	115	34.8	39.1	26.1
	R2D/+ BR2	34	41	29	104	32.7	39.4	27.9
	R2D/+ mean				219	33.7	39.3	27.0
	IPD/IPD BR1	41	39	20	100	41.0	39.0	20.0
	IPD/IPD BR2	38	37	29	104	36.5	35.6	27.9
	IPD/IPD mean				204	38.8	37.3	23.9
	+/IPD BR1	35	43	32	110	31.8	39.1	29.1
	+/IPD BR2	38	36	34	108	35.2	33.3	31.5
	+/IPD mean				218	33.5	36.2	30.3
	IPD/+ BR1	35	40	30	105	33.3	38.1	28.6
	IPD/+ BR2	33	39	33	105	31.4	37.1	31.4
IPD/+ mean				210	32.4	37.6	30.0	

Table continued on the next page

probe	cell type	SS	SD	DD	cells scored	SS%	SD%	DD%
cosOT1	+/AirT BR1	52	36	25	113	46.0	31.9	22.1
	+/AirT BR2	50	35	25	110	45.5	31.8	22.7
	+/AirT mean				223	45.7	31.8	22.4
	R2D/+ BR1	20	39	43	102	19.6	38.2	42.2
	R2D/+ BR2	24	44	42	110	21.8	40.0	38.2
	R2D/+ mean				212	20.7	39.1	40.2
cos9G	+/AirT BR1	50	11	44	105	47.6	10.5	41.9
	+/AirT BR2	48	14	40	102	47.1	13.7	39.2
	+/AirT mean				207	47.3	12.1	40.6
	AirT/+ BR1	62	12	28	102	60.8	11.8	27.5
	AirT/+ BR2	57	15	32	104	54.8	14.4	30.8
	AirT/+ mean				206	57.8	13.1	29.1
	AirT/AirT BR1	62	14	35	111	55.9	12.6	31.5
	AirT/AirT BR2	59	14	35	108	54.6	13.0	32.4
	AirT/AirT mean				219	55.2	12.8	32.0
	+/R2D BR1	51	14	36	101	50.5	13.9	35.6
	+/R2D BR2	53	12	39	104	51.0	11.5	37.5
	+/R2D mean				205	50.7	12.7	36.6
	R2D/+ BR1	52	18	56	126	41.3	14.3	44.4
	R2D/+ BR2	50	16	48	114	43.9	14.0	42.1
	R2D/+ mean				240	42.6	14.2	43.3
	R2D/R2D BR1	60	11	36	107	56.1	10.3	33.6
	R2D/R2D BR2	55	13	39	107	51.4	12.1	36.4
	R2D/R2D mean				214	53.7	11.2	35.0

Table 3: DNA FISH asynchrony in mutant pMEFs. For each used probe and each genotype two biological replicates (BR) were analysed. Details as in Table 2. Please note that the rows with grey text were already shown in (Koerner, 2006).

2.1.4 Can the system detect loss of DNA FISH asynchrony? – Using the *Igf2/H19* region as control

It was shown previously that the maternal transmission of a 13kb deletion comprising the ICE of the *Igf2/H19* imprinted cluster as well as the ncRNA *H19* leads to loss of DNA FISH asynchrony (Greally et al., 1998; Gribnau et al., 2003; Leighton et al., 1995). I therefore used pMEFs of this genotype (phil/+) and of the reciprocal cross (+/phil) which I obtained from Karl Pfeifer (Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland, USA) to see if my system can detect a loss of DNA FISH asynchrony. In addition I used cells harbouring a smaller deletion, which leaves *H19* intact but deletes the ICE, which I

obtained from Marisa Bartolomei (University of Pennsylvania, Philadelphia, USA) (Thorvaldsen et al., 2002). Also here I used pMEFs, both from a maternal transmission of the deletion (3.8kb Δ /+) and a paternal one (+/3.8kb Δ). An effect of this smaller deletion on DNA FISH asynchrony has not been analysed previously.

For DNA FISH I used all those 4 pMEF cell lines as well as wildtype cells as control. As probe I used plgf2P6, which covers *Igf2*. In wildtype cells as well as in cells with a paternal deletion, +/phil and +/3.8kb Δ , an SD pattern was observed in 29.4-33.36% of BrdU-positive cells, therefore this locus showed DNA FISH asynchrony. However, the maternally deleted phil/+ and 3.8kb Δ /+ cells showed an SD pattern in only 13.8-15.2% of BrdU-positive cells and therefore DNA FISH synchrony (see Fig. 14, Table 4). My system is therefore able to observe a loss of DNA FISH asynchrony and my results confirm a role for the unmethylated *Igf2* ICE in DNA FISH asynchrony.

probe		SS	SD	DD	cells scored	%SS	%SD	%DD
plgf2P6	+/+ BR1	40	36	32	108	37.0	33.3	29.6
	+/+ BR2	42	34	31	107	39.3	31.8	29.0
	+/+ mean				215	38.1	32.6	29.3
	+/phil BR1	44	35	39	118	37.3	29.7	33.1
	+/phil BR2	38	34	32	104	36.5	32.7	30.8
	+/phil mean				222	36.9	31.2	31.9
	phil/+ BR1	51	16	49	116	44.0	13.8	42.2
	phil/+ BR2	46	15	42	103	44.7	14.6	40.8
	phil/+ mean				219	44.3	14.2	41.5
	+/3.8kb Δ BR1	43	41	41	125	34.4	32.8	32.8
	+/3.8kb Δ BR2	40	30	32	102	39.2	29.4	31.4
	+/3.8kb Δ mean				227	36.8	31.1	32.1
	3.8kb Δ /+ BR1	48	17	47	112	42.9	15.2	42.0
	3.8kb Δ /+ BR2	44	15	41	100	44.0	15.0	41.0
3.8kb Δ /+ mean				212	43.4	15.1	41.5	

Table 4: DNA FISH asynchrony in wildtype pMEFs and pMEFs carrying a deletion of either the ICE and the *H19* gene (phil) or the ICE only (3.8kb Δ). Details as in Table 2.

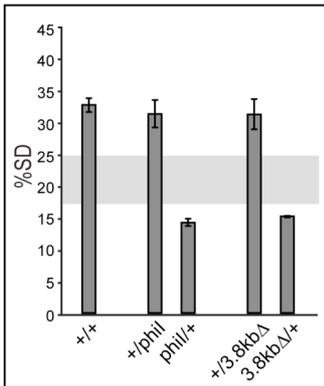


Fig. 14: DNA FISH asynchrony in the *Igf2/H19* imprinted cluster in wildtype pMEFs or pMEFs carrying a *phil* or 3.8kb Δ allele on one or both parental chromosomes. As probe *plgf2P6* was used. For each genotype two biological replicates (BR) were analysed. Details as in Fig. 12C.

2.1.5 Modulators of DNA FISH asynchrony: Is histone acetylation involved?

Several publications showed that treatment of cells with histone deacetylase inhibitors like sodium butyrate or Trichostatin A (TSA) lead to the loss of DNA FISH asynchrony in the *Igf2* and *Kcnq1* imprinted gene clusters (Bickmore and Carothers, 1995; Kagotani et al., 2002). I wanted to see if one of this chemicals also has an effect on the DNA FISH asynchrony observed in the *Igf2r/Airn* cluster. Trichostatin A is not only a general inhibitor of class I and class II histone deacetylases but also can lead to cell cycle arrest and apoptosis. In addition, also several genes show up- or downregulation upon TSA treatment (Marks et al., 2001; Yoshida et al., 1990). First, I therefore had to determine a TSA concentration which leads to deacetylation but still allows the cells to replicate, as for the DNA FISH experiments described above only cells are used which are in S phase.

To determine which TSA concentration still allowed cycling of pMEFs, I incubated them for 16 hours with various concentrations of TSA ranging from 5ng/ml to 75ng/ml or dimethyl sulfoxid (DMSO; the solvent used for TSA). During the last 50 minutes they were additionally incubated with BrdU, afterwards cells were harvested, fixed and the percentage of BrdU-positive cells was determined by immunofluorescence. Two biological replicates have been performed. In untreated and DMSO-treated cells, 21 to 27% of cells were BrdU-positive and therefore in S phase. TSA treatment in general led to a decrease in the number of BrdU-positive

cells. Upon treatment with 5ng/ml TSA 18-19% of cells were BrdU positive, when 7.5ng/ml TSA were used, 9-13% of cells were BrdU-positive. Higher concentrations (10ng/ml, 15nm/ml, 25ng/ml, 75ng/ml) did not show consistently more than 9% of BrdU-positive cells (see Fig. 15A). Surprisingly, after treatment with 50ng/ml of TSA, in both biological replicates more than 10% of BrdU-positive nuclei were found. However, as all but one concentrations higher than 7.5ng/ml showed either in one or in both biological replicates less than 9% of BrdU-positive cells, 7.5ng/ml TSA seems to be the highest concentration which allows pMEFs to sufficiently and consistently replicate.

Next I assayed, if TSA treatment led to changes in histone deacetylation. I did not measure histone deacetylation levels directly but rather assayed *Prss11* steady state level. *Prss11* is an HDAC1 target gene which is upregulated upon an increase in histone acetylation (Zupkovitz et al., 2006). After TSA treatment for 16 or 23 hours, RNA was isolated and analysed by RNA blots (see Fig. 15B, E). Methylene blue staining of rRNA as well as *Gapdh* served as a control for equal loading. Band intensities were quantified to measure changes of RNA levels, *Prss11* and *Igf2r* were quantified relative to *Gapdh*. Expression levels in DMSO-treated cells were set to 100%. *Prss11* showed an increase in expression with increasing concentrations of TSA and prolonged treatment time (see Fig. 15C, F). However, *Igf2r* did not show a change in expression levels upon TSA treatment, neither in cells in which both alleles of *Igf2r* are present, nor in cells which just have the maternal (+/Thp) or the paternal (Thp/+) *Igf2r* allele (see Fig. 15D, G).

pMEFs treated with 7.5ng/ml TSA showed an upregulation of *Prss11* between 10 and 39% compared to the DMSO-treated control, which indicated that 7.5ng/ml TSA shows a biological effect indicative of increased histone acetylation levels. Furthermore, the published studies mentioned above where the effect of TSA on DNA FISH asynchrony was analysed, used 4.8ng/ml or 10.0ng/ml of TSA respectively, my chosen concentration is therefore in the range used previously (Bickmore and Carothers, 1995; Kagotani et al., 2002).

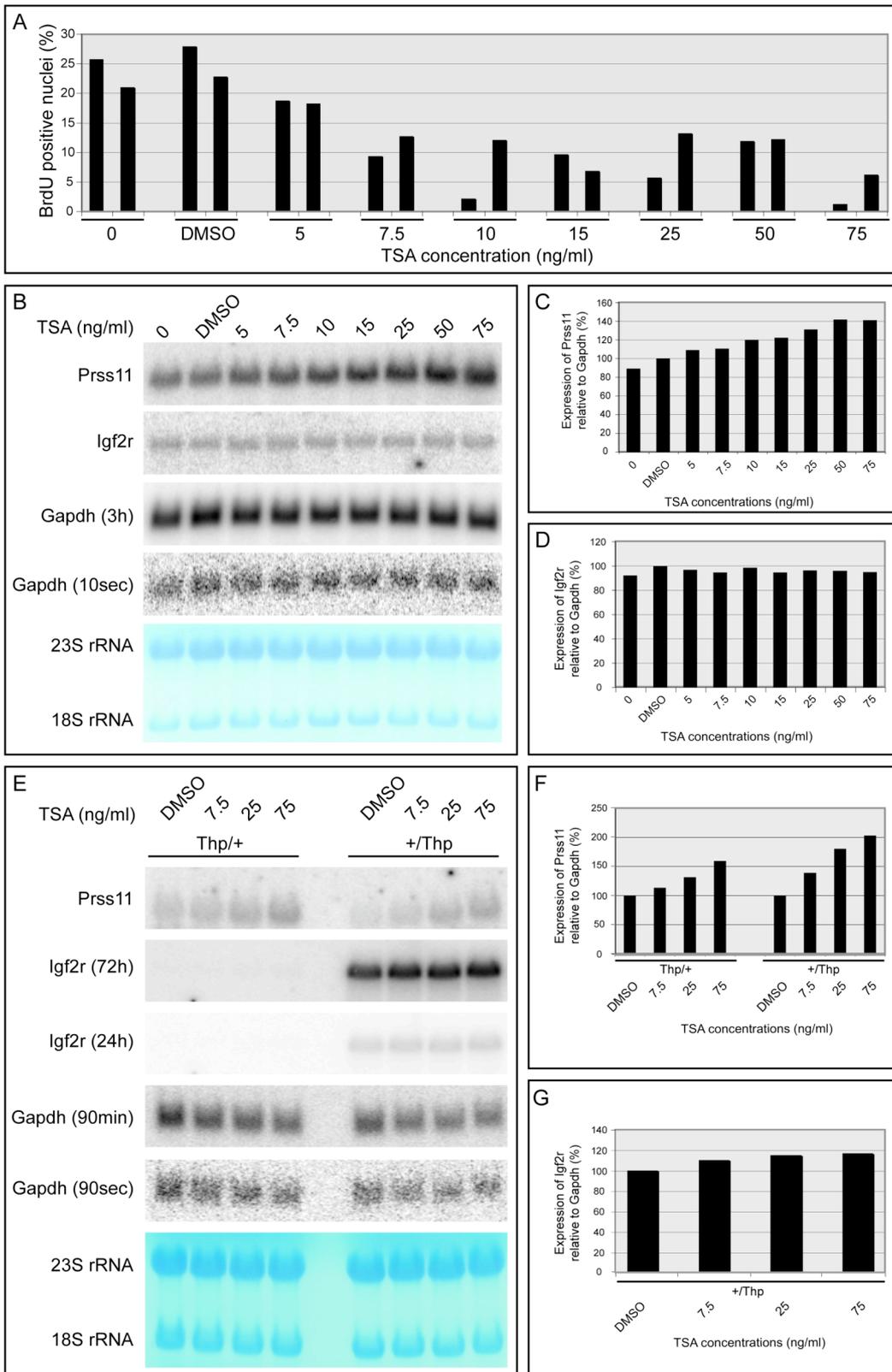


Fig. 15: TSA treatment of pMEFs. (A) Percentage of BrdU-positive nuclei after TSA treatment. Cells were incubated with TSA for 16 hours and during the last 50 min in addition with BrdU. Afterwards, cells were subjected to BrdU immunofluorescence and the percentage of BrdU-positive cells was determined. Two biological replicates were performed for each TSA concentration. (B) TSA treatment

for 16 hours. RNA blot showing the expression of *Prss11*, a gene which is upregulated upon TSA treatment, *Igf2r* and *Gapdh* in wildtype pMEFs upon 16 hours treatment with different concentrations of TSA or DMSO (the solvent for TSA) and in untreated cells (0). For *Gapdh* two different exposure times are shown, 3 hours and 10 seconds. Methylene blue staining of rRNAs as loading control. (C) Quantification of *Prss11* levels relative to *Gapdh* from the RNA blot shown in (B). Note, that for the quantification of *Gapdh* the 10 seconds exposure was used. Expression in DMSO-treated cells was set to 100%. (D) Quantification of *Igf2r* levels relative to *Gapdh* (10 seconds exposure) from the RNA blot shown in (B). Expression in DMSO-treated cells was set to 100%. (E) RNA blot showing the expression of *Prss11*, *Igf2r* and *Gapdh* in Thp/+ and +/Thp cells after 23 hours treatment with different concentrations of TSA or DMSO. For *Igf2r* and *Gapdh* two different exposure times are shown (72 hours and 24 hours for *Igf2r*, 90 minutes and 90 seconds for *Gapdh*). Methylene blues staining of rRNA as loading control. (F) Quantification of *Prss11* levels relative to *Gapdh* (90 seconds exposure) from the RNA blot shown in (E). Expression in DMSO-treated cells was set to 100%. (G) Quantification of *Igf2r* levels in +/Thp cells (24 hours exposure) relative to *Gapdh* (90 seconds exposure) from the RNA blot shown in (E). Expression in DMSO-treated cells was set to 100%.

I performed a DNA FISH experiment on wildtype pMEFs which were either untreated, or treated for 16 hours with DMSO or 7.5ng/ml TSA. Two to three biological replicates were performed per experiment using cos940PS as a probe. The untreated and DMSO-treated cells behaved in a very similar way and showed an SD pattern in 34.8-41.7% of BrdU-positive nuclei, demonstrating that the DMSO-treatment did not disturb the DNA FISH asynchrony. Upon treatment with 7.5ng/ml TSA I observed a decrease in the number of BrdU-positive nuclei showing an SD pattern, which was now between 15.2-21.9%. This percentage is below 25%, which is the minimal percentage for DNA FISH asynchrony. However, it is still considerably higher than the percentages I normally observed for regions showing DNA FISH synchrony. TSA treatment therefore led to a reduction of the DNA FISH asynchrony, but did not abolish it completely (see Fig. 16, Table 5).

probe	substance added	SS	SD	DD	cells scored	SS%	SD%	DD%
cos940PS	0 BR1	51	39	22	112	45.5	34.8	19.6
	0 BR2	43	37	21	101	42.6	36.6	20.8
	0 mean				213	44.1	35.7	20.2
	DMSO BR1	41	43	19	103	39.8	41.7	18.4
	DMSO BR2	43	36	23	102	42.2	35.3	22.5
	DMSO BR3	50	43	30	123	40.7	35.0	24.4
	DMSO mean				328	40.9	37.3	21.8
	7.5ng/ml TSA BR1	48	16	41	105	45.7	15.2	39.0
	7.5ng/ml TSA BR2	49	23	33	105	46.7	21.9	31.4
	7.5ng/ml TSA BR3	45	21	37	103	43.7	20.4	35.9
	7.5ng/ml TSA mean				313	45.4	19.2	35.5

Table 5: DNA FISH asynchrony in wildtype pMEFs either untreated, treated with DMSO or 7.5ng/ml TSA. For each treatment two to three biological replicates (BR) were analysed. Details as in Table 2.

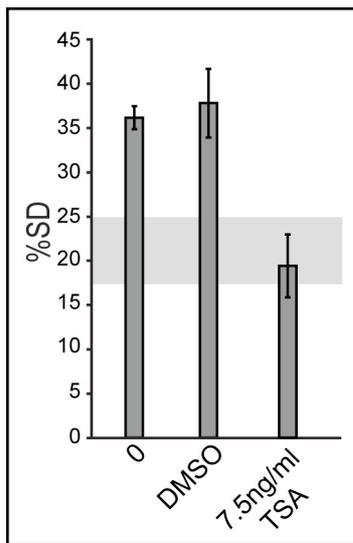


Fig. 16: DNA FISH asynchrony in wildtype pMEFs which were untreated (0), treated with DMSO or 7.5ng/ml TSA. As probe cos940PS was used. For each treatment two (untreated) or three (DMSO/TSA-treated) biological replicates were analysed. Details as in Fig. 12C. Bars from untreated and DMSO-treated cells clearly show DNA FISH asynchrony. The bar from the TSA-treated cells however resides within the grey area, indicating a reduction but not a complete loss of DNA FISH asynchrony.

2.1.6 3D FISH to analyse chromatin compaction

To see, if differences in chromatin compaction between the parental alleles could explain the observed DNA FISH asynchrony, I used 3D DNA FISH. For this, I used cells with a Thp allele on either the paternal or the maternal allele, which enable me to specifically just look on either the maternal or paternal *Igf2r/Airn* cluster. I used two probes which are spaced by 500kb, cos5B (*Plg*) and cosLAI (*Tcp1*) which were labelled in green or red respectively and co-hybridised them to Thp/+ or +/Thp pMEFs (see Fig. 17A). After three-dimensional images have been created and processed, the distance between the centers of homogenous masses of the green (cos5B) and red (cosLAI) signal was measured for the maternal and the paternal allele. For each genotype two biological replicates were performed, at least 25 nuclei were scored per biological replicate (see Fig. 17B, C).

of the interquartile range (difference of the 75th quartile and the 25th quartile, indicative of the data distribution) slightly vary between different cell types and biological replicates, however, there is no obvious difference between the different cell types and replicates. The pooled biological replicates demonstrate the similarity between the maternal and the paternal allele with respect to chromatin compaction in an even more striking way, as the median and the interquartile range are highly similar between +/Thp and Thp/+ cells. The median distance on the maternal allele is 0.51µm, on the paternal allele it is 0.47µm. The maximum value (not being an outlier) for the maternal allele is 1.13µm, on the paternal allele 1.25µm. The minimum value on the maternal allele is 0.10µm, on the paternal allele 0.07µm. The low number of outliers (2 for +/Thp and 1 for Thp/+) demonstrates the high homogeneity of the analysed distances. Therefore, although the region around the imprinted *Igf2r/Airn* cluster shows DNA FISH asynchrony, there is no significant difference in the compaction of the parental alleles in a 500kb region.

2.1.7 Does DNA FISH asynchrony represent DNA replication asynchrony?

It was generally assumed for a long time, that the asynchrony observed by DNA FISH represents DNA replication asynchrony (Kitsberg et al., 1993). However it was suggested later, that the DNA FISH asynchrony often rather represents a difference in sister chromatid cohesion instead of DNA replication (Azura et al., 2003). To test if the DNA FISH asynchrony in the *Igf2r/Airn* cluster is really a result of DNA replication asynchrony I used S phase fractionation followed by qPCR detection of newly replicated DNA. The amount of newly synthesised DNA per fraction is quantified for the genomic locus and normalised against mitochondrial DNA which replicates randomly throughout the cell cycle (Bogehagen and Clayton, 1977).

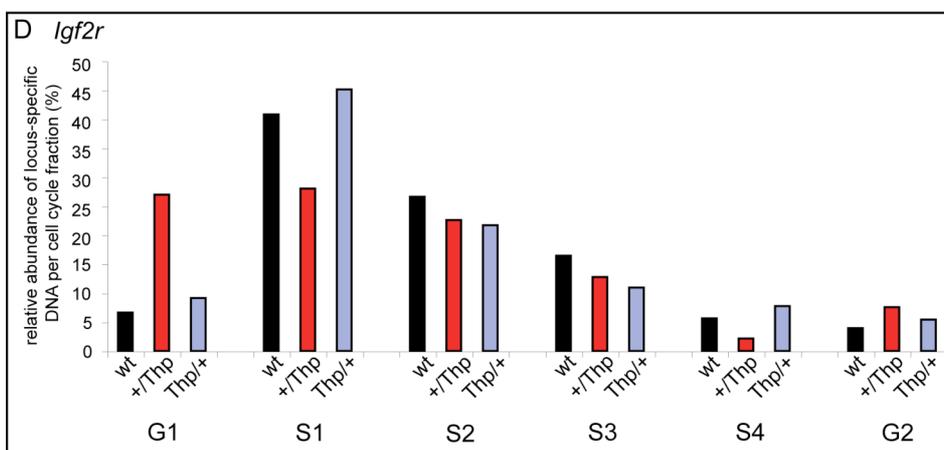
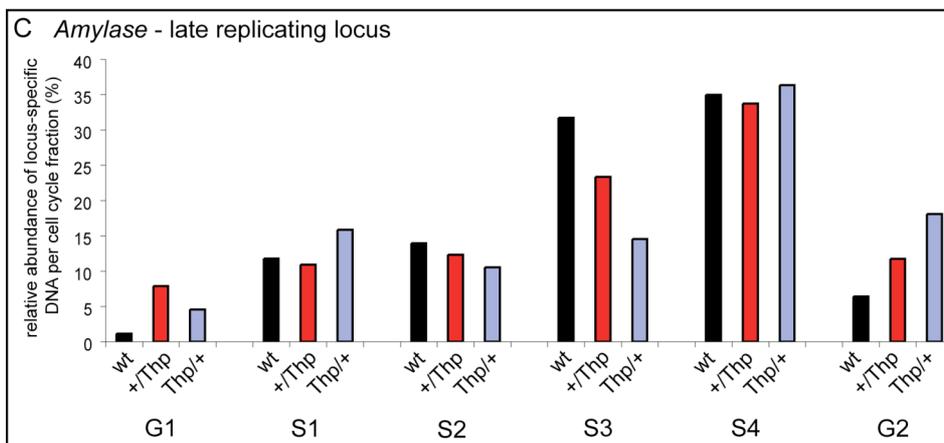
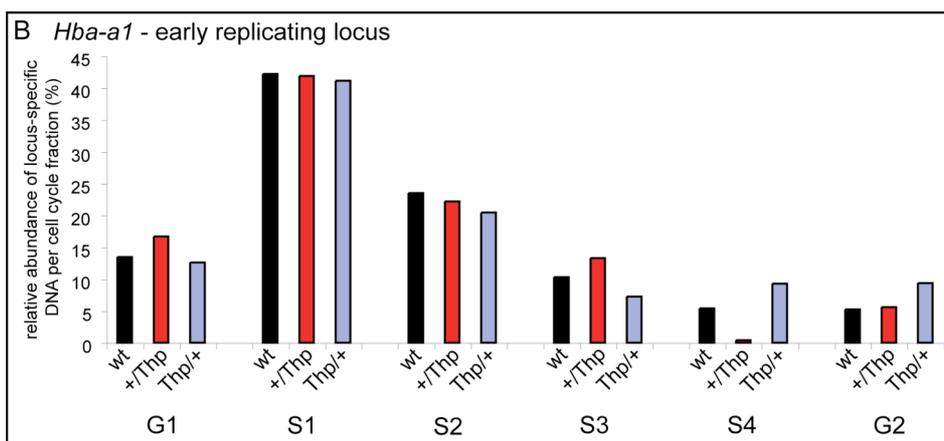
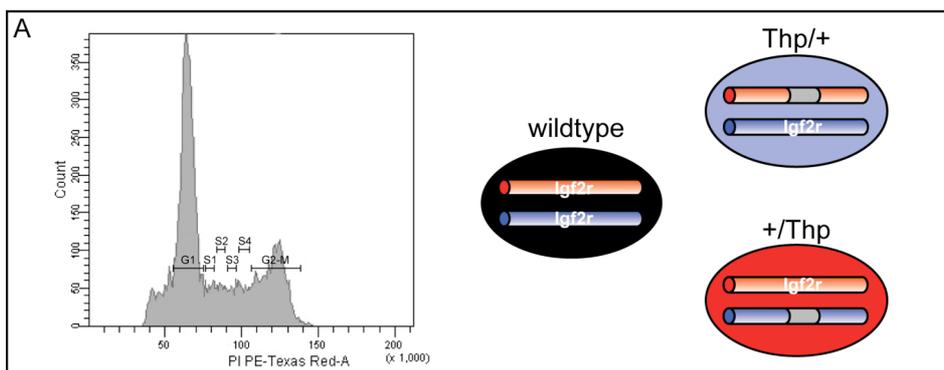


Fig. 18: S phase fractionation. (A) On the left a typical histogram with the gating done for the different fractions (G1, S1, S2, S3, S4, G2-M) is shown. The histogram shows counts on the y-axis and Propidium iodide fluorescence intensity on the x-axis. On the right the used genotypes of pMEFs are shown. Wildtype cells are diploid for the whole genome, Thp/+ cells are haploid for the paternal *Igf2r* region, +/Thp cells for the maternal *Igf2r* region. Please note that the two Thp cells are diploid for *Amylase* and *Hba-a1*. (B) Relative abundance of *Hba-a1* DNA per fraction. *Hba-a1* is a marker gene for early replication. (C) Relative abundance of *Amylase* DNA per fraction. *Amylase* is a marker gene for late replication. (D) Relative abundance of *Igf2r* DNA per fraction. *Igf2r* DNA shows enrichment in early S phase independent of its parental origin.

I did one experiment with wildtype pMEFs and pMEFs carrying the Thp deletion on either the paternal or the maternal allele. To check if the fractionation worked, I used *Hba-a1* which is a marker gene for early replication and *Amylase*, which is a late replicating gene. Thp/+ and +/Thp genes are diploid for those loci, but haploid for the *Igf2r* locus. *Hba-a1* DNA shows main enrichment in the S1-phase, and minor enrichment in the S2-phase (see Fig. 18B). *Amylase* shows main enrichment in the S4-phase and medium enrichment in the S3 phase (see Fig. 18C). All 3 pMEFs show main enrichment for *Igf2r* in S1, indicating that both alleles of *Igf2r* replicate early during S phase (see Fig. 18D). As no obvious difference can be detected between the cells having just the maternal *Igf2r* allele (+/Thp) or only the paternal *Igf2r* allele (Thp/+), these data derived from three independent cell lines indicate, that the parental alleles of the *Igf2r* cluster do not replicate asynchronously.

2.2 The role of the tandem direct repeats and the CpG island within the *Igf2r* ICE

As shown above, the *Igf2r* ICE does not play a role in the regulation of DNA FISH asynchrony. Next, I asked if there are special DNA elements present in the *Igf2r* ICE and if they could have other functions in genomic imprinting in this cluster.

2.2.1 Elements within the *Igf2r* ICE

Tandem direct repeats are a widespread feature of imprint control elements and they can also be found in the ICE of the *Igf2r/Airn* cluster in mice and in the syntenic

region in human (see Fig. 19A, C) (Neumann et al., 1995). The CpG island at the *Igf2r* promoter however, does not show a high abundance of tandem direct repeats, neither in mice nor human (see Fig. 19B, D).

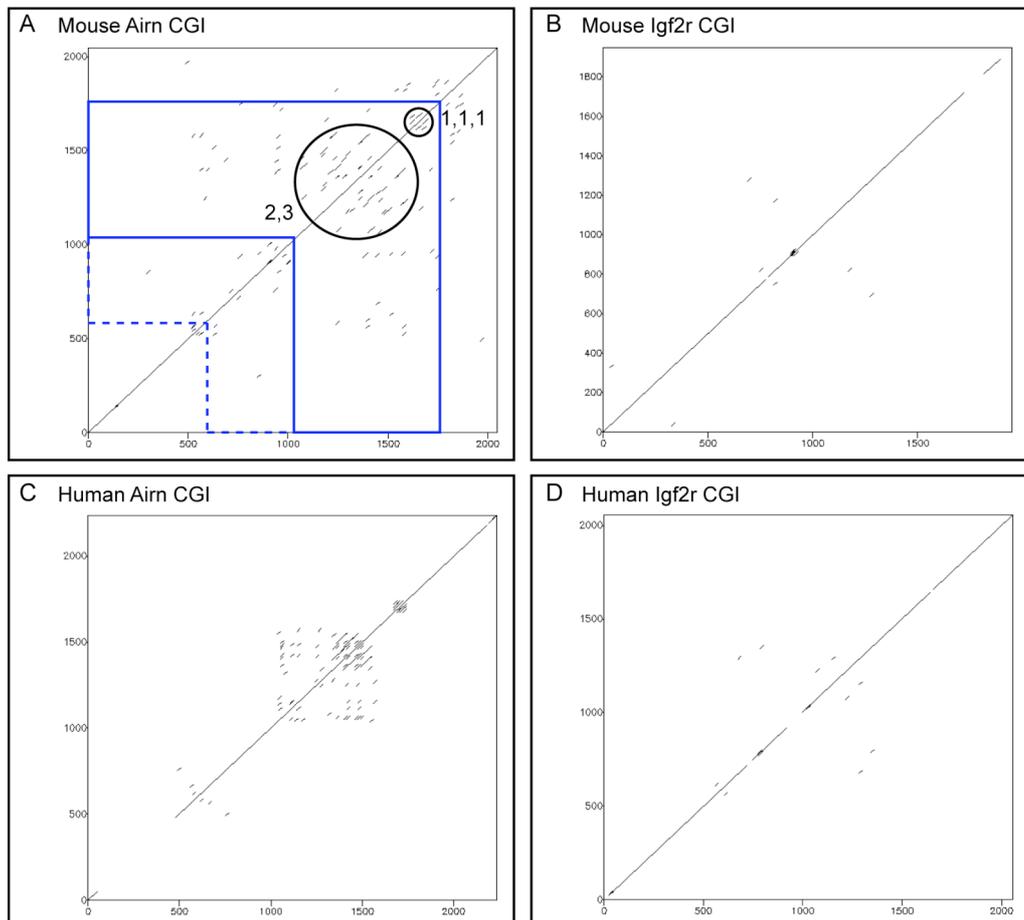


Fig. 19: Dotplots identifying tandem direct repeats as black lines in CpG islands (CGIs) of imprinted genes. For all genomic regions, the CGI as defined by the UCSC genome browser and 500bp up- and downstream have been analysed. (A) Shown is the murine *Airn* CGI, the ICE for this imprinted cluster. The 1,1,1 and 2,3 repeats identified in (Neumann et al., 1995) are encircled in black. The region which was deleted in the R Δ 800 ES cells is framed in blue. The region additionally deleted in the CpG Δ ES cells is framed by a dashed blue line. (B) Shown is the mouse *Igf2r* CGI where no widespread tandem direct repeats are detected. (C) Shown is the human *Airn* CGI which also shows a large amount of tandem direct repeats. (D) Shown is the human *Igf2r* CGI which similarly to the mouse does not show a widespread appearance of tandem direct repeats.

Besides the tandem direct repeats, further potentially interesting elements have been identified in the *Airn* CGI. First, a study which used the injection of DNA fragments into male or female pronuclei followed by analysis of their methylation status identified a *de novo* methylation sequence (DNS) and an allele-discriminating sequence (ADS) (Birger et al., 1999) (see Fig. 20). Furthermore, a stretch of triple-

stranded H-DNA structures has been identified (Davey and Allan, 2003) (see Fig. 20). The R2 Δ allele described above deletes 3.6kb of genomic sequence comprising the whole CGI, the tandem direct repeats, the DNS, ADS and the H-DNA as well as the *Airn* promoter (see Fig. 20) (Birger et al., 1999; Davey and Allan, 2003; Lyle et al., 2000; Stricker et al., 2008; Wutz et al., 2001). In this section I analysed the effect of two smaller subdeletions in this region on the function of the *Airn* ncRNA using the *in vitro* ES cell differentiation system. First, a deletion of the tandem direct repeats (R Δ 800) and second a deletion of the CpG island (CpG Δ).

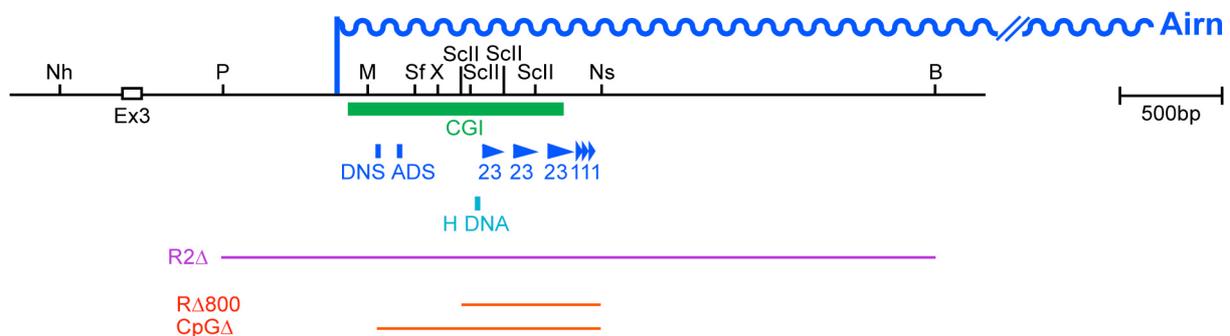


Fig. 20: Schematic overview of the *Igf2r/Airn* imprint control element (ICE). The ICE is located in intron 2 of *Igf2r* and was genetically defined by the R2 Δ deletion (purple line) which removed 3.6kb of genomic sequence from the PaeI to the BamHI site (Wutz et al., 2001). Exon 3 of *Igf2r* is shown as a white box. The CGI is shown as green bar. The transcriptional start site of *Airn* (blue wavy line) is located 5' to the CGI. The allele discriminating signal (ADS) and the *de novo* methylation signal (DNS) discovered in (Birger et al., 1999) are shown as blue bars. The 2,3 and the 1,1,1 tandem direct repeats discovered in (Neumann et al., 1995) are shown as blue triangles. The stretch of H-DNA is shown as a turquoise bar (Davey and Allan, 2003). A 700bp deletion comprising the genomic region between the first SacII site to the NsiI site removes all the repeats and is termed R Δ 800. CpG Δ removes 1.1kb of genomic sequence and contains the ADS, DNS, the tandem direct repeats and the H-DNA. Restriction enzymes: Nh: NheI, P: PaeI, M: MluI, Sf: SfuI, X: XbaI, ScII: SacII, Ns: NsiI, B: BamHI.

2.2.2 Generation of repeat deletion ES cells

To test the role of these tandem direct repeats within the ICE, I generated ES cells carrying a deletion of those repeats by homologous recombination. The targeting vector, pR Δ 800, for a deletion comprising both, the 2,3 repeats as well as the 1,1,1 repeats, has been generated by Stefan Stricker and Yvonne Schichl. Although the vector deletes 700bp, including all SacII sites to the NsiI site, it was termed pR Δ 800, as at the vector was designed according to the GenBank accession number

AJ249895, where the region between all SacII and the NsiI site spans 780bp due to a sequencing error. A selection cassette containing a neomycin resistance gene driven by an HSV-*thymidin-kinase* (HSV-*tk*) promoter and polyadenylated by a SV40-pA sequence and a HSV-*tk* gene driven by an HSV-*tk* promoter stopped by an HSV-*tk*-pA sequence flanked by loxP sites was inserted into the NheI site in intron 3 of *Igf2r*. Yvonne Schichl targeted D3 wildtype ES cells, I targeted DLS30A10 and A9 ES cells (see Fig. 21). DLS30A10 cells are D3 ES cells which carry a SNP in Ex12 of *Igf2r*. Together with Florian Pauler I created the targeting vector for generating DLS30A10, Laura Steenpass then targeted D3 ES cells with it and obtained a homologously targeted clone which now has a change in a single base pair on the maternal allele compared to the wildtype paternal allele (Latos et al., 2009). A9 ES cells are C57BL6/129Sv intraspecies F1 ES cells already successfully used for blastocyst injection to generate knockout mice (Anton Wutz, Wellcome Trust Centre for Stem Cell Research, Cambridge, UK, personal communication).

After targeting ES cells with the pR Δ 800 targeting vector, homologous recombinants carrying an R Δ 800+cas allele were identified by DNA blotting. Genomic DNA was digested with EcoRI and probed with AirT, resulting in a 6.2kb wildtype band and a 5.5kb R Δ 800 band. For D3 ES cells, 1 positive clone (DMK1B11) was obtained from 192 picked neomycin-resistant clones. For DLS30A10 ES cells, 4 positive clones were obtained from 384 picked colonies, two of them (DMK3A2, DMK3D5) were used for further work. For A9 ES cells, in the first targeting (AMK5) two positive clones were obtained from 384 picked colonies, however as they were not kept under selection during thawing, they drifted towards wildtype and could not be used for further experiments. The second targeting of A9 cells resulted in one positive clone (AMK9-302) from 384 picked colonies and this one was kept under selection during thawing (see Table 6, Fig. 21).

The selection cassette was removed by transient transfection with a Cre-recombinase expressing plasmid (pCre, a gift from Anton Wutz, Wellcome Trust Centre for Stem Cell Research, Cambridge, UK) to generate an R Δ 800-cas allele. Removal of the selection cassette from DMK1B11 resulted in 4 clones out of 192

picked colonies with a wildtype and an R Δ 800-cas allele, one clone (DMK2H9) was used for further experiments. Removal of the selection cassette from DMK3A2 and DMK3D5 resulted in 6 and 2 clones out of 96 picked colonies each with a wildtype and a R Δ 800-cas allele, respectively. One clone each (DMK4A2-24, DMK4A2-38) was used for further experiments. After removal of the selection cassette from AMK9-302 9 clones out of 96 picked colonies showed a wildtype and a R Δ 800-cas allele, 4 clones (AMK10B11, AMK10D10, AMK10F12, AMK10H10) were used for further experiments (see Table 6, Fig. 21).

name	parental cell line	targeting vector	number of picked clones	number of recovered DNAs	number of positive clones	names of clones used further
DMK1	D3	pR Δ 800	192	148	1	DMK1B11
DMK2	DMK1B11	pCre	192	170	4	DMK2H9
subclone	DMK2H9	-	48	47	45	DMK2H9-4 DMK2H9-5 DMK2H9-23
DMK3	DLS30A10	pR Δ 800	384	378	4	DMK3A2 DMK3D5
DMK4A2	DMK3A2	pCre	96	89	6	DMK4A2-24
subclone	DMK4A2-24	-	72	68	61	DMK4A2-24/2 DMK4A2-24/30
DMK4D5	DMK3D5	pCre	96	93	2	DMK4D5-38
subclone	DMK4D5-38	-	72	69	68	DMK4D5-38/1 DMK4D5-38/3
AMK5	A9	pR Δ 800	384	334	2	drifted to wildtype after thawing
AMK9	A9	pR Δ 800	384	308	1	AMK9-302
AMK10	AMK9-302	pCre	96	95	9	AMK10B11 AMK10D10 AMK10F12 AMK10H10

Table 6: ES cell lines having a 700bp deletion of the tandem direct repeats in the *Airn* CGI. Given are the name of the created cell line, the name of the parental cell line, the used targeting vector, the number of picked clones, the number of clones from which DNA was recovered, the number of positive clones and the names of clones used for further experiments.

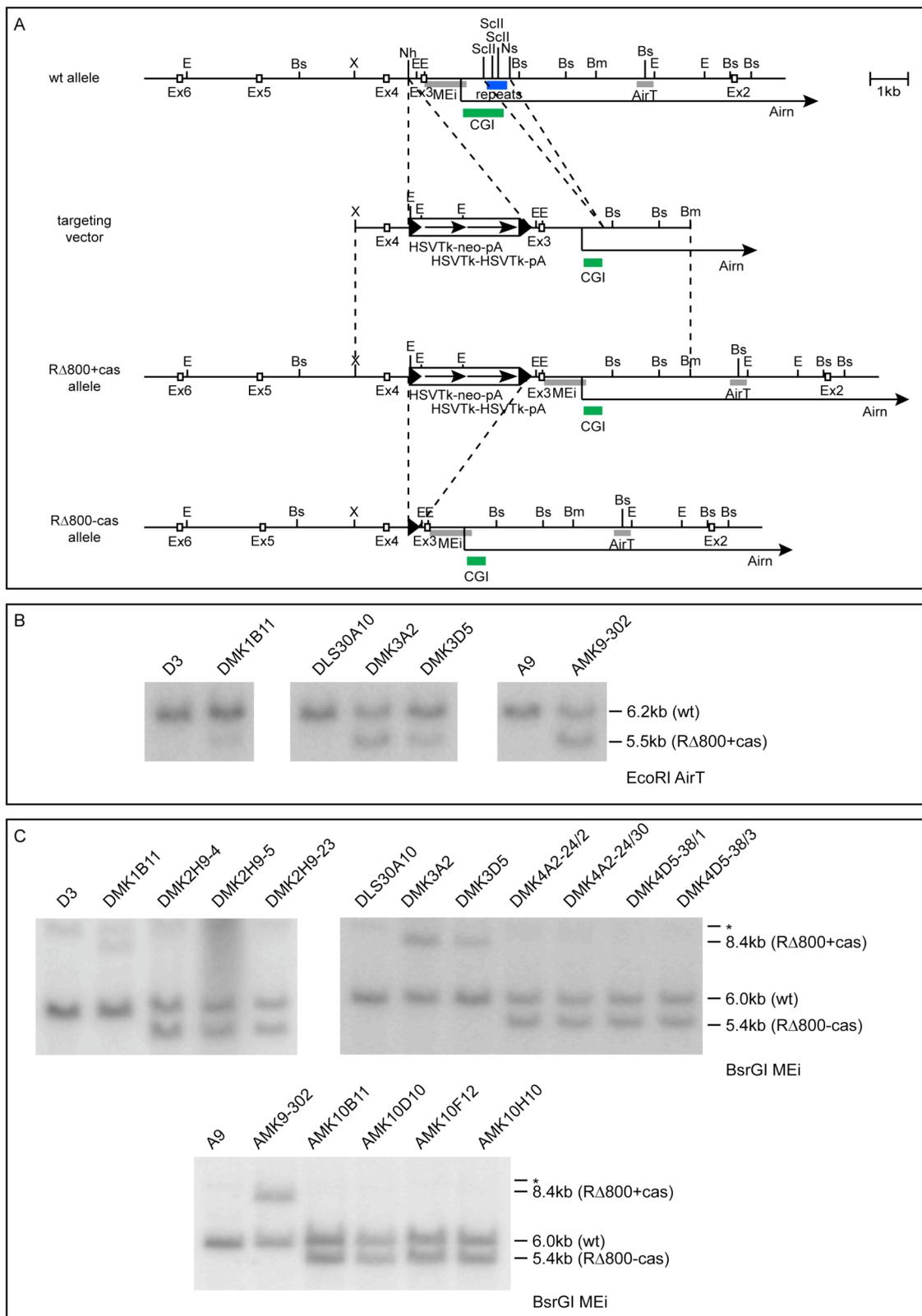


Fig. 21: Generation of R Δ 800 ES cells. (A) Targeting strategy for generating R Δ 800 ES cells. On top a scheme of the wildtype (wt) allele surrounding the transcriptional start site of *Airn* is shown. The targeting vector has a homology region from the XbaI site in intron 3 of *Igf2r* until the BmiI site in intron 2 of *Igf2r*. The tandem direct repeats downstream of *Airn* have been deleted in the targeting

vector, from the first SacII site until the NsiI site. A floxed selection cassette with an HSV-*tk* promoter driving a neomycin resistance gene stopped by a SV40 polyadenylation (pA) signal and an HSV-*tk* promoter driving the HSV-*tk* gene stopped by a HSV-*tk* pA signal was inserted into the NheI site in intron 3 of *Igf2r*. Homologous recombination created the R Δ 800+cas allele. Genotyping was performed by DNA blotting using an EcoRI digest and hybridisation with probe AirT. Transient transfection with a Cre-recombinase expressing plasmid resulted in recombination of the two loxP sites and deletion of the selection cassette, leaving a single loxP site in intron 3 of *Igf2r* and creating the R Δ 800-cas allele. Genotyping was performed by DNA blotting using a digest with BsrGI and hybridisation with probe MEi. Exons of *Igf2r* are shown as white boxes. The tandem direct repeats are shown as a blue bar. The *Airm* CGI is shown as a green bar. Transcriptional orientation of *Airm* is shown as an arrow. Hybridisation probes are shown as grey bars. LoxP sites are shown as black triangles, transcriptional orientations of the neomycin resistance gene and the HSV-*tk* gene are shown as arrows. Only restriction enzymes crucial for targeting or genotyping are given. E: EcoRI, Bs: BsrGI, X: XbaI, Nh: NheI, ScII: SacII, Ns: NsiI, Bm: BmiI. (B) Genotyping by DNA blotting of ES cells carrying an R Δ 800+cas allele. Genomic DNA was digested using EcoRI and hybridised using probe AirT. Shown are the targeted ES cells (DMK1B11, DMK3A2, DMK3D5, AMK9-302) and their respective parental ES cell line (D3, DLS30A10, A9). The wildtype band is 6.2kb in length, the targeted band 5.5kb. (C) Genotyping by DNA blotting of ES cells carrying an R Δ 800+cas allele (DMK1B11, DMK3A2, DMK3D5, AMK9-302) or an R Δ 800-cas allele (DMK2H9-4, DMK2H9-5, DMK2H9-23, DMK4A2-24/2, DMK4A2-24/30, DMK4D5-38/1, DMK4D5-38/3, AMK10B11, AMK10D10, AMK10F12, AMK10H10) and the original parental ES cell line (D3, DLS30A10, A9). Genomic DNA was digested using BsrGI and hybridised with probe MEi. A wildtype band is 6.0kb in length, an R Δ 800+cas allele gives a 8.4kb long band, an R Δ 800-cas allele a 5.4kb long band. Note that hybridisation also resulted in a faint unspecific band with approximately 10kb in size (marked by an asterisk).

DMK2H9 was subcloned, 46 out of 48 picked clones showed a wildtype band and an R Δ 800-cas band. Three subclones (DMK2H9-4, DMK2H9-5, DMK2H9-23) were used for further experiments (see Table 6, Fig. 21). DMK4A2-24 and DMK4A2-38 were subcloned as well. Two subclones each (DMK4A2-24/2, DMK4A2-24/30, DMK4D5-38/1, DMK4D5-38/3) were used for further experiments (see Table 6, Fig. 21).

As AMK10B11, AMK10D10, AMK10F12 and AMK10H10 are intended to be used for blastocyst injection, they were not subcloned to reduce the time in culture and thus to preserve their ability of generating chimeras and contributing to the germline.

As the HSV-*tk* gene induces male sterility (Braun et al., 1990), AMK9-302 could not be used directly for blastocyst injection followed by removal of the selection cassette by crossing to a Cre-recombinase expressing mouse. Instead, the selection cassette had to be removed in the ES cell system. As each transformation process and additional passaging in culture potentially might decrease the capability of ES cells to create a knock out mouse, I decided to re-clone the targeting vector (pR Δ 800). I

replaced the existing selection cassette with another one which has a neomycin resistance gene driven by a *Pgk1*-promoter and stopped by a *Pgk1* polyadenylation signal, flanked by loxP sites (pR Δ 800A) (the selection cassette was a gift of Maria Sibila, Department of Dermatology, Medical University of Vienna). I electroporated A9 ES cells with the new targeting vector, one clone out of 552 picked colonies was homologously targeted (AMK11A10) (see Table 7, Fig. 22). Afterwards, the positive clone was subcloned, four of the subclones were used further.

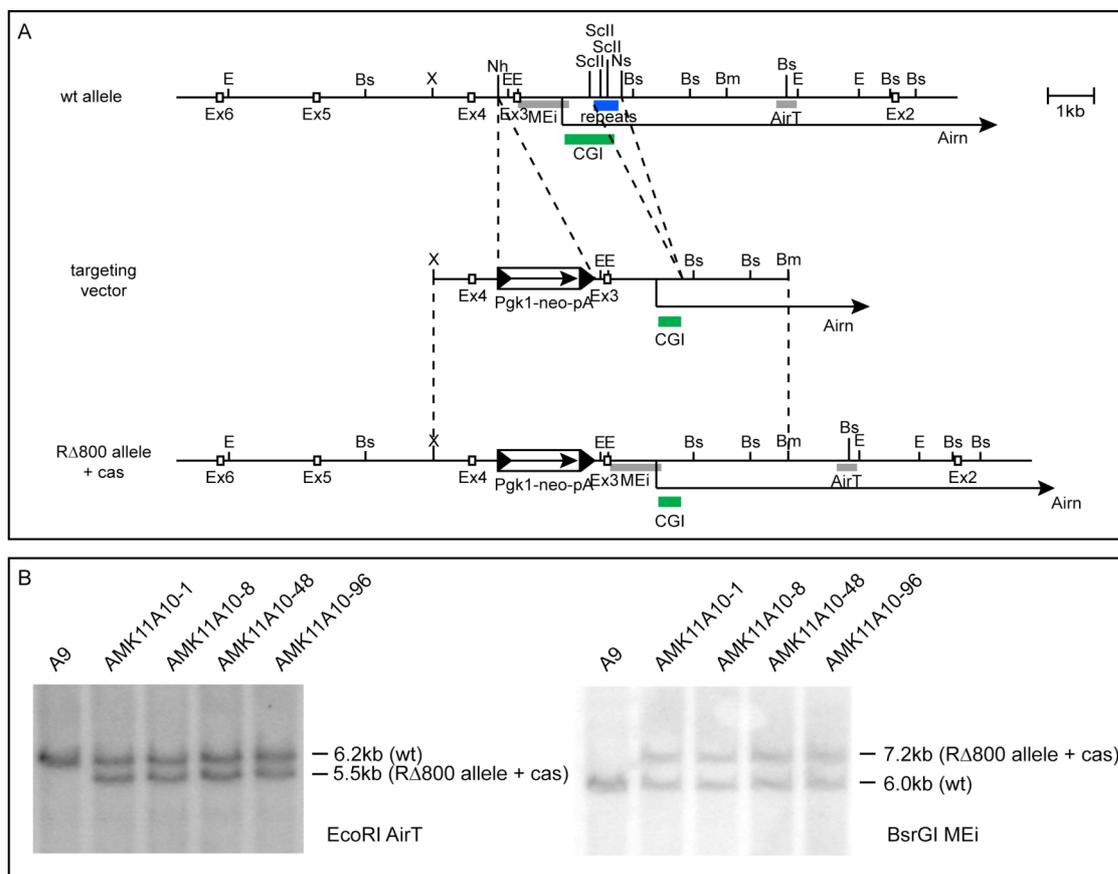


Fig. 22: Generation of AMK11 ES cells. (A) As in Fig. 21 but here the selection cassette contains a *Pgk1* promoter driving a neomycin resistance gene stopped by a *Pgk1* polyadenylation (pA) signal flanked by loxP sites. (B) Genotyping by DNA blotting of ES cells carrying a R Δ 800+cas allele. Genomic DNA was digested using EcoRI and hybridised using the external probe AirT or BsrGI and using the internal probe MEi. Shown are the targeted ES cells (AMK11A10) and the parental ES cell line (A9). The wildtype band is 6.2kb in length, the targeted band 5.5kb.

name	parental cell line	targeting vector	number of picked clones	number of recovered DNAs	number of positive clones	names of clones used further
AMK11	A9	pR Δ 800A	552	501	1	AMK11A10
subclone	AMK11A10	-	96	92	91	AMK11A10-1 AMK11A10-8 AMK11A10-48 AMK11A10-96

Table 7: ES cell line having a 700bp deletion of the tandem direct repeats in the *Airn* CGI. For details see Table 6. In contrast to the targeting vector used to create the cell lines in Table 5, (pR Δ 800) which contains a selection cassette with a neomycin resistance gene and an HSV-*tk* gene, the targeting vector used here (pR Δ 800A) contains a selection cassette with a neomycin resistance gene only.

To assay which of the parental alleles was targeted, I made use of the differential DNA methylation present on the ICE. I digested genomic DNA of targeted cells without the selection cassette (DMK2H9-4, DMK2H9-5, DMK2H9-23, DMK4A2-24/2, DMK4A2-24/30, DMK4D5-38/1, DMK4D5-38/3, AMK10B11, AMK10D10, AMK10F12, AMK10H10) or with the selection cassette (AMK11A10-1, AMK11A10-8, AMK11A10-48, AMK11A10-96) and the original parental cell line (D3, DLS30A10, A9) with EcoRI and MluI. DNA blots were hybridised with probe AirT. A wildtype maternal allele will be detected as a 6.2kb fragment, a wildtype paternal allele as a 5.0kb fragment. A targeted allele will result in a 4.3kb fragment, irrespective of the parental allele. This is, because the DNA of the homology region which was grown in bacteria does not have CpG DNA methylation. Therefore, integration on the maternal allele will result in a replacement of the endogenous DNA methylated region with unmethylated DNA and will therefore give the same fragment as in the case of integration on the paternal allele, which is free of methylation, as MluI now can cut on both parental alleles. Both, DMK2 and DMK4 ES cell lines showed a 6.2kb long fragment corresponding to a methylated maternal wildtype allele, and a 4.3kb long fragment corresponding to a targeted allele. Therefore, targeting of D3 ES cells resulted in 5 out of 5 achieved homologous recombination events in a targeting of the paternal allele (in DMK2, DMK4A2, DMK4D5 as well as DMK4A1 and DMK4D4, data not shown) (see Fig. 23). AMK10 and AMK11 ES cell lines

however showed a 5.0kb long fragment corresponding to an unmethylated paternal wildtype allele and a 4.3kb band correlating with a targeted allele. Therefore, homologous recombination in A9 ES cells resulted in two out of two cases in a maternal targeting (see Fig. 23).

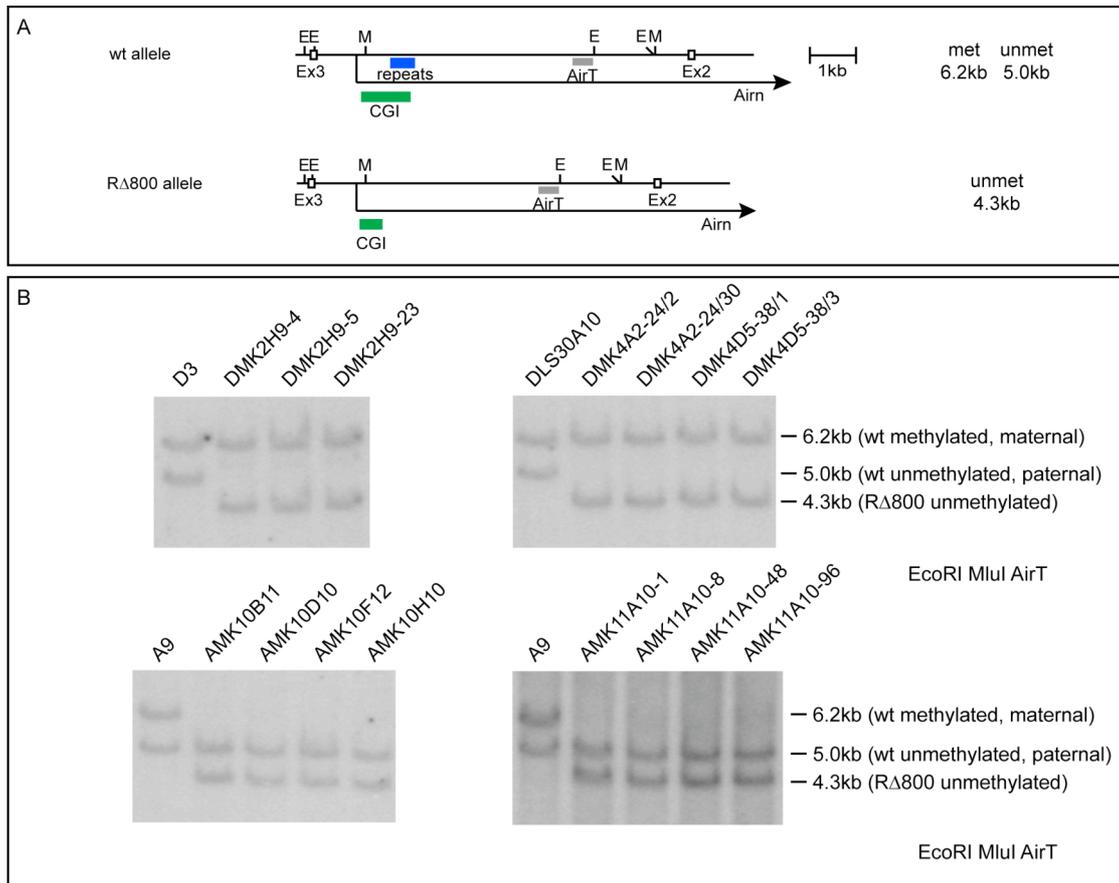


Fig. 23: Analysis of parental targeting of repeat deletion ES cells. (A) Overview of the genomic locus and DNA blot strategy. Details as in Fig. 21. (B) DNA blots of wildtype parental ES cells and ES cells carrying a repeat deletion allele. Genomic DNA was digested using EcoRI and MluI and hybridised using probe AirT. When using D3 and DLS30A10 ES cells, homologous recombination took place on the paternal allele in all cases. When using A9 ES cells, homologous recombination happened on the maternal allele. Please note that recombination on the maternal allele will result in the replacement of the methylated endogenous sequence with unmethylated DNA, as the targeting vector was amplified in bacteria, therefore the MluI site of the R Δ 800 allele will always be unmethylated, irrespective of the parental allele.

The preferential targeting of the paternal allele when using D3 ES cells is consistent with previous experiments (Wang et al., 1994, Latos P.A. and Stricker S.H unpublished). Therefore it was surprising to see, that when using A9 cells, two out of two homologous recombinations took place on the maternal allele. A9 ES cells are

F1 ES cells between C57/BL6 and 129Sv, however we did not know the direction of the cross, meaning who is the father and who the mother. As the homology region of the targeting vector was derived from the cosmid cos940PS which contains genomic DNA of 129Sv origin, the preferential maternal targeting of A9 ES cells could be explained by the mother being 129Sv. As A9 ES cells are male (Anton Wutz, personal communication), I could make use of the X chromosome to sort out the parental origin of A9 ES cells. Using the Mouse SNP (single nucleotide polymorphism) Database (<http://mousesnp.roche.com/cgi-bin/msnp.pl>) I found out that *Cybb*, a gene located on the X chromosome, has several SNPs between 129Sv and C57/BL6. I designed PCR primers for two amplicons containing two SNPs each, performed a PCR on genomic DNA of A9 ES cells, DLS30A10 ES cells (which are of 129Sv origin) and of C57/BL6 adult liver and obtained the sequences of the PCR amplicons. Using those sequences I was able to show that at all four positions the A9 sequence corresponds to the C57/BL6 but not to the 129Sv sequence. Therefore the X chromosome of A9 ES cells is of C57/BL6 origin, showing that the mother was C57/BL6, the father 129Sv (see Fig. 24). Therefore the preferential maternal targeting observed in A9 cells cannot be explained by the maternal allele fitting the strain from which the targeting vector originated.

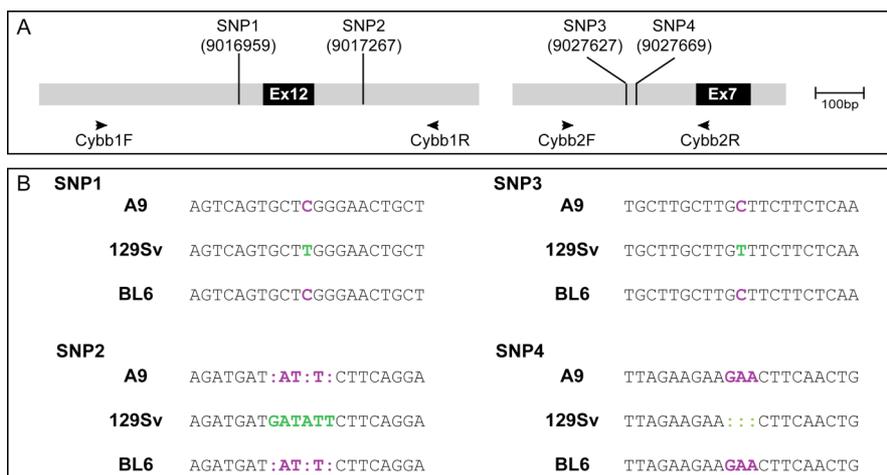


Fig. 24: Using SNPs in an X-linked gene to determine the parental origin of A9 ES cells. (A) Schematic overview of the two chosen PCR amplicons containing two SNPs each. Exons are shown as black boxes. Positions of SNPs are given by vertical black lines. The numbers in brackets give the first bp of the SNP according to the UCSC genome browser. PCR primers (Cybb1F, -R, Cybb2F, -R) are shown as arrows. (B) Sequencing result of the SNP and the surrounding region. A9: genomic DNA from A9 ES cells. 129Sv: DLS30A10 genomic DNA, BL6: adult liver from C57/BL6 mouse. SNPs are shown by green or purple font. Deleted bases are shown as colons.

AMK10 and AMK11A10 cells were generated to obtain a repeat deletion mouse by blastocyst injection. Aneuploidy is the major problem in obtaining chimeras and contribution to the germ line, therefore only ES cells containing more than 70% of euploid nuclei should be used for blastocyst injection (Nagy et al., 2003). I analysed the chromosome numbers of AMK10 and AMK11A10 by metaphase spreads to see, how many cells have 40 chromosomes (see Fig. 25). AMK10 ES cells showed complete aneuploidy, as no metaphase spread showed 40 chromosomes, instead, they mainly had between 42 and 43 chromosomes. Therefore AMK10 ES cells are not suited for blastocyst injection. For AMK11A10 ES cells however, all four subclones showed in more than 80% of metaphases 40 chromosomes and therefore can be used for blastocyst injection.

Targeted ES cells frequently show a loss of the unmethylated ICE of the *Dlk1* cluster (Ru Huang, personal communication). A methylated ICE correlates with expression of *Dlk1* which acts as a growth promoter (Moon et al., 2002). To additionally check the quality of the repeat deletion ES cells created for blastocyst injection, I analysed DNA methylation of the *Dlk1* ICE using a methylation-sensitive restriction enzyme and DNA blotting (see Fig. 26). Whereas A9 ES cells and the four AMK11A10 subclones (AMK11A10-1, AMK11A10-8, AMK11A10-48, AMK11A10-96) showed two bands corresponding to a methylated and an unmethylated allele, AMK10 ES cells (AMK10B11, AMK10D10, AMK10F12, AMK10H10) showed only the larger band corresponding to a methylated allele. As all four AMK10 ES cell lines are clones generated by removing the selection cassette from the same parental ES cell (AMK9-302), it is likely that already this clone either gained DNA methylation of the maternal *Dlk1* ICE or lost the whole maternal chromosome 12 or parts of it and doubled up the corresponding regions from the paternal allele. This result gives an additional indication, that AMK11A10 subclones are suited for blastocyst injection, as they show a normal *Dlk1* DMR, whereas AMK10 clones are clearly not suited for injection, as they are aneuploid and have a disturbed DNA methylation pattern at the *Dlk1* DMR.

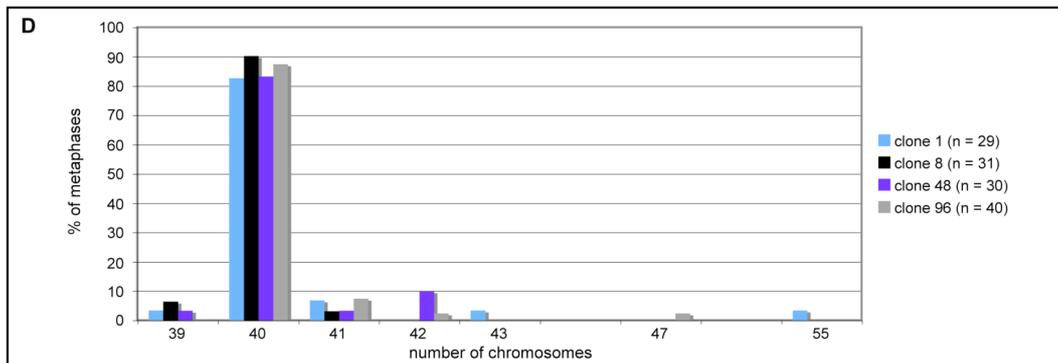
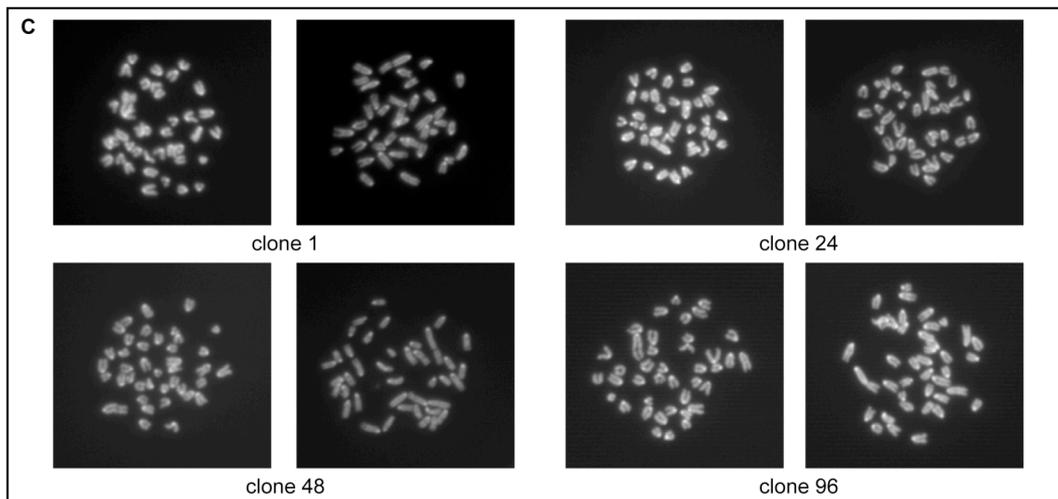
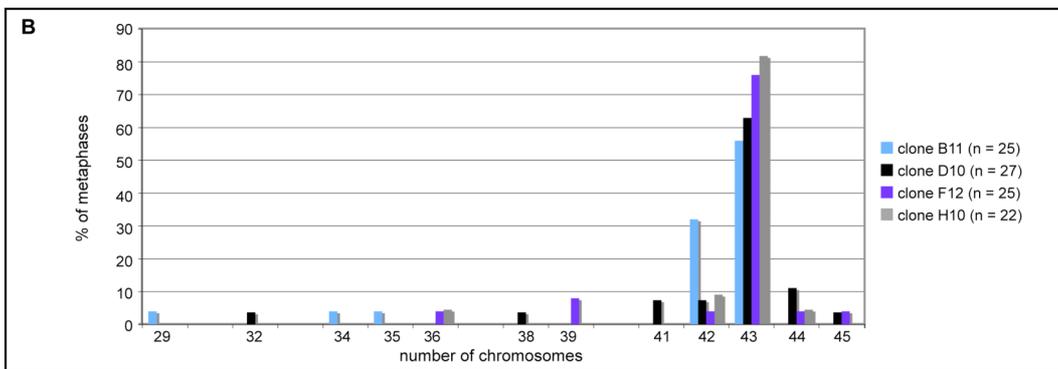
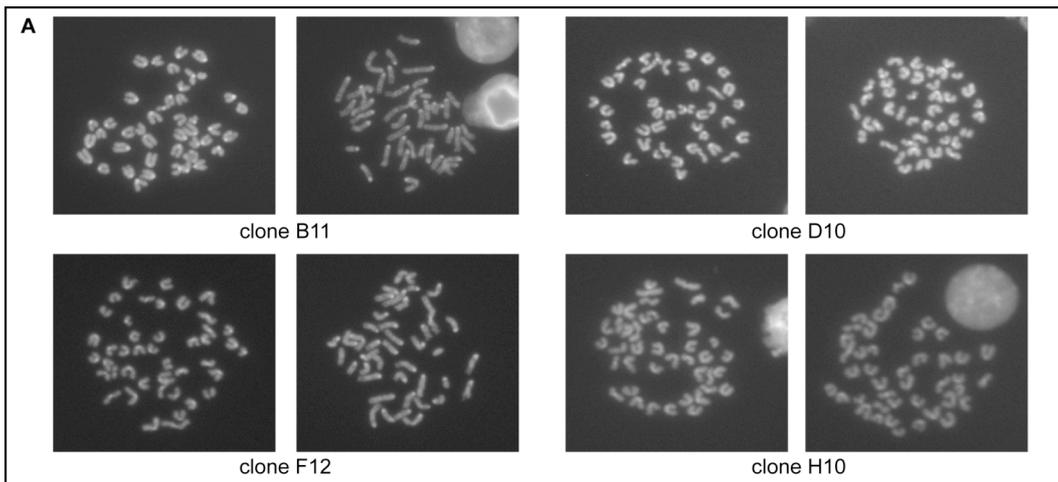


Fig. 25: Metaphase spreads of AMK10 (A, B) and AMK11A10 (C, D) ES cells. (A) Representative pictures of metaphases for the four cassette-removed clones of AMK10 (clone B11, clone D10, clone F12, clone H10). (B) Percentage of metaphases from (A) with different chromosome numbers. All four clones show complete aneuploidy. (C) Representative pictures of metaphases for the four subclones of AMK11A10 (clone 1, clone 24, clone 48, clone 96). (D) Percentage of metaphases from (C) with different chromosome numbers. All four subclones have more than 80% metaphases with 40 chromosomes.

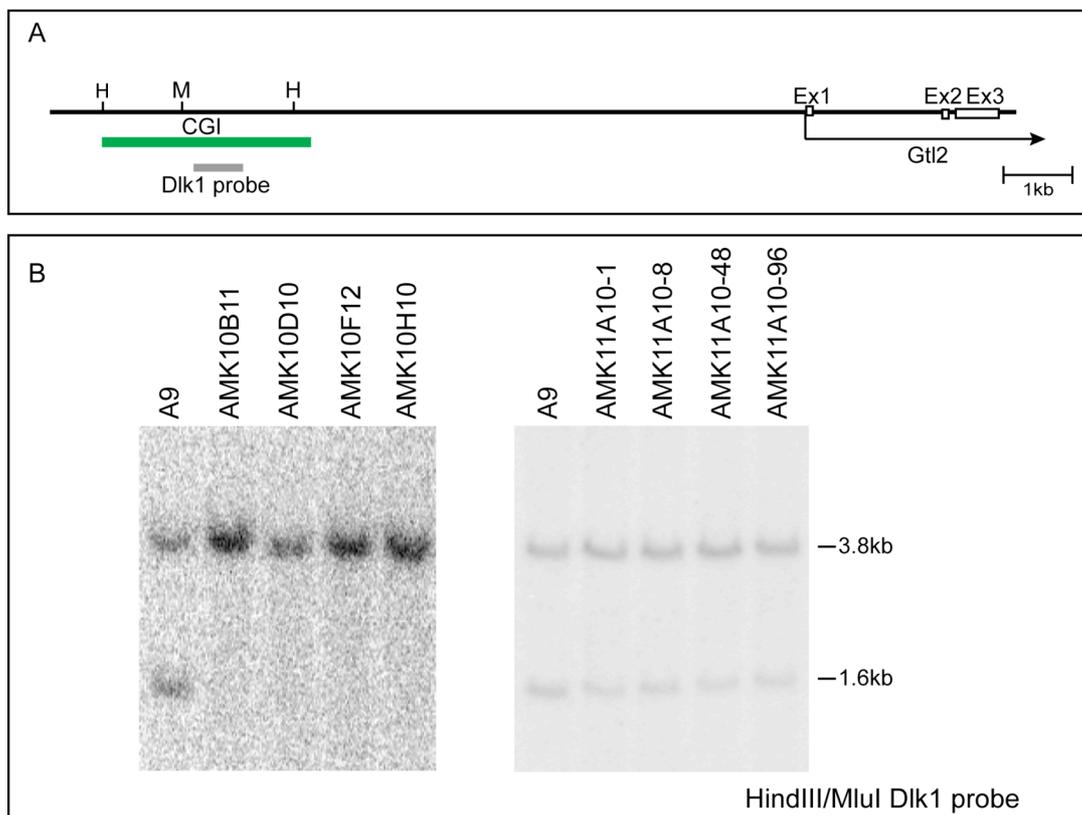


Fig. 26: Methylation status of the *Dlk1* ICE by DNA blotting. (A) Schematic overview of the analysed region. The ICE of this region located 10kb upstream of *Gtl2* and contains a CpG island (CGI) which is normally methylated on the paternal allele but free of methylation on the maternal allele. (B) Genomic DNA of ES cells was subjected to a HindIII/MluI digest, DNA blots were probed using the *Dlk1* probe. The left handed DNA blot was done by Ru Huang, the right one by myself. Please note that MluI is methylation-sensitive. A methylated allele will result in a 3.8kb band, an unmethylated allele in a 1.6kb band.

2.2.3 Does deletion of the *Airn* tandem direct repeats result in a loss of the differential DNA methylation on the ICE?

By using the *in vitro* differentiation system described in the introduction I analysed the gametic differential DNA methylation mark present on the ICE. I first analysed wildtype and DMK2 ES cells (R Δ 800 in D3 wildtype cells) and harvested genomic

DNA of undifferentiated ES cells (d0) and of differentiated ES cells after 14 days of retinoic acid treatment (d14). The DNA was digested using EcoRI and MluI, the DNA blot was hybridised with probe AirT. As controls, wildtype tail DNA was digested either with EcoRI only or EcoRI and MluI. Furthermore I also used genomic DNA of +/-Thp and Thp/+ 13.5dpc embryos to be able to confirm, which fragment comes from which parental allele. The DNA blot result shows, that the direct tandem repeats on the paternal allele are not needed to keep the paternal ICE free of DNA methylation, as all DMK2 ES cells have a strong 4.3kb band. Surprisingly, in all 3 analysed DMK2 ES cell lines (DMK2H9-4, DMK2H9-5, DMK2H9-23), at d0 an additional faint fragment correlating with a methylated paternal allele was visible but absent at d14 (see Fig. 27). I cannot conclude, if this low level of paternal ICE methylation is due to the lack of the tandem direct repeats or is a general feature of undifferentiated ES cells, as in wildtype ES cells methylated parental alleles can not be distinguished by size. However, at d14 of differentiation, this additional band is not present anymore, indicating that differentiation or other changes associated with the differentiation process (e.g. changes in transcription patterns) lead to a loss of the low level methylation on the paternal allele. Besides that, this result demonstrates that a deletion of the tandem direct repeats on the paternal allele does not disturb the maintenance of the differential DNA methylation on the ICE, as both, a methylated and an unmethylated band can be seen at d0 and d14 of ES cell differentiation.

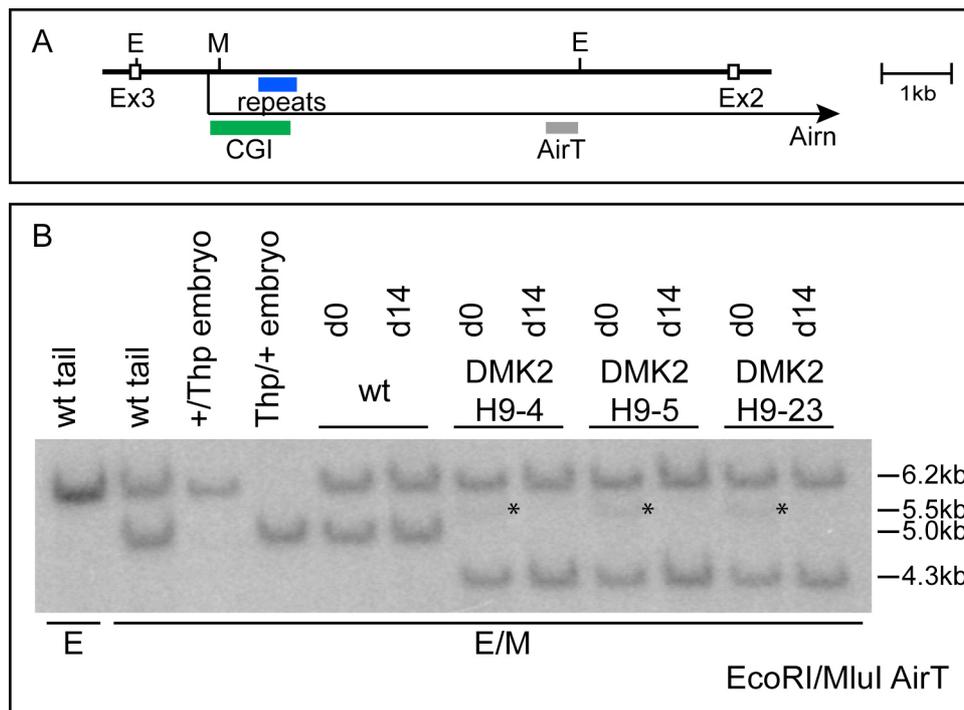


Fig. 27: Analysis of the DNA methylation imprint mark in D3 wildtype cells carrying an R Δ 800 allele. (A) Scheme for the DNA blotting assay analysing a diagnostic MluI site in the ICE. White boxes show exons of *Igf2r*. Restriction enzymes crucial for the assay are given (E: EcoRI, M: MluI). The location of the repeats is shown as a blue bar, the CGI as a green bar. Transcriptional direction of *Airn* is shown as an arrow. The probe AirT used for the DNA blotting is shown as a grey bar. (B) DNA blotting result. Wildtype mouse tails were used as controls and digested either with EcoRI alone or with EcoRI and MluI. +Thp and Thp/+ 13.5dpc embryos were used to confirm the parental origin of the detected fragments: The top 6.2kb fragment correlates with a wildtype methylated maternal allele, the lower 5.0kb fragment with a wildtype unmethylated paternal allele. Genomic DNA of ES cells (wildtype (wt) or DMK2H9-4, DMK2H9-5, DMK2H9-23) was harvested at day 0 (d0) or after 14 days (d14) of retinoic acid differentiation. Wildtype ES cells show the same bands as wildtype tails, both at d0 and d14. DMK2 ES cells show the same 6.2kb fragment as wildtype ES cells, as the maternal allele is also wildtype in those cells. The 5.0kb fragment is not present in those cells, instead they show a 4.3kb fragment, as 700bp of genomic DNA containing the repeats have been deleted on the paternal allele. In addition, at d0 an additional faint band of approximately 5.5kb is visible (asterisk), correlating with a methylated paternal R Δ 800 allele.

DMK4 cells (R Δ 800 in DLS30A10) were differentiated three times independently and also assayed for the differential DNA methylation on the ICE. For the third differentiation set, genomic DNA was also harvested at d5 of differentiation additionally to d0 and d14. As controls served again wildtype tails or wildtype pMEFs and +Thp and Thp/+ 13.5dpc embryos or pMEFs. Again, the differential DNA methylation on the ICE was not disturbed in the repeat deletion cells, and again a faint 5.5kb band correlating with a low level of paternal DNA methylation was present at d0 in all four analysed clones (DMK4A2-24/2, DMK4A2-24/30, DMK4D5-38/1, DMK4D5-38/3) and for all three biological replicates (see Fig. 28).

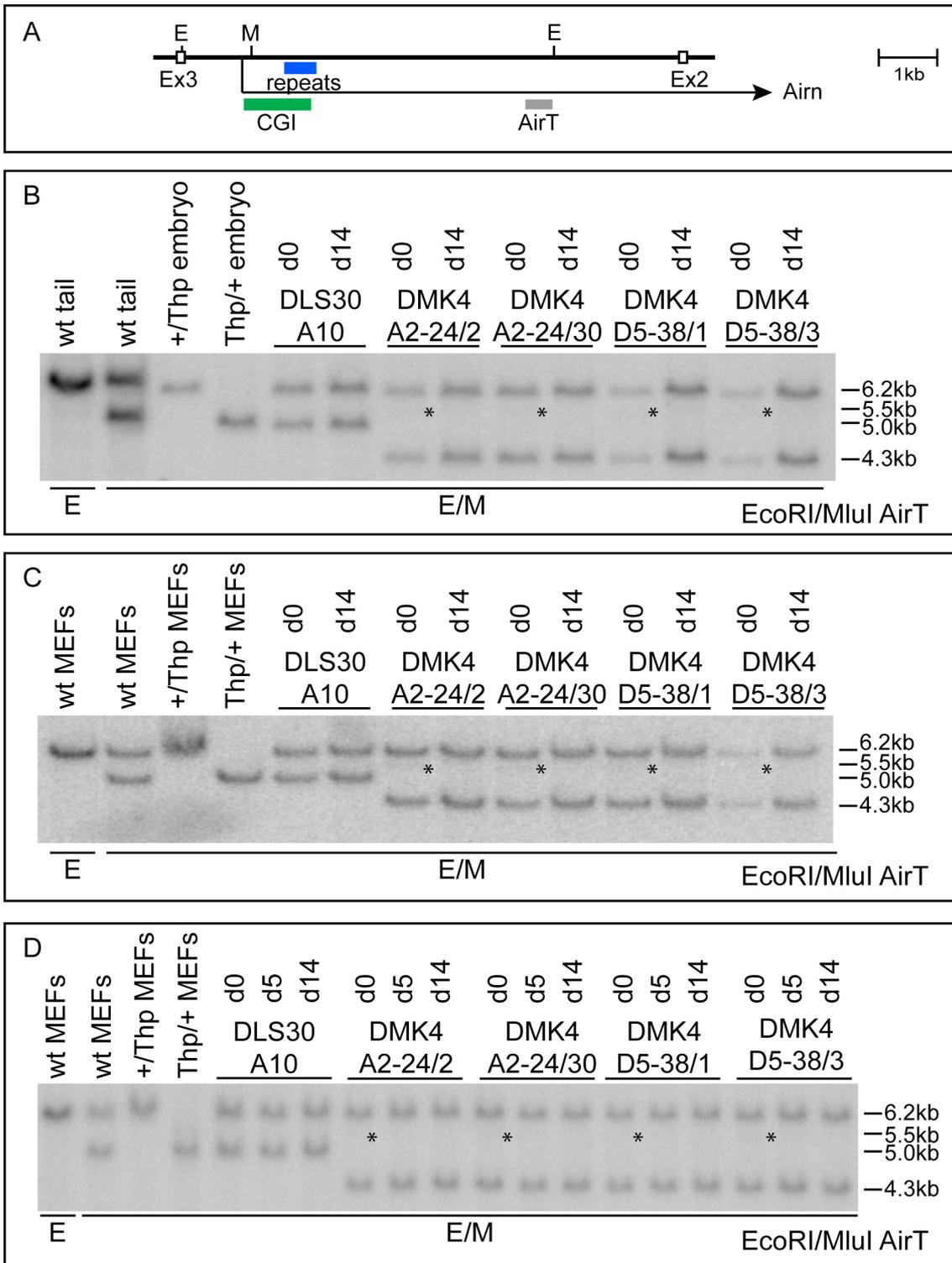


Fig. 28: As Fig. 27 but for DMK4 (R Δ 800 in DLS30A10 cells). (B)-(D) Three independent differentiation sets. As controls served either wildtype pMEFs or wildtype tails and +Thp and Thp/+ 13.5dpc embryos or pMEFs. In (D) additionally to d0 and d14 of differentiation, also d5 was assayed.

2.2.4 Does deletion of the *Airn* tandem direct repeats lead to changes in *Airn* expression?

In the next step I assayed, if the deletion of the *Airn* tandem direct repeats led to changes in *Airn* expression. To obtain a global impression of *Airn* appearance in R Δ 800 ES cells, I isolated total RNA from DMK2H9-4 and DMK4D5-38/1 ES cells differentiated for 5 days and prepared cDNA which was hybridised to an RNA expression tiling array (RETA) against genomic DNA. As a comparison I used cDNA hybridisations of DLS30A10 cells prepared by Basak Senergin and of +/Thp and Thp/+ pMEFs prepared by Ru Huang. The results of the RETA can be seen in Fig. 29. Generally, no drastic changes in the appearance of *Airn* seem to occur upon the deletion of the tandem direct repeats. However, a closer comparison of *Airn* in DLS30A10 cells and DMK2H9-4 and DMK4D5-38/1 ES cells reveals that R Δ 800 *Airn* seems to be slightly shorter compared to wildtype *Airn*. Around bp position 13014100-1302900 there is a big gap in the tiling array which is due to the presence of mainly LINE elements in this region where no single copy tiles could be spotted onto the array. After this gap, the tiles in the hybridisation of DLS30A10 still show a clear enrichment of cDNA versus genomic DNA, however the repeat deletion ES cells do not show enrichment at this position anymore. Instead, they rather adopt a more pMEF-like appearance of *Airn*, which is shorter compared to ES cells. Furthermore, although also in DLS30A10 cells *Airn* exhibits a gradual decline in signal intensities from the 5' to the 3' end, this decline seems to be enhanced in the R Δ 800 ES cells. *Igf2r* does not seem to be changed in DMK2 or DMK4 cells compared to DLS30A10 cells, however as signals on the tiling array are easily saturated, a quantitative comparison of highly expressed mRNA genes like *Igf2r* cannot be made from the RETA.

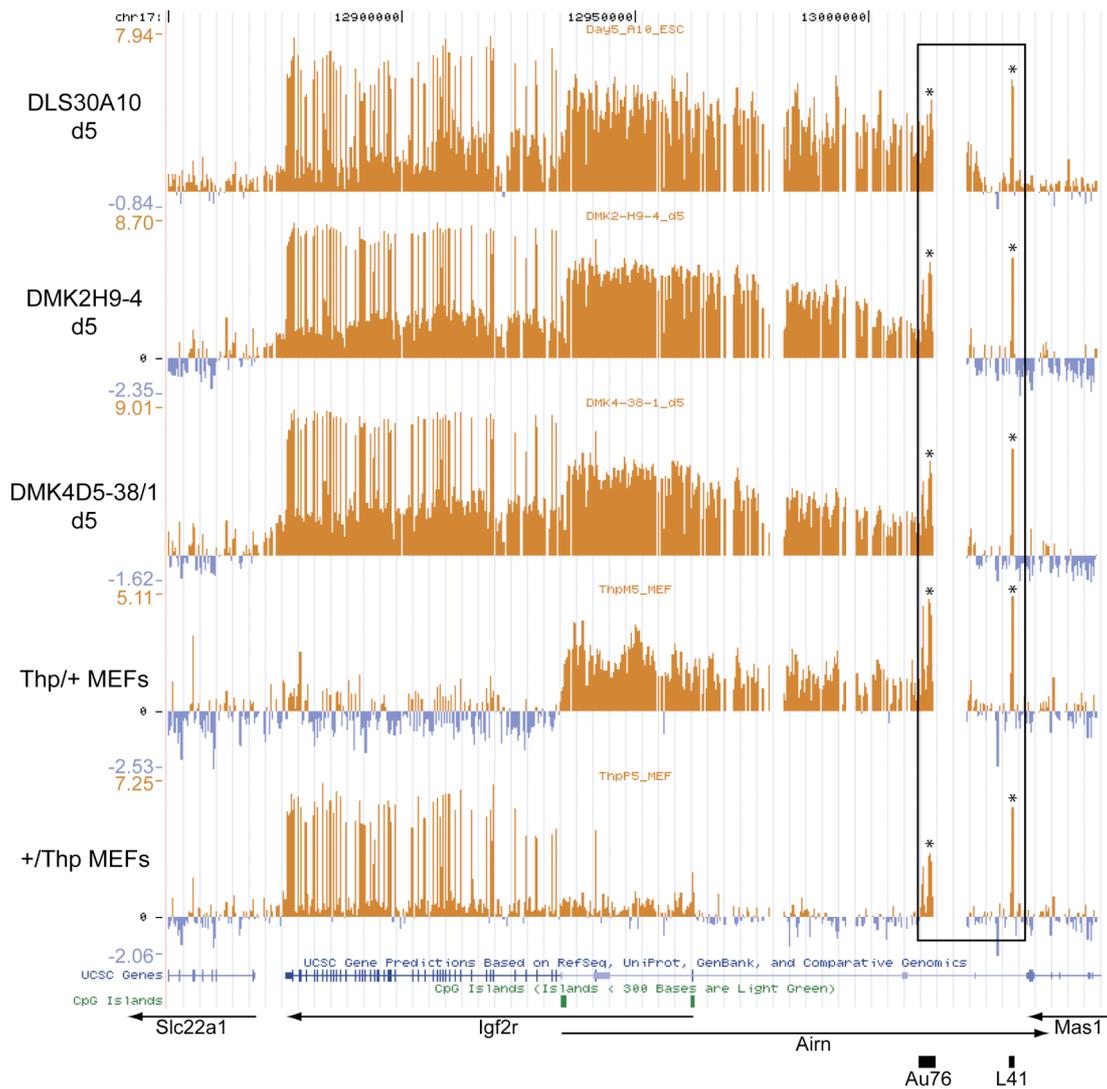


Fig. 29: RNA expression tiling array (RETA). Shown is a screenshot of the UCSC genome browser. The top five rows show cDNA hybridised against genomic DNA of DLS30A10, DMK2H9-4, DMK4D5-38/1 (all three ES cell lines were differentiated for 5 days (d5) with retinoic acid), Thp+/+ and +/Thp pMEFs. On the x-axis the bp position is given, on the y-axis the signal intensities in a \log_2 scale. Each orange signal corresponds to cDNA enriched compared to genomic DNA and therefore to gene expression. Below, UCSC Genes are shown in blue, CpG islands are shown in green, directions of genes are given as arrows below. The positions of the pseudogenes Au76 and L41 are shown as black bars. As *Airn* and *Igf2r* overlap each other at their respective 5' ends, hybridisations from Thp+/+ and +/Thp pMEFs are shown as a comparison (cDNA preparation from those cells was done by Ru Huang). Thp+/+ cells show only expression of *Airn*. In the gene body of *Airn* each tile shows enrichment of cDNA, reflecting the typical appearance of macro ncRNAs (Ru Huang, unpublished). Note the gap present in the framed region close to the 3' end of *Airn* around bp position 13014100-1302900 which is due to the fact, that only single copy sequences are present on the tiling array and the corresponding genomic region shows a high enrichment of mainly LINE elements. The high peak directly in front of this gap marked with an asterisk comes from the *Au46* pseudogene or crosshybridisation with its original *Rangap1* gene which is localised on chromosome 15. The other high signal on a single tile close to the 3' end of *Mas1* marked with an asterisk correlates with the pseudogene L41. +/Thp pMEFs show no expression of *Airn* but expression of *Igf2r*. Note that the high signals within *Igf2r* correspond to exons, the lower signals to introns, indicating a high transcription rate (Ru Huang, personal communication). *Slc22a2* and *Mas1* are not expressed neither in pMEFs nor in ES cells. The ES cells at d5 of differentiation show expression of both, *Airn* and *Igf2r*.

Next, I assayed steady-state levels of *Airn* in wildtype and DMK2 ES cells which were undifferentiated (d0), or differentiated for 3 days (d3), 5 days (d5) or 14 days (d14) with retinoic acid. Total RNA was isolated at the indicated timepoints, DNaseI treated and reverse transcribed into cDNA. Steady-state levels of *Airn* were measured at four different positions along the gene body of *Airn* using quantitative PCR (qPCR). The 'START' assay is a SybrGreen assay where both primers are localised within the first 53bp of *Airn*, in front of the splice donor used for all known splice variants of *Airn* and therefore measures both, spliced and unspliced *Airn* and covers bp positions -1 to 40 with respect to the main transcriptional start site T1. The 'RP11' assay is localised at the 5' end of *Airn*, covering bp positions 8 to 155 with respect to the main transcriptional start site T1. The 'Air middle' assay is localised in the middle of *Airn*, covering bp positions 53246-53325 with respect to the main transcriptional start site T1. The 'Air end' assay is localised at the end of *Airn*, covering bp positions 98675-98747 with respect to the main transcriptional start site T1. 'RP11', 'Air middle' and 'Air end' detect unspliced *Airn* only (see Fig. 30A). In undifferentiated ES cells, *Airn* is not expressed, neither in wildtype cells nor in R Δ 800 ES cells, indicating that the repeats are not necessary to repress *Airn* at this stage. Upon differentiation, wildtype ES cells show upregulation of *Airn* at all 4 positions assayed with increasing amounts during time. Also R Δ 800 *Airn* is upregulated upon differentiation, at least at the three most 5' assayed positions. *Airn* levels in differentiated R Δ 800 ES cells analysed using 'START', 'RP11' and 'Air middle' reach 20-80% of wildtype *Airn*, dependent on the assayed position and the clone. The 'Air end' assay however, could not detect any substantial amounts of R Δ 800 *Airn* at any of the examined timepoints during differentiation (see Fig. 30B). This result, that R Δ 800 *Airn* is shorter than wildtype *Airn* is consistent with the results obtained from RETA.

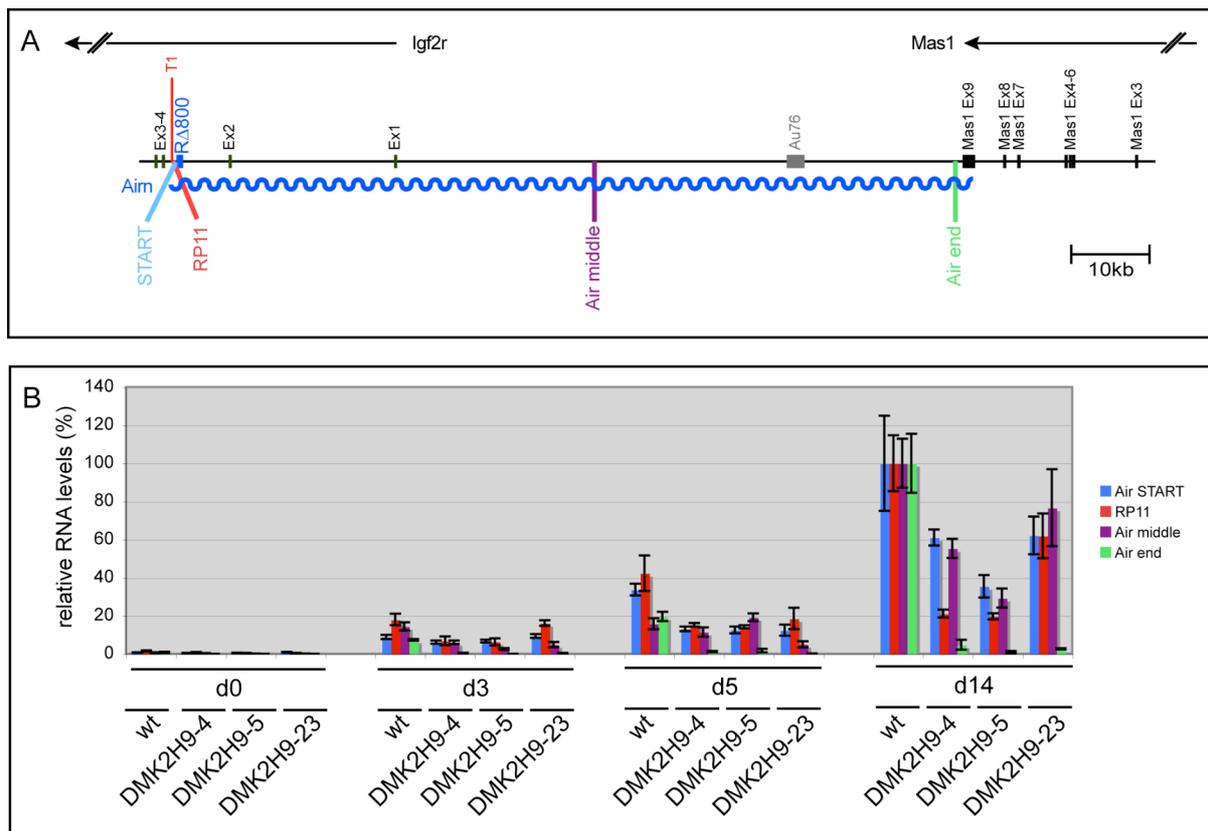


Fig. 30: Measurement of steady-state levels of *Airn* during ES cell differentiation using qPCR. (A) Schematic overview of the region examined. Exons of *Igf2r* and *Mas1* are shown as black boxes (note that those genes extend beyond the region shown here). Transcriptional orientations of *Igf2r* and *Mas1* are shown by arrows. The pseudogene *Au76* is shown as a grey box. *Airn* is shown as a blue wavy line, the main transcriptional start site is marked by T1. Location of the repeats is shown in blue. Locations of the qPCR assays ('START', 'RP11', 'Air middle', 'Air end') are shown below. (B) qPCR assays of RNA harvested at d0, d3, d5 and d14 of ES cell differentiation using the assays shown in (A). *Airn* levels were normalised to *Cyclophilin A*, wildtype levels at d14 were set to 100%. Bars give the mean value, error bars the standard deviation of three technical replicates.

The same experiment was repeated using DMK4 ES cells. Three differentiation sets have been performed using this cell line. Again, *Airn* is undetectable at d0 but upregulated upon differentiation with increasing amounts over time. Wildtype *Airn* and R Δ 800 *Airn* behave in a very similar way with the exception for the 'Air end' assay, which shows reduced steady state levels of the 3' end of *Airn* in R Δ 800 ES cells compared to DLS30A10 ES cells (see Fig. 31).

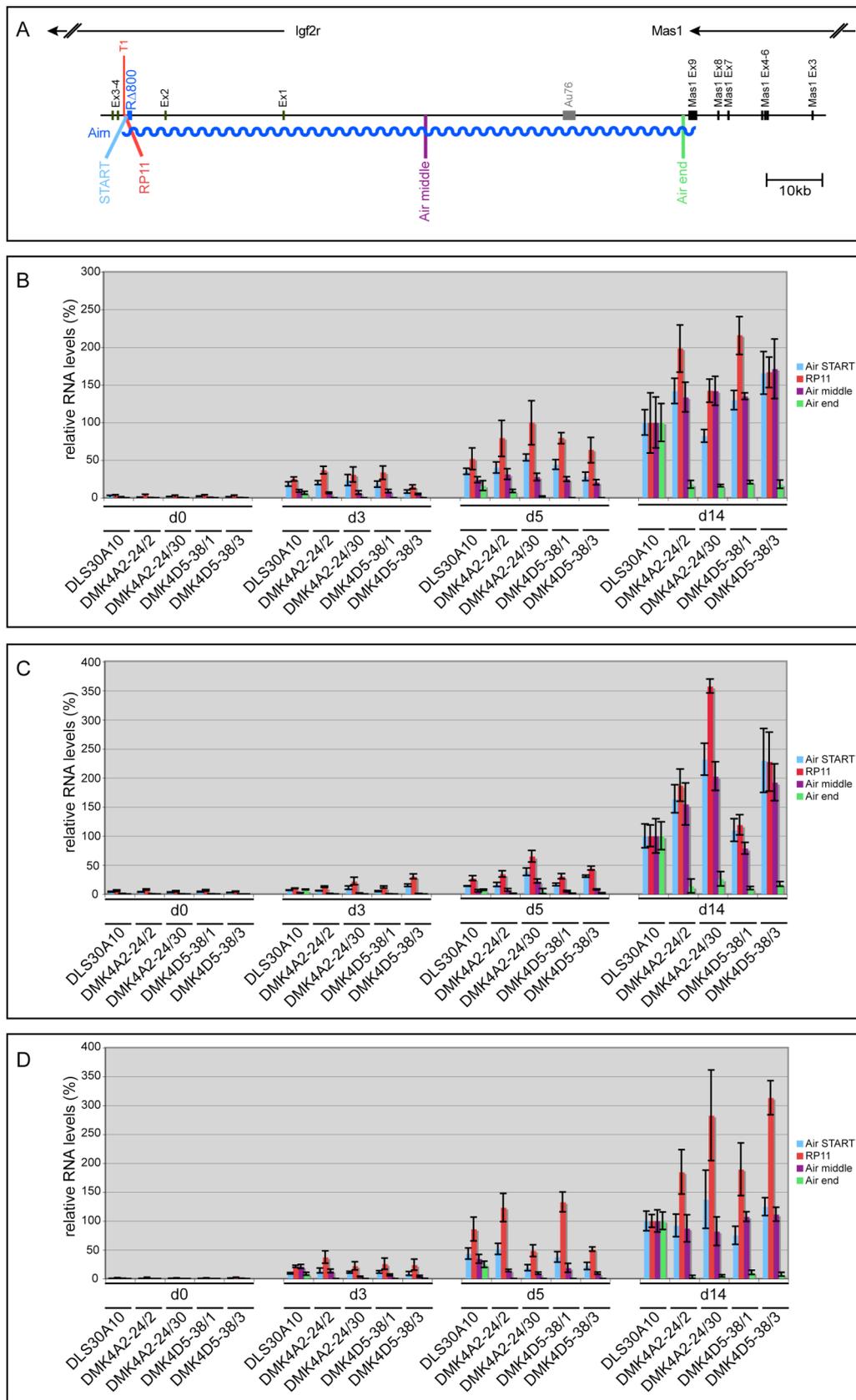


Fig. 31: As Fig. 30 but for DMK4 ES cells. (B)-(D) show three differentiation sets.

2.2.5 Does deletion of the *Airn* tandem direct repeats lead to changes in the low splicing capability of *Airn*?

Next, I assayed steady state levels of the known splice variants of *Airn*, again first in DMK2 and wildtype ES cells. Only approximately 5% of *Airn* transcripts are subjected to splicing and five different splice variants have been characterised in detail (Seidl et al., 2006). d0, d3, d5 and d14 of retinoic acid differentiation steady-state levels of four of those *Airn* splice variant, SV1a, SV1, SV2 and SV3, were assayed by qPCR. All those splice variants have the same exon 1 corresponding to the first 53bp of *Airn* but have different 3' exons (see Fig. 32A). The Taqman assays used have the same forward primer and probe localised in exon 1 but reverse primers specific for the different splice variants (Seidl et al., 2006). At d0, *Airn* splice variants are undetectable but upregulated upon differentiation with increasing amounts over time, both in wildtype ES cells as well as in DMK2 ES cells. However, the two splice variants assayed at the largest distance from the transcriptional start (SV2, SV3) show a consistent reduction in the three examined DMK2 cell lines compared to wildtype ES cells. At d14, SV2 reaches approximately 20% of wildtype levels in DMK2 ES cells, SV3 5-9%. Also this result indicates that *Airn* deficient for the repeats is shorter than wildtype *Airn*. The other two splice variants (SV1a, SV1) show a similar level in at least one of the DMK2 ES cell lines as in wildtype ES cells, indicating that their expression is not changed by a deletion of the tandem direct repeats (see Fig. 32B).

The same experiment was repeated with three differentiation sets of DMK4 ES cells. As in DMK2 ES cells, DMK4 ES cells did not show expression of *Airn* splice variants at d0 but showed upregulation during differentiation. In the first two differentiation sets of DMK4 at d5 and d14, both SV2 and SV3 were notably reduced compared to wildtype (see Fig. 33B-C). In the third differentiation set, this reduction of both SV2 and SV3 splice variants was seen only at d5 but at d14 SV2 showed in DMK4A2-24/30 and DMK4D5-38/3 levels comparable to DLS30A10 cells (see Fig. 33D). However, in the third differentiation set, the increase in *Airn* expression from d5 to d14 is much lower than in the preceding differentiation sets which could indicate, that

this last differentiation set is maybe not completely comparable to the other two. The results therefore indicate that SV2 and SV3, the *Airn* splice variant which are assayed at the largest distance away from the transcriptional start are notably reduced in RΔ800 ES cells compared to wildtype ES cells.

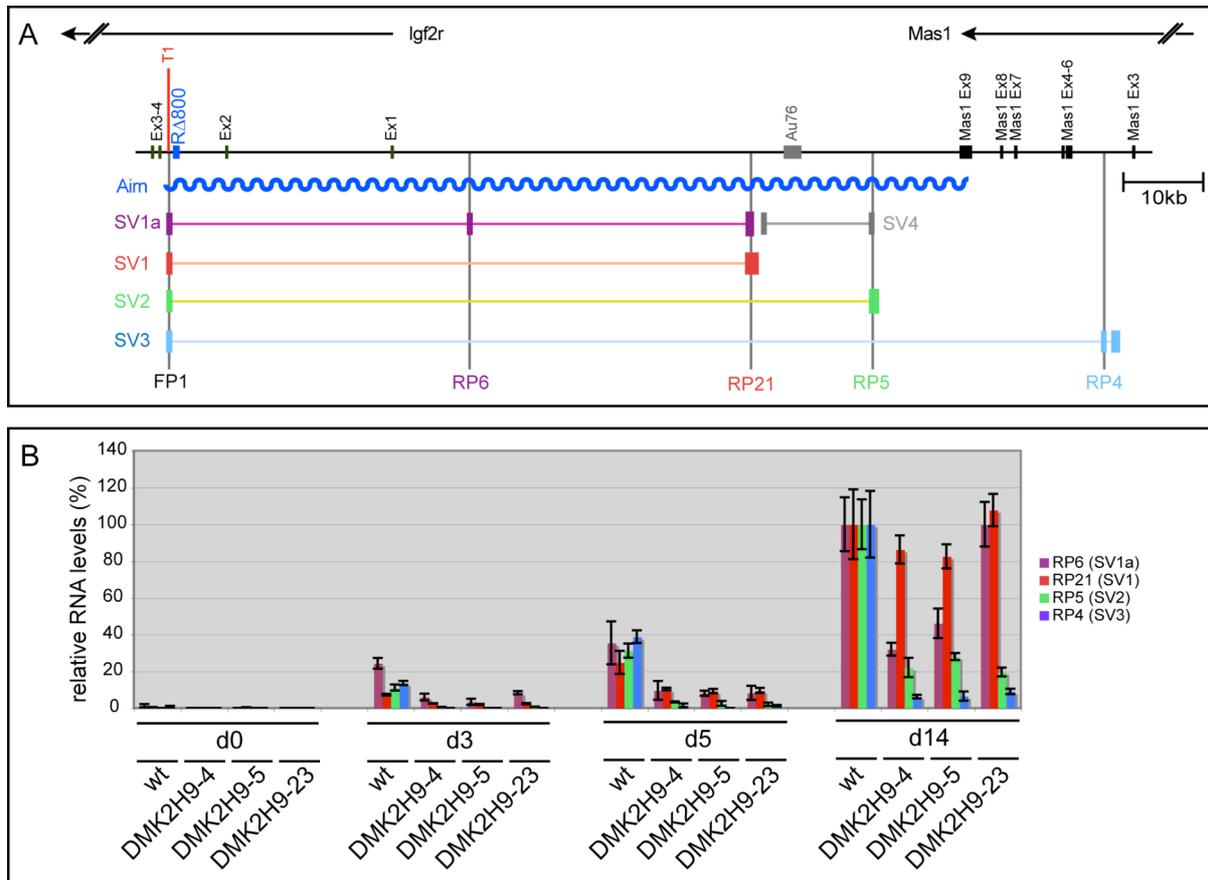


Fig. 32: Measurement of steady-state levels of known *Airn* splice variants during ES cell differentiation using qPCR. (A) Known splice variants (SV1, SV1a, SV2, SV3, SV4) are shown below full-length *Airn*, exons are marked by thick bars. Locations of the qPCR assays are shown below. All assays share the same forward primer FP1 and the same probe (not shown) localised within the first 53bp of *Airn* but use different reverse primers (RP6, RP21, RP5, RP4) specific for the different splice variants. (B) qPCR assays of RNA harvested at d0, d3, d5 and d14 of ES cell differentiation using the assays shown in (A). Other details as in Fig. 30.

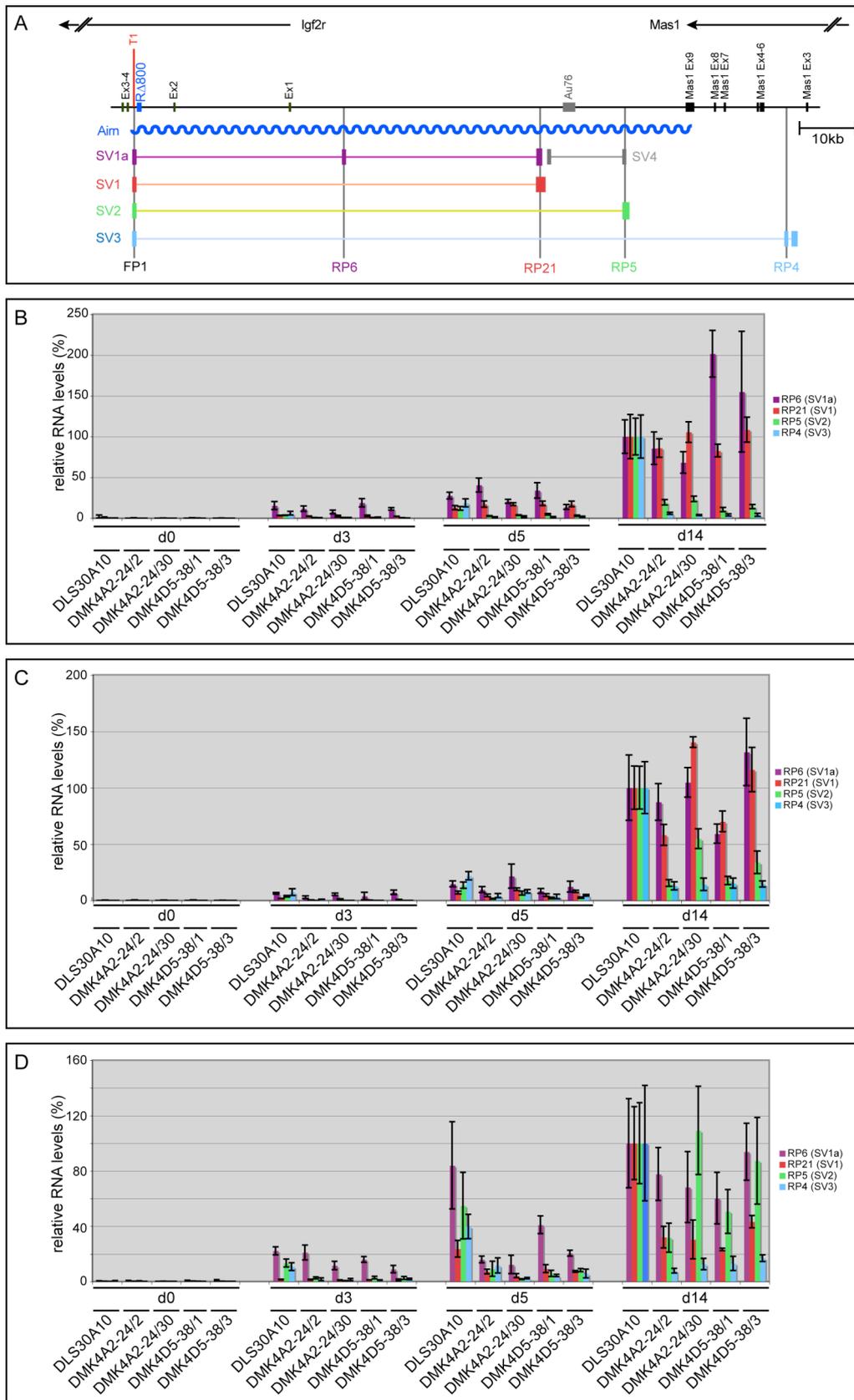


Fig. 33. As Fig. 32 but for DMK4 ES cells. (B)-(D) show three independent differentiation sets.

2.2.6 Does deletion of the *Airn* tandem direct repeats lead to changes in the nuclear retention of *Airn*?

It was reported that in MEF cells unspliced *Airn* largely fails to be exported to the cytoplasm but stays retained in the nucleus with a cytoplasmic to nuclear ratio of 1:30 to 1:80 relative to *Cyclophilin A* (Seidl et al., 2006). To analyse, if this behaviour is changed in the R Δ 800 ES cells, I performed cytoplasmic-nuclear fractionation at d14 of retinoic acid differentiation of DLS30A10 cells as controls and DMK4 cells and assayed *Airn* levels in the nucleus and the cytoplasm relative to *Cyclophilin A* by qPCR. As controls I also assayed *Igf2r* which was reported to show a cytoplasmic to nuclear ratio of 1:2.1 to 1:2.4 relative to *Cyclophilin A*, *45S-pre-rRNA* for which a cytoplasmic to nuclear ratio of 1:342 to 1:901 relative to *Cyclophilin A* was shown and *Gapdh* mRNA for which an almost equal cytoplasmic to nuclear ratio relative to *Cyclophilin A* was reported (Seidl et al., 2006). During one differentiation, I assayed only DMK4A2-24/2 and DMK4D5-38/1 together with DLS30A10 ES cells, in another differentiation, I assayed additionally DMK4A2-24/30 and DMK4D5-38/3. No substantial differences could be found between R Δ 800 and DLS30A10 ES cells. In all cases, *Airn* assayed with the 'Air middle' assay was localised 52-194 fold more in the nucleus than in the cytoplasm with respect to *Cyclophilin A*. *Igf2r* assayed with the 'Igf2r Ex48' assay showed a slightly higher enrichment in the nucleus compared to the cytoplasm with a cytoplasmic to nuclear ratio of 1:1.5 to 1:5.7. *45S-pre-rRNA* showed a 527-28924 fold enrichment in the nucleus compared to the cytoplasm whereas *Gapdh* was enriched 1.1-5.5 fold more in the cytoplasm compared to the nucleus (see Fig. 34). Therefore, deletion of the *Airn* repeats does not change the nuclear localisation of *Airn*.

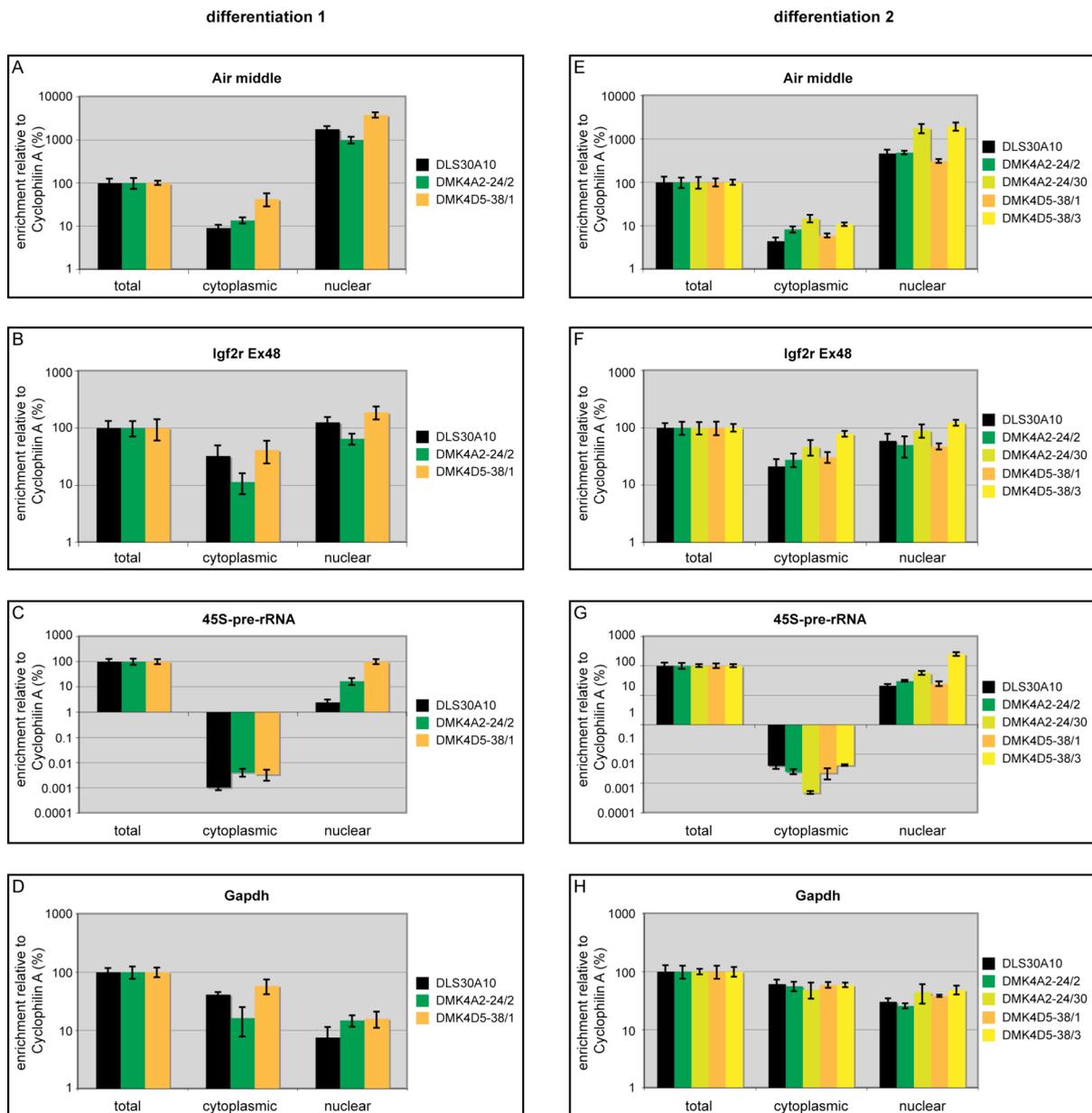


Fig. 34: Cytoplasmic-nuclear fractionation of DLS30A10 and DMK4 ES cells after 14 days of retinoic acid differentiation. The left column shows the first differentiation set, the second column the second differentiation set. Bars show relative RNA levels assayed for the indicated genes. All assays were normalised to *Cyclophilin A*. Levels in the total RNA fraction were set to 100%. Bars give the mean value, error bars the standard deviation of three technical replicates. (A), (E) *Airn* levels assayed using 'Air middle'. (B), (F) *Igf2r* assayed using 'Igf2r Ex48'. (C), (G) *45S-pre-rRNA*. (D), (H). *Gapdh* mRNA.

2.2.7 Can *Airn* in R Δ 800 ES cells still lead to imprinted expression of *Igf2r*?

Imprinted expression of *Igf2r* can be analysed in several ways. First, it can be assayed indirectly by the gain of DNA methylation on the paternal *Igf2r* promoter. In undifferentiated ES cells and pre-implantation embryos, *Igf2r* shows a low level of expression from both parental alleles and both parental promoters are unmethylated. Upon *in vitro* differentiation or after implantation, the maternal *Igf2r* allele is upregulated whereas the paternal one stays expressed at low levels, often referred to as being silenced (Latos et al., 2009; Lerchner and Barlow, 1997; Szabo and Mann, 1995; Terranova et al., 2008). Furthermore, the paternal *Igf2r* promoter gains DNA methylation, and therefore exhibits a so called somatic differentially methylated region (Latos et al., 2009; Stoger et al., 1993).

I assayed genomic DNA of DMK2 ES cells and wildtype ES cells at d0 and d14 of retinoic acid differentiation. The DNA was digested using EcoRI and the methylation-sensitive restriction enzyme NotI, the DNA blot was hybridised using probe EEi. As controls served wildtype tail DNA digested with EcoRI only or EcoRI in combination with NotI. Furthermore, also DNA from +/Thp and Thp/+ 13.5dpc embryos was used to verify the parental origin of the restriction fragments. +/Thp embryonic DNA showed a 4.0kb band only corresponding to an unmethylated maternal allele. In Thp/+ embryonic DNA both, a 4.0kb and a 5.0kb band are visible, as DNA methylation of the paternal *Igf2r* allele is not always complete at this stage. In wildtype ES cells at d0, only a 4.0kb band is present, correlating with both alleles being unmethylated. At d14 however, a 5.0kb and a 4.0kb are present, indicating a methylated paternal and an unmethylated maternal allele. All three DMK2 cell lines showed the same result as wildtype ES cells: Two unmethylated alleles at d0 and a methylated and a unmethylated allele at d14, indicating that during differentiation DMK2 ES cells gain imprinted expression of *Igf2r* (see Fig. 35).

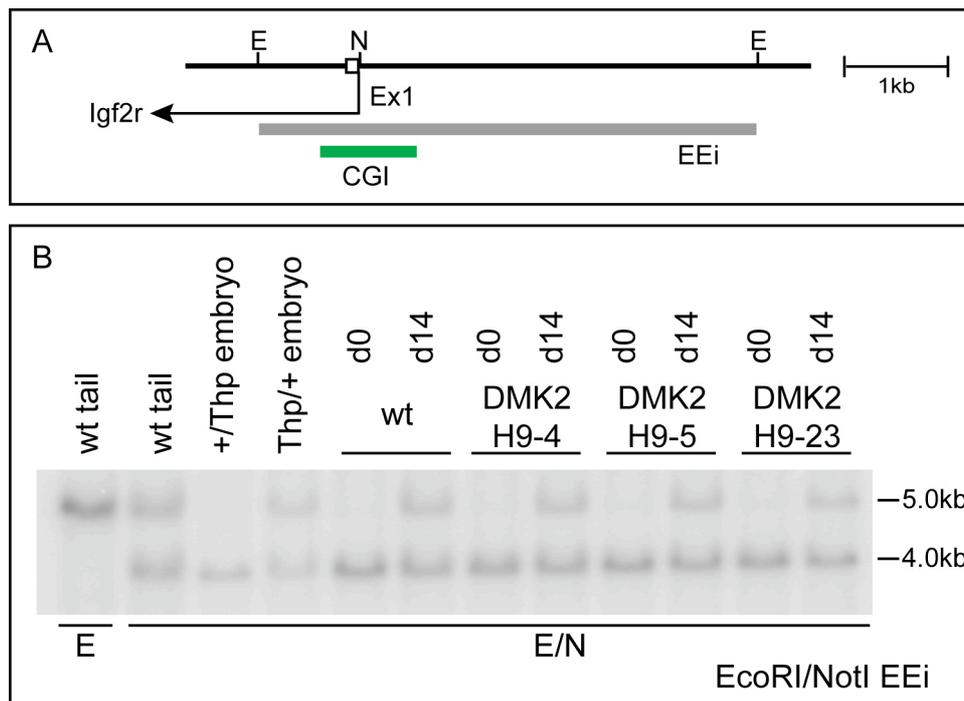


Fig. 35: Assay of the gain of DNA methylation on the paternal *Igf2r* promoter during ES cell differentiation. (A) Schematic overview of the examined region. Exon 1 of *Igf2r* is shown as a white box, transcriptional orientation of *Igf2r* is shown by an arrow. Given are only restriction enzymes necessary for the experiment (E: EcoRI, N: NotI). Probe EEi is shown as a grey bar, the *Igf2r* CGI is shown as a green bar. (B) DNA blot result of genomic DNA digested either with EcoRI alone or with EcoRI in combination with NotI and probed with EEi. Wildtype tail DNA represents an adult tissue and shows two bands with 5.0kb and 4.0kb in length. Please note that unmethylated alleles do not only give a 4.0kb band, but in addition also a 1.0kb band, which is not shown here, as it does not add any additional information. +/Thp 13.5dpc embryo DNA shows a 4.0kb band only correlating with an unmethylated maternal allele. Thp/+ 13.5dpc embryo DNA shows both, a 4.0kb and a 5.0kb band, as methylation of the paternal *Igf2r* promoter is not always complete at this embryonic stage. At d0, wildtype ES cells show only a 4.0kb band correlating with two unmethylated parental alleles, at d14 a 5.0kb band is gained correlating with the gain of DNA methylation on the paternal allele. DMK2H9-4, DMK2H9-5 and DMK2H9-23 ES cells also show only unmethylated *Igf2r* promoters at d0 and show one DNA methylated and one unmethylated allele at d14 of differentiation.

The same experiment was repeated with three differentiation sets of DMK4 ES cells. During the third differentiation, d5 in addition to d0 and d14 was analysed. As in DMK2 cells also DMK4 cells gain DNA methylation of the *Igf2r* promoter on one allele during ES cell differentiation. From the third differentiation set where d0, d5 and d14 were analysed, it can be seen, that the gain of DNA methylation is a gradual process with a lower level of DNA methylation at d5 and a higher level of DNA methylation at d14 (see Fig. 36). I provisionally conclude from these results that in RΔ800 ES cells *Airn* still leads to imprinted expression of *Igf2r*.

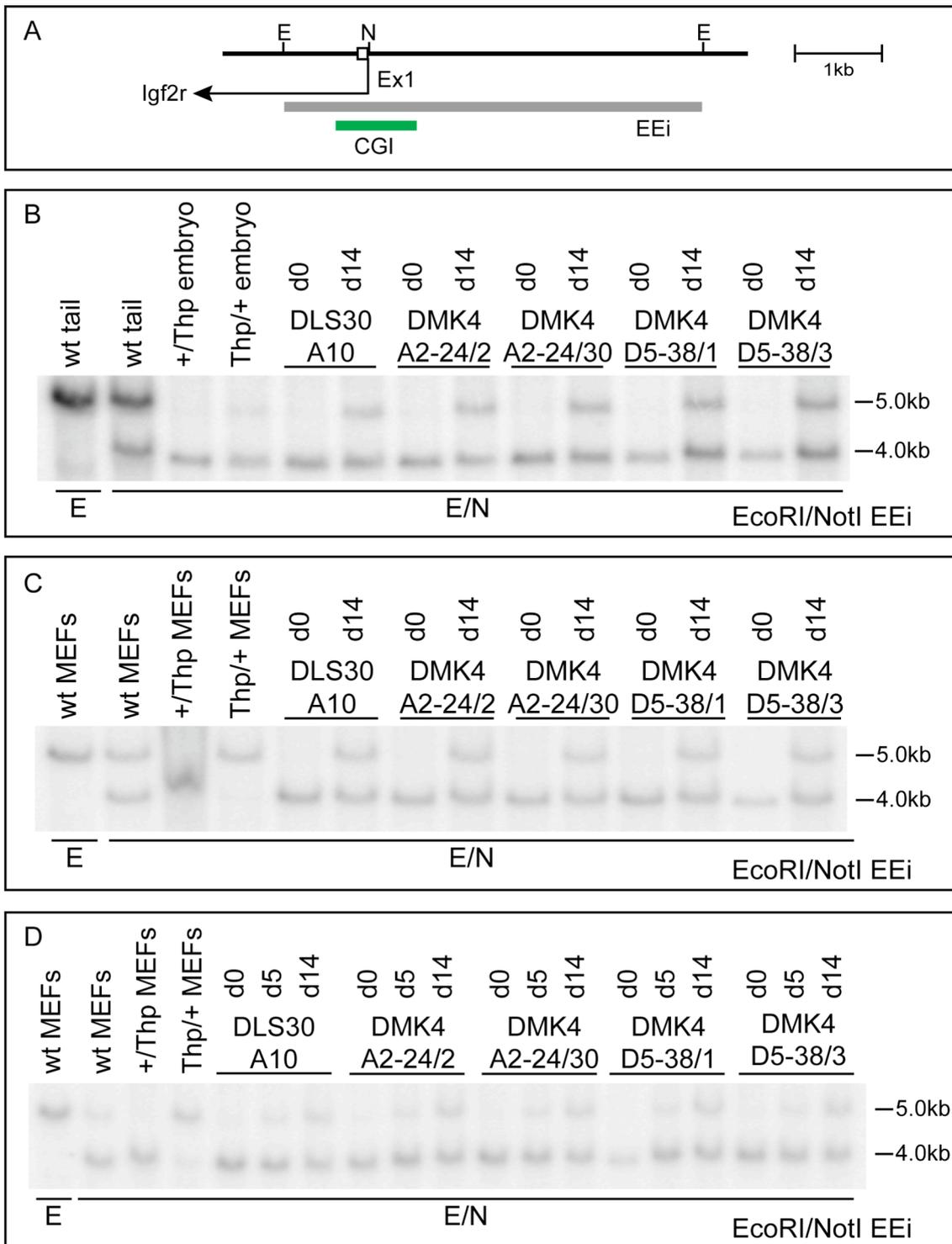


Fig. 36: As Fig. 35 but for DMK4 ES cells. (B)-(D) show three differentiation sets. In (D) additionally genomic DNA at d5 was analysed. Please note that DNA methylation of the paternal *Igf2r* promoter in Thp/+ 13.5dpc embryos and pMEFs is not always complete.

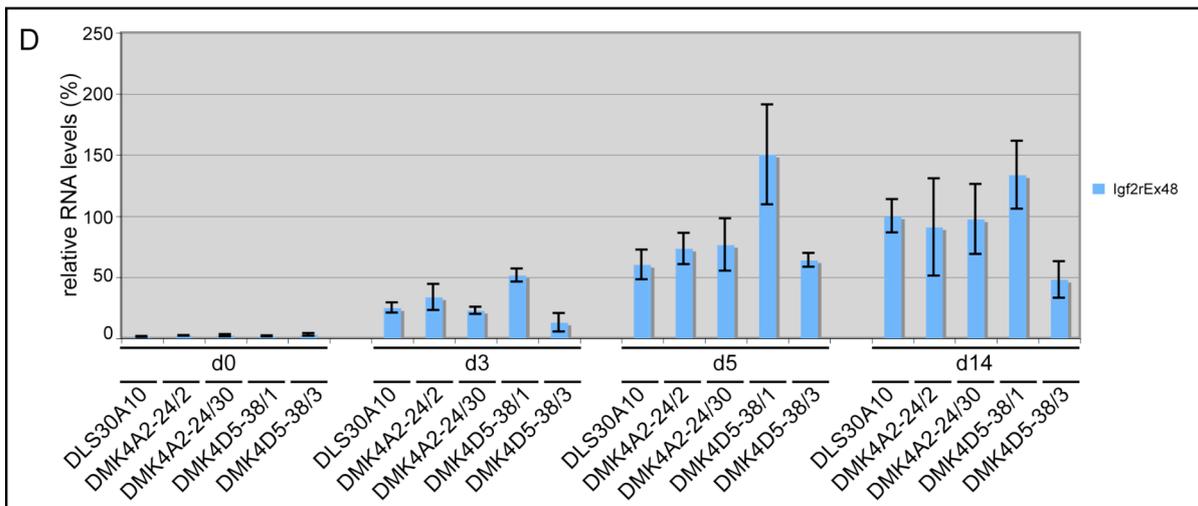
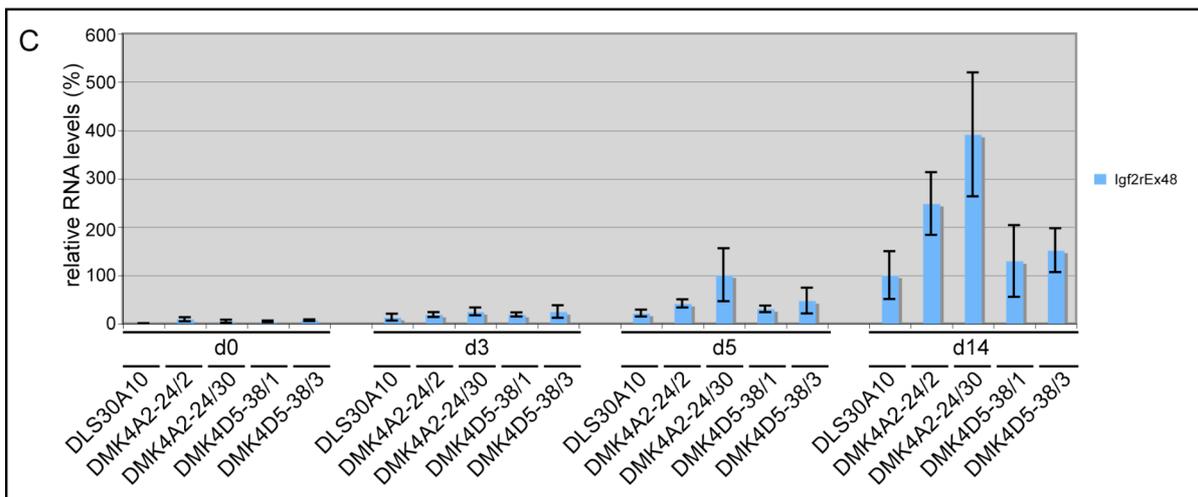
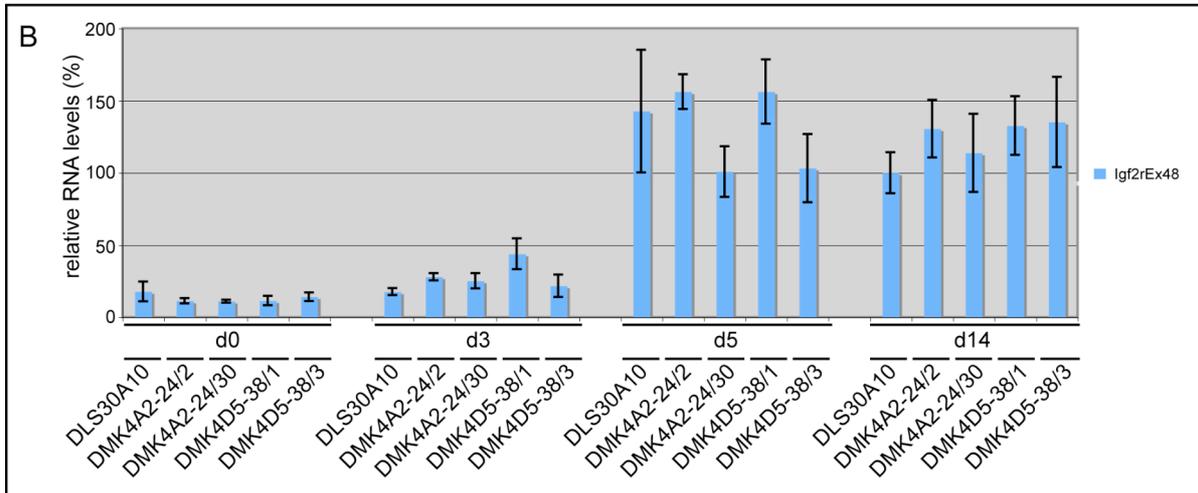


Fig. 37: Quantification of total *Igf2r* expression during ES cell differentiation by qPCR. (A) Shown is a schematic drawing of the used assay. The forward primer spans the junction between exon 47 and exon 48, the probe and the reverse primer are both localised within exon 48. (B)-(D) Three differentiation sets using DLS30A10 and DMK4 ES cells. *Igf2r* shows low steady state levels at d0 and is upregulated upon ES cell differentiation, in both DLS30A10 and DMK4 ES cells. Bars give the mean value, error bars the standard deviation of three technical replicates. No obvious or consistent differences between DLS30A10 and DMK4 ES cells can be observed.

Next, I directly examined steady-state levels of total *Igf2r* mRNA by qPCR at d0, d3, d5 and d14 of differentiation. I used DMK4 ES cells and DLS30A10 ES cells as control and performed three differentiation sets. At d0, *Igf2r* shows only a low expression level but gets upregulated progressively during ES cell differentiation, both in DLS30A10 and DMK4 ES cells. Total *Igf2r* levels showed a high variability between differentiation sets and in some cases also between co-differentiated ES cell clones. No obvious differences between DLS30A10 and DMK4 ES cells could be observed from these experiments, indicating that the deletion of the tandem direct repeats did not affect overall *Igf2r* expression levels (see Fig. 37).

To answer the question, if the deletion of the tandem direct repeats influenced the ability of *Airn* to prevent upregulation of the paternal allele during differentiation, I made use of the introduced SNP in Ex12 of *Igf2r* which is present in both, DLS30A10 and DMK4 ES cells. I performed an RT-PCR with primers in *Igf2r* exon 11 and exon 14, flanking the SNP in exon 12 which destroyed a PstI site on the maternal allele. The PCR fragments were subjected to a restriction digest with PstI, resulting in an uncut band for the maternal allele and two smaller fragments for the paternal allele. RNA at d0, d5 and d14 of retinoic acid differentiation was analysed from three differentiation sets, the results are shown in Fig. 38. All cell lines show in all replicates at d0 both, a large fragment corresponding to the maternal allele, and two smaller bands corresponding to the paternal allele. During differentiation, in the first and third differentiation set, for DLS30A10 cells, the two smaller bands are very faint but visible at d5 and completely gone at d14. All DMK4 lines however showed at d5 still two small bands and they were still faint but visible at d14. In the second differentiation set, also in DLS30A10 ES cells at d5 the two small fragments were still visible, although fainter than in the DMK4 cells. These results were a first indication, that in cells where the tandem direct repeats have been deleted from the

paternal allele, *Airn* fails to prevent the upregulation of *Igf2r* from the paternal allele to the same extent as in wildtype cells. However, as this assay is based on a non-quantitative RT-PCR, the results have to be treated with care.

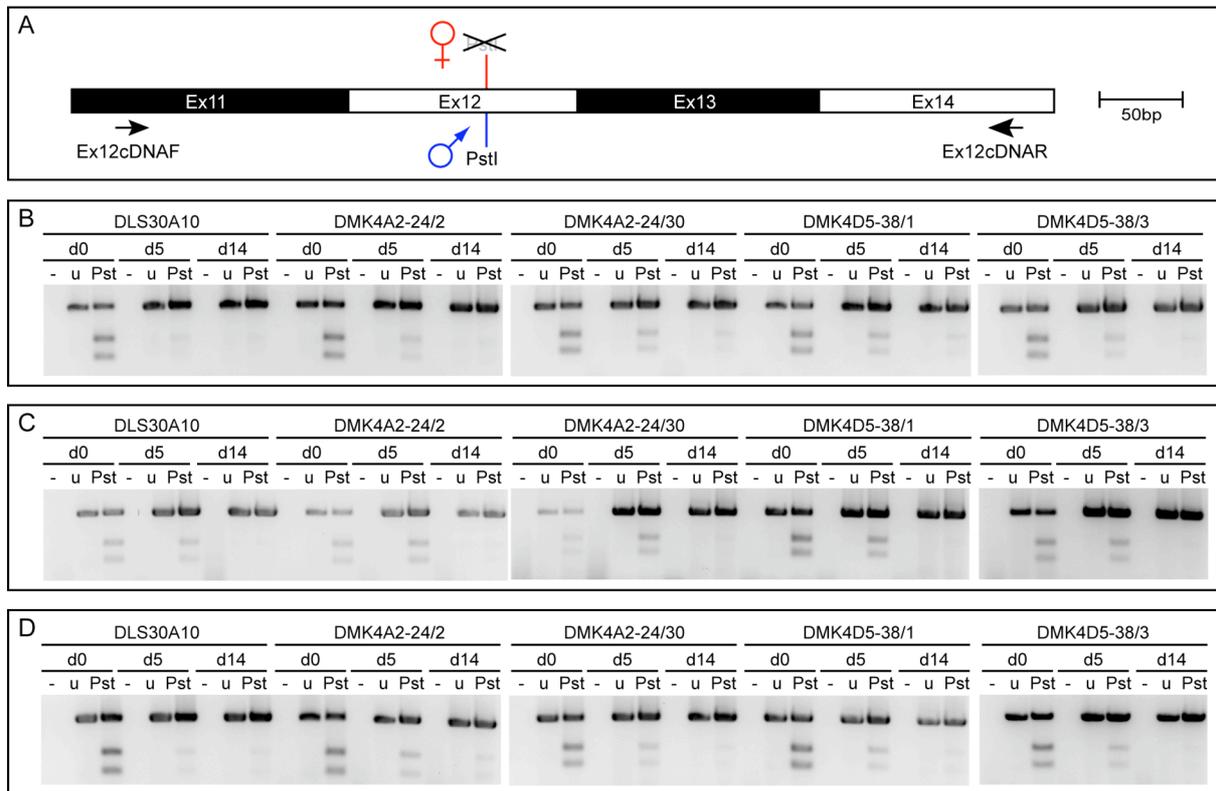


Fig. 38: RT-PCR to detect allelic expression of *Igf2r*. (A) Schematic overview of the assay. Shown are the primers (Ex12cDNAF, -R) located in exons 11 and 14 of *Igf2r* respectively as well as the SNP in exon 12 which destroyed a PstI site on the maternal allele. (B)-(D) Three differentiation sets of DLS30A10 and DMK4 ES cells. RNA was analysed at d0, d5 and d14 of differentiation. -: -RT reaction. u: uncut PCR fragment. Pst: PCR fragment subjected to PstI digestion.

A qPCR assay which was able to discriminate between the allelic *Igf2r* expression levels in a quantitative way was established earlier (Latos et al., 2009). This assay uses a reverse primer common to both alleles but a forward primer specific for either the maternal or the paternal SNP in exon 12 of *Igf2r* (Latos et al., 2009). However, in my hands this assay was not very specific as I saw 9.5% crosshybridisation of the primer for the mutant (maternal) allele with a plasmid carrying the wildtype (paternal) allele. In the reverse direction, the crosshybridisation of the wildtype primer with the mutant allele was at 0.5% and not as severe (see Fig. 39B). To make the assay more specific, I introduced the following three changes. A lowering of the MgCl₂

concentration from 9mM to 5mM and increasing the annealing/extension temperature from 64°C to 65°C resulted in a reduction of crosshybridisation to 5.3% and 0.1% respectively (see Fig. 39C). Furthermore, I tested two modified forward primers which have in their sequence body either a T→A (WtSeFTA, MutSeFTA) or a C→G (WtSeFCG, MutSeFCG) conversion. This mismatch together with the additional mismatch at the 3' end should destabilise the primer binding to the wrong allele to an extent, that only the correct allele can be recognised. Using the original conditions (9mM MgCl₂ and 64°C), crosshybridisation indeed was reduced to 2.6% (MutFCG on wildtype plasmid), 1.8% (MutFTA on wildtype plasmid) and 0.3% (WtFCG on mutant plasmid), 0.2% (WtFTA on mutant plasmid) (see Fig. 39B). Using the changed conditions (5mM MgCl₂ and 65°C) reduced the crosshybridisation further to 0.4% (MutFCG on wildtype plasmid), 0.5% (MutFTA on wildtype plasmid) and 0.0% (WtFCG on mutant plasmid), 0.0% (WtFTA on mutant plasmid) (see Fig. 39C).

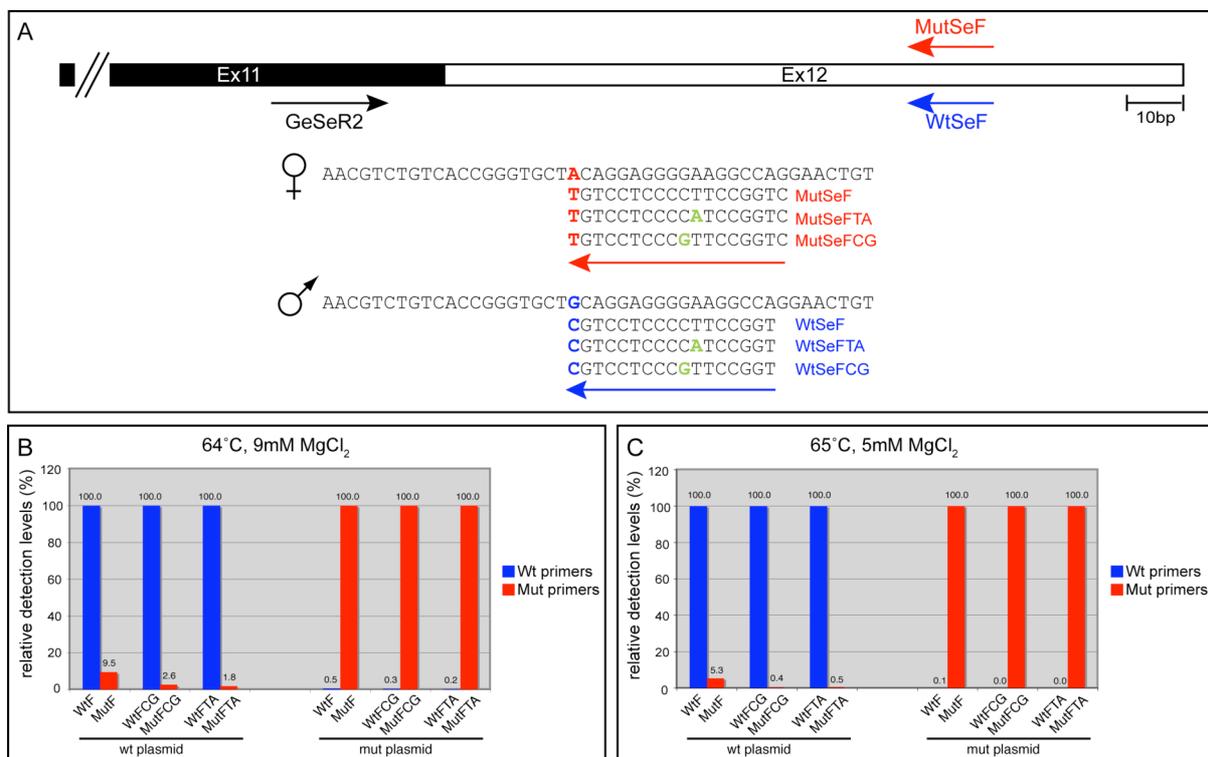


Fig. 39: (A) Schematic overview of the *Igf2r* exon 12 qPCR SNP assay. Shown are exons 11 and 12 of *Igf2r*, the common reverse primer (GeSeR2) and forward primers specific for either the mutant maternal allele (MutSeF) or the wildtype paternal allele (WtSeF). Below, the sequence of the maternal and the paternal alleles are shown with the respective mutant and wildtype forward primers used in this experiment. Bases specific for the maternal or paternal allele are marked in red or blue

respectively. Additionally introduced mismatches are marked in green. (B) qPCR using the different primer combinations on plasmids containing the wildtype or the mutant cDNA. MgCl₂ concentration was 9mM, annealing/extension temperature was 64°C. Detection levels of the correct primer were set to 100%, detection levels due to crosshybridisation of the wrong primers are displayed relative to the 100% detection levels of the correct primer. (C) As in (B) but with 5mM MgCl₂ and annealing/extension temperature of 65°C.

As the C→G conversion in the forward primers gave less crosshybridisations than the T→A conversion, I decided to use MutSeFCG and WtSeFCG in combination with GeSeR2 for further experiments. I analysed total RNA from three differentiation sets of DLS30A10 and DMK4 using this SYBR green assay at d0, d3, d5 and d14 of differentiation (see Fig. 40). The ratio of the maternal/paternal allele in DLS30A10 cells at d0 was set to 1, as at this stage *Igf2r* does not show imprinted expression but is expressed biallelically (Braidotti et al., 2004; Latos et al., 2009; Wang et al., 1994). In DLS30A10 cells, in all three differentiation sets, the ratio increased successively during differentiation, showing that the maternal allele is upregulated up to 9-22 fold compared to the paternal allele. In DMK4 ES cells, the ratio also increased. In two differentiation sets at d3, in all three differentiation sets at d5 and in two differentiation sets at d14 however the ratios of DMK4 cells did not reach the same height as the ratio in DLS30A10 cells. At d5 for example, the ratio of maternal to paternal expression was only 30% to 50% of the ratio observed in DLS30A10 cells. This indicates, that in the cells carrying a deletion of the tandem direct repeats, the prevention of upregulation of *Igf2r* from the paternal allele is compromised or delayed (see Fig. 40).

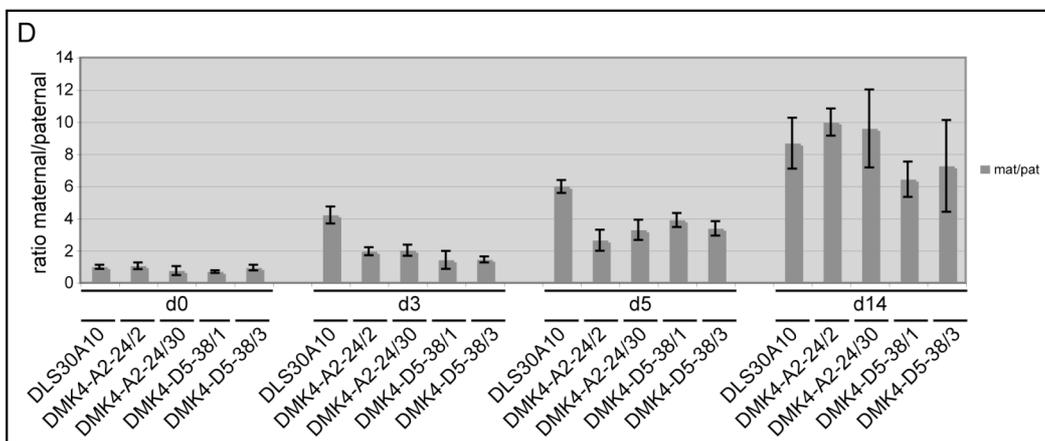
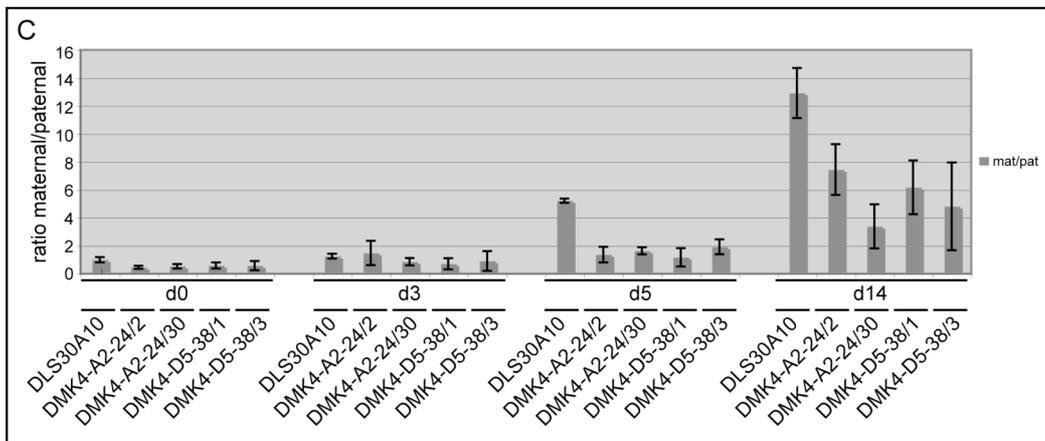
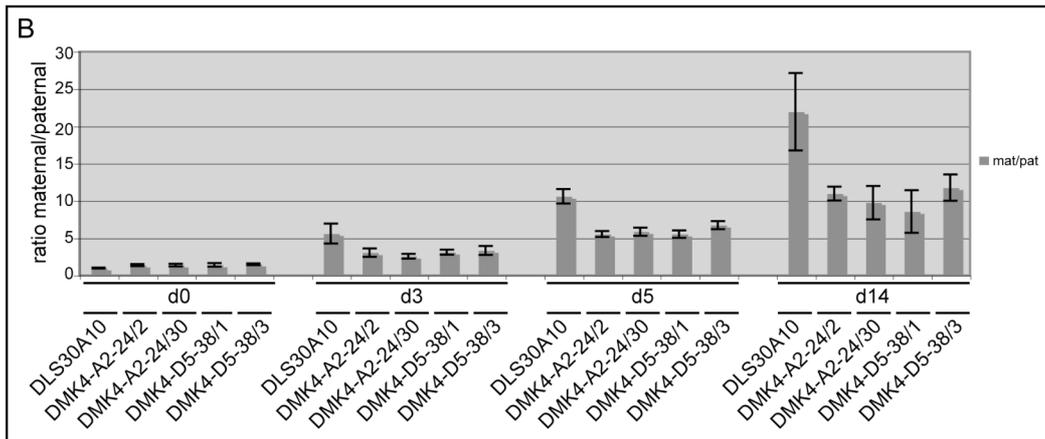
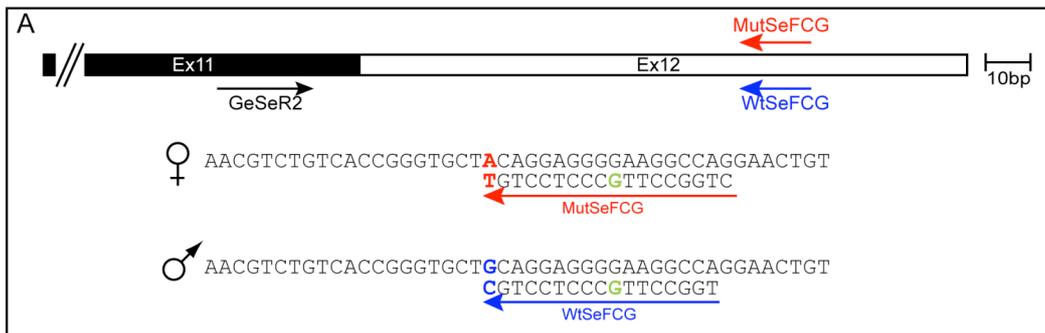


Fig. 40: qPCR analysis of the parental *Igf2r* expression in DLS30A10 and DMK4 ES cells. (A) Schematic overview of the used assay. See Fig. 39 for further details. (B)-(D) Three differentiation sets of DLS30A10 and DMK4 ES cells. Shown is the ratio of the maternal to the paternal allele. In DLS30A10 cells at d0 the ratio was set to 1. Bars give the mean value, error bars the standard deviation of three technical replicates.

2.2.8 Generation of CpG island deletion ES cells

It was shown previously that for the expression of *Airn*, the promoter region upstream of the transcriptional start site is crucial (Stricker et al., 2008). However, it is not known, if also the CpG island that is localised downstream of the transcriptional start of *Airn* is necessary for its expression and its transcriptional features. To test this, I generated ES cells carrying a deletion of the CpG island by homologous recombination. The pCpG Δ targeting vector contains a deletion which starts 176bp downstream of the main transcriptional start site, leaving the MluI site diagnostic for DNA methylation of the CpG island intact. The deletion extends over 1130bp, ending at the same position as the R Δ 800 deletion (see Fig. 20). In addition, the targeting vector contains a selection cassette with a neomycin resistance gene driven by a *Pgk1* promoter and stopped by a *Pgk1* polyadenylation signal, which was inserted into the NheI site in intron 3 of *Igf2r*. Although 107bp of the CpG island as defined by the UCSC genome browser stay behind, this remaining region is too small to reach the criteria for a CpG island (see Fig. 41).

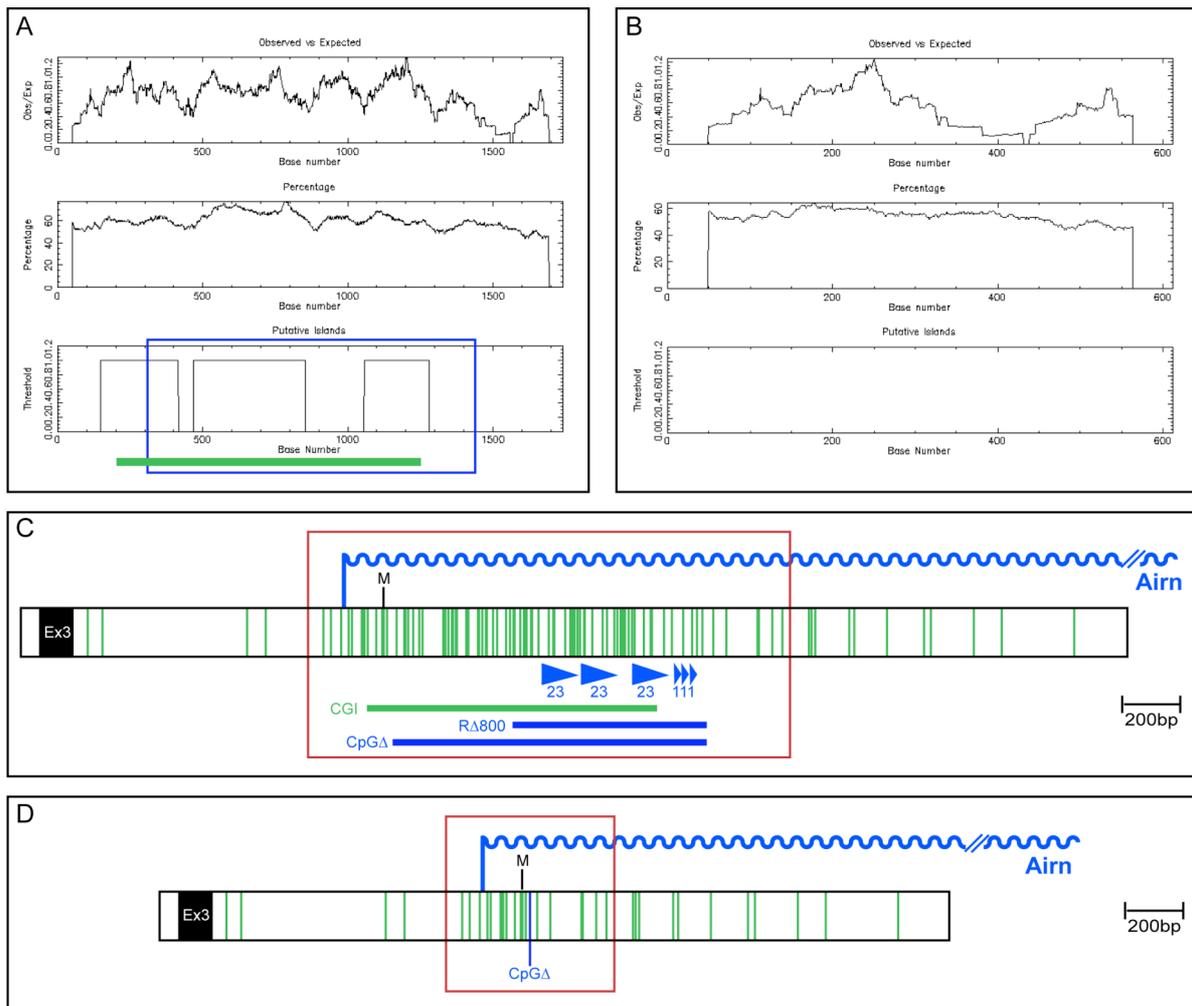


Fig. 41: CpG distribution before and after the CpG Δ . (A) CpG plot of 12934107-12935849bp on chr. 17 (UCSC genome browser, assembly July 2007) before the CpG Δ . Shown are the ratio of the observed to the expected number of CpG dinucleotides (top), the percentage of G and C nucleotides (middle), and CpG islands reaching the default parameters (ratio observed/expected greater than 0.6; percentage of G and C greater than 50%; bottom). Three regions reach the criteria for a CpG island. The green bar marks the region defined by the UCSC genome browser as the CpG island. The blue frame marks the CpG Δ . (B) As in (A) but after the CpG Δ . No region reaches the criteria for a CpG island. (C) Schematic representation of CpG abundance around the transcriptional start site of *Airn*. Each CpG is shown by a green vertical bar. *Airn* is shown as a blue wavy line. Exon 3 of *Igf2r* is shown as a black box. The tandem direct repeats are shown as blue triangles. Regions deleted in R Δ 800 and CpG Δ are shown as blue horizontal bars. The CGI as defined by the UCSC genome browser is shown as a green horizontal bar. The region framed in red is the region shown in (A). (D) As in (C) but after the CpG Δ . The region framed in red is the region shown in (B).

I targeted DLS30A10 ES cells using the pCpG Δ targeting vector, homologous recombinants were identified by DNA blotting. Genomic DNA was digested using EcoRI and probed using AirT, resulting in a 6.2kb wildtype fragment and a 5.1kb CpG Δ fragment. Unusually in comparison to previous experiments, the first two targetings where I picked 192 and 384 neomycin-resistant clones respectively did

not result in any homologous recombinants. Therefore I repeated the targeting, this time picking 576 neomycin-resistant clones and obtained two homologously targeted clones (DMK14C12, DMK14D8), which were used for further experiments (see Table 8, Fig. 42).

The selection cassette was removed by transient transfection with a Cre-recombinase expressing plasmid (pCre, gift from Anton Wutz, Wellcome Trust Centre for Stem Cell Research, Cambridge, UK). Removal of the selection cassette from DMK14C12 and DMK14D8 ES cells resulted in 10 and 15 clones, respectively, out of 96 picked ones with a wildtype and a CpG Δ allele. One clone each (DMK15C12-12, DMK15D8-14) was subjected to subcloning. All analysed subclones showed a wildtype and a CpG Δ allele, two subclones each (DMK15C12-12/1, DMK15C12-12/24, DMK15D8-14/25, DMK15D8-14/36) were used for further experiments (see Table 8, Fig. 42).

name	parental cell line	targeting vector	number of picked clones	number of recovered DNAs	number of positive clones	names of clones used further
DMK12	DLS30A10	pCpG Δ	192	178	0	-
DMK13	DLS30A10	pCpG Δ	384	363	0	-
DMK14	DLS30A10	pCpG Δ	576	550	2	DMK14C12 DMK14D8
DMK15C12	DMK14C12	pCre	96	93	10	DMK15C12-12
subclone	DMK15C12-12	-	96	96	96	DMK15C12-12/1 DMK15C12-12/24
DMK15D8	DMK14D8	pCre	96	89	15	DMK15D8-14
subclone	DMK15D8-1	-	96	96	96	DMK15D8-14/25 DMK15D8-14/36

Table 8: ES cell lines created having a 1.1kb deletion of the *Airn* CGI. Details as in Table 6.

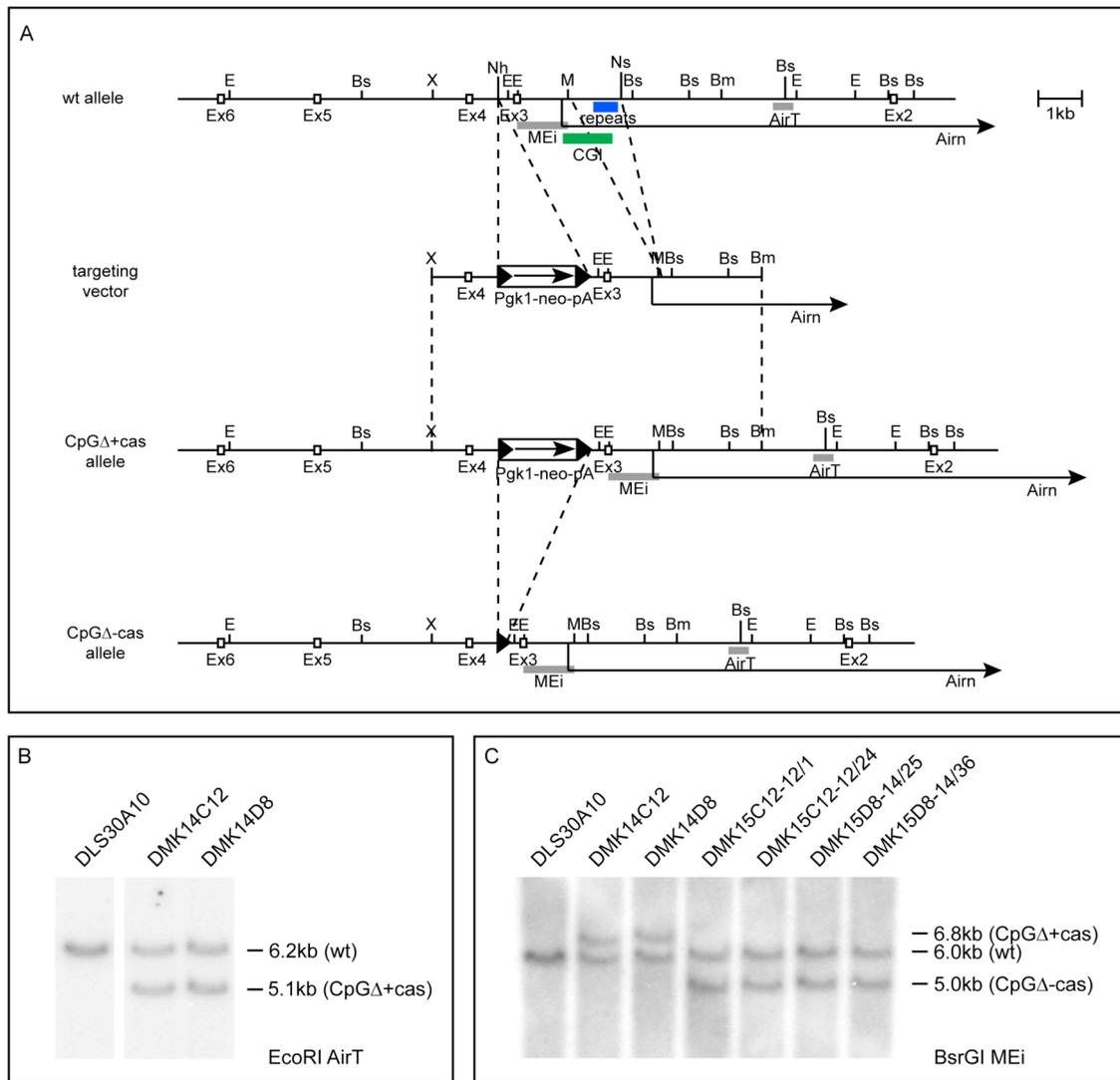


Fig. 42: Generation of CpG Δ ES cells. (A) Targeting strategy for generating CpG Δ ES cells. The CpG island downstream of the *Airn* transcriptional start site has been deleted in the targeting vector, from 29bp downstream of the MluI site until the NsiI site. A floxed selection cassette with a *Pgk1* promoter driving a neomycin resistance gene stopped by a *Pgk1* polyadenylation signal (pA) was inserted into the NheI site in intron 3 of *Igf2r*. Homologous recombination created the CpG Δ +cas allele. Transient transfection with a Cre-recombinase expressing plasmid resulted in recombination of the two loxP sites and deletion of the selection cassette, leaving behind a single loxP site in intron 3 of *Igf2r* and creating the CpG Δ -cas allele. Only restriction enzymes crucial for targeting or genotyping are given. E: EcoRI, Bs: BsrGI, X: XbaI, Nh: NheI, ScII: SacII, Ns: NsiI, Bm: BmiI. (B) Genotyping by DNA blotting of ES cells carrying a CpG Δ +cas allele. Genomic DNA was digested using EcoRI and hybridised using probe AirT. Shown are the targeted ES cells (DMK14C12, DMK14D8) and their parental ES cell line (DLS30A10). The wildtype band is 6.2kb in length, the targeted band 5.1kb. (C) Genotyping by DNA blotting of ES cells carrying a CpG Δ +cas allele (DMK14C12, DMK14D8) or an CpG Δ -cas allele (DMK15C12-12/1, DMK15C12-12/24, DMK15D8-14/25, DMK15D8-14/36) and the original parental ES cell line (DLS30A10). Genomic DNA was digested using BsrGI and hybridised with probe MEI. A wildtype band is 6.0kb in length, a CpG Δ +cas allele gives a 6.8kb long band, an CpG Δ -cas allele a 5.0kb long band. Further details as in Fig. 21.

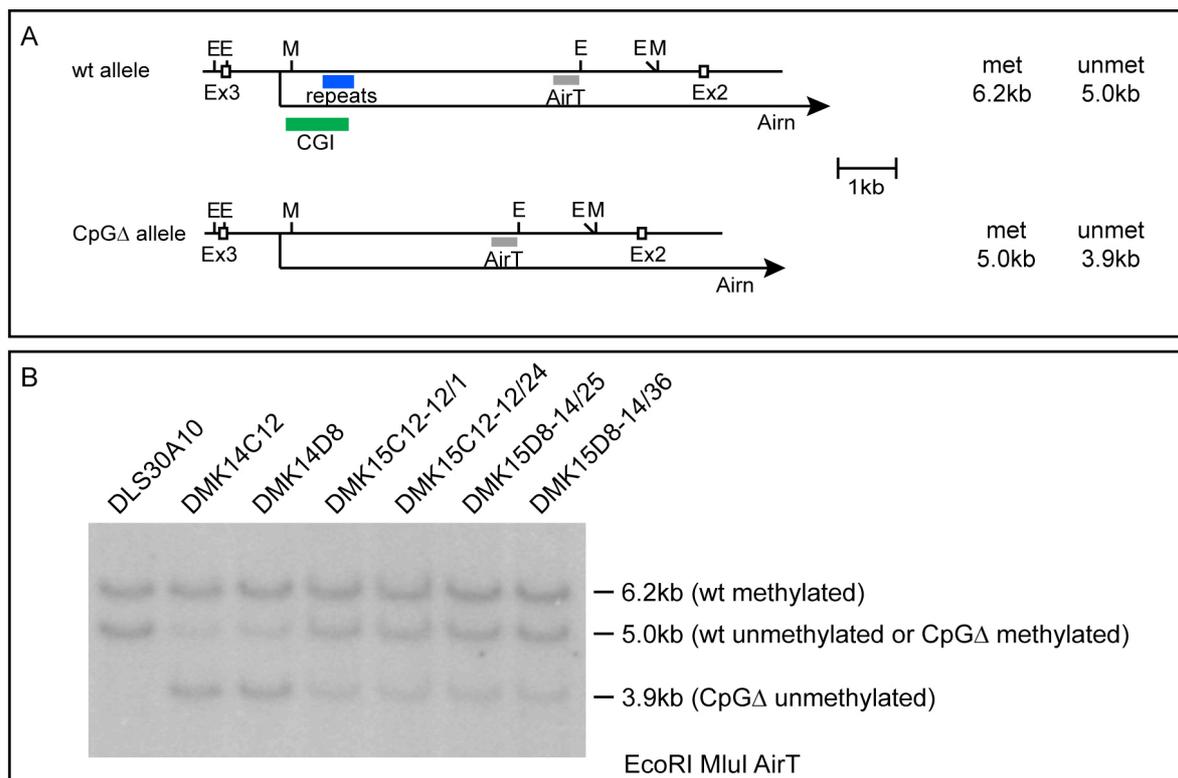


Fig. 43: Analysis of parental targeting of CpGΔ ES cells. (A) Overview of the genomic locus and DNA blot strategy. Details as in Fig. 42. met: methylated, unmet: unmethylated (B) DNA blots of the parental ES cells (DLS30A10) and ES cells carrying a CpGΔ allele (DMK14C12, DMK14D8, DMK15C12-12/1, DMK15C12-12/24, DMK15D8-14/25, DMK15D8-14/36). Genomic DNA was digested using EcoRI and MluI and hybridised using probe AirT. Homologous recombination occurred on the paternal allele and the targeted allele seems to gain DNA methylation over passaging.

To assay which parental allele was targeted, I digested genomic DNA of targeted cells with and without selection cassette (DMK14C12, DMK14D8, DMK15C12-12/1, DMK15C12-12/24, DMK15D8-14/25, DMK15D8-14/36) and the original parental cell line (DLS30A10) with EcoRI and MluI. DNA blots were hybridised with probe AirT. A wildtype methylated allele will be detected as a 6.2kb fragment, a wildtype unmethylated or a targeted methylated allele as a 5.0kb fragment and a targeted unmethylated allele will result in a 3.9kb fragment. Both, DMK14C12 and DMK14D8 ES cell lines as well as their cassette-removed subclones showed a 6.2kb, a 5.0kb and a 3.9kb long fragment. DMK14C12 and DMK14D8 showed strong 6.2kb and 3.9kb fragments and a weak 5.0kb fragment. DMK15C12-12/1, DMK15C12-12/24, DMK15D8-14/25 and DMK15D8-14/36 showed strong 6.2kb and 5.0kb fragments and a weaker 3.9kb fragment. This indicates, that it was the paternal unmethylated allele, where homologous recombination happened, however, loss of the CpG island

seems to lead to a gain of DNA methylation *in cis* with an effect increasing over passaging, as the 'younger' DMK14 ES cells show a stronger 3.9kb and a weaker 5.0kb band compared to 'older' DMK15 ES cells (see Fig. 43).

2.2.9 Does deletion of the *Airn* CGI lead to an additional gain of DNA methylation on the paternal ICE during *in vitro* differentiation?

As described above, deletion of the CpG island led to a gain of DNA methylation on the paternal allele *in cis* during culturing (see Fig. 43). To see, if this effect is increased during *in vitro* differentiation of ES cells, I analysed genomic DNA of CpGΔ ES cells (DMK15) and control ES cells (DLS30A10) in the undifferentiated and retinoic acid treatment-induced differentiated state. In the first differentiation set, genomic DNA after 14 days (d14) of differentiation was analysed, for the second and the third differentiation set, additionally genomic DNA after 5 days (d5) of differentiation was analysed. DNA was digested with EcoRI and MluI, the blot was hybridised with probe AirT (first differentiation set) or MEi (second and third differentiation set). As additional controls, I used genomic DNA of wildtype embryos digested either with EcoRI alone or with EcoRI plus MluI, as well as +/-Thp and Thp/+ 13.5dpc embryos to be able to confirm the parental origin of the detected restriction fragments. The DNA blots show, that the MluI site on the paternal allele in CpGΔ ES cells is largely methylated but this does not change further during 14 days of *in vitro* differentiation as the intensities of the unmethylated 3.9kb band (see Fig. 44B) or the 1.1kb band (see Fig. 44C-D) remain constant.

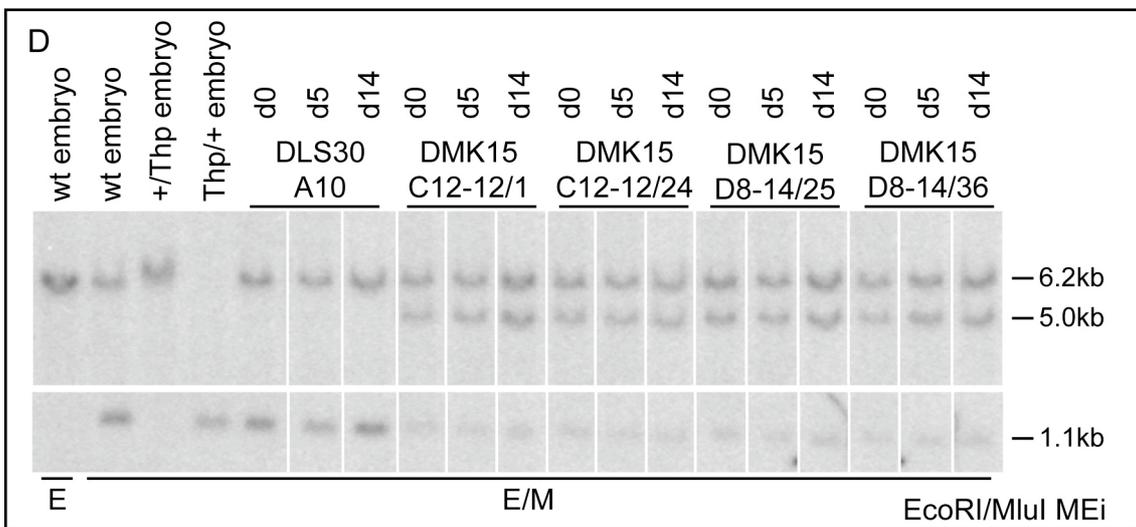
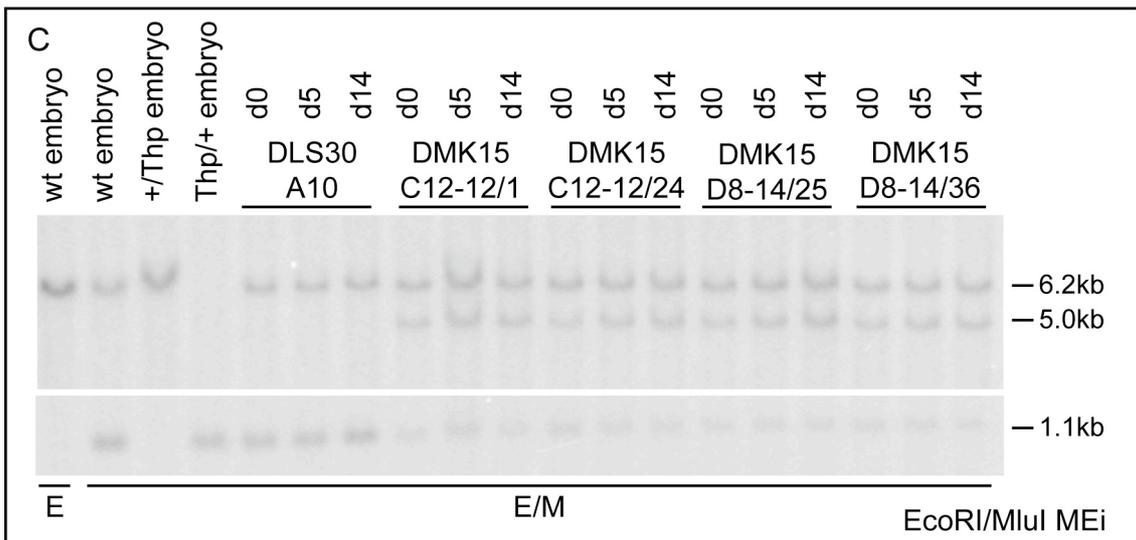
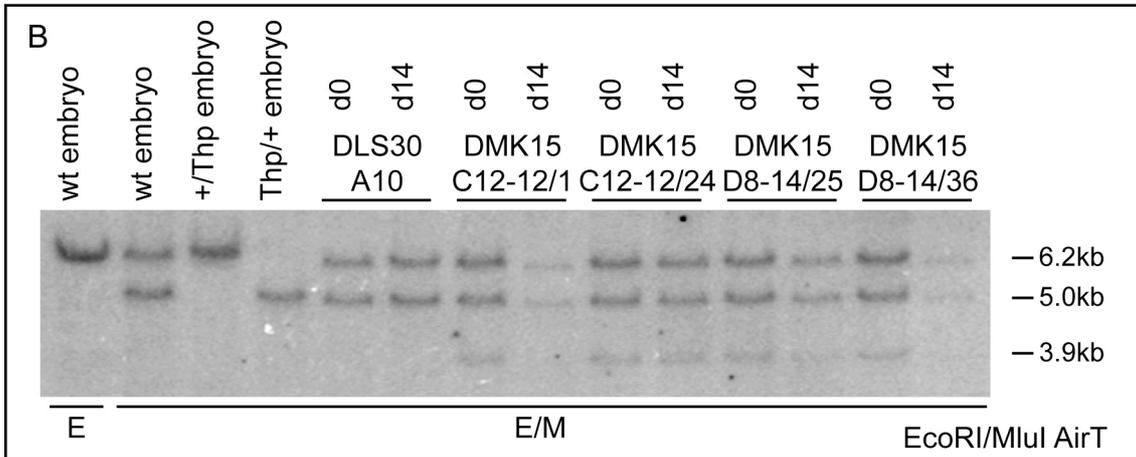
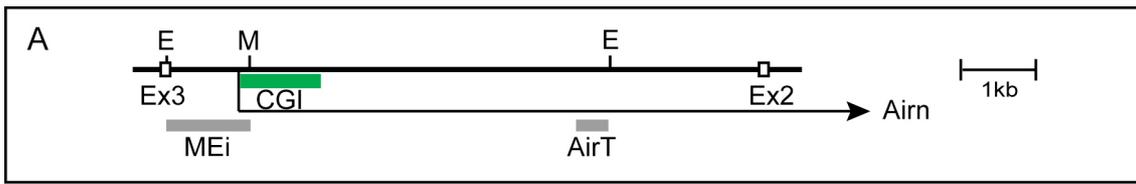


Fig. 44: Analysis of DNA methylation on the ICE in DLS30A10 and DMK15 ES cells. Details as in Fig. 27. (B)-(D) show the results of three independent differentiation sets. The 6.2kb fragment correlates with a methylated wildtype (maternal) allele. After hybridisation with AirT, the 5.0kb band correlates with either a methylated targeted (paternal) or an unmethylated wildtype (paternal) allele, the 3.9kb band correlates with an unmethylated targeted (paternal) allele. After hybridisation with MEi, the 5.0kb fragment correlates with a methylated targeted (paternal) allele, the 1.1kb fragment with an unmethylated wildtype or targeted (paternal) allele. The DNA blots show, that the level of DNA methylation on the paternal CpGΔ allele at the diagnostic MluI site does not change during *in vitro* differentiation.

2.2.10 Does deletion of the *Airn* CGI lead to changes in *Airn* expression?

Next I assayed, if the deletion of the *Airn* CpG island leads to changes in *Airn* expression. For this, I isolated total RNA from undifferentiated ES cells (d0) and ES cells differentiated for 5 days (d5) and 14 days (d14) by retinoic acid treatment and assayed steady-state levels of *Airn* in wildtype (DLS30A10) and CpGΔ (DMK15) ES cells during three differentiation sets by qPCR (see Fig. 45). I used the 'START', 'RP11', 'Air middle' and 'Air end' assays described above and in addition the 'AirT3' assay, which is localised 4kb downstream of the *Airn* transcriptional start. At d0, in both, DLS30A10 and DMK15 ES cells, all assays showed only 0-10% of *Airn* levels compared to *Airn* at d14 in DLS30A10 ES cells, indicating that the cells were largely undifferentiated. In DLS30A10, *Airn* showed upregulation at d5 and at d14. At d14, *Airn* levels were detected as followed: In all four DMK15 clones, the 'START' assay showed 10-40% of *Airn* compared to wildtype *Airn*. 'RP11' showed 20-80% in DMK15 clones compared to DLS30A10. 'AirT3' showed approximately 10% in DMK15 compared to DLS30A10. 'Air middle' showed 2-15% in DMK15 compared to DLS30A10. 'Air end' showed 0-0.8% in DMK15 compared to DLS30A10. This shows, that upon deletion of the CpG island, *Airn* levels are largely reduced 3' of the deletion, but 5' of the deletion still a substantial amount of *Airn* can be detected.

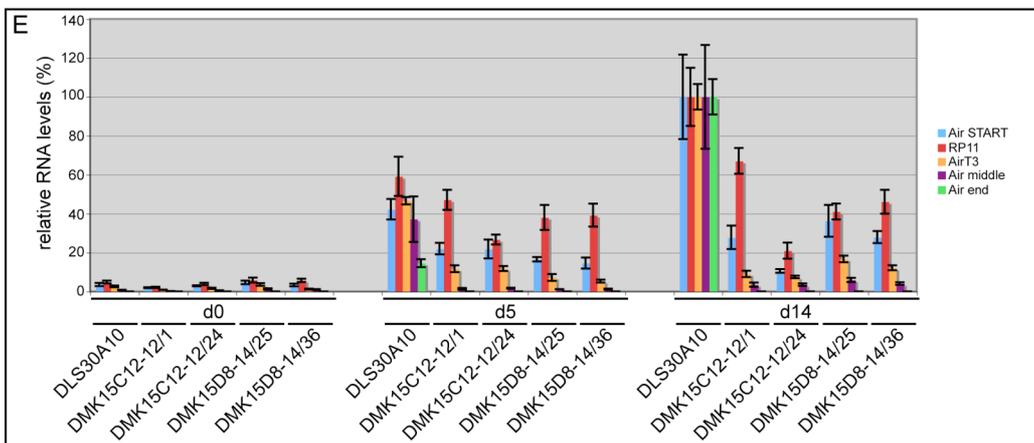
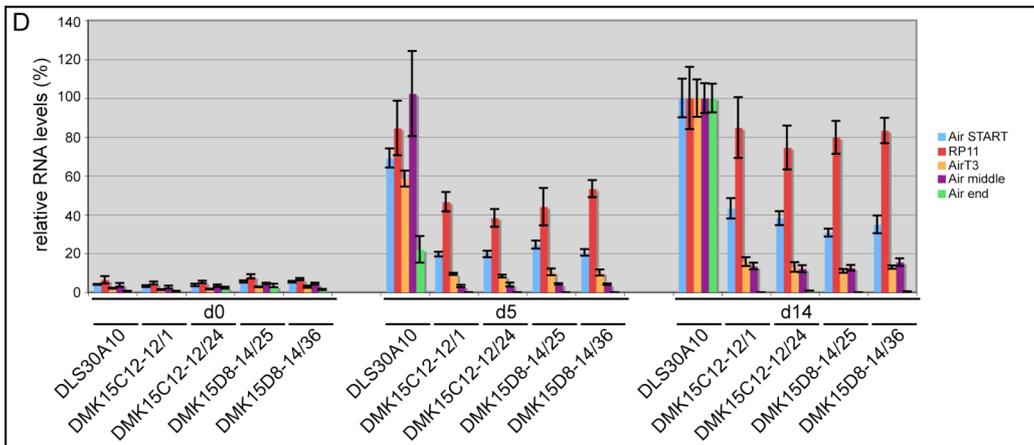
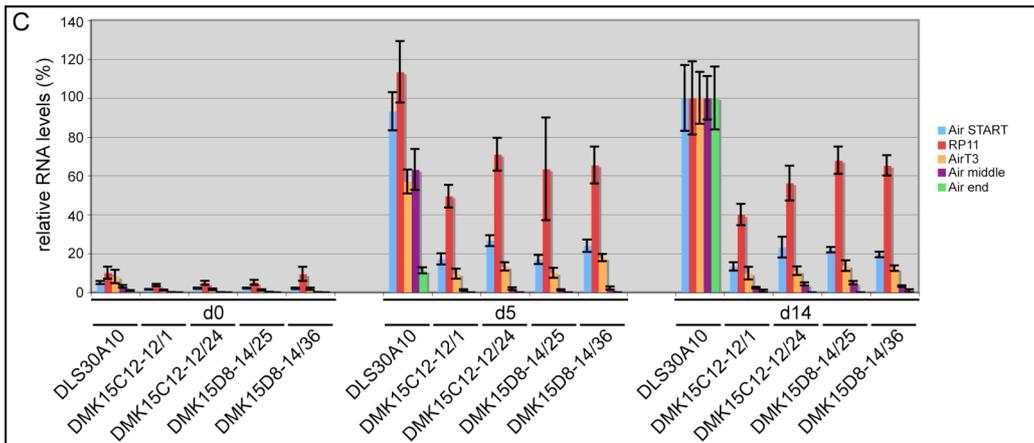
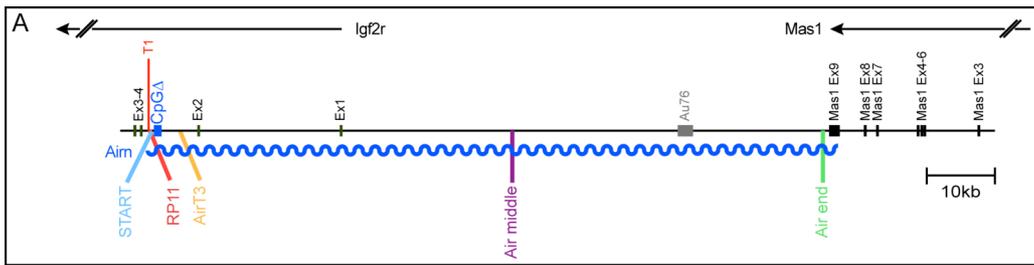


Fig. 45: Measurement of steady-state levels of *Airn* during ES cell differentiation using qPCR in DLS30A10 and DMK15 ES cells. Details as in Fig. 30. (B) Schematic overview of the 5' region of *Airn*. T1, T2, T3 show the three transcriptional start sites identified for *Airn* (Seidl et al., 2006). The 53bp donor is the splice donor used for all known *Airn* splice variants (Seidl et al., 2006). The CpGΔ is shown as a dark blue bar. The forward (START-F) and reverse (START-R) primers of the 'START' assay are shown as light blue arrows. Note that the 5' end of START-F extends 5' of T1. Primers (FP1, RP11) and probe (AS) for the 'RP11' assay are shown as red arrows/line. (C)-(E) show three differentiation sets.

2.2.11 Do CpGΔ cells show imprinted expression of *Igf2r*?

To answer the question, if the CpGΔ ES cells still show imprinted expression of *Igf2r*, I first analysed the gain of DNA methylation on the sDMR on the *Igf2r* promoter. I assayed genomic DNA of DMK15 and DLS30A10 ES cells at d0 and d14 for three differentiation sets and for two differentiation sets I also assayed genomic DNA at d5. The DNA was digested using EcoRI and NotI and hybridised using probe EEi. As controls I used 13.5dpc wildtype, Thp/+ and +/Thp embryonic DNA (see Fig. 46). Wildtype embryonic DNA showed two bands, a 4.0kb one corresponding to an unmethylated maternal and a 5.0kb band corresponding to a methylated paternal allele. In all three differentiation sets, DLS30A10 ES cells showed a successive gain of the 5kb fragment corresponding to a gain of DNA methylation indicative of a gain of imprinted expression of *Igf2r* during differentiation. In DMK15 ES cells, at all differentiation stages, a strong 4.0kb band was visible. In the second and third differentiation set, at d0, additionally a faint 5.0kb band correlating to a methylated allele was visible, which likely derives from feeder cells still present at d0, but this band disappeared upon differentiation. This indicates that in differentiated DMK15 cells *Igf2r* does not gain a differential DNA methylation on the *Igf2r* promoter, indicating a lack of imprinted expression.

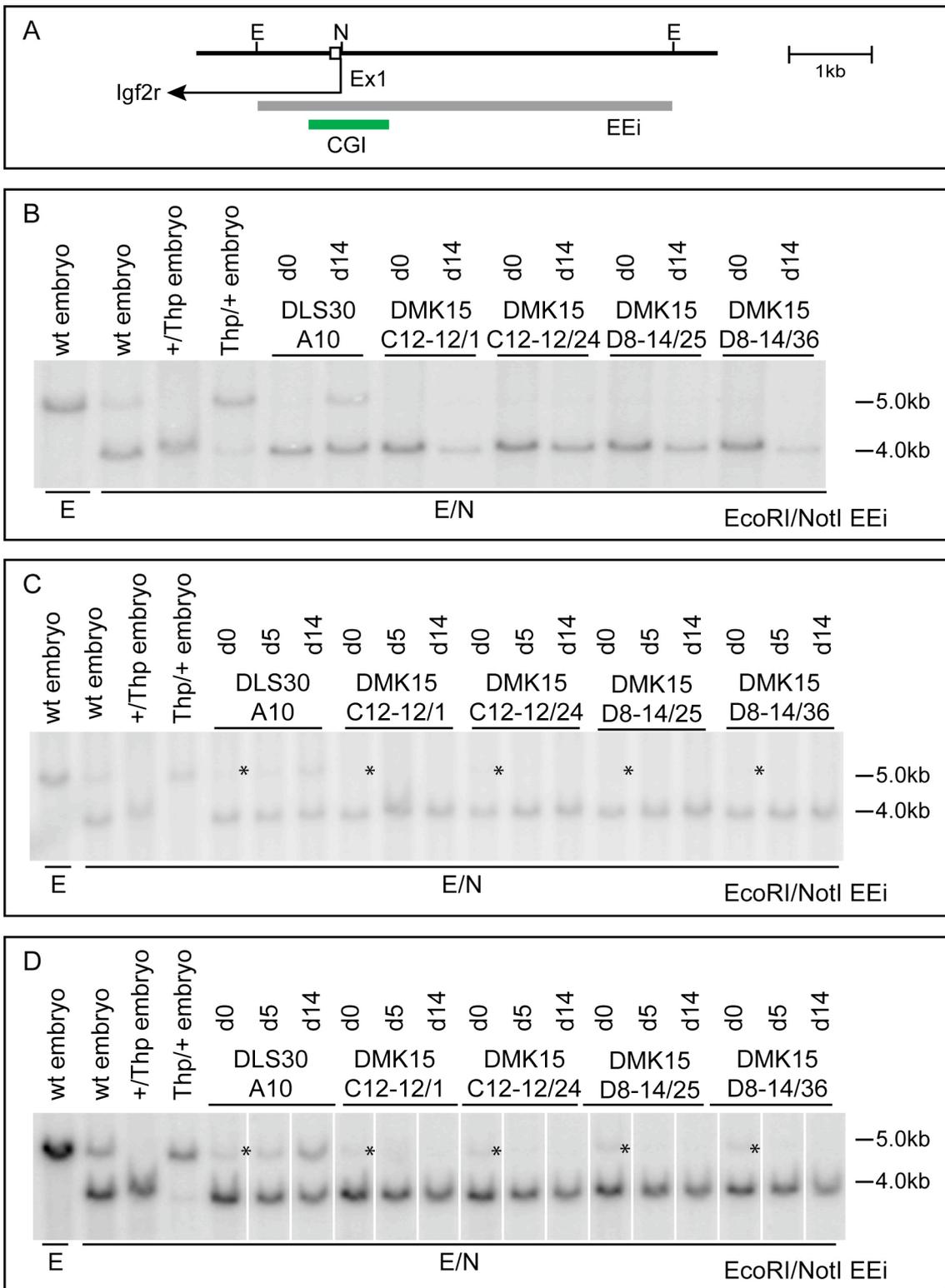


Fig. 46: DNA methylation assay of the *Igf2r* promoter during ES cell differentiation. Details as in Fig. 35. (B)-(D) show three independent differentiation sets. The asterisks mark the presence of a faint 5.0kb band present in d0 cells of all cell lines.

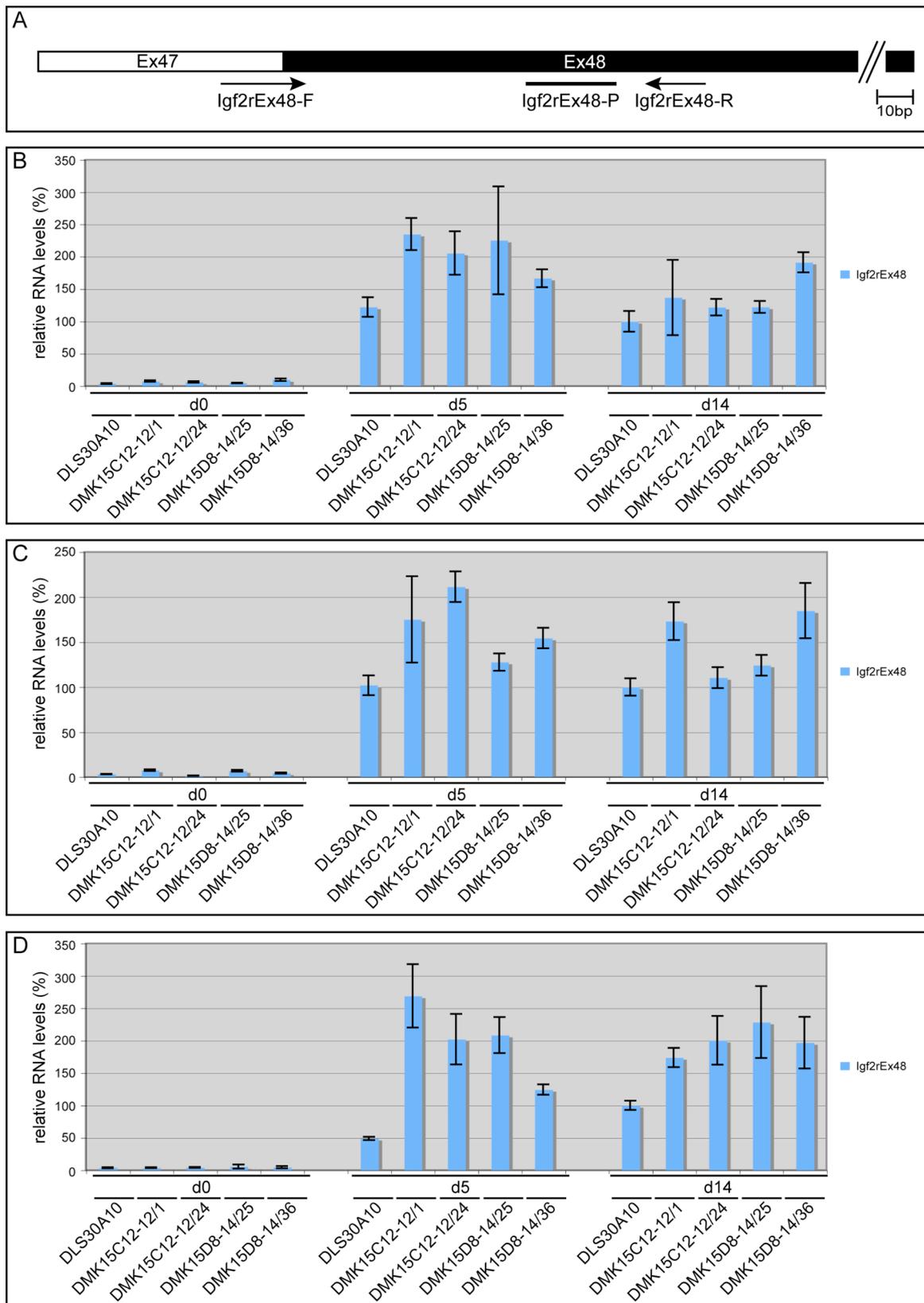


Fig. 47: Analysis of steady-state levels of *Igf2r* during differentiation of DLS30A10 and DMK15 ES cells. Details as in Fig. 37.

Next, I examined steady-state levels of total *Igf2r* mRNA by qPCR at d0, d5 and d14 in DLS30A10 and DMK15 cells. In undifferentiated ES cells, *Igf2r* shows only a low steady-state level but is upregulated during differentiation in both DLS30A10 and DMK15 ES cells. In differentiated DMK15 ES cells, total *Igf2r* levels are 1.1-6 fold compared to *Igf2r* in DLS30A10 ES cells, indicating increased *Igf2r* expression upon deletion of the CpG island (see Fig. 47).

Finally, to test if the increase in total *Igf2r* levels resulted from a loss of imprinted expression, I assayed the ratio of maternal *Igf2r* to paternal *Igf2r* in DLS30A10 and DMK15 ES cells at d0, d5 and d14 of three differentiation sets by qPCR (see Fig. 48). For DLS30A10 at d0, the ratio was set to 1, as those cells were previously shown to equally express *Igf2r* from both parental alleles (Braidotti et al., 2004; Latos et al., 2009; Wang et al., 1994). Undifferentiated DMK15 ES cells showed a ratio of 0.5-0.8% and therefore close to 1. In DLS30A10 ES cells, *Igf2r* showed upregulation during differentiation up to 1.8-12fold, indicating increased maternal versus paternal expression. In DMK15 ES cells however, the ratio stayed close to 1, indicating biallelic expression of *Igf2r*.

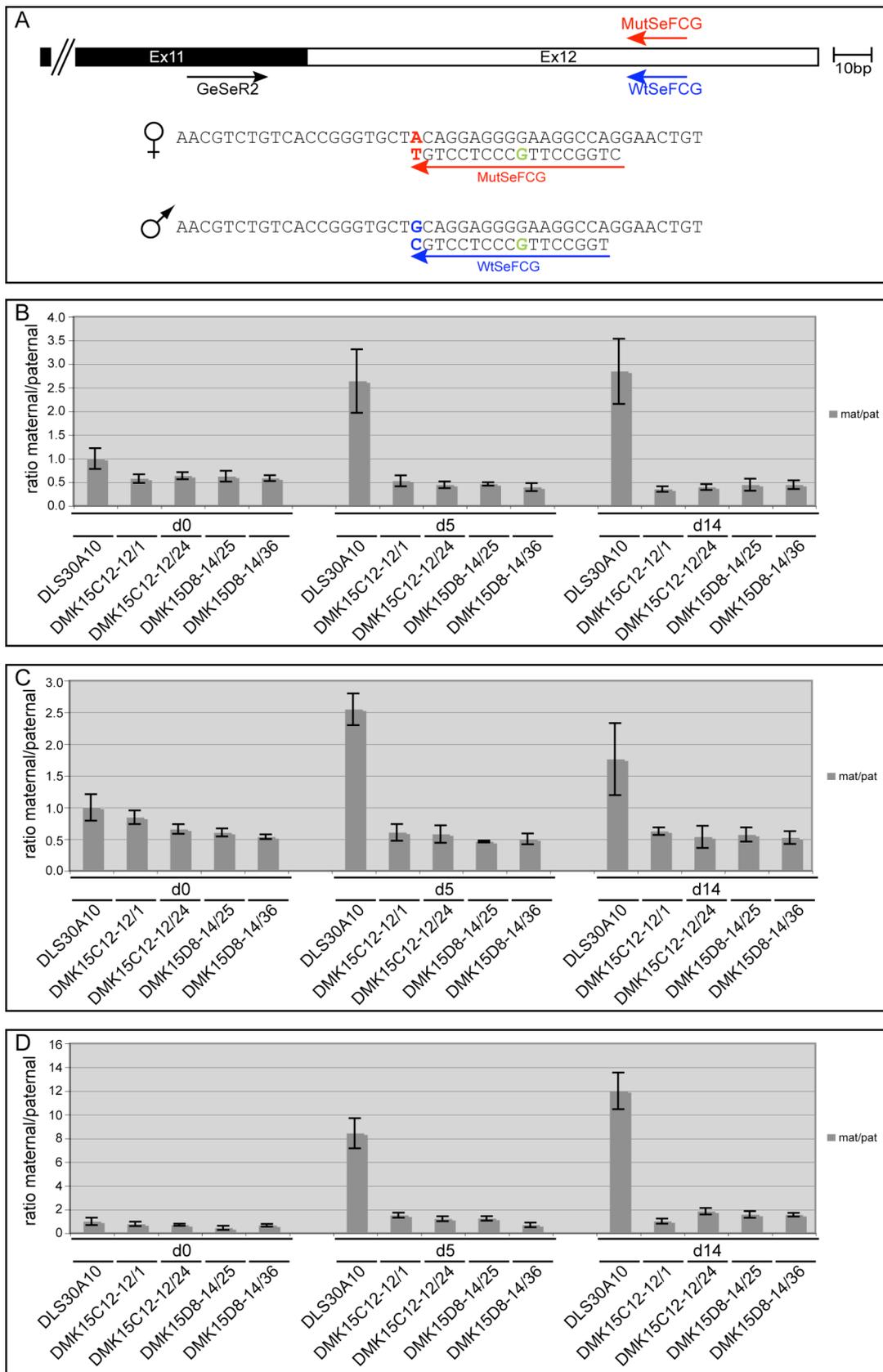


Fig. 48: qPCR analysis of the parental *Igf2r* expression in DLS30A10 and DMK15 ES cells. Details as in Fig. 40.

3. DISCUSSION

3.1 Summary of results

In the first part of this thesis I analysed the role of the *Igf2r* ICE on the chromosomal level and showed that it does not play a role in regulating DNA FISH asynchrony. In the second part of this thesis I analysed the effect of two subdeletions of the ICE removing either the tandem direct repeats or the CpG island. I showed that the tandem direct repeats regulate the length of *Airn* and have a moderate effect on the ability of *Airn* to induce imprinted expression of *Igf2r*. The CpG island is crucial for full-length *Airn* expression, either indirectly by protecting the paternal ICE from DNA methylation or directly by enhancing *Airn* expression.

3.2 The role of the *Igf2r* ICE on the chromosomal level

3.2.1 DNA FISH asynchrony in the imprinted *Igf2r* cluster extends over 3Mbp

It was shown previously, that the genomic region containing the imprinted *Igf2r* cluster exhibits DNA FISH asynchrony and that it is the paternal allele, which first displays two double spots (Kitsberg et al., 1993). The size of this region however had not been determined during this study. This thesis together with my diploma project (Koerner, 2006) demonstrates that in pMEFs DNA FISH asynchrony extends over 3Mbp, starting between *Papbc3* and *Qk* and ending between *Sod2* and *Tcte2*. In contrast, the region showing imprinted expression in pMEFs is only 167kb long, as it only contains *Igf2r* and *Airn*, since *Slc22a2* and *Slc22a3* are not expressed in pMEFs. The region showing DNA FISH asynchrony is also much larger than the size of the whole imprinted cluster including *Slc22a2* and *Slc22a3*, which extends over 321kb. This shows that the region showing DNA FISH asynchrony contains several biallelically expressed genes, of which at least *Agpat4* and *Map3k4* are expressed in pMEFs.

Although several imprinted clusters have been tested for and were found to show DNA FISH asynchrony, to my knowledge this is the first study that accurately maps the extent of the region showing this phenomenon at high resolution (Alexander et al., 2007; Bickmore and Carothers, 1995; Grealley et al., 1998; Gribnau et al., 2003; Gunaratne et al., 1995; Kagotani et al., 2002; Kitsberg et al., 1993; Knoll et al., 1994; Simon et al., 1999; Smrzka et al., 1995). It is therefore impossible to conclude, if the size of 3Mbp determined in this study is unusually large. For human chromosome 11 it was shown by a low resolution mapping, that 14 probes spread over a region of at least 40Mbp containing the imprinted *Igf2* cluster all showed DNA FISH asynchrony (Bickmore and Carothers, 1995). However, as the distance between the probes analysed in this study spans up to several Mbp, it cannot be excluded, that in the intervening sequences DNA FISH asynchrony shifts to synchrony.

3.2.2 DNA FISH asynchrony in the imprinted *Igf2r* cluster is independent of *Airn*, the ICE and the sDMR

After determining the size of the region exhibiting DNA FISH asynchrony I looked for elements which control this behaviour. DNA FISH asynchrony associated with imprinted genes is erased in the germ line before meiosis and reset during late gametogenesis, as DNA FISH asynchrony can be detected already in the pronuclei of the zygote (Simon et al., 1999) and it is also present in cell types which do not show imprinted expression of the genes in the respective cluster, like ES cells for the *Igf2r/Airn* cluster (Gribnau et al., 2003). Therefore the onset of DNA FISH asynchrony and imprinted expression are not correlated. However it was shown, that a biparental origin is necessary for DNA FISH asynchrony, as parthenogenetic ES cells do not display DNA FISH asynchrony for imprinted genes (Gribnau et al., 2003).

The imprint which discriminates the maternal and the paternal allele in imprinted clusters was shown to be DNA methylation and loss of DNA methylation leads to loss of imprinted expression (Li et al., 1993). Therefore it is also obvious to suggest,

that DNA FISH asynchrony is controlled by DNA methylation as well. However, DNA methylation does not seem to play a role in DNA FISH asynchrony. ES cells deficient for DNMT1 or for both *de novo* methyltransferases DNMT3A and DNMT3B as well as ES cells derived from a mother deficient for DNMT3L, which lack oocyte deposits of DNMT3L, still show DNA FISH asynchrony (Gribnau et al., 2003). Furthermore, treatment with 5-azacytidine which reduces DNA methylation levels, does not abolish DNA FISH asynchrony (Bickmore and Carothers, 1995). In addition for the *Igf2* cluster it was demonstrated, that that the *Mnt* mutation still shows DNA FISH asynchrony (Cerrato et al., 2003). The *Mnt* mutation is a 3Mbp inversion on distal mouse chromosome 7 with its breakpoint locating to the mesodermal enhancers of the *Igf2* cluster. Upon maternal transmission, the allele acquires a paternal type DNA methylation (Davies et al., 2002). Therefore both parental alleles are methylated on their ICEs (Pant et al., 2003). All those experiments indicate, that there is no direct role for DNA methylation, also no role for the imprint on the ICE, in regulating DNA FISH asynchrony. However, there still could be other marks different from DNA methylation present on the ICE which have not been identified yet and which regulate DNA FISH asynchrony.

As the *Igf2r* ICE regulates imprinted expression by serving as a methylation-sensitive promoter for *Airn* expression I first analysed, if *Airn* has a function not only in regulating imprinted expression but also in regulating DNA FISH asynchrony. A paternal transmission of a truncation of *Airn* from 108kb to 3kb abolishes imprinted expression of the imprinted protein-coding genes in the *Igf2r* cluster, whereas the shortened version of *Airn* still shows imprinted expression (Sleutels et al., 2002). I used pMEFs carrying an AirT allele on the maternal, the paternal or both parental alleles and showed, that none of the genotypes leads to a loss of DNA FISH asynchrony. Therefore a role of *Airn* in regulating DNA FISH asynchrony in this region can be excluded.

Next I analysed a possible role for the ICE itself in the regulation of DNA FISH asynchrony. A paternal transmission of the deletion of the ICE abolishes imprinted expression of all genes in the *Igf2r* cluster including *Airn* (Wutz et al., 2001). By

using pMEFs lacking the ICE on the maternal, the paternal or both parental alleles I showed, that in contrast to imprinted expression, the DNA FISH asynchrony in this maternally imprinted cluster is not controlled by the ICE. This result clearly contrasts with the results obtained for the paternally imprinted *Igf2* cluster (Greally et al., 1998; Gribnau et al., 2003). In the *Igf2* cluster the maternal transmission of a 13kb deletion comprising *H19*, the ICE and a downstream region, leads to a loss of DNA FISH asynchrony (Greally et al., 1998; Gribnau et al., 2003; Leighton et al., 1995). I could not only confirm these data in this thesis, but I showed in addition, that also a smaller 3.8kb deletion removing the ICE but leaving *H19* intact, had the same effect. Furthermore, it was shown that also a mutation of the CTCF binding sites on the maternal *Igf2* ICE abolishes DNA FISH asynchrony (Sandhu et al., 2009). However, for the maternally imprinted *Snrpn* cluster it was shown in human, that a 5-30kb deletion (the size of the deletion was not mapped more accurately (Sutcliffe et al., 1994)) comprising the ICE on the paternal allele does not abolish DNA FISH asynchrony (Gunaratne et al., 1995). On one hand, this could indicate that DNA FISH asynchrony is differentially regulated in paternally and maternally imprinted clusters. However, the 3.8kb and 13kb deletions in the *Igf2* cluster as well as the 5-30kb deletion in the *Snrpn* cluster, include also sequences outside of the ICE which might play a different role in the two clusters.

The only other region in the *Igf2r* cluster known to show an epigenetic difference between the parental alleles is the sDMR which gains DNA methylation during development on the *Igf2r* promoter. Up to now, the role of an sDMR in regulating DNA FISH asynchrony has not been analysed. As described above a direct role for DNA methylation in controlling DNA FISH asynchrony has been excluded already, however also on the sDMR potentially unidentified additional elements or modifiers could be present regulating DNA FISH asynchrony. I therefore also analysed pMEFs with a deletion of the *Igf2r* promoter region on the maternal, the paternal or both parental alleles (Sleutels et al., 2003) and found, that also this region does not control DNA FISH asynchrony.

3.2.3 TSA leads to a relaxation of DNA FISH asynchrony in the imprinted *Igf2r* cluster

Besides DNA methylation, also histone modifications were shown to be involved in genomic imprinting (Lewis et al., 2004b; Regha et al., 2007; Umlauf et al., 2004). I tested, if treatment with Trichostatin A, a general inhibitor of histone deacetylases, has an influence on DNA FISH asynchrony in the *Igf2r* cluster and showed, that TSA treatment leads to a relaxation but not a complete loss of DNA FISH asynchrony. Treatment of cells with TSA or sodium butyrate, another histone deacetylation inhibitor, also reduced DNA FISH asynchrony in the *Igf2* cluster (Bickmore and Carothers, 1995; Kagotani et al., 2002). However, chemicals like TSA also have a range of off-target effects, as they lead to the transcriptional activation or silencing of various genes and their gene products which in turn might have an influence on DNA FISH asynchrony. Furthermore, TSA can lead to differentiation, cell cycle arrest and apoptosis (Marks et al., 2001). Therefore a direct role of histone acetylation in controlling DNA FISH asynchrony cannot be concluded from these results.

Another histone modification, methylation of H3K27, also seems to have an influence on DNA FISH asynchrony. Examination of ES cells deficient for *Eed*, a protein necessary for mono-, di- and trimethylation of H3K27 revealed, that DNA FISH asynchrony in the *Igf2* cluster is reduced (Alexander et al., 2007; Montgomery et al., 2005). *Eed* mutant ES cells show derepression of several developmentally regulated genes and tend to differentiate rapidly under conditions in which they should stay undifferentiated (Boyer et al., 2006). *Eed* mutant embryos die before E9.5, and the deregulation of several imprinted genes like *Cdkn1c*, *Ascl2*, *Grb10* and *Meg3* found in these mutants might contribute to this early lethality. However, not all imprinted genes are affected in *Eed* mutant embryos, as *Igf2r*, *Kcnq1*, *Snrpn*, *Kcnq1ot1* and others retain their imprinted expression (Faust et al., 1995; Mager et al., 2003). As *Eed* is not involved in the regulation of imprinted expression in the *Igf2r* cluster, it is unlikely that it plays a role in regulating DNA FISH asynchrony in this cluster. Furthermore, a study mapping chromatin modifications in the *Igf2r* cluster showed, that the only allele-specific modifications are localised to the *Igf2r*

and the *Airn* promoter regions (Regha et al., 2007). As I have shown that both regions are not involved in DNA FISH asynchrony, I can rule out, that those modifications play a role in regulating DNA FISH asynchrony. In summary, my results show that no known epigenetic feature, with the possible exception of histone acetylation, play a role in the regulation of DNA FISH asynchrony in the *Igf2r* cluster.

3.2.4 The parental alleles containing the *Igf2r* cluster do not display differences in chromatin compaction as determined by 3D DNA FISH

Besides DNA FISH asynchrony also other features of imprinted regions are indicative of a different chromatin state between the two parental alleles. Different frequencies of meiotic recombination are found in or near to imprinted clusters with a higher recombination rate during male meiosis (Paldi et al., 1995; Robinson and Lalande, 1995). Furthermore, during homologous targeting experiments using ES cells a preferential targeting of the paternal allele in the region of transcriptional overlap of *Igf2r* and *Airn* can be detected (Wang et al., 1994, Paulina A. Latos, Stefan H. Stricker, Florian M. Pauler, Martha V. Koerner unpublished). Chromatin fractionation by centrifugation revealed, that the paternal *Igf2r* allele is enriched in the more compacted heterochromatic fraction whereas the maternal *Igf2r* allele is found in the less compacted euchromatic fraction (Watanabe et al., 2000). In contrast, the *Snrpn* gene located in the *Pws/As* cluster was not found to display differences in chromatin packaging by the same method (Watanabe et al., 2000), therefore differential chromatin compaction might not be a unique feature of all imprinted regions.

3D FISH is a technique which can be used to measure chromatin compaction of genomic regions larger than 50kb (Chambeyron and Bickmore, 2004; Fuxa et al., 2004; Yokota et al., 1997). By using 3D FISH I measured the distance on the maternal and the paternal allele of two probes spaced by 500kb and found no difference in chromatin compaction. This indicates, that either DNA FISH asynchrony is not reflected by or due to chromatin compaction or that 3D FISH cannot detect a

chromatin compaction as measured by chromatin fractionation by centrifugation. This result indicates that there are no global chromatin organisation differences between the parental alleles, which is in agreement with studies done previously which showed that with the exception of the *Airn* and *Igf2r* promoter regions no allele-specific chromatin marks like DNaseI hypersensitive sites or histone modifications can be identified in MEFs which only show imprinted expression of *Igf2r* and *Airn* (Pauler et al., 2009; Pauler et al., 2005; Regha et al., 2007).

3.2.5 DNA FISH asynchrony in the *Igf2r* cluster is independent of DNA replication

Originally, the DNA FISH assay as used in this thesis to analyse DNA FISH asynchrony was thought to be a measurement of DNA replication with the idea, that in nuclei with two single spots no allele has undergone DNA replication yet, in nuclei with a single and a double spot one allele has undergone DNA replication already whereas the other did not, and in nuclei with two double spots both alleles have been replicated already (see Fig. 48A). However, as the detection of two spots depends on the separation of sister chromatids after replication, it is also possible that not always true DNA replication is detected but rather differences in sister chromatid cohesion (Azuara et al., 2003) (see Fig. 48B). Therefore a second method was developed to directly monitor DNA replication. For this technique, unsynchronised BrdU-pulsed cells are fractionated according to their cell cycle stage using fluorescence-activated cell sorting (FACS). Alternatively, cells are synchronised, BrdU-pulsed and harvested at different time points after S phase entry. Subsequently, the DNA of these cells is immunoprecipitated using an anti-BrdU antibody. This allows the isolation of DNA fragments, which replicated during a specific timeframe during S phase. Then, the isolated DNA is subjected to PCR detection of newly replicated DNA (Azuara, 2006; Hansen et al., 1993). By comparing these two techniques it was found that delayed sister chromatid segregation can lead to false positive results analysing replication using DNA FISH

(Azuara et al., 2003) and that SD patterns can appear without asynchronous DNA replication (Alexander et al., 2007).

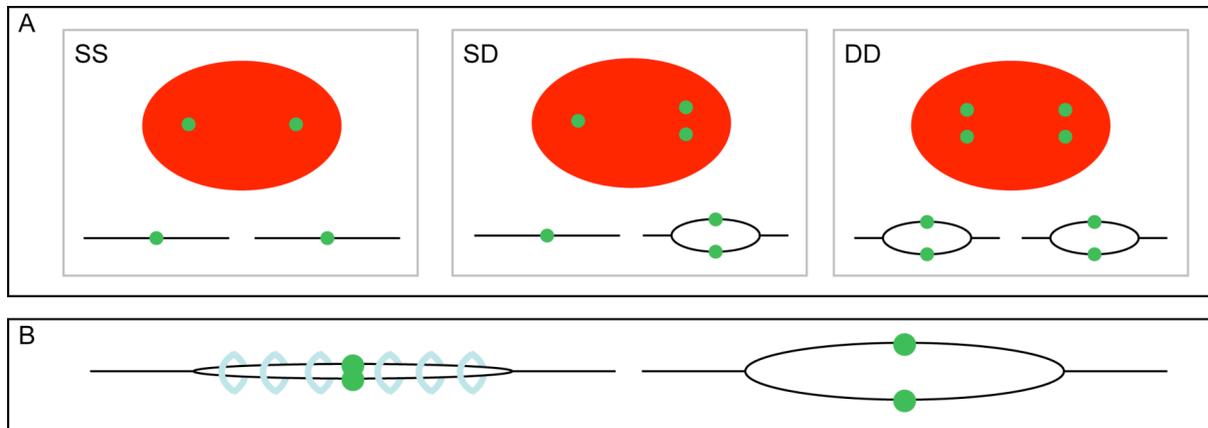


Fig. 48: Possible explanations for the appearance of different DNA FISH patterns in diploid cells. (A) DNA FISH patterns might arise from differences in allelic DNA replication. Shown are nuclei (red ellipse) with the three possible FISH patterns, single-single (SS), single-double (SD) or double-double (DD) (FISH probe shown in green). Below each nucleus the corresponding DNA replication status is shown. The black line represents a DNA double strand, the black oval represents a DNA replication bubble. (B) DNA FISH patterns might arise from differences in sister chromatid cohesion. Shown is the possible explanation for an SD pattern in nuclei, where both alleles have replicated already. The blue rings represent COHESIN.

I analysed allelic replication timing by S phase fractionation of the maternal and the paternal allele of the *Igf2r* cluster and found that both replicate in early S phase. This indicates that the DNA FISH asynchrony present in the *Igf2r* cluster is not due to differences in replication timing of the two parental alleles. It is therefore more likely, that the DNA FISH asynchrony in the *Igf2r* cluster is due to differences in sister chromatid cohesion. In contrast to the results presented here, where I used inbred cells for the analysis of allelic replication timing, a previous study which used an interspecies cross (*Mus musculus/spretus*), found replication asynchrony in the *Igf2r* cluster (Simon et al., 1999). However, in the study by Simon *et al.*, just one direction of the cross with the male being *Mus spretus* and the female being *Mus musculus* has been analysed and it is known, that interspecies crossings can disrupt genomic imprinting (Shi et al., 2004; Vrana et al., 1998). It is therefore difficult to conclude, if the difference in allelic replication timing found in the study by Simon et al. reflects differences due to genomic imprinting or due to strain-specific differences.

For the *Igf2* cluster many more but also partially contradictory data exist about the presence of replication asynchrony as there are publications which find replication asynchrony (Gribnau et al., 2003), and others which find that both alleles replicate synchronously (Windham and Jones, 1997). In addition it was also demonstrated, that mutations of the CTCF binding sites on the maternal ICE of the *Igf2* cluster abolish DNA FISH and replication asynchrony whereas cells with a mutation of the CTCF binding sites on the paternal ICE show DNA FISH and replication asynchrony (Bergstrom et al., 2007; Sandhu et al., 2009). Ignoring the study which did not find replication asynchrony in the *Igf2* cluster (Windham and Jones, 1997), these results indicate, that in the *Igf2* cluster the maternal ICE not only regulates DNA FISH asynchrony, but also replication asynchrony. It is however still not clear how this regulation is achieved, as CTCF binding to the ICE is methylation-sensitive and as described above it was shown that loss of DNA methylation does not abolish DNA FISH asynchrony and that the *Mnt* mutation which has DNA methylation on both parental ICEs also does not lose DNA FISH asynchrony (see Fig. 49). However, it is possible, that replication asynchrony and DNA FISH asynchrony are not regulated by CTCF directly. It could be, that another yet unknown protein which binds allele-specifically, but is independent of DNA methylation and sensitive to mutations of the CTCF binding sites, regulates DNA FISH and replication asynchrony. There are indeed other proteins known, which bind the *Igf2* ICE allele-specifically. COHESIN binds to the maternal allele like CTCF, however recruitment of COHESIN depends on CTCF, therefore COHESIN is an unlikely candidate to regulate the phenomena described above (Wendt et al., 2008). Furthermore it was demonstrated, that in brain binding of ATRX, a member of the SNF2 family of chromatin remodelling proteins, to the maternal *Igf2* ICE leads to the recruitment of COHESIN and CTCF (Kernohan et al., 2010). Therefore, a protein like ATRX could be a good candidate for regulating DNA FISH and replication asynchrony. However, it has not been demonstrated yet if ATRX binding to the ICE is methylation-sensitive. Furthermore, loss of ATRX does not change expression of genes in the *Igf2* cluster in embryonic brain but leads to increased *H19* expression in post-natal brain, and this increase is not due to loss of imprinted expression, instead *H19* shows upregulation from the maternal allele (Kernohan et al., 2010). Additionally, ATRX involvement in the *Igf2* cluster has only

been demonstrated in the brain but not in other tissues. Therefore additional work is still needed to determine, how DNA FISH and replication asynchrony in the *Igf2* cluster are controlled.

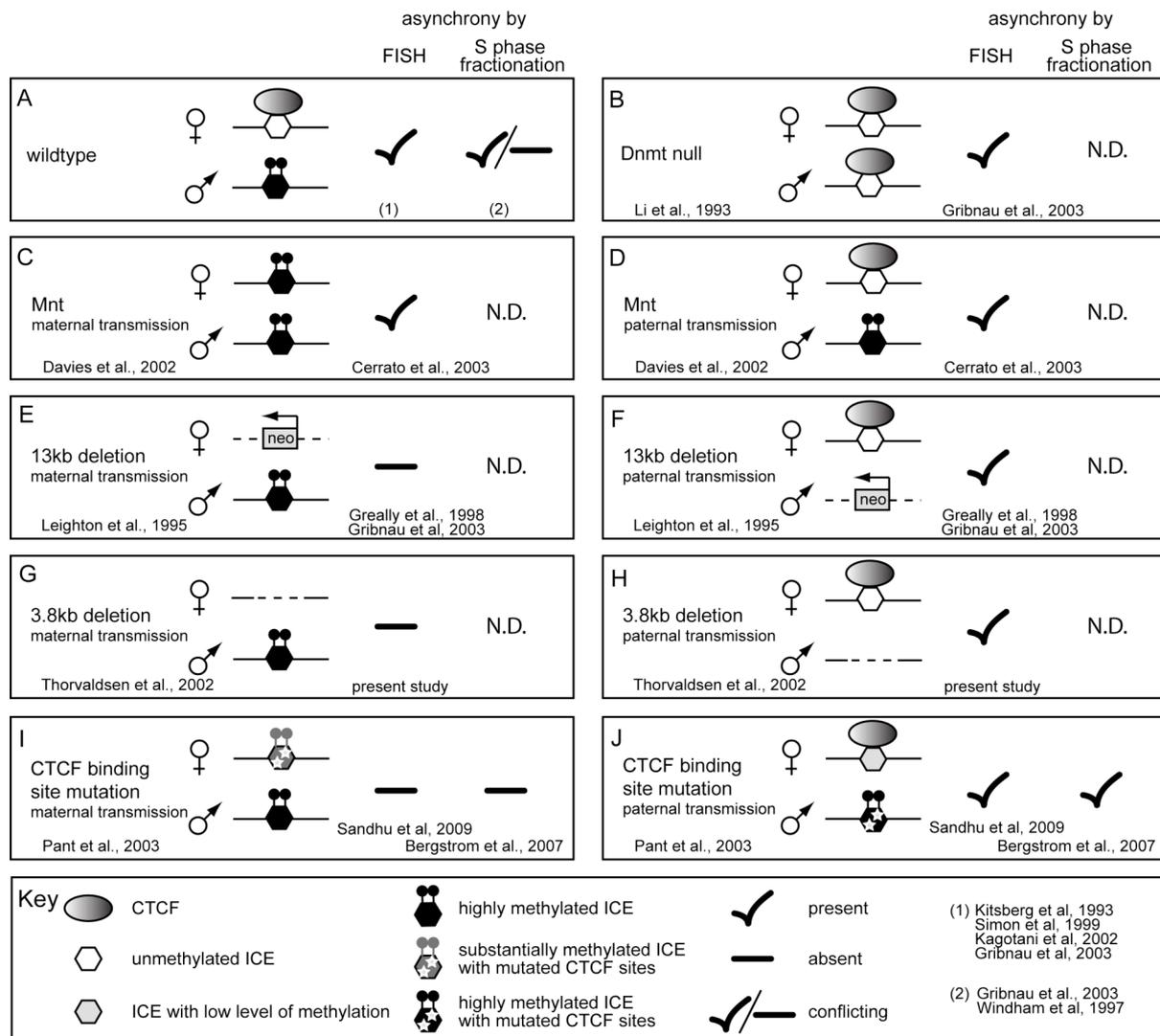


Fig. 49: Summary of replication asynchrony and S phase fractionation results for the *Igf2* imprinted cluster. (A)-(J) Schemes of the *Igf2* ICE in wildtype and mutant cells. References for the creation and primary analysis of the effect of the mutation on imprinted expression are given below the genotype. Next to the scheme the results of the analyses for DNA FISH and replication asynchrony are presented, references are given below. N.D.: not determined. See key for further details. (A) In wildtype cells, the maternal unmethylated ICE is bound by CTCF, the paternal methylated ICE is free of CTCF binding. In these cells, DNA FISH asynchrony is present, but data for S phase fractionation are conflicting. (B) Targeted disruption of *Dnmt1* leads to loss of DNA methylation on the ICE, thus both alleles bind CTCF, but DNA FISH asynchrony is not abolished. (C) A maternal transmission of the *Mnt* mutation leads to DNA methylation on both parental ICEs. DNA FISH asynchrony is still present. (D) A paternal transmission of the *Mnt* mutation does not change ICE methylation nor DNA FISH asynchrony. (E) Maternal transmission of the 13kb deletion (deleted region shown as dashed line) where the ICE and *H19* are replaced by a neomycin resistance gene leads to loss of DNA FISH asynchrony. (F) Paternal transmission of the 13kb deletion does not disturb DNA FISH asynchrony.

(G) Maternal transmission of the 3.8kb deletion (deleted region shown as dashed line) removing the ICE and flanking sequence but leaving *H19* intact removes DNA FISH asynchrony. (H) Paternal transmission of the 3.8kb deletion does not have an effect on DNA FISH asynchrony. (I) Mutation of the CTCF binding sites on the maternal allele leads to a substantial gain of DNA methylation and loss of CTCF binding on the maternal allele and loss of DNA FISH and replication asynchrony. (J) Mutation of the CTCF binding sites on the paternal allele leads to a low amount of DNA methylation *in trans* on the maternal allele, however CTCF still can bind. DNA FISH and replication asynchrony are still present.

A more detailed analysis is additionally needed as it was shown, that in the cells which carry the mutated CTCF sites on the maternal allele, not only DNA FISH asynchrony in the *Igf2* cluster is lost but also in the *Mest* and *Dlk1* clusters as well as at the solo imprinted *Gatm*, *Impact*, *Ins1* and *Htr2a* genes (Sandhu et al., 2009), indicating that the *Igf2* ICE controls DNA FISH asynchrony of many if not all imprinted clusters.

3.2.6 DNA FISH asynchrony and imprinted expression – cause, consequence or correlation?

DNA FISH asynchrony in the *Igf2*, *Igf2r*, *Pws/As* and *Kcnq1* clusters is still present in the DNA methyltransferase mutants where imprinted expression is lost (Gribnau et al., 2003). It can also be detected in cells where imprinted genes do not show imprinted expression, like in ES cells for the *Igf2r* cluster (Gribnau et al., 2003). Furthermore, a truncation of *Airn* or a deletion of the *Igf2r* ICE abolishes imprinted expression (Sleutels et al., 2002; Wutz et al., 2001), however replication asynchrony is still present. Therefore, imprinted expression is clearly not a prerequisite for DNA FISH asynchrony.

Is DNA FISH asynchrony necessary but not sufficient for imprinted expression? Also this assumption is unlikely as the imprinted RSV-*Pgk-myc* transgene, a variant of the RSV-*Ig-myc* transgene, shows imprinted methylation and parental-specific expression, but no DNA FISH asynchrony could be detected (Shuster et al., 1998). Furthermore, the study by Sandhu *et al.* which showed that mutation of the CTCF binding sites in the *Igf2* ICE abolishes DNA FISH asynchrony in several imprinted

clusters, gives a further indication, that DNA FISH asynchrony is not needed for imprinted expression (Sandhu et al., 2009). Up to now it was not shown, that a mutation which abolishes imprinted expression of one cluster has an effect on imprinted expression of another cluster. At least for the two imprinted clusters present on distal chromosome 7, the *Igf2* and the *Kcnq1* cluster, it was demonstrated that they do not influence each other. A maternal transmission of the 13kb deletion comprising the *Igf2* ICE and *H19* does not have an effect on imprinted expression of genes in the *Kcnq1* cluster (Casparly et al., 1998). However, in the study by Sandhu et al. the effect of the mutations of the CTCF binding sites in the *Igf2* ICE on the DNA FISH asynchrony of the *Kcnq1* cluster, was not analysed (Sandhu et al., 2009). Therefore it still cannot be formerly excluded, that the *Igf2* ICE regulates DNA FISH asynchrony of many but not all imprinted clusters and in those clusters where it regulates DNA FISH asynchrony it also influences imprinted expression.

3.3 The function of tandem direct repeats

As mentioned in the introduction, tandem direct repeats are commonly found in ICEs and thus it was hypothesised early, that they could play a role in the regulation of genomic imprinting, especially with respect to controlling the DNA methylation imprint mark (Neumann et al., 1995). However, not all imprinted clusters seem to need tandem direct repeats for controlling their DNA methylation imprint. Deletion of the tandem direct repeats from *U2afbp-rs* did not have an effect on DNA methylation (Sunahara et al., 2000). *Impact*, which shows imprinted expression in mouse, rabbit, cottontail and lemming, has tandem direct repeats in mouse only (Okamura and Ito, 2006; Okamura et al., 2008). However, in other imprinted clusters, tandem direct repeats seem to play a role in the regulation of DNA methylation. Tandem direct repeat sequences from the *Igf2r*, *Snurf/Snrpn* and *Kcnq1* ICEs were shown to be critical in the maintenance of the differential DNA methylation during embryogenesis using the RSV-*Ig-myc* transgenic model (Reinhart et al., 2002; Reinhart et al., 2006). The RSV-*Ig-myc* transgene contains a Rous-Sarcoma-Virus (RSV) long terminal

repeat (LTR) and a 5' truncated *c-myc* fused to the alpha constant and switch region of the *Ig* heavy chain locus. It gains DNA methylation during oogenesis and shows expression upon paternal inheritance only (Swain et al., 1987). Its differentially methylated region can be deleted and replaced by other regions to test, if this region can confer imprinting of the transgene (Reinhart et al., 2002). In the ICE of the *Igf2r* imprinted cluster, two distinct classes of direct tandem repeats have been identified, the first repeats are 30-32bp in length, repeated three times and called the 1-repeats, and the repeats of the second class are 172-180bp in length, also repeated three times and termed the 2,3-repeats (Neumann et al., 1995; Reinhart et al., 2002). Using the RSV-*Ig-myc* transgenic model it has been shown that two copies of the second repeat class are sufficient to maintain the differential DNA methylation of the ICE, the first class however does not play a role in this aspect (Reinhart et al., 2006). It was further shown, that it is the CpG content of the second repeat class which plays a crucial role in DNA methylation which also might explain the lack of function of the first repeat class as they are relatively poor in CpGs (Reinhart et al., 2006). The role of the tandem direct repeats in the endogenous *Igf2r* ICE is not yet known; however, the *in vivo* deletion of a subset of tandem repeats from the *Kcnq1ot1* ICE or of direct repeats that flank the *Igf2* ICE did not change ICE DNA methylation (Lewis et al., 2004a; Mancini-Dinardo et al., 2006; Reed et al., 2001; Thorvaldsen et al., 2002). However, these deletion experiments do not contradict the results of the RSV-*Ig-myc* experiments. The tandem direct repeats of the *Igf2* ICE could not induce imprinted DNA methylation in the RSV-*Ig-myc* transgenic system either (Reinhart et al., 2002). Furthermore, the region which was deleted in the endogenous *Kcnq1* ICE is smaller and was missing additional tandem direct repeats compared to the region which was shown to induce imprinted DNA methylation using the RSV-*Ig-myc* transgenic system (Mancini-Dinardo et al., 2006; Reinhart et al., 2006).

A direct evidence for the importance of tandem direct repeats in controlling the DNA methylation imprint comes from experiments performed in the paternally imprinted *Rasgrf1* cluster located on mouse chromosome 9. A mouse strain-specific loss of methylation was observed following the deletion of a repeat region from the paternal

allele (Yoon et al., 2002). Furthermore, a replacement of the *Rasgrf1* repeats with the *Igf2r* ICE repeats was reported to restore the DNA methylation and in addition to lead to a gain of DNA methylation on the maternal allele *in trans* (Herman et al., 2003). This range of experiments demonstrates, that until now no general function can be assigned to tandem direct repeats for the regulation of imprinting. Whereas some loci seem to need them for their differential DNA methylation, others do not. Additionally it is still not clear if also in those clusters where the tandem direct repeats have a function in regulating DNA methylation, this function depends on the tandem direct repeats *per se* or rather on the CpG content and CpG periodicity, and that the repeats are rather a consequence of this than the cause. In support of this idea, using crystallography it was shown, that DNMT3A and DNMT3L interact directly and form a tetrameric complex with two active sites, which are separated from each other by the size of about one DNA helical turn. Further analysis indicated, that DNMT3A methylates DNA in a periodic pattern of 8 to 10 basepairs (Jia et al., 2007). Thus, the tandem nature of the repeats could support DNA methylation by DNMT3A. This also would explain, why tandem direct repeats, which are rather poor in CpG dinucleotides do not have a function in regulating DNA methylation.

Tandem direct repeats are not only found in imprinted regions but also in genes regulating X chromosome inactivation. Two of those tandem direct repeats were shown to have a function in the regulation of X chromosome inactivation, however, not in the regulation of DNA methylation. For *Tsix*, a ncRNA involved in X chromosome inactivation, close to and downstream of its transcriptional start, a region called DXPas34 has been identified, which contains tandem direct repeats and gains DNA methylation on the active X chromosome after implantation (Courtier et al., 1995; Prissette et al., 2001). DXPas34 has been shown to be essential for initiation of *Tsix* expression in undifferentiated ES cells, and its deletion diminishes (but does not completely abolish) the strict control that normally prevents up-regulation of *Xist* expression in differentiated male ES cells (Vigneau et al., 2006). Furthermore it has been demonstrated, that *in vivo* DXPas34 is crucial for imprinted and random XCI (Cohen et al., 2007). Also *Xist* which itself is regulated by *Tsix* has

conserved tandem direct repeats at its 5' end (Brown et al., 1992) which when deleted do not affect localization of the *Xist* ncRNA to the inactive X chromosome but specifically the silencing ability of *Xist* (Wutz et al., 2002).

3.3.1 The role of the tandem direct repeats in the *Igf2r* ICE

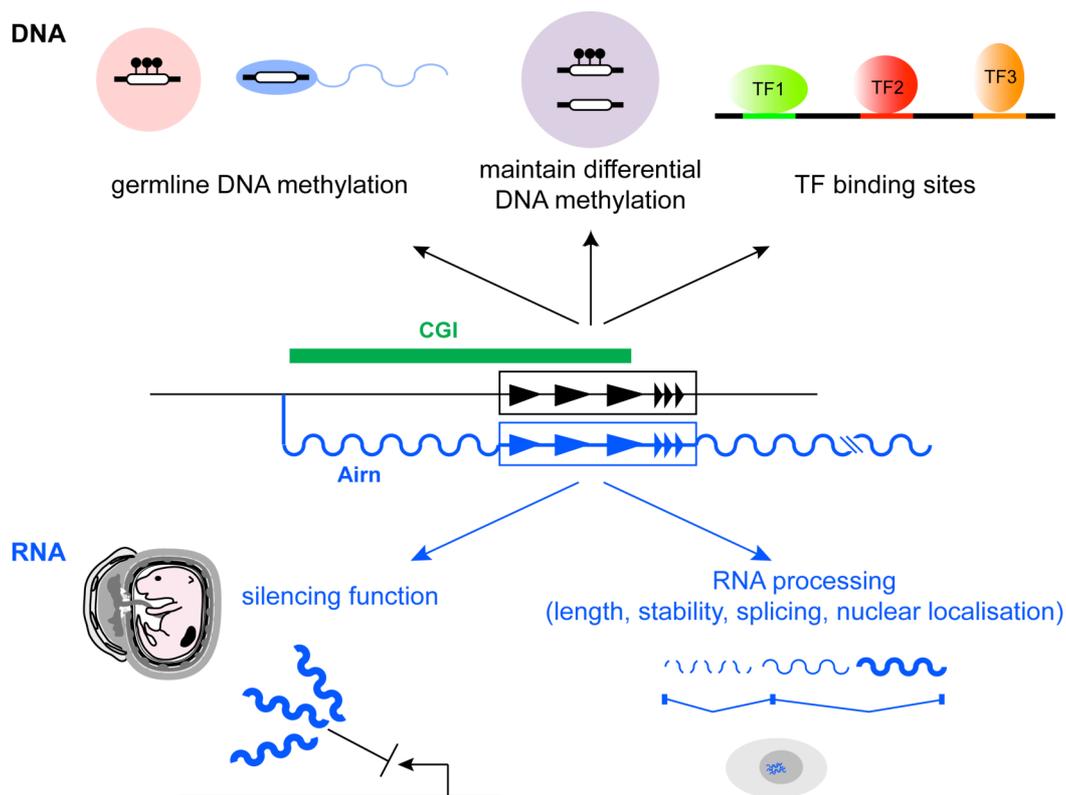


Fig. 50: Features of the imprinted *Igf2r* cluster which might be controlled by the tandem direct repeats. On the DNA level (top), the repeats could play a role in setting the differential DNA methylation in oocytes or preventing it in sperm (left). In diploid somatic cells they might help to maintain this differential methylation (middle). As the repeats are within the CpG island, they could contain transcription factor binding sites necessary for *Airn* expression (right). On the RNA level (bottom), the tandem direct repeats could effect the silencing ability of *Airn*, and this could be different in embryonic and extraembryonic tissues (left). Furthermore, they might contribute to the features of the *Airn* RNA, like nuclear localisation, splicing or the low stability of *Airn* (right). *Airn* ncRNA: wavy blue line. genomic DNA: black line. CpG island (CGI): green bar. tandem direct repeats: boxed black and blue triangles. oocyte: pink filled circle. somatic cell: violet filled circle. transcription factors: green, red, orange filled ellipses (TF). cytoplasm and nucleus: light grey/ dark grey ellipse.

Although not all tandem direct repeats seem to have a function, several of them can control diverse aspects of gene regulation like DNA methylation, transcriptional gene regulation or RNA-mediated gene silencing. To analyse, if the tandem direct repeats

of the *Igf2r* ICE control one or several of the features listed above (see Fig. 50), I generated ES cells with a targeted deletion of the tandem direct repeats in the endogenous locus on the paternal allele (R Δ 800), and analysed them using the ES cell *in vitro* differentiation system.

3.3.2 The tandem direct repeats are not needed for the maintenance of the differential DNA methylation on the ICE

I analysed three subclones with an R Δ 800 allele in wildtype D3 cells and four subclones from two independent homologous targetings with an R Δ 800 allele in DLS30A10 cells for DNA methylation of the ICE in undifferentiated and differentiated ES cells. I showed, that deletion of the tandem direct repeats does not lead to a gain of DNA methylation on the normally unmethylated paternal allele. Therefore the tandem direct repeats are not needed for the maintenance of the differential DNA methylation on the ICE, at least not at developmental stages later than 3.5dpc, as this is the timepoint, at which ES cells are isolated from the blastocyst of a developing mouse embryo.

It is possible, that once the differential DNA methylation is established, it is self-propagating. CpG islands are with some exceptions normally devoid of DNA methylation. The exact mechanism, how CpG islands are protected from DNA methylation remains unclear. One possibility is an active DNA demethylase acting specifically on CpG islands, however, such an enzyme has not been identified so far in somatic cells (Illingworth and Bird, 2009). Furthermore, transcription factors could play a role in the prevention of DNA methylation. It was shown, that deletion of binding sites for the transcription factor SP1 in the murine *Aprt* gene led to a gain of DNA methylation (Brandeis et al., 1994; Macleod et al., 1994; Mummaneni et al., 1998). However, also CpG islands associated with silent genes are normally free of DNA methylation, and also *Airn* is not expressed during early development, but its paternal allele is unmethylated (Latos et al., 2009; Stoger et al., 1993). It was shown, that DNMT3L, a protein which interacts and stimulates the *de novo*

methyltransferases, cannot bind to regions which show H3K4me3 (Hata et al., 2002; Jia et al., 2007). As the unmethylated ICE is marked by H3K4me3 (Mikkelsen et al., 2007; Regha et al., 2007), this chromatin signature could prevent the unmethylated ICE from gaining DNA methylation. H3K4me3 on the unmethylated *Airn* promoter is spread over approximately 3kb (Regha et al., 2007). As in the R Δ 800 allele only 700bp have been deleted from this region, the remaining 2.3kb of H3K4me3 could be enough to prevent DNMT3L-mediated DNA methylation. In summary, the unmethylated form of a CpG island seems to be the default state, also if it is still unclear, how these regions are protected from DNA methylation. The other ICE allele gains DNA methylation during oogenesis or spermatogenesis (Coombes et al., 2003; Engemann et al., 2000; Shemer et al., 1997; Stoger et al., 1993; Takada et al., 2002; Tremblay et al., 1995) by a not yet identified mechanism, and DNMT1, the maintenance methyltransferase, propagates the methylated state during each cell division (Li et al., 1992). Therefore it is possible, that a role of the tandem direct repeats in influencing DNA methylation on the ICE could be present during the establishment phase only. To see, if the tandem direct repeats play a role during the establishment of the differential DNA methylation on the ICE, I generated and characterised ES cells for blastocyst injection. These ES cells will be used later in the Barlow lab for the generation of a repeat deletion knock out mouse, to assay the role of the tandem direct repeats at early stages of development.

3.3.3 Deletion of the tandem direct repeats leads to a reduction in *Airn* length, but upregulation of *Airn* during *in vitro* differentiation and the low splicing capability of *Airn* are unaffected

Using two different methods, RNA expression tiling array (RETA) and qPCR, I have shown that *Airn* in R Δ 800 ES cells is shorter compared to *Airn* in wildtype ES cells. Analysis of steady-state levels by qPCR of mapped *Airn* splice variants showed that they are unaffected with the exception of the two splice variants, SV2 and SV3, whose specific exons are localised at the 3' end of *Airn*. qPCR analysis further revealed, that deletion of the tandem direct repeats did not lead to a de-repression of

Airn in undifferentiated ES cells, and during ES cell differentiation, RΔ800 *Airn* showed similar upregulation kinetics as wildtype *Airn*. Therefore surprisingly, deletion of the tandem direct repeats on the 5' end of *Airn* led to a truncation of *Airn* on the 3' end, and to an *Airn* with an appearance similar to that seen in pMEFs.

Generally, the length of *Airn* seems to differ between different tissues as determined by RETA (Ru Huang, unpublished) and whereas the qPCR assay 'Air end' readily detects *Airn* in ES cells, the same assay is unable to detect *Airn* in pMEFs and RΔ800 ES cells. If the different appearance of *Airn* in different tissue has functional consequences still has to be determined in detail. For example it could be, that during the establishment phase of imprinted expression, which takes place during ES cell differentiation or at the time of implantation during embryonic development, a longer *Airn* is necessary compared to the maintenance phase, like in midgestation embryos, adults and tissues harvested from those developmental stages like pMEFs. Furthermore, *Airn* leads to imprinted expression of different genes in different tissues. This might require different lengths of *Airn* as well. As described in the introduction it could be, that in tissues where *Airn* leads to imprinted expression of *Igf2r* only, it just has to cross its promoter to suppress it by transcriptional interference, or that regulatory regions within the *Airn* ncRNA crucial for the imprinting process of *Igf2r* are contained in the more 5' end. In extraembryonic tissues, where *Airn* leads to imprinted expression of not only *Igf2r* but additionally of *Slc22a2* and *Slc22a3*, *Airn* might have to cross a putative transcription-sensitive enhancer upstream of *Igf2r* or other regulatory regions within the more 3' end of *Airn* are crucial. But why should then differentiated ES cells have a longer *Airn* than pMEFs? Cells from the inner cell mass, the cell lineage from which ES cells are generated, do not only give rise to the embryo proper but also to the extraembryonic tissues of the amnion, the extra-embryonic mesoderm of the visceral yolk sac and the labyrinth layer of the placenta (see Fig. 51). Therefore one can speculate, that also an ES cell culture contains precursor cells of extraembryonic tissues which have to express a longer *Airn*.

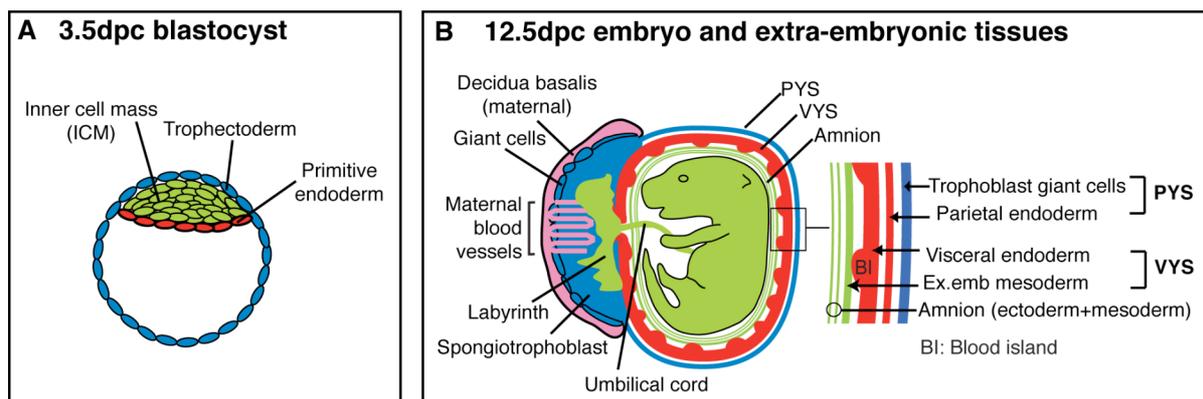


Fig. 51: Lineage relationships in embryonic development. (A) Shown is a blastocyst at 3.5dpc. The trophoblast (blue) contributes to the parietal yolk sac (PYS) and placenta of the 12.5dpc embryo (B). The inner cell mass (ICM, green) gives rise to the embryo proper and contributes to the visceral yolk sac (VYS) and amnion that are extraembryonic membranes. The primitive endoderm (red) differentiates into the endoderm layer of the PYS and VYS. (B) Shown is a 12.5dpc embryo and its extraembryonic tissues. The embryo (green) is surrounded by the amnion, which consists of ICM-derived ectoderm and mesoderm. The middle extraembryonic membrane is the VYS, which consists of ICM-derived mesoderm (green) and endoderm (red). The outer membrane, the PYS, is lost after 13.5dpc and consists of ICM-derived parietal endoderm (red) and trophoblast giant cells (blue) which are not ICM derived. The placenta consists of distinct layers, the inner labyrinth (green), the spongiotrophoblast (blue) and giant cells (blue). The outmost part of the placenta, the decidua basalis (pink), is derived from maternal tissue. The intermingling of maternal blood vessels with the placenta is indicated (Figure modified from Koerner et al., 2009).

It is possible, that macro ncRNAs in general do not have a strictly determined 3' end. Interestingly, also the mapped splice variants of *Airn* have different 3' ends, and the 3' end of splice variant SV3 maps 17kb downstream of the end of the originally mapped full-length *Airn* (Lyle et al., 2000; Seidl et al., 2006). For *Kcnq1ot1* also different lengths have been reported, ranging from 61kb to 91.5kb (Mancini-Dinardo et al., 2006; Pandey et al., 2008). In the study which determined a length of 91.5kb, the authors however state, that steady-state levels of *Kcnq1ot1* dropped after 61kb (Pandey et al., 2008). It is technically challenging to determine the exact length of macro ncRNAs, as a detection by RNA blots does not result in a discrete band but rather in a high molecular weight smear due to their length. Therefore, length determination is mostly limited to focal analyses by PCR or RNase protection assays and a quantitative comparison of different PCR or RNase protection assays is difficult due to differences in assay length or amplification or hybridisation efficiencies. By RETA, macro ncRNAs exhibit a gradual decline in signal intensities from the 5' to the 3' end (Ru Huang, unpublished). Although this 5' to 3' slope could be due to technical reasons, it also could reflect a real property of macro ncRNAs.

For both, *Kcnq1ot1* and *Airn*, a whole range of polyadenylation sites within the transcribed region have been reported (Pandey et al., 2008; Seidl, 2006). They could be used with varying efficiency in general, and also with different specificities in different tissues. For *Airn* in RΔ800 ES cells I showed a more pronounced 5' to 3' slope which could indicate, that deletion of the tandem direct repeats led to a more favoured usage of the polyadenylation sites located more 5' compared to wildtype ES cells (see Fig. 52).

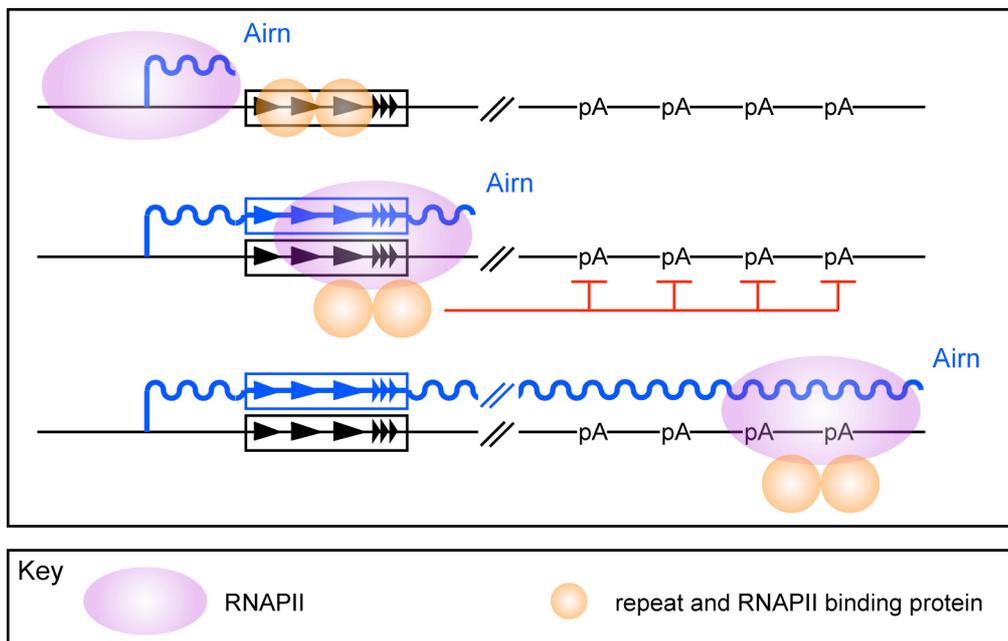


Fig. 52: Model for the length determination of *Airn* by the tandem direct repeats. A hypothetical protein could bind to the tandem direct repeats (boxed black arrows) which is then loaded onto RNA polymerase II (RNAPII) transcribing *Airn*. This in turn enables RNAPII to ignore the polyadenylation signals (pA) within the gene body of *Airn*, thus leading to full-length *Airn* transcripts. See key for further details.

3.3.4 Deletion of the tandem direct repeats does not influence the nuclear localisation of *Airn*

Wildtype unspliced *Airn* is localised in the nucleus as shown by nuclear-cytoplasmatic fractionation followed by qPCR detection as well as RNA FISH (Braidotti et al., 2004; Seidl et al., 2006; Terranova et al., 2008). Using nuclear-cytoplasmatic fractionation I showed, that the cellular localisation of wildtype *Airn*

and R Δ 800 *Airn* is unchanged. As already described in the introduction, it is unclear, if the lack of export to the cytoplasm is of functional consequence, or, if it is only due to the lack of splicing, as *Airn* splice variants are efficiently exported to the cytoplasm and mRNA genes mutated to inhibit splicing show nuclear retention (Custodio et al., 1999; Ryu and Mertz, 1989; Seidl et al., 2006). The results of an unchanged nuclear localisation of *Airn* in the R Δ 800 cells are in agreement with the fact, that neither changes in the steady-state levels of unspliced *Airn* nor increased splicing was detectable.

3.3.5 Deletion of the tandem direct repeats leads to a less efficient but not completely abolished imprinted expression of *Igf2r*

To analyse imprinted expression of *Igf2r* in R Δ 800 ES cells, I used the acquisition of the DNA methylation on the sDMR as an indirect indication. Furthermore I analysed total *Igf2r* levels by qPCR as well as allelic *Igf2r* levels by non-quantitative PCR and qPCR. The differential DNA methylation on the sDMR was still gained during *in vitro* differentiation of R Δ 800 ES cells, indicating that *Igf2r* still showed imprinted expression. Quantification of total *Igf2r* levels did not give a conclusive result, as no consistent difference between *Igf2r* in wildtype and R Δ 800 ES cells could be found. Non-quantitative and quantitative PCR using the SNP in exon 12 of *Igf2r* however revealed, that imprinted expression of *Igf2r* seems to be compromised. In the non-quantitative RT-PCR followed by restriction digestion with PstI, the paternal allele was seen more strongly in differentiated R Δ 800 ES cells or during later stages of differentiation. The quantitative assay showed, that in R Δ 800 ES cells the ratio of the maternal to the paternal allele still increases during differentiation, indicative of an expression bias between the parental alleles, as reported previously (Latos et al., 2009). The increase was however at least twofold lower at d5 in all three differentiation sets and at d14 in two differentiation sets compared to wildtype cells. This indicates, that the tandem direct repeats play a role in imprinted expression of *Igf2r*. As we still do not know, if *Airn* regulates imprinted expression of *Igf2r* by

transcriptional interference or by RNA-directed targeting, it is still not possible, to determine the exact role of the tandem direct repeats, however, models can be designed for both situations.

First I discuss the potential role the tandem direct repeats could play in the transcriptional interference model. Shearwin et al. list in their review several models for transcriptional interference (Shearwin et al., 2005). In the 'promoter competition model', binding of RNA polymerase II (RNAPII) to one promoter inhibits binding of RNAPII to the second promoter. This model however can be excluded due to two reasons. First, in AirT mice, both, the truncated *Airn* as well as *Igf2r* can be expressed from the same allele (Sleutels et al., 2002). Second, in wildtype cells the paternal *Igf2r* allele is not completely silenced but shows a low level expression when it shows imprinted expression (Latos et al., 2009). In the 'occlusion model' RNAPII cannot bind to the interfered promoter due to its transient occupation by the interfering RNAPII, when it transcribes through. Also this model can be excluded, because as mentioned above, the paternal *Igf2r* allele shows a low but readily detectable expression levels *in cis* to *Airn* (Latos et al., 2009). In the 'sitting duck interference model', the interfered promoter is slow in the transition from the initiation to the elongation RNAPII complex and gets displaced by the arriving interfering RNAPII. This model could be attributable to the situation of *Igf2r* and *Airn*, assuming that at a low frequency the RNAPII on the *Igf2r* promoter is fast enough to transit into the elongation complex before it gets displaced by the RNAPII transcribing *Airn*. In the 'collision model', the clashing of the two RNAPII in the region of the transcriptional overlap leads to a stalling of both RNAPII. One of the RNAPII could be helped by host factors to win the situation and continue transcription. Also the 'collision model' would fit the situation of *Igf2r* and *Airn*, if the *Airn* transcribing RNAPII has better host factors, allowing it to finish the transcription process more often than the RNAPII transcribing *Igf2r*. Theoretically, a role for the tandem direct repeats could be imagined in both, the 'sitting duck interference' as well as in the 'collision model'. In the 'sitting duck interference', the influence of the tandem direct repeats on the length of *Airn* as discussed above, could directly lead to the situation, that in R Δ 800 ES cells fewer RNAPII reach the *Igf2r* promoter, thus allowing more

RNAPII bound to the *Igf2r* promoter to transit from the initiation to the elongation form of RNAPII and thus to transcribe *Igf2r*. No differences in the steady state levels of *Airn* were seen when using the 'Air middle' qPCR assay, which is located 25kb downstream of the *Igf2r* promoter. It therefore has to be postulated that the deletion of the tandem direct repeats led to a reduced transcription rate combined with a higher stability of *Airn* transcripts, which I consider unlikely. Therefore, a similar amount of *Airn* transcribing RNAPII transits the *Igf2r* promoter in wildtype as well as in RΔ800 cells, which means that the 'sitting duck interference model' is not applicable. However, for the 'collision model' it is easy to imagine that the tandem direct repeats play a crucial role in the battle of the two stalled RNAPII. The model shown in Fig. 53A would predict, that certain proteins bind to the tandem direct repeats and are then loaded onto the RNAPII transcribing *Airn*. Binding of those proteins renders the RNAPII transcribing *Airn* into a variant, which is stronger than the RNAPII transcribing *Igf2r*, thus displacing it in most but not all cases, still allowing for low level expression of *Igf2r*. Deletion of the tandem direct repeats however weakens the *Airn* transcribing RNAPII, so it loses the battle more often, resulting in a higher level of *Igf2r* from the paternal allele and thus in a relaxed imprinted expression. However, the *Airn* transcribing RNAPII in the RΔ800 ES cells does not lose the 'fight' with the *Igf2r* transcribing polymerase too often, as no reduction of *Airn* steady-state levels at the 5' end was detectable. However, as *Igf2r* has a sevenfold longer half-life than *Airn* (Seidl, 2006), also small changes in the amount of succeeding RNAPII transcribing *Igf2r* will have a substantial impact on *Igf2r* steady state levels.

Now I discuss the potential role of the tandem direct repeats in imprinted expression, within the framework of a model whereby *Airn* acts by RNA-directed targeting. As mentioned above, *Xist*, the ncRNA involved in X chromosome inactivation, acts by RNA-directed targeting. *Xist* induces the formation of an RNAPII-deficient nuclear compartment (Chaumeil et al., 2006; Okamoto et al., 2004). At the beginning of X chromosome inactivation, the genes are still active and located to the periphery of this compartment in contact with the transcription machinery. As soon as gene silencing occurs, the genes are localised into the repressive compartment (Chaumeil

et al., 2006). *Xist* furthermore leads to the recruitment of Polycomb group proteins (PcG) of the Polycomb repressive complex 1 (PRC1) and PRC2, catalysing H2AK119ub1 (monoubiquitination of lysine 119 of histone H2A) and H3K27me3, respectively (de Napoles et al., 2004; Fang et al., 2004; Leeb and Wutz, 2007; Plath et al., 2004; Silva et al., 2003). Further proteins recruited by *Xist* are the Trithorax group protein ASH2L, the scaffold attachment factor SAF-A and the histone variant macroH2A (Costanzi and Pehrson, 1998; Helbig and Fackelmayer, 2003; Mermoud et al., 1999; Mietton et al., 2009; Pullirsch et al., 2010; Rasmussen et al., 2000). Surprisingly, the establishment of the repressive chromatin compartment as well as the recruitment of the proteins described above can also occur in cells carrying *Xist* with a deletion of the repeat A (Pullirsch et al., 2010; Schoeftner et al., 2006). As described above, *Xist* deleted for the repeat A cannot induce gene silencing and those cells also cannot relocate genes into the repressive compartment, indicating that the main function of the repeat A within *Xist* is, to lead to this relocation. How this relocation is achieved exactly and which role repeat A plays, is however still unknown. Another ncRNA shown to play a role in gene regulation is *Hotair*, a spliced ncRNA transcribed from the HoxC locus which binds PcG proteins and leads to the repression of genes from the HoxD cluster *in trans* (Rinn et al., 2007). Furthermore it was shown in the placenta, that *Airn* binds G9A, which in turn catalyses H3K9me2, and both, *Airn* as well as H3K9me2 were found to be present at the promoter region of *Slc22a3*, indicating that *Airn* silences *Slc22a3* in placental tissues by the recruitment of G9A (Nagano et al., 2008). In an analogous way, the tandem direct repeats within *Airn* could serve as binding platform for RNA-binding proteins, which then in turn directly or indirectly by recruiting other proteins lead to silencing of *Igf2r* (see Fig. 53B). The effect of the deletion of the tandem direct repeats on imprinted expression of *Igf2r* is however not completely penetrant. This effect could be explained by the existence of other regions within *Airn* binding further proteins which partially could compensate the loss of the tandem direct repeats.

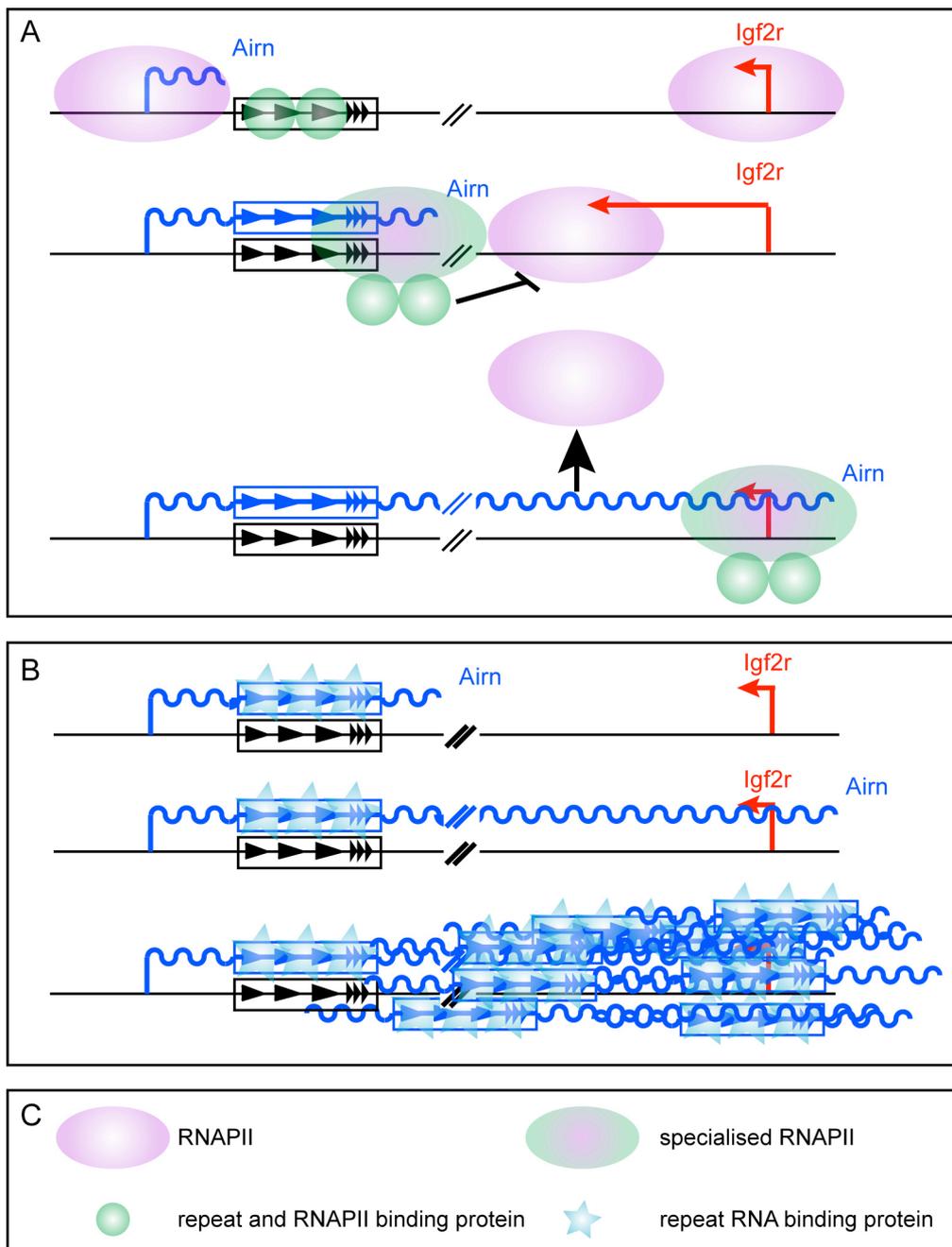


Fig. 53: Models how the tandem direct repeats could influence imprinted expression of *Igf2r*. (A) If *Airn* acts by transcriptional interference, proteins could bind to the tandem direct repeats and be loaded onto the RNAPII transcribing *Airn*. Also *Igf2r* is transcribed by RNAPII. The two RNAPII collide and become stalled. The proteins loaded onto the *Airn* transcribing RNAPII act as host factors to allow this polymerase to continue transcription, whereas the RNAPII transcribing *Igf2r* becomes displaced. (B) If *Airn* acts by RNA-directed targeting, an RNA binding protein binds to the tandem direct repeats within *Airn*, influencing either directly or indirectly the suppressing function of *Airn* on *Igf2r*.

As described in the introduction, there are several indications, that *Airn* could act in different ways in embryonic compared to extraembryonic tissues. In embryos it could

act by transcriptional interference to induce imprinted expression of *Igf2r*, whereas in extraembryonic tissues it could act by RNA-directed targeting to additionally lead to imprinted expression of *Slc22a2* and *Slc22a3*. It is therefore conceivable, that the tandem direct repeats also have different roles in the embryo compared to extraembryonic tissues. To analyse an effect of the deletion of the tandem direct repeats on imprinted expression of *Slc22a2* and *Slc22a3*, the knock out mouse for whose generation, I made and characterised R Δ 800 ES cells for blastocyst injection, will provide an important tool in the future.

3.4 The role of the CpG island in the *Igf2r* ICE

CpG islands are defined as a region longer than 200bp with a G + C content of 50% and a CpG frequency of observed to expected of at least 0.6 (Gardiner-Garden and Frommer, 1987; Larsen et al., 1992). Approximately 60% of all promoters are associated with a CpG island, and this includes all housekeeping genes as well as approximately half of all tissue-specifically expressed genes (Antequera, 2003; Illingworth and Bird, 2009). Not all CpG islands map to annotated transcriptional start sites but might serve as promoter regions for yet unknown transcripts which could not only be protein-coding genes but also ncRNAs (Illingworth and Bird, 2009). CpG islands normally cover the DNA sequence upstream of the transcriptional start site, the first exon and the first introns (Gardiner-Garden and Frommer, 1987). A situation where the CpG island is localised entirely downstream of the transcriptional start site like for *Airn* is therefore unusual (Stricker et al., 2008). Whereas non-CpG island promoters normally have just one precisely defined transcriptional start site, CpG island promoters initiate transcription from multiple positions (Sandelin et al., 2007). Whereas CpG dinucleotids in the bulk genome are DNA methylated, CpG dinucleotids in CpG islands are normally free of DNA methylation and display an open chromatin state with acetylated histones H3 and H4, deficiency of the linker histone H1, positioned nucleosomes and a nucleosome-free region (Antequera, 2003). In certain cases, CpG islands however can get DNA methylated during normal development, these are CpG islands in imprinted clusters, associated with X

chromosome inactivation, and MAGE genes, the latter are only active and unmethylated in the male germline (Antequera, 2003). In cancers, also other CpG islands can acquire DNA methylation which correlates with transcriptional repression of the associated genes (Antequera et al., 1990; de Bustros et al., 1988; Esteller, 2002). Furthermore, CpG islands contain many binding sites for transcription factors like SP1, which binds to CCCGCC sequences (Suske, 1999) and their binding is often independent of the transcriptional state of the associated gene (Antequera, 2003).

In this work I addressed the function of CpG islands, and my model is *Airn*, where the CpG island is localised downstream of the transcriptional start site and for which a promoter-region upstream of the transcriptional start site crucial for expression has been determined (Lyle et al., 2000; Stricker et al., 2008). To do this, I generated ES cells with a deletion of the CpG island (CpG Δ) on the paternal allele and analysed them using the ES cell *in vitro* differentiation system.

3.4.1 Deletion of the CpG island leads to DNA methylation during cell culturing

The original CpG island of *Airn* as defined by the UCSC genome browser comprises 1049bp and contains 82 CpG dinucleotids. Cells with a CpG Δ allele still have 107bp and 9 CpG dinucleotids of the original CpG island left which do not fulfil the criteria for a CpG island anymore. The CpG Δ led to a gain of DNA methylation on the paternal allele during culturing. Comparison of the 'younger' DMK14 ES cells, which have the CpG Δ , but where the selection cassette has not been removed, with the 'older' DMK15 ES cells after selection cassette removal, revealed, that the longer the cells were in culture, the higher was their degree of DNA methylation on the paternal allele. During *in vitro* differentiation over 14 days, the DNA methylation did not increase any further. The lack of an additional gain of DNA methylation could have two reasons. First, the two weeks in culture were not long enough to see an additional increase in DNA methylation, especially as differentiated ES cells proliferate slower than undifferentiated ES cells. Second, the remaining alleles which

are free of methylation could be, due to an unknown reason, resistant to a gain of DNA methylation. It would be interesting to see, if the alleles which are detected as the unmethylated fragment by DNA blotting are really completely free of DNA methylation, as here only a single NotI site was used for the analysis. However it also could be, that all paternal alleles gained a substantial amount of DNA methylation but the diagnostic NotI site is still unmethylated in some cells. In the future, this question could be answered using bisulfite sequencing, which assays multiple CpGs.

The gain of DNA methylation upon deletion of the CpG island indicates, that one function of the *Airn* CpG island is, to keep the *Airn* promoter region free of DNA methylation. But why do the CpG Δ ES cells gain DNA methylation? This question is not easy to answer, because as described above, it is still not clear, which mechanisms protect CpG islands from DNA methylation. For the *Aprt* gene it was shown, that deletion of SP1 binding sites led to a gain of DNA methylation on its promoter (Brandeis et al., 1994; Macleod et al., 1994; Mummaneni et al., 1998). Also within the CpG island of *Airn*, several SP1 binding sites can be predicted using bioinformatical analysis which were deleted in the CpG Δ cells (Seidl et al., 2006). However, a homozygous deletion of the *Sp1* gene does not lead to general gain of DNA methylation on CpG islands, nor to a gain of DNA methylation on the *Aprt* CpG island, indicating, that SP1 is not crucial for keeping CpG islands free of DNA methylation (Marin et al., 1997). It would be interesting to see, if deletion of the CpG island also led to a decrease in H3K4me3, as this histone modification inhibits interaction with DNMT3L (Jia et al., 2007). Loss of this modification therefore could explain the gain of DNA methylation on the *Airn* ICE upon deletion of the CpG island.

3.4.2 Deletion of the CpG island reduces expression of *Airn*

I showed, that CpG Δ ES cells show 10-80% of *Airn* steady state levels 5' of the deletion and 0-15% 3' of the deletion. Whereas the three qPCR assays 3' of the deletion, detected successively decreasing amounts of *Airn* with the distance from

the 5' end, the two qPCR assays 5' of the deletion gave a complex picture. 'START', the assay localised more 5', showed a lower amount of *Airn* compared to 'RP11', which is localised more 3' but partially overlaps 'START' (see Fig. 45B). The higher levels for 'RP11' compared to 'START' could indicate, that the transcriptional start site is shifted more 3' in the CpG Δ cells compared to wildtype ES cells, leading to a less efficient detection of *Airn* using the 'START' assay compared to 'RP11'. As described above, CpG island promoters often have several transcriptional start sites, which is also the case for *Airn*, for which three transcriptional start sites have been mapped (Seidl et al., 2006). Therefore it is also possible, that the deletion of the CpG island influences the position of the transcriptional start site.

The 2-15% of *Airn* steady state levels detected using the 'AirT3' and 'Air middle' assay which detect *Airn* at 4kb and 53kb respectively could originate from those cells with an unmethylated *Airn* promoter. The drastic decrease in steady-state levels 3' of the deletion could indicate, that in the majority of cells, RNAPII still can initiate transcription of *Airn* in CpG Δ cells, however it cannot successfully elongate those transcripts. If this inability in elongation is due to deletion of the CpG island or due to the gain of DNA methylation on the *Airn* promoter cannot be concluded from my results. Future experiments could include an RNAi-mediated knock down of the maintenance methyltransferase DNMT1 or a re-targeting of the CpG Δ in *Dnmt1*^{-/-} ES cells (Lei et al., 1996). If a reduction in the DNA methylation level on the *Airn* promoter carrying a CpG deletion would result in an increased amount of *Airn* 3' of the deletion, it would indicate, that the function of the CpG island is to protect the *Airn* promoter from DNA methylation but that it is not needed for transcriptional elongation. If the DNA methylation levels are reduced without an increase in *Airn* levels 3' of the deletion, the CpG island would have additional functions in *Airn* transcription, for example by containing binding sites for crucial transcription factors enabling full-length transcription of *Airn*.

3.4.3 Deletion of the CpG island leads to biallelic expression of *Igf2r*

CpGΔ cells, in contrast to wildtype cells and RΔ800 cells, did not gain an sDMR on the *Igf2r* promoter during differentiation. Furthermore, total *Igf2r* levels were increased in CpGΔ cells compared to wildtype cells. Analysis of the allelic ratios of the parental alleles showed biallelic expression not only at d0, where in wildtype cells *Igf2r* shows biallelic expression as well, but also at d5 and d14, where in wildtype ES cells *Igf2r* shows imprinted expression. Taken together, these results show that in CpGΔ cells *Airn* fails to induce imprinted expression of *Igf2r*. As the majority of *Airn* transcripts in CpGΔ cells are shorter than 4kb, this result is in agreement with the fact, that a truncation of *Airn* to 3kb resulted in loss of imprinted expression of *Igf2r* (Sleutels et al., 2002). The 2-15% of *Airn* levels detected using 'AirT3' and 'Air middle' could either be not sufficient to induce imprinted expression of *Igf2r*, in case they represent a low level of *Airn* expression in the entire cell population. If the 2-15% of *Airn* levels represent normal *Airn* expression only in a subset of cells, for example only in those, which have an *Airn* promoter allele completely free of methylation, it would still be possible, that those cells show imprinted expression of *Igf2r*, which however would be masked by the majority of cells showing biallelic expression of *Igf2r*. Also here, further information potentially could be gained by an analysis in ES cells lacking DNMT1. Further insights could be gained by the analysis of *Airn* expression on a single cell level, for example by using RNA FISH, to see if a small percentage of cells consistently expresses *Airn* or if all cells express a low level of *Airn*.

3.5 The roles of the *Igf2r* ICE

In this work I demonstrated, that on the chromosomal level the ICE does not play a role, at least not with respect to DNA FISH asynchrony. Certain elements within the ICE however seem to control various aspects of the biology of genomic imprinting. The tandem direct repeats have an influence on the length of *Airn* as well as on the *Airn*-mediated silencing of *Igf2r*. The CpG island is clearly necessary for functional

Airn expression. If this role is due to the prevention of DNA methylation or due to the presence of regulatory elements like transcription factor binding sites cannot be concluded from this work. Nevertheless, I clearly demonstrated, that despite the defined promoter region upstream of the transcriptional start site (Lyle et al., 2000; Stricker et al., 2008), the downstream CpG island is mandatory for *Airn* expression and its role in inducing imprinted expression.

4. MATERIALS AND METHODS

Please note that manufacturers and buffer/solution compositions are stated at the end of this section.

4.1 Methods

4.1.1 Mini prep of plasmid DNA from bacteria by alkaline lysis

Single colonies were inoculated in 3ml LB broth medium with appropriate selection antibiotic (50µg/ml ampicillin, 50µg/ml kanamycin), the liquid cultures were incubated overnight at 37°C. On the next morning, 1.5ml of the liquid culture were transferred into 1.5ml microfuge tubes and spun for 30sec at 13.2krpm in a microcentrifuge. The supernatant was discarded and the pellet was resuspended on ice in 100µl of ice cold Alk-1. 200µl of Alk-2 were added to the suspension, mixed by rapid inversion and incubated for 5min on ice. The lysis reaction was stopped by addition of 150µl of Alk-3, rapid inversion and 5min incubation on ice. After 10min centrifugation at 13.2krpm at 4°C, the supernatant was transferred into a fresh 1.5ml microfuge tube. The DNA was precipitated by addition of 0.6 volumes isopropanol, vortexing, 2min incubation at room temperature followed by a centrifugation for 20min at room temperature at 13.2krpm. The DNA was washed two times with 70% ethanol and dissolved in an appropriate volume of TE buffer.

4.1.2 Midi prep of plasmid and cosmid DNA from bacteria

Single colonies were inoculated in a liquid 2-step-culture of LB broth medium with appropriate selection antibiotic. In the morning, a 3ml culture was inoculated and incubated for 8-10 hours. Afterwards, this 3ml culture was used to inoculate a larger culture with a final volume of 50ml (plasmids) or 250ml (cosmids). The plasmid DNA was isolated using the Promega PureYield™ Plasmid Midiprep System using vacuum purification according to the kit manual. For the larger cosmid cultures, the alternative lysate clearing protocol was used.

4.1.3 Maxi prep of plasmid DNA from bacteria for preparing targeting vectors

Single colonies were inoculated in a liquid 2-step-culture of LB broth medium with appropriate selection antibiotic in a final volume of 250ml (see above). The plasmid DNA was isolated using the EndoFree Plasmid Maxi kit according to the kit manual.

4.1.4 Mini and maxi preps of BAC DNA from bacteria

Those were done using the BACMAX™ DNA Purification Kit according to the kit manual.

4.1.5 Transformation using CaCl₂ competent bacteria

DH5 α CaCl₂ competent *E. coli* were thawed for 30min on ice. After addition of the ligation solution (see later) they were incubated on ice for additional 30min. After a 1min long heat shock in a 42°C waterbath and 2min incubation on ice 250ml LB broth medium pre-warmed to 37°C were added. Then, the bacteria were incubated for 1h at 37°C shaking at 300rpm in a thermomixer and afterwards plated on Circlegrow agar plates with appropriate selection antibiotic (50 μ g/ml ampicillin, 50 μ g/ml kanamycin or 12.5 μ g/ml chloramphenicol respectively). For blue-white selection, 100 μ l 0.1M IPTG and 20 μ l of 50mg/ml X-Gal in N,N'-dimethyl-formamide were spread over the surface of an LB plate and incubated for 30min at 37°C prior to use. Plated bacteria were incubated overnight at 37°C.

4.1.6 Restriction digests

1-35 μ g of DNA were digested using an appropriate enzyme with an amount of units dependent on the incubation time and the DNA amount in the supplied buffer for 3h or overnight at working temperature of the enzyme.

4.1.7 Gel electrophoresis

DNA was loaded onto a 0.8%-2% agarose gel together with a DNA-marker. Gel electrophoresis was carried out in 1xTAE at 7V/cm. The gel was stained in ethidium bromide solution (1mg/l). Pictures were taken using the Alphamager.

4.1.8 DNA sequencing

DNA was sequenced using standard (T7, SP6, T3) or custom sequencing primers at VBC-Biotech Service GmbH, Vienna, Austria or AGOWA GmbH, Berlin, Germany.

4.1.9 Enzymatic DNA modifications for cloning of DNA

Blunt-ending of 5' protruding ends was done using 5u of Klenow fragment in either the supplied buffer or a compatible restriction enzyme buffer and 0.05mM dNTP mix for 10min at 37°C. The reaction was stopped by heating at 75°C for 10min.

Blunt-ending of 3' protruding ends was done using 1u of T4 DNA Polymerase in either the supplied buffer or a compatible restriction enzyme buffer and 0.1mM dNTP mix at room temperature for 5min. The reaction was inactivated by heating at 75°C for 10min.

5'-termini of vector backbones were dephosphorylated using Calf Intestine Alkaline Phosphatase in either the supplied buffer or an appropriate restriction buffer. The reaction was incubated for 30min at 37°C and stopped by an incubation for 15min at 85°C.

4.1.10 Gel elution

DNA was eluted from agarose gels using the Wizard® SV Gel and PCR Clean-Up System according to the kit manual in a suitable volume of nuclease-free water.

4.1.11 Ligation

Ligation reactions were done using either the pGEM®-T Easy Vector System according to the kit manual or using T4 DNA ligase (1u for sticky end ligation, 5u for blunt end ligation) and the supplied buffer in 10-20µl reactions at 16°C overnight. Ligation reactions were set in 1:1 and 1:3 vector-insert molar ratios.

4.1.12 PCR

Primers were designed using Primer3 (Rozen and Skaletsky, 2000) (see Table 9) and synthesised by VBC-Biotech Service GmbH (Vienna, Austria) at Standard DNA-Oligonucleotide-Synthesis – RP-HPLC-Purification. The primers were delivered in lyophilised form and dissolved in TE buffer to a stock concentration of 100pmol/µl. 1:10 dilutions were used as working solutions.

If primers had to be 5'-termini phosphorylated, this was done using 10u of T4 polynucleotide kinase in the supplied reaction buffer A and 0.2mM ATP at 37°C for 20min. The reaction was stopped by heating to 75°C for 10min.

name	sequence	start (bp)	stop (bp)	chr./ Acc.	product size (bp)
120970A-F	AGCCTCAGCCACTCAAGATG	12938949	12938968	17	442
120970A-R	AAGCACCACAGTTCATGCTG	12939372	12939391	17	
H19P13kb-F	CACCACTCGAAGATGGTGTCTAAGG	149775638	149775662	7	467
H19P13kb-R	CCAGACCTCCTAGAACCTGAGGAAA	149775165	149775189	7	
KODEL F	ACGTTTTGGGGTTGCTGGGTG	12957553	12957573	17	325
KODEL R	CTCCAGATTGGCAGTTCACAC	12957249	12957269	17	
PrF	TTGTGATGGGTGTGAACCAC	436	455	NM_008084.2	796
PrR	GTGGGTGCAGCGAACTTTAT	1213	1232	NM_008084.2	
Dk1F	GCTTTGGAATTCCTGATGGA	110766865	110766884	12	975
Dk1R	CACCTGGCTAACCAATGAGCA	110767820	110767839	12	

Table 9: Used PCR-primers. The bp positions for start and stop of the oligos refer to the Mouse July 2007 (mm9) assembly from the UCSC Genome Browser or to the indicated accession numbers.

PCR reactions were performed in 25µl reaction volume with 0.1µl GoTaq DNA polymerase (5U/µl), 0.5µl dNTP mix (10mM), 5µl 5x GoTaq Flexi Buffer, 2.5µl MgCl₂ (25mM), 1.25µl forward primer (10pmol/µl), 1.25µl reverse primer (10pmol/µl) and 4µl betaine (5M). As template 1ng of plasmid, 2ng of cosmid/BAC or 10ng of genomic DNA were used.

The reaction was carried out in a Peltier Thermal Cycler PTC-200.

Initial denaturation at 94°C for 3min was followed by 35 cycles of 96°C for 10sec, 94°C for 30sec, 58-60°C for 1min and 72°C for 1min/kb and a final extension step of 72°C for 5min.

4.1.13 RT-PCR

RT-PCR reactions were performed in 50µl reaction volume with 0.2µl GoTaq DNA polymerase (5U/µl), 1µl dNTP mix (10mM), 10µl 5x GoTaq Flexi Buffer, 4µl MgCl₂ (25mM), 2.5µl forward primer (10pmol/µl), 2.5µl reverse primer (10pmol/µl) (see Table 10). As template 2µl of cDNA or 2µl of the –RT reaction (cDNA reaction in the absence of reverse transcriptase) were used.

The reaction was carried out in a Peltier Thermal Cycler PTC-200.

Initial denaturation at 95°C for 3min was followed by 35 cycles of 95°C for 45sec, 59°C for 45sec, 72°C for 45sec and a final extension step of 72°C for 5min.

If PCR products were subjected to a restriction digest, they were cleaned using the Wizard[®] SV Gel and PCR Clean-Up System.

name	sequence	start (bp)	stop (bp)	chr./ Acc.	product size (bp)
Ex12cDNAF	TTCACAGGTGAGGTGGACTG	1500	1519	NM_010515.2	541
Ex12cDNAR	CCGTGCAGTTCTCTCCTTCT	2021	2040	NM_010515.2	

Table 10: Used RT-PCR primers. Details as in Table 9.

4.1.14 Long PCR/ Proof Reading PCR

Long PCRs and PCRs where a proof reading function of the DNA polymerase was necessary were done using the Long PCR Enzyme Mix (Fermentas) according to the kit manual. For used primers see Table 11. For long PCRs (>8kb), cycling conditions were the following, using a combined annealing and elongation step of 68°C. Initial denaturation at 94°C for 2min was followed by 10 cycles of 96°C for 20sec, 68°C for 30sec + 1min/kb. Afterwards, 25 cycles with 96°C for 20sec, 68°C for 30sec + 1min/kb +10sec/cycle were performed. The last step was a final extension step of 68°C for 10min.

For proof reading PCRs (<1kb), the following cycle conditions were used: The initial denaturation at 94°C for 2min was followed by 35 cycles of 96°C for 10sec, 94°C for 30sec,

61°C for 1min, 68°C for 1min. Afterwards a final extension step of 68°C for 5min was carried out. Before sequencing, PCR products were cleaned using the Wizard® SV Gel and PCR Clean-Up System.

name	sequence	start (bp)	stop (bp)	chr.	product size (bp)
CpGdel-F2	TGGAACCCCTTCCTTTGCGGAATC	12934391	12934413	17	8296
CpGdel-R2	TGCATGAGGGTGCCACACTCCT	12935544	12935565	17	
Igf2P-F6	CTTAGTAAGATCTATCTGGCTGCTATTC GAGGTG	149845972	149846005	7	9764
Igf2P-R6	GACAGACTCTAGCATAGCCTTTATAGAG GGAACA	149836242	149836275	7	
Cybb1F	AGGGTTTCATGATGGCTAAAACCTG	9016604	9016627	X	850
Cybb1R	GGTCTCTTTCTCAACACCCGAAC	9017431	9017453	X	
Cybb2F	GGCCCAACAGATCTATCCAGTTT	9027465	9027487	X	362
Cybb2R	CGCTGGAAACCCTCCTATGG	9027807	9027826	X	

Table 11: Used PCR-primers for long PCR. Details as in Table 9.

4.1.15 Cloning of pH19P13kb (probe for Southern blotting)

This plasmid was generated by doing a PCR using as primers H19P13kb-F and H19P13kb-R (see Table 9) and genomic DNA of pMEFs (FVB background) as template, PCR conditions as described above and 60°C annealing temperature. The PCR product was cloned into pGEM-T Easy (see above).

4.1.16 Cloning of plgf2P6 (for DNA FISH)

This construct was cloned by doing a long PCR using as primers Igf2P-F6 and Igf2P-R6 (see Table 11) and the BAC RP23-50N22 as template. The PCR product was cloned into pGEM-T Easy (see above).

4.1.17 Cloning of pNS13DTAmNeoinpBS (targeting vector for Ex12 SNP)

This targeting vector was cloned in collaboration with Florian M. Pauler and Laura Steenpass. cos940PS was digested with BstXI and ClaI. A 7.1kb fragment containing *Igf2r* exons 12-18 was isolated, blunt-ended with T4 DNA polymerase and ligated into pBSIIKS(-), which was digested with EcoRV and HincII, to give rise to pNSin13. pNSin13 was digested using XhoI and Eco47III to remove some cosmid vector backbone and to further shorten the homology arm (*Igf2r* exons 15-18 were removed), blunt-ended using Klenow fragment and re-ligated, to give rise to pNSin13real-5k. Now, the homology region comprises 5kb, corresponding to 1290988-12914937bp of chr.17 (July 2007 mm9 assembly, UCSC genome browser). In the next step, the Diphtheria toxin antigen gene from pDTA-in-pBSKS was isolated using a KpnI HincII digest, blunt-ended with T4 DNA polymerase and insert into the

SmaI site of NSin13, giving rise to pNS13DTA. A 1kb HindIII fragment from pNSin13-real5k surrounding *Igf2r* exon 12 was subcloned into the HindIII site of pBSIIKS(-). The C- to T-change that mutates the PstI site in *Igf2r* exon 12 was generated by site-directed mutagenesis using the QuikChange XL Site-Directed Mutagenesis Kit. The mutated HindIII fragment was cloned back into NSin13DTA, generating pNS13DTAm. The selection cassette from pTKneoloxP511 was isolated using a NotI KpnI digest, blunt-ended with T4 DNA polymerase and inserted into pNSin13DTAm which was digested with BsrGI and blunt-ended with T4 DNA polymerase, generating pNS13DTAmNeoinpBS. Before electroporation into ES cells, the vector was linearised using NotI.

4.1.18 Cloning of p Δ 800 (targeting vector for repeat deletion)

The original p Δ 800 targeting vector used to generate DMK1/2, DMK3/4 and AMK9/10 was cloned by Paulina Latos, Yvonne Schichl and Stefan Stricker. A 7.3kb XbaI-XbaI fragment from cos940PS containing the *Airn* transcriptional start and *Igf2r* exons 3-4 was subcloned into pBSIIKS(-) which was digested with BamHI and SmaI and blunt-ended, generating pBSXbaXba. pBSXbaXba was digested using BsmI and SacI, re-ligated, thereby reducing the homology arm on one side to a final length of 6.5kb (corresponding to 12931302-12937790bp of chr. 17, July 2007 mm9 assembly, UCSC genome browser), generating pXSB2. A selection cassette (containing a loxP site followed by an HSV-*tk* promoter driving a neomycin gene and stopped by a SV40pA signal followed by an HSV-*tk* promoter driving the HSV *thymidine kinase* gene stopped by an HSV-*tk*-pA signal followed by a second loxP site) was cloned into the NheI site in intron 3 of *Igf2r*, generating pXSB2Y2. pXSB2Y2 was digested using NsiI and SacII to delete 700bp of the *Airn* repeats (12934848-12935547bp of chr. 17, July 2007 mm9 assembly, UCSC genome browser), blunt-ended using T4 DNA polymerase and re-ligated, creating p Δ 800. Before electroporation into ES cells, the targeting vector was linearised with ClaI.

I cloned a modified version of p Δ 800 to generate AMK11 ES cells. I digested pXSB2 using SacII and NsiI, to create the 700bp repeat deletion. After blunt-ending using T4 DNA polymerase the plasmid was re-ligated, creating p Δ 800-cas. The loxP-*Pgk1*-neo-*Pgk1*pA-loxP selection cassette was isolated from pKSloxPNT (a gift from Maria Sibilla) using EcoRI and Sall, blunt-ended using Klenow fragment and inserted into the NheI site (blunt-ended using Klenow fragment) in intron 3 of *Igf2r* of p Δ 800-cas, creating p Δ 800A (= p Δ 800+cas). Before electroporation into ES cells, the targeting vector was linearised with XmiI.

4.1.19 Cloning of pCpGΔ (targeting vector for CpG island deletion)

An inverted long PCR was made using pXSB2 (see above) as template and CpGdel-F2 and CpGdel-R2 (see Table 11) as primers (primers were phosphorylated) to generate a 1129bp deletion of the *Airn* CpG island (12934414-12935543bp of chr. 17, July 2007 mm9 assembly, UCSC genome browser). The PCR fragment was blunt-ended using T4 DNA polymerase and ligated together, generating pCpGΔ-casPCR. The homology region for this targeting vector therefore is the same as for pXSB2. A 1.5kb BstXI fragment surrounding the deletion was completely sequenced to exclude polymerase errors. For further steps a clone carrying a single mismatch at position 12935945 was used. pXSB2 was cut using BstXI and the sequenced BstXI fragment from pCpGΔ-casPCR was inserted into the 6.8kb BstXI backbone fragment of pXSB2, generating pCpGΔ-cas. Insertion of a selection cassette (containing a loxP site followed by a murine *Pgk1* promoter driving a neomycin resistance gene stopped by a *Pgk1*pA followed by a loxP site) was done as for pRΔ800+cas, generating pCpGΔ+cas. Before electroporation into ES cells, the targeting vector was linearised with XmiI.

4.1.20 Preparation of primary MEFs

For the preparation of feeder cells for ES cell culture, wildtype FVB, DR4 (Tucker et al., 1996) or IPΔ (Sleutels et al., 2003) and Thp (Johnson, 1974; Johnson, 1975) mice were crossed.

For DNA-FISH and S phase fractionation, heterozygous mice carrying a mutant allele (either Thp, R2Δ, AirT or IPΔ) (Johnson, 1974; Johnson, 1975; Sleutels et al., 2003; Sleutels et al., 2002; Wutz et al., 2001) of FVB background were crossed with each other or to wildtype mice.

At 13.5dpc pregnant mice were sacrificed and the embryos were dissected. For S phase fractionation and DNA-FISH, head and liver were removed. The remaining trunks or whole embryos (for feeder cells) were finely minced through a G20 Sterican needle and seeded embryo-wise onto 10cm or 3 embryos per 15cm (wildtype and DR4 feeders) cell culture dish. The cells were maintained in MEF media at 37°C in 5% CO₂ atmosphere. After 3 days cells were harvested by trypsinisation and frozen in 50% MEF media, 40% FBS and 10% dimethyl sulfoxide (DMSO).

The embryonic membranes or heads were used for genotyping by DNA blotting (see below).

4.1.20.1 Maintenance of primary MEFs

Primary MEFs were maintained in MEF media. They were passaged every third day and split at a 1:3 ratio. Cells were washed once in pre-warmed D-PBS, trypsinised at 37°C and resuspended in fresh medium.

4.1.20.2 Preparation of feeder cells for ES cell culture

Primary MEFs were expanded for three passages as described above. Afterwards, they were harvested, gamma-irradiated for 5min with 5gray/min and frozen as described above.

4.1.21 ES cell culture

ES cell lines were grown in ES cell media at 37°C in 5% CO₂ atmosphere. The media was changed every day. Cells were split according to their density every second to third day. ES cell lines were kept on gelatinised dishes with feeder layers. Wildtype feeders were used for standard growth. Before differentiation, cells were cultured for at least one passage on IPΔ/Thp feeders. During geneticin-selection, cells were grown on DR4 feeders. Differentiation in monolayers was induced by feeder-depletion for 20min, LIF-withdrawal and addition of 0.27μM retinoic acid. Cells were seeded at appropriate densities to be confluent on the day of harvesting. Differentiating cells were fed every second day and harvested by trypsinisation.

4.1.22 Gene targeting of ES cells by homologous recombination

Prior to electroporation, ES cells were feeder-depleted. 800μl of 5x10⁶ to 1x10⁷ cells per ml D-PBS were electroporated with 35μg of linearised targeting vector with 0.24kV and 500μF using the Gene Pulser®II. Afterwards, cells were plated onto DR4 feeders. Selection was started after 24 hours in ES cell media supplemented with 400μg/ml geneticin for D3 cells and 300μg/ml for A9 cells (selection media). The selection media was changed every day and selection was performed for 6-8 days. Single colonies were picked, trypsinised, transferred to a 24-well plate with wildtype feeder cells and grown for 2-3 days in non-selective media. Clones were trypsinised, half of the suspension was frozen in a 96-well plate with 50% ES cell media, 40% FBS and 10% DMSO. The second half of the cell suspension was transferred to gelatinised 24-well plates and grown to confluency for DNA isolation followed by genotyping using DNA blots (see below).

4.1.23 Removal of the selection cassette by transient transfection

For selection cassette removal, 800µl of 5×10^6 cells per ml D-PBS were electroporated with 50µg pCre plasmid (0.25 kV, 500µF), diluted 1:300 to 1:5000 and plated on wildtype feeder dishes. Colonies were allowed to grow for 5-7 days without selection. Picking, freezing and expansion for DNA isolation was performed as described above.

4.1.24 Subcloning of ES cell clones

After cassette removal, cells were subcloned to obtain pure clones. Cells from a 10cm dish on the second to third day after splitting were harvested in growth media and seeded in a 1:50 to 1:1000 dilution onto wildtype feeders. Cells were grown for 5-7 days. Picking, freezing and expansion for DNA isolation was performed as described above.

4.1.25 Preparation of genomic DNA

Cells, embryonic membranes or embryonic heads were lysed overnight with an appropriate amount of DNA lysis buffer (700µl for embryonic membranes, heads, cells from 10cm cell culture dishes; 400µl for ES cell clones on 24-well plates) at 55°C. Cellular debris was precipitated by addition of 300µl of saturated NaCl solution followed by centrifugation at 13.2krpm for 10' in a microcentrifuge. The DNA in the supernatant was precipitated with 0.6x vol isopropanol. After one further centrifugation at 13.2krpm for 10' the DNA was washed twice with 70% ethanol and dissolved in a suitable amount of TE overnight at 55°C.

4.1.26 Isolation of total RNA

Cultured cells were trypsinized and resuspended in an appropriate amount of TRI[®] reagent. RNA was isolated according to the manufacturer's protocol. Pellets were dissolved in 100-200µl RNA storage solution and precipitated with 1/10 vol 3M NaAc and 2.5 vol 100% EtOH. Samples were stored as precipitate at -20°C.

4.1.27 DNaseI treatment of RNA

To avoid DNA contamination, RNA samples for reverse transcription were treated with DNaseI using the DNA-free[™] kit according to the manufacturer's protocol. DNaseI treatment was performed for 30min.

4.1.28 Reverse transcription of RNA

4µg of DNaseI-treated RNA samples (2µg for +reverse transcriptase, 2µg for – reverse transcriptase reaction) were used for reverse transcription using the RevertAid™ First Strand cDNA Synthesis Kit with random hexamer primers according to the manufacturer's protocol.

4.1.29 DNA and RNA blots

4.1.29.1 DNA blots

15-20µg of genomic DNA (for ES cell screening: 20µl of 50µl DNA solution) were digested overnight with 2U/µg of appropriate enzymes in a suitable buffer. The whole digests were loaded onto 0.8% agarose gels. Gel electrophoresis was performed in 1xTBE at 5.3V/cm together with a suitable DNA marker. The gel was stained in ethidium bromide solution (1mg/l). Pictures were taken using the Alphamager. Next, the gel was denatured two times for 30min in denaturing solution and then put upside down on 3MM chromatography paper located on a glass plate with its ends soaking in denaturing solution. A Hybond-XL nylon membrane was prewashed in denaturing solution and put onto the gel. Another two sheets of 3MM chromatography paper soaked in denaturing solution were put on top. Remaining areas of the gel and the bottom 3MM paper were covered with plastic stripes to prevent circuit shortcuts. A staple of paper towels was put on top of the blot and weighed down using a glass plate and a blot weight. Capillary transfer was allowed to go on for at least 18 hours, afterwards the blot was disassembled. The membrane was neutralised in 20mM Na₂HPO₄ for 5min.

4.1.29.2 RNA blots

General rules for working with RNA were observed. Gel chamber, gel tray and gel combs were soaked in sterile water + 1/1000 (v/v) DEPC for 30min. 15µg of RNA were precipitated and dissolved in 6µl of RNA storage solution. 18µl of formaldehyde loading dye were added. The solution was incubated at 65°C for 15min. Samples were kept on ice until loaded. 1% (w/v) denaturing agarose gels were prepared using NorthernMax denaturing gel buffer and the gel electrophoresis was performed in NorthernMax Running Buffer according to the manufacturer's recommendations. Gel electrophoresis was performed with 6.6V/cm for 3hrs at 4°C. Gels were blotted in 50mM Na₂HPO₄ + 1/1000 (v/v) DEPC. After blotting, the membrane was dried at 55°C for 15min, crosslinked for 18sec with 120000µJ/cm² in the UV Stratalinker 1800 and stained with methylene blue. The membrane was washed with 25mM Na₂HPO₄ and the picture was scanned.

4.1.29.3 Probe labelling for DNA and RNA blots

20ng of probe fragment were denatured (5min at 99°C, 2min on ice) and added to 18µl of LS buffer and 5.5µl of CGT. Water was used to adjust to a final volume of 37.5µl. Subsequently 2µl of $\alpha^{32}\text{P}$ -ATP (10µCi/µl) and 2U of Klenow fragment were added and the reaction was incubated for at least 6h or overnight at room temperature. The probe was cleaned using a Sephadex G-50 spin column (packed by centrifugation in a 1ml syringe). For DNA blot probes see Table 12, for RNA blot probes see Table 13).

name	start	end	chr.	primers/restriction enzymes
MEi	12933234	12934384	17	MluI/EcoRI
AirT	12938949	12939391	17	120970A-F, 120970A-R
EEi	12961305	12966278	17	EcoRI/EcoRI
OT2.4	12937224	12939421	17	EcoRI/BamHI
MSi	12934381	12935395	17	MluI/SfuI
kodel	12957249	12957573	17	KODEL F, KODEL R
H19P13kb	149775165	149775662	7	H19P13kb-F, H19P13kb-R
Dlk1 probe	110766865	110767839	12	Dlk1F, Dlk1R

Table 12: Used DNA blot probes. Details as in Table 9. Primers/restriction enzymes list the PCR primers used for probe amplification or the restriction enzymes for probe isolation.

name	start	end	Acc number	primers/restriction enzymes
Prss11	gift from C. Seiser (MFPL, Vienna, Austria)			
Gapdh	436	1232	NM_008084.2	PrF, PrR
Igf2r (HX)	exons 3-48 from plasmid with <i>Igf2r</i> cDNA (pHXIgf2r)			

Table 13: Used RNA blot probes. Details as in Table 9.

4.1.29.4 Membrane hybridisation

For RNA-blot hybridisation tubes were treated with autoclaved water + 1/1000 DEPC (v/v) for 30min. Membranes were prehybridised for 30min to 2h in Church buffer in a turning hybridisation tube at 65°C. The prehybridisation solution was discarded and the denatured (5min at 99°C, 2min on ice) random primed probe diluted in Church buffer was added. The hybridisation was done for at least 18h at 65°C in a turning hybridisation tube. Then the membrane was washed two times for 30min in pre-warmed 65°C Church wash. The membrane was sealed in plastic foil and exposed to a PhosphorImager screen. The screen was scanned using the Typhoon Scanner 5600.

4.1.30 Real-time qPCR

Primers and Taqman probes were designed using PrimerExpress. See Table 14 for details of Taqman assays and Table 15 for details of SybrGreen assays. 5'-FAM/3'-TAMRA labelled probes were synthesised by Microsynth AG, Balchag, Switzerland. 5'-FAM/3'-MGB labelled

probes were synthesised by Applied Biosystems Inc. Primers were synthesised by VBC-Biotech Service GmbH, Vienna, Austria.

assay	name	sequence	start (bp)	stop (bp)	chr./Acc. number
Igf2r-Ex48	Igf2r-Ex48-F	TCCTACAAGTACTCAAAGGTCAGCA	7191	7215	NM_0105 15.2
	Igf2r-Ex48-P	FAM-CCAAGACTAGGCAAGGACGGGCAAGA-TAMRA	7278	7303	
	Igf2r-Ex48-R	CGCCTTGGTGGTGATATGG	7310	7328	
Air-5'	FP1	AAGCACAGCACCGCCAGT	12934245	12934262	17
	AS-P	FAM-CCACGCAGACATC-MGB	12934265	12934277	
	RP11	TCCTCTAACGCGTGGAATCC	12934373	12934392	
Air SV1	FP1	AAGCACAGCACCGCCAGT	12934245	12934262	17
	AS-P	FAM-CCACGCAGACATC-MGB	12934265	12934277	
	RP21	CCATGTCTCTTTCTTTTCCACTACC	13007378	13007401	
Air SV1a	FP1	AAGCACAGCACCGCCAGT	12934245	12934262	17
	AS-P	FAM-CCACGCAGACATC-MGB	12934265	12934277	
	RP6	AGGCCTTTGTTCACATCTCTTCA	12972165	12972187	
Air SV2	FP1	AAGCACAGCACCGCCAGT	12934245	12934262	17
	AS-P	FAM-CCACGCAGACATC-MGB	12934265	12934277	
	RP5	CAAAGGTGCTTGCCCTCCAA	13022721	13022739	
Air SV3	FP1	AAGCACAGCACCGCCAGT	12934245	12934262	17
	AS-P	FAM-CCACGCAGACATC-MGB	12934265	12934277	
	RP4	CAGGACCTCAAGTCAGGAACCT	13051981	13052002	
AirT3	AirT3TQF	CCCTAGGAAGGCACAGATGC	12938294	12938313	17
	AirT3	FAM- CCGCTTCCAGCAGCTGTTACATCTAGTGC-TAMRA	12938264	12938292	
	AirT3TQR	ACAGCGATCCTCCAGAAGAGTG	12938241	12938262	
Air middle	Air-TQF	GACCAGTTCCGCCCCGTTT	12987545	12987562	17
	AirTQ	FAM- TACAAGTGATTATTAACCTCCACGCCAGCCTCA-TAMRA	12987483	12987543	
	Air-TQR	GCAAGACCACAAAATATTGAAAAGAC	12987483	12987508	
Air end	AirEnd-Fwd	GGACTGGCTCAGGCAAGCT	13032966	13032984	17
	AirEnd-P	FAM-CCTGCTCGAGTTGCCATTCCCAGA-TAMRA	13032938	13032961	
	AirEnd-Rev	TTCAGTCAAAAATCCAAAACATGT	13032912	13032936	
Cyclophilin A (Ppia)	CypA-F	AGGGTTCCTCCTTTTACAGAATT	178	200	BC08307 6.1
	CypA-P	FAM-TCCAGGATTCATGTGCCAGGGTGG-TAMRA	203	226	
	CypA-R	GTGCCATTATGGCGTGTAAGTC	228	250	
Gapdh-Ex5	Gapdh-Ex5-F	CATGGCCTTCCGTGTTCTTA	734	753	NM_0080 84.2
	Gapdh-Ex5-P	FAM-TCGTGGATCTGACGTGCCGCC-TAMRA	769	789	
	Gapdh-Ex5-R	TGTCATCATACTTGGCAGGTTT	795	816	

Table 14: Used qPCR primers and probes for Taqman assays. Details as in Table 9. FAM and TAMRA refer to fluorescent dyes, MGB to a non-fluorescent quencher.

qPCR was performed using an ABI PRISM 7000 Sequence Detection System with 900nM primers, 200nM probe and qPCR Mastermix Plus for Taqman assays or 100nM primers and Mesa Green qPCR Master Mix Plus for SybrGreen assays. Cycling conditions: 2min 50°C, 10min 95°C, 40 cycles of 15sec 95°C and 1min 60-61°C (for Taqman probes) and 5min 95°C, 40 cycles of 15sec 95°C and 1min 60-65°C (for SybrGreen assays). For the *Igf2r* exon 12 SNP assay, the Power SYBR® Green PCR Master Mix was used and various MgCl₂ concentrations were tested (for final experiments, 5mM MgCl₂ were used). RNA quantification was done by the standard curve method using serial dilutions of cDNA or plasmid DNA. Relative quantification and statistics were performed as described in the manual of the sequence detection system.

assay	name	sequence	start (bp)	stop (bp)	chr./Acc. number
45S pre-rRNA	45SpreRNA-F	TGCTGCCCCGTATCAGTAACTGTC	6201	6224	X82564.1
	45SpreRNA-R	CCCTGGCCCGAAGAGAACT	6311	6329	
Air START	START-F1	GCAGCAAGAAGCACAGCAC	12934237	12934255	17
Igf2r-Ex48	START-R1	GATGTCTGCGTGGTAACTGG	12934258	12934277	NM_010515.2
Ex12-SNP assays	WtSeF	TGGCCTTCCCCTCCTGC	1724	1740	NM_010515.2
	MutSeF	CTGGCCTTCCCCTCCTGT	1724	1741	
	WtSeFTA	TGGCCTACCCCTCCTGC	1724	1740	
	MutSeFTA	CTGGCCTACCCCTCCTGT	1724	1741	
	WtSeFCG	TGGCCTTGCCCTCCTGC	1724	1740	
	MutSeFCG	CTGGCCTTGCCCTCCTGT	1724	1741	
	GeSeR2	GCTATGACCTGTCTGTGTTGGCT	1606	1628	
Amylase	Amyrt-F2	GGCATTAGCTATGCACAAACCA	113239875	113239896	3
	Amyrt-R2	AAGCTATCAAGATTGGCAATTGAA	113239923	113239946	
Hba-a1	Hba-a1rtF3	AGCATCTTATCCTACTTTATTTTCATATCCA	32183255	32183284	11
	Hba-a1rtR3	CCTGTTTCACACTTTCTTTCTCAA	32183318	32183342	NM_010515.2
Igf2r S phase	Igf2r-95F	GTTATTGAGACCTTGTGACTTAATAATTCT G	12964901	12964931	17
	Igf2r-95R	TTGGTCCCCTTTTCTCACCAT	12964843	12964863	
P6Hu	P6Hu-F	CCTCTTGATGACCCCGTTTAAAC	160351552	160351573	6
	P6Hu-R	AAAATAAAGAAAGCAGCAGCAAGAA	160351597	160351621	
mitochondrial DNA	mitort-F2	GCCTGCCAGTGACTAAAGTTT	1979	2000	NC_005089.1
	mitort-R2	AACAAGTGATTATGCTACCTTTGCA	2023	2047	

Table 15: Used qPCR primers for SYBRGreen assays. Details as in Table 9. P6Hu which refers to the Human March 2006 (hg18) assembly.

4.1.31 RNA expression tiling array (RETA)

4.1.31.1 Sample preparation

6.5µg of DNase I (DNA-free™ kit; see above) treated total RNA was used as template for first-strand cDNA synthesis. 2µl of random hexamer primers (2.5µg/µl) were added, heated at 70°C for 10min and incubated on ice for 5min. After addition of 7µl 5x First strand cDNA buffer, 3.5µl 0.1M DTT, 3µl 10mM dNTP mix, 1µl RiboLock™ RNase inhibitor (40U/µl), 2µl of Superscript II RNaseH Reverse Transcriptase and nuclease-free water to a final volume of 35µl, the solution was incubated for 1h at 42°C and afterwards placed on ice. Second-strand cDNA synthesis was performed by adding 76µl nuclease-free water, 30µl 5x Second strand cDNA buffer, 3µl 10mM dNTP mix, 4µl DNA polymerase I, 1µl E. coli DNA ligase and 1µl RNase H, followed by incubation at 16°C for 2h. 2µl of T4 DNA polymerase were added followed by an additional 10min incubation at 16°C. Afterwards, the reaction was incubated at 70°C for 10min. The double-stranded cDNA was purified using the QIAquick PCR Purification Kit according to the manufacturer's protocol. After precipitation in ethanol at -20°C overnight, the cDNA was recovered by centrifugation at 13.2krpm for 30min at 4°C. The pellet was re-suspended in 14-18µl nuclease-free water.

For input sample preparation (done by Ru Huang), 100µg of isolated CCE genomic DNA were first treated with RNaseA (final concentration: 25µg/ml) overnight at 37°C. Then proteins in the DNA sample were digested by 15µl ProteinaseK (10mg/ml) at 55°C for 2-3hours. After phenol-chloroform extraction, 1ml of genomic DNA was sonicated for 5min at 20% power with 20sec on/1min off pulse. The size distribution of the sonicated DNA was checked on a 2% agarose gel. Sonicated genomic DNA was purified using the QIAquick PCR Purification Kit according to the manufacturer's protocol. After ethanol precipitation, the input DNA was dissolved in nuclease-free water.

4.1.31.2 Sample quantification and quality control

The concentration of each sample was measured on a NanoDrop ND-1000 spectrophotometer and adjusted to 300ng/µl to 500ng/µl. The minimal total amount of each sample was 4.5µg. The 260/280 absorbance ratio of DNA samples was between 1.85 to 1.95. The size distribution of the samples was measured on the 2100 Bioanalyzer using the DNA 7500 kit according to the manufacturer's protocol. The size range of DNA samples was between 100bp to 800bp.

4.1.31.3 Sample hybridisation

Both, the double-stranded cDNA and input samples were sent to ImaGenes GmbH, Berlin, Germany. Hybridization and scanning was performed by the NimbleGen service department of ImaGenes. For hybridization, the double-stranded cDNA labelled with Cy5 and the input sample labelled with Cy3 were co-hybridized to a NimbleGen MIRTA chip. The NimbleGen MIRTA chip contains isothermally designed (T_m : 76°C) oligos covering 39 imprinted and control regions with a 100bp resolution designed according to the Mouse July 2007 (mm9) assembly of the UCSC genome browser.

4.1.31.4 Expression chip analysis

The normalized \log_2 (double-stranded cDNA/input) ratios were gained by the subtraction of a robust average, calculated by a one step Tukey bi-weight function across all the experimental probes, from each raw \log_2 (double-stranded cDNA/input) ratio value.

4.1.32 Nuclear/cytoplasmatic RNA extraction

This method was adapted from Sambrook and Russel (Sambrook and Russel, 2001). For this, confluent cells from 3xT80 cell culture flasks were used. Media was removed and cells were washed three times with ice-cold D-PBS on ice. Cells were scraped off in D-PBS using a cell scraper. The cell suspension was centrifuged for 5min at 2000g at 4°C in a tabletop centrifuge. The supernatant was removed by aspiration, the cell pellet was re-suspended in ice-cold C/N lysis buffer. The cell suspension was underlayered with an equal volume of C/N lysis buffer containing sucrose (24% w/v) and 1% NP-40. After an incubation for 5min on ice and a centrifugation for 5min at 10000g for 20min at 4°C using the swing-out rotor HB4 (Sorvall centrifuge), the upper turbid cytoplasmatic layer was transferred to a new reaction tube. The clear sucrose phase was discarded and the dark nuclear pellet was resuspended in C/N lysis buffer. To both, the nuclear and the cytoplasmic phase, an equal volume of 2x C/N ProteinaseK buffer and ProteinaseK to a final concentration of 200µg/ml were added, mixed and incubated at 37°C for 30min. In addition, in the nuclear phase nuclei were disrupted and genomic DNA was sheared by repeated squirting of the viscous solution through a G19 Sterican needle before incubation for ProteinaseK digestion. Afterwards, proteins from both fractions were removed by extraction with acid phenol/chloroform. An equal volume of acid phenol/chloroform was added, the sample was mixed well by vortexing and was centrifuged at 13000g for 10min at 4°C. The upper, aqueous phase was recovered and RNA was precipitated by addition of 2.5vol 96% ethanol and 0.1vol 3M sodium acetate overnight at -20°C. The RNA was recovered by centrifugation at 13.2krpm at 4°C for 20min.

The pellet was washed twice with 70% ethanol. Afterwards, the pellet was resuspended in an appropriate amount of RNA storage solution and stored as precipitate at -20°C. The DNaseI treatment was done as described above, however, the nuclear fraction was digested twice with DNaseI for 30min at 37°C.

4.1.33 Metaphase spreads

Cells were incubated with 0.1µg/ml colcemid (KaryoMAX) in growth medium for 3-6h at 37°C in 5% CO₂ atmosphere. The cells were washed twice with pre-warmed D-PBS and harvested by trypsinisation. After resuspension in growth medium they were centrifuged for 5min at 1000rpm in a tabletop centrifuge, the supernatant was removed and cells were resuspended in 300µl of growth medium. 5ml 0.4% KCl were added and the suspension was incubated in the waterbath at 37°C for 10min. 100µl of fresh fixative (MeOH:acetic acid = 3:1) were added. The cells were spun at 1000rpm for 5min, the supernatant was discarded. The cell pellet was resuspended in 5ml fixative and incubated at room temperature for 20min. After one further centrifugation (1000rpm for 7min) the supernatant was removed and the pellet was resuspended in 200-800µl (dependent on the cell density) of fixative. 15µl of cell suspension were dropped onto microscope slides pre-chilled to -20°C and dried overnight. On the next day the slides were mounted in Vectashield containing DAPI and covered with a coverslip. The slides were analysed using a Axioplan2 epifluorescence microscope with a DAPI (359/461nm) fluorescence filter. For the 60x objective immersion-oil was used. The spreads were magnified 600x. Images were acquired using a coloured digital CCD camera.

4.1.34 DNA fluorescence in situ hybridisation (DNA FISH)

4.1.34.1 Cell preparation for DNA FISH

Prior to sample preparation, cells were pulsed with 10µM BrdU in growth medium for 50min at 37°C in 5% CO₂ atmosphere. The medium was removed and cells were washed twice with pre-warmed D-PBS. After trypsinisation cells were harvested in growth medium, counted using a cell counter, centrifuged at 1200rpm for 5min in a tabletop centrifuge and resuspended in D-PBS at 2x10⁶ cells per 1ml. 20µl of cell suspension were applied onto positively charged SuperFrost®Plus microscope slides in a spot of about 1cm in diameter. The cells were allowed to adhere to the slides for 2min. The slides were treated with CSK buffer for 30sec, CSK buffer + 0.5% Triton X-100 for 3min and CSK buffer for 30sec again.

Cells were rinsed in 1xPBS, fixed in 4% paraformaldehyde in 1xPBS pH 7.4 for 10min and washed with 70% ethanol. Slides were stored in 70% ethanol at -20°C.

4.1.34.2 Probe preparation for DNA-FISH

Cosmid/BAC/plasmid DNA (see Table 16 for details) was sonicated on ice to a fragment size of 100 – 500 bp using the Sonicator HD70, cycle 60, 60% power, five times for 1min with 1min breaks in between. Fragment size was checked on a 2% agarose gel. Single nucleotides were removed using the Wizard[®] SV Gel and PCR Clean-Up System according to the kit manual. 1µg of DNA was labelled and cleaned using the PlatinumBrigt[™] Nucleic Acid Labeling Kit according to the kit manual or the ULYSIS[®] Nucleic Acid Labeling Kit Alexa Fluor[®] 568 according to the kit manual and purified using Micro Bio-Spin[®] P-30 Tris Chromatography Columns according to the manufacturer's protocol. 210ng of labelled probe were co-precipitated with 10000ng of mouse C₀t1-DNA. After overnight precipitation and centrifugation for 20min at 13.2krpm, the pellet was washed twice with 70% ethanol and afterwards dissolved in 45µl of Hybrisol VII. The probe was placed at 37°C for 1h to dissolve.

name	chr.	start	end
cos4L	17	12980173	13001928
cos5B	17	12551220	12588463
cos940PS	17	12907561	12947481
cos940	17	12879561	12913164
cos9G	17	13819551	13853850
cosLA1	17	13106293	13143058
cosMS6	17	12994819	13033035
cosOT1	17	12922095	12965701
cosRP17B3	17	12408297	12444524
pSod2	17	13200891	13207824
RP24-223O4	17	8600228	8761831
RP24-173F10	17	9789470	9920936
RP24-222A12	17	10227717	10372729
RP24-362H17	17	10933562	11086865
RP24-281L7	17	11075498	11212469
RP24-99O8	17	11581400	11715844
plgf2P6	7	149836242	149840005

Table 16: Used probes for DNA FISH. Details as in Table 9.

4.1.34.3 Prehybridisation treatment for DNA FISH

Slides were dehydrated with 70%, 80%, 96% ethanol for 2min each at room temperature. For 3D DNA FISH the dehydration series was omitted. They were treated with RNaseA 0.1mg/ml in 2xSSC for 30min at 37°C in a humid environment and washed three times in 2xSSC for 5min at 37°C. Samples were dehydrated in 70%, 80%, 96% ethanol for 2min each at room temperature. For 3D DNA FISH the dehydration series was omitted.

4.1.34.4 Hybridisation for DNA FISH

15µl of labelled probe in Hybrisol VII were added per slide and covered with a coverslip. The coverslip was sealed with rubber cement and the slides were placed in the HYBrite hybridisation oven. The probes and samples were co-denatured at 80°C for 5min, hybridisation was carried out at 37°C for at least 17h. The rubber cement was removed, the coverslip was floated off in 2xSSC/50% formamide. The slides were washed three times in 2xSSC/50% formamide at 39.5°C for 5min, then three times in 2xSSC at 39.5°C for 5min followed by one wash in 1xSSC at room temperature for 10min. Finally, slides were washed for 2min in 1xPBD (Qbiogene).

4.1.34.5 BrdU-immunofluorescence for DNA-FISH

Slides were washed two times for 5min in BrdU-wash solution at room temperature. Blocking was performed for 30min in BrdU-detection solution at room temperature. Detection was carried out for 40min with BrdU-detection solution + antiBrdU AlexaFluor 546 conjugate (1:50). Slides were washed three times for 5min in BrdU-wash solution at room temperature, shaking. Slides were mounted in Vectashield containing DAPI and covered with a coverslip. The coverslip was fixed with clear nail polish. The slides were stored at 4°C in the dark.

4.1.34.6 Image acquisition and analysis

The slides were analysed using a Axioplan2 epifluorescence microscope with following fluorescence filters: DAPI (359/461nm), FITC (495/519nm), Cy3 (558,568nm). For the 60x objective immersion-oil was used. Images were acquired using a coloured digital CCD camera. Exposure time depended on used filters and the experiment and was between 80ms and 3000ms, afterwards the images were pseudocoloured using MetaVue v.0r6 (Universal Imaging Corporation).

For 3D analysis, 0.15µm z-stacks were acquired using a DeltaVision (Olympus 1X71) microscope system with following filters: DAPI (457nm), FITC (528nm) and RD-TR-Pe

(617nm) using the softWoRx[®] 2.50 software. Cell magnification performed using a 60x objective. Exposure time was dependent on the used filter and on the experiment and was between 200 and 2000ms. Images were acquired using a cooled, digital CCD camera (Coolsnap_Hq/ICX285) and processed with softWoRx[®] 2.50. Acquired image stacks were deconvolved using 3D iterative constrained deconvolution on softWoRx[®] 2.50 with 15 restrictive iterations. Three-dimensional reconstruction was done by voxel imaging, rendering the acquired FISH signals into geometric surfaces using Imarisx64 6.0.1. Three-dimensional coordinates of the centers of homogenous mass were identified and the distance was calculated according to the following formula $=\text{SQRT}((x1-x2)^2+(y1-y2)^2+(z1-z2)^2)$. Box plots showing the distance distribution were generated using Microsoft Excel (see below).

4.1.35 TSA treatment of cells

Cells were treated with 5-7ng/ml Trichostatin A (TSA) in DMSO in growth media at 5% CO₂-atmosphere at 37°C. In case of a following BrdU pulse, BrdU was directly added to the growth media + TSA after 16h and 10min and incubated for further 50min. DNA FISH was described as above.

4.1.36 S phase fractionation

This method was done according to (Azulara, 2006).

Cells were pulsed with 50µm BrdU in growth medium for 30min at 5% CO₂ atmosphere and 37°C. Cells were washed twice in prewarmed D-PBS, trypsinised and harvested in growth medium. The suspension was centrifuged at 1000rpm for 5min in a tabletop centrifuge, the supernatant was removed and the cell pellet was resuspended in 5ml of cold D-PBS. Cells were centrifuged at 1200rpm for 5min at 4°C, the supernatant was removed. The cell pellet was resuspended in 5ml of D-PBS. Cells were counted using a cell counter, aliquots of 10⁷ cells were prepared for subsequent steps. Cells were centrifuged at 1200rpm for 5min at 4°C, the supernatant was removed and cells were fixed with 10ml ice-cold 70% EtOH, added dropwise with mixing. Cells were incubated at 4°C for 30min or overnight. Cells were pelleted by centrifugation at 1200rpm for 5min at 4°C and washed twice with cold D-PBS. The cell pellet was resuspended in 2ml of FACS staining buffer, cells were declumped by pipetting and incubated for 10min at room temperature followed by 20min incubation on ice. After staining, cell clumps were removed by filtration through a cell strainer into a 5ml tube and sorted with a FACSAria. 20000-50000 cells were collected per fraction directly into 200µl of FACS lysis buffer I. After fractionation, cells were incubated for 2h at 55°C and

stored at -20°C for subsequent processing. Samples were thawed, BrdU-labelled human HS-27 cell DNA was added at a concentration of 7.5ng per 10⁴ sorted cells. 1x volume of phenol was added, samples were mixed by shaking for 1min and centrifuged for 5min at 13.2krpm. The phenol was removed from the bottom of the tube, 1x volume of chloroform was added. Samples were mixed by shaking and centrifuged for 5min at 13.2krpm. The aqueous phase was transferred into a new tube and precipitated with 2x volume of ice cold 96% EtOH for at least two hours at -20°C. DNA was recovered by centrifugation for 30min at 4°C at 13.2krpm. The EtOH was removed, the DNA pellet was washed with 70% EtOH and centrifuged for 10min, 13.2krpm at 4°C. The pellet was air dried, resuspended in 480µl TE buffer and dissolved for 1 hour at 37°C. 20µl of salmon sperm DNA (10mg/ml) were added, the sample was incubated for 30min at 37°C. DNA was sonicated for 30sec, cycle 60, 60% power on ice. Samples were denatured for 5min at 95°C followed by a 2min incubation on ice. 50µl of adjuster buffer were added, followed by the addition of 80µl anti-BrdU antibody (25µg/ml) and gentle mixing. Samples were incubated for 20min at room temperature, rotating at low speed. After addition of 35µg rabbit anti-mouse IgG antibody and gentle mixing, samples were incubated again for 20min at room temperature, rotating at low speed. Samples were put on ice and centrifuged for 20min at 4°C. The supernatant was removed with a pipette, the pellet was washed with 1ml FACS wash buffer. The sample was centrifuged for 20min, 13.2krpm at 4°C. Afterwards, the wash buffer was removed again using a pipette. The pellet was resuspended in 200µl FACS lysis buffer II and incubated for 1 hour at 37°C. The sample was vortexed and incubated overnight at 37°C. Another 100µl of FACS lysis buffer II were added, the sample was incubated for 1 hour at 50°C. The DNA was purified using the QIAGEN Gel Extraction kit according to the kit manual. Elution was performed using buffer EB with 200 cell equivalents/µl. 2.5µl of DNA were used for following qPCR analysis. Human HS-27 DNA served as a control for successful immunoprecipitation. Samples were normalised to mitochondrial DNA.

4.1.37 Bioinformatics

For genomic sequences, the UCSC genome browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>) was used. For mouse sequences, the July 2007 (mm9) assembly was used, for human sequences the March 2006 (hg18) assembly.

Genomic localisations of constructs was obtained by a BLAT search against the Mouse July 2007 (mm9) assembly of the UCSC genome browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>).

Interspersed repeats were identified using RepeatMasker (<http://www.repeatmasker.org/>).

Dotplots for the identification of tandem direct repeats were generated using EMBOSS dotmatcher (<http://emboss.bioinformatics.nl/>). Window size was set to 15, threshold to 50. The submitted sequence was repeatmasked and contained the CpG island as determined by the UCSC genome browser and 500bp up- and downstream.

CpG plots were performed using EMBOSS CpGPlot (<http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html>) using default parameters.

Quantification of Northern blot signals was done using ImageJ (<http://rsbweb.nih.gov/ij/>).

p-values were calculated with an unpaired t-test using QuickCalcs (www.graphpad.com/quickcalcs).

4.1.38 Box plots

Box plots showing the distance distribution in the 3D DNA FISH experiment were generated using Microsoft Excel. Each data point was put in a separate cell of a column. Then the following table was created (see Table 17). Note that in this example A:A means that all data points were put in the Excel worksheet into column A. The numbers in column 3 are the x-values and had to be changed in order to show several plots next to each other.

min	MIN(A:A)	3
min	MIN(A:A)	5
min	MIN(A:A)	4
25 th	PERCENTILE(A:A;0.25)	4
25 th	PERCENTILE(A:A;0.25)	7
25 th	PERCENTILE(A:A;0.25)	1
75 th	PERCENTILE(A:A;0.75)	1
75 th	PERCENTILE(A:A;0.75)	7
25 th	PERCENTILE(A:A;0.25)	7
med	MEDIAN(A:A)	7
med	MEDIAN(A:A)	1
75 th	PERCENTILE(A:A;0.75)	1
75 th	PERCENTILE(A:A;0.75)	4
max	MAX(A:A)	4
max	MAX(A:A)	5
max	MAX(A:A)	3

Table 17: Table used for the generation of box plots. See text for details.

This table was displayed as 'scatter with data points connected with lines' plot, where column 3 were the x-values and column 2 the y-values. The distance between the 25th and the 75th percentile defines the inter-quartile range. Data points lying outside of the borders of the box + 1.5x the inter-quartile range were plotted separately as outliers.

4.2 Materials

4.2.1 Cells

D3	129Sv/129Sv	provided by Erwin Wagner
A9	C57BL6/129Sv	provided by Arabella Meixner

4.2.2 Mice

FVB (wildtype)

DR4 (Tucker et al., 1996)

AirT (Sleutels et al., 2002)

IP Δ (Sleutels et al., 2003)

R2 Δ (Wutz et al., 2001)

Thp (Johnson, 1974; Johnson, 1975)

All mutant mouse strains were kept on FVB background.

4.2.3 Vectors/Plasmids

BACs	Children's Hospital Oakland Research Institute (CHORI)
pGEM-T Easy	Promega
pBSIIKS(-)	Stratagene

4.2.4 Consumables

α - ³² P-d[ATP]	PerkinElmer
β -mercaptoethanol	Sigma
β -mercaptoethanol for cell culture	Gibco
5-Bromo-2'-deoxyuridine (BrdU)	Sigma
5ml tube with cell strainer cap	BD Biosciences
3MM chromatography paper	Whatman
Acid phenol chloroform	Ambion
Agar	AppliChem
Agarose for DNA work	Biozym
Agarose for RNA work	Ambion
Ampicillin	Roche
anti-BrdU AlexaFluor 546 conjugate	Molecular Probes
anti-BrdU antibody	Becton Dickinson
ATP	Fermentas
Betaine	Sigma
Boric acid	AppliChem
Bovine Serum Albumin (BSA)	Sigma
Cell scraper	BD Falcon
Chloramphenicol	AppliChem
Chloroform	Merck
CircleGrow broth	MP Biomedicals

Coverslips	Roth
Diethyl pyrocarbonate (DEPC)	Sigma
Dimethyl sulfoxide (DMSO)	Sigma
Dithiothreitol (DTT)	Invitrogen
DNA ladder, 100bp	Fermentas
DNA ladder, 1kb	Fermentas
DNA Polymerase I	Invitrogen
dTTP	Bioron
dCTP	Bioron
dGTP	Bioron
dNTP mix (10mM)	Fermentas
Dulbecco's Modified Eagle Medium (DMEM)	Gibco
Dulbecco's Modified Eagle Medium (DMEM) + Hepes	Gibco
Dulbecco's phosphate buffered saline (D-PBS)	Gibco
E. coli DNA Ligase	Invitrogen
Ethanol	Merck
Ethidium bromide	Merck
Ethylenediaminetetraacetic acid (EDTA)	Merck
Fetal bovine serum (FBS)	Gibco
Fetal bovine serum (FBS), ES cell tested	PAA
First strand cDNA buffer, 5x	Invitrogen
Formaldehyde loading dye	Ambion
Formamide	Fluka
G19 Sterican needle	Braun
G20 Sterican needle	Braun
Gelatin	Sigma
Geneticin	Gibco
Gentamicin	Gibco
Glacial acetic acid	VWR
Glucose	Gibco
Glycerol	Qbiogene
GoTaq DNA polymerase	Promega
HCl	Merck
Hepes	Roth
Hybond-XL (nylon membrane)	Amersham
Hybrisol VII	Qbiogene
Immersion oil	Fluka
Isopropanol	Merck
Isopropyl- β -D-thiogalactopyranoside (IPTG)	AppliChem
KAc	Sigma
Kanamycin	AppliChem
KaryoMAX	Gibco
Klenow Fragment	Fermentas
LB broth medium	Lab M Limited
L-Glutamin	Gibco
Mesa Green qPCR Mastermix Plus	Eurogentec
Methanol	Roth
Methylene blue	Merck
MgCl ₂	Sigma
MgCl ₂ , 25mM (for PCR)	Fermentas
Microfuge tubes, 1.5ml	Sarstedt
Microfuge tubes, 1.5ml for RNA work	Ambion
Microfuge tubes, 2ml	Roth
Microscope slides (for metaphase spreads)	Roth
Mouse C ₀ t1-DNA	Invitrogen

N,N'-dimethyl-formamide	Sigma
NaCl	neo Lab
NaOH	AppliChem
Non-essential amino acids	Gibco
NorthernMax™ 10x denaturing gel buffer	Ambion
NorthernMax™ 10x MOPS gel running Buffer	Ambion
NP-40	Calbiochem
Nuclease-free water	Ambion
Paraformaldehyde	Sigma
PBD	Qbiogene
Penicillin-Streptomycin-Glutamine (PSG)	Gibco
Phenol	AppliChem
Phosphate buffered saline (PBS)	Qbiogene
Polypropylene reaction tubes, 15ml	Greiner Bio-one
Polypropylene reaction tubes, 50ml	Greiner Bio-one
Power SYBR® Green PCR Master Mix	Applied Biosystems
Propidium iodide	Sigma
Proteinase K	Qbiogene
qPCR Mastermix Plus	Euorgenec
Rabbit anti-mouse IgG antibody	Sigma
Random hexamer primers for RT PCR	Invitrogen
Random hexamer primers for labelling radioactive probes	Pharmacia
Restriction enzymes	Fermentas/Roche
Retinoic acid	Sigma
RiboLock™ RNase inhibitor	Fermentas
Ribonucleoside Vanadyl Complex	New England BioLabs
RNA millennium marker	Ambion
RNA storage solution	Ambion
RNaseA	Fermentas
RNaseH	Invitrogen
Rubber cement Fixogum	Marabu
Second strand cDNA buffer, 5x	Invitrogen
Sephadex™ G-50	Amersham
Sodium acetate for DNA work	VWR
Sodium acetate for RNA work	Ambion
Sodium citrate	Sigma
Sodium dodecyl sulfate (SDS)	AppliChem
Sodium pyruvate	Sigma
Sucrose	Sigma
SuperFrost® Plus microscope slides	Thermo Scientific
Superscript II RNaseH Reverse Transcriptase	Invitrogen
T4 DNA ligase	Fermentas
T4 DNA polymerase	Invitrogen
T4 polynucleotide kinase	Fermentas
TRI® reagent	Sigma
Trichostatin A (TSA)	Sigma
Tris	AppliChem
Triton X-100	Sigma
Trypsin-EDTA	Gibco
Tween®-20	Sigma
Ultrapure salmon sperm DNA solution	Invitrogen
Vectashield containing DAPI	Vector Labs
X-Gal	Roth
Xylenol orange	Sigma

4.2.5 Kits

BACMAX™ DNA Purification Kit	Epicentre
DNA 7500 kit	Agilent
DNA-free™-kit	Ambion
EndoFree Plasmid Maxi kit	Qiagen
Long PCR Enzyme Mix	Fermentas
Micro Bio-Spin® P-30 Tris Chromatography Columns	BioRad
pGEM®-T Easy Vector System	Promega
PlatinumBriq™ Nucleic Acid Labeling Kit	Kreatech Biotechnology
PureYield™ Plasmid Mini Prep System	Promega
QIAgen Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
QuikChange® XL Site-Directed Mutagenesis Kit	Stratagene
RevertAid™ First Strand cDNA Synthesis Kit	Fermentas
ULYSIS® Nucleic Acid Labeling Kit Alexa Fluor® 488	Molecular Probes
Wizard® SV Gel and PCR Clean-Up System	Promega

4.2.6 Solutions

Alk-1

50mM glucose
25mM Tris pH8.0
10mM EDTA pH8.0

Alk-2

200mM NaOH
1% SDS

Alk-3

3M KAc
11.5% glacial acetic acid

DNA lysis buffer

1xTEN pH9.0
50mM Tris pH9.0
20mM EDTA pH8.0
40mM NaCl
1% SDS
0.5mg/ml Proteinase K

5xTEN pH9.0

50mM Tris pH9.0
20mM EDTA pH8.0
200mM NaCl

TE buffer

10mM Tris-Cl pH8.0
1mM EDTA

CGT mix for radioactive probes

100 μ M dTTP
100 μ M dCTP
100 μ M dGTP
2mg/ml BSA

LS buffer for radioactive probes

25ml 1M Hepes pH6.6
25ml 250mM Tris-Cl pH8.0 / 25mM MgCl₂·6H₂O / 50mM β -mercaptoethanol
1ml 30 OD U/ml random hexamer primers in TE, pH8.0

Denaturing solution for Southern blots

0.5M NaOH
1.5M NaCl

Church buffer

0.25M Na₂HPO₄
7%SDS
1mM EDTA

Church wash

1%SDS
20mM Na₂HPO₄

CSK buffer

3mM MgCl₂
100mM NaCl
0.1% Sucrose
10mM PIPES

ES cell media

15% FBS
50µg/ml Gentamicin
1× non-essential amino acids
1mM sodium pyruvate
2mM L-Glutamin
0.1mM β-mercaptoethanol
appropriate amount of LIF
Hepes-buffered D-MEM (L-Glutamine, 4500mg/l D-glucose, 25mM HEPES buffer,
without sodium pyruvate)

MEF media

10% FBS
1% PSG or 50µg/ml Gentamicin
2mM L-Glutamin
D-MEM (L-Glutamine, 4500mg/l D-glucose, 110mg/l Sodium pyruvate)

Loading buffer for agarose gels

0.5% Xylenol orange
30% glycerol
in TAE

TAE buffer

40mM Tris
0.1142% glacial acetic acid
1mM EDTA pH8.0

TBE buffer

89.1mM Tris
89.0mM Boric acid
2mM EDTA pH 8.0

20x SSC

3M NaCl
300mM Sodium citrate
pH7.0 with HCl

BrdU-wash solution

0.1M Tris
0.15M NaCl
0.005% Tween20

BrdU-detection solution

0.1M Tris
0.15M NaCl
BSA (2mg/ml)

C/N Lysis buffer

0.14M NaCl
1.5mM MgCl₂
10mM Tris HCl pH 8.6
0.5% NP-40
10mM Vanadyl-Ribonucleoside complex

C/N Proteinase K buffer

0.2M Tris HCl pH 7.5
25mM EDTA pH 8.0
0.3M NaCl
2% SDS

FACS staining buffer

3mM EDTA pH8.0
0.05% NP-40
50µg/ml propidium iodide
1mg/ml RNaseA
in D-PBS

FACS lysis buffer I

1M NaCl
10mM EDTA pH8.0
50mM Tris-HCl pH8.0
0.5% SDS
0.4mg/ml proteinase K
0.5mg/ml salmon sperm DNA

FACS lysis buffer II

50mM Tris-HCl pH 8.0
10mM EDTA pH 8.0
0.5% SDS
0.25mg/ml proteinase K

FACS adjuster buffer

110mM sodium phosphate buffer pH 7.0
1.54M NaCl
0.55% Triton X-100

FACS wash buffer

10mM sodium phosphate buffer pH 7.0
0.14M NaCl,
0.05% Triton X-100

BrdU-labeled DNA from HS-27 cells

Pulse-label cells culture with 100mM BrdU for at least 2 h.
Isolate total DNA from cells.

4.2.7 Machines

2100 Bioanalyzer	Agilent
ABI PRISM 7000 Sequence Detection System	Applied Biosystems
Alphamager	Alpha Innotech
Axioplan2 epifluorescence microscope	Zeiss
Casy counter	Schärfe System
Coloured digital CCD camera	SPOT, Vistron
Coolsnap_Hq/ICX285	Photometrix
DeltaVision microscope system	Applied Precision
FACSAria	BD Biosciences
Gene Pulser®II	BioRad
Heraeus Biofuge primo	Thermo Electron Corporation
Heraeus Megafuge 1.0R	Thermo Electron Corporation
High speed cooling centrifuge RC5C	Sorvall Instruments
HYBrite hybridisation oven	VYSIS
Microcentrifuge 5415 R	Eppendorf
NanoDrop ND-1000 spectrophotometer	Peqlab
Peltier Thermal Cycler PTC-200	MJ Research
PhosphorImager screen	Fuji Photo Film
Sonicator Sonoplus GD70	Bandelin
Thermomixer comfort	Eppendorf
Typhoon Scanner 5600	Amersham
UV Stratalinker 1800	Stratagene
Waterbath	Gesellschaft für Labortechnik mbH

4.2.8 Software

Imarisx64 6.0.1	Bitplane
PrimerExpress	Applied Biosystems
Primer3	http://frodo.wi.mit.edu/ (Rozen and Skaletsky, 2000)
softWoRx®	Applied Precision

4.3 Abbreviations

bp	basepairs
BrdU	5-bromo-2-deoxyuridine
CGI	CpG island
dpc	days post coitum
g	gram
gDMR	gametic differentially methylated region
h	hour
ICE	imprint control element
kb	kilobasepairs
(k)rpm	(kilo)rounds per minute
kV	kiloVolts
l	litre
LIF	leukaemia inhibitory factor
Mbp	megabasepairs

mg	milligram
min	minute(s)
ml	millilitre
ms	milliseconds
ncRNA	non-coding RNA
ng	nanogram
u	units
wt	wildtype
sec	second(s)
sDMR	somatic differentially methylated region
μ F	microfarad

5. REFERENCES

- Alexander, M. K., Mlynarczyk-Evans, S., Royce-Tolland, M., Plocik, A., Kalantry, S., Magnuson, T. and Panning, B.** (2007). Differences between homologous alleles of olfactory receptor genes require the Polycomb Group protein Eed. *J Cell Biol* **179**, 269-76.
- Allis, C. D., Berger, S. L., Cote, J., Dent, S., Jenuwein, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhatar, R. et al.** (2007a). New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-6.
- Allis, C. D., Jenuwein, T. and Reinberg, D.** (2007b). Overview and concepts. In *Epigenetics*, (ed. C. D. Allis T. Jenuwein D. Reinberg and M. Caparros), pp. 23-61. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Antequera, F.** (2003). Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci* **60**, 1647-58.
- Antequera, F., Boyes, J. and Bird, A.** (1990). High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* **62**, 503-14.
- Azuara, V.** (2006). Profiling of DNA replication timing in unsynchronized cell populations. *Nat Protoc* **1**, 2171-7.
- Azuara, V., Brown, K. E., Williams, R. R., Webb, N., Dillon, N., Festenstein, R., Buckle, V., Merckenschlager, M. and Fisher, A. G.** (2003). Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication. *Nat Cell Biol* **5**, 668-74.
- Barlow, D. P. and Bartolomei, M. S.** (2007). Genomic imprinting in mammals. In *Epigenetics*, (ed. D. C. Allis T. Jenuwein and D. Reinberg), pp. 357-375. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Barlow, D. P., Stoger, 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**, 84-7.
- Baroux, C., Spillane, C. and Grossniklaus, U.** (2002). Evolutionary origins of the endosperm in flowering plants. *Genome Biol* **3**, reviews1026.
- Bartolomei, M. S., Zemel, S. and Tilghman, S. M.** (1991). Parental imprinting of the mouse H19 gene. *Nature* **351**, 153-5.
- Bell, A. C. and Felsenfeld, G.** (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**, 482-5.
- Berger, S. L., Kouzarides, T., Shiekhatar, R. and Shilatfard, A.** (2009). An operational definition of epigenetics. *Genes Dev* **23**, 781-3.
- Bergstrom, R., Whitehead, J., Kurukuti, S. and Ohlsson, R.** (2007). CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain. *Cell Cycle* **6**, 450-4.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K. et al.** (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-26.
- Bickmore, W. A. and Carothers, A. D.** (1995). Factors affecting the timing and imprinting of replication on a mammalian chromosome. *J Cell Sci* **108 (Pt 8)**, 2801-9.
- Bielinska, B., Blaydes, S. M., Buiting, K., Yang, T., Krajewska-Walasek, M., Horsthemke, B. and Brannan, C. I.** (2000). De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. *Nat Genet* **25**, 74-8.
- Birger, Y., Shemer, R., Perk, J. and Razin, A.** (1999). The imprinting box of the mouse Igf2r gene. *Nature* **397**, 84-8.
- Bogenhagen, D. and Clayton, D. A.** (1977). Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. *Cell* **11**, 719-27.

Bourc'his, D. and Proudhon, C. (2008). Sexual dimorphism in parental imprint ontogeny and contribution to embryonic development. *Mol Cell Endocrinol* **282**, 87-94.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K. et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-53.

Braidotti, G., Baubec, T., Pauler, F., Seidl, C., Smrzka, O., Stricker, S., Yotova, I. and Barlow, D. P. (2004). The Air noncoding RNA: an imprinted cis-silencing transcript. *Cold Spring Harb Symp Quant Biol* **69**, 55-66.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A. and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature* **371**, 435-8.

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. *Mol Cell Biol* **10**, 28-36.

Braun, R. E., Lo, D., Pinkert, C. A., Widera, G., Flavell, R. A., Palmiter, R. D. and Brinster, R. L. (1990). Infertility in male transgenic mice: disruption of sperm development by HSV-tk expression in postmeiotic germ cells. *Biol Reprod* **43**, 684-93.

Bressler, J., Tsai, T. F., Wu, M. Y., Tsai, S. F., Ramirez, M. A., Armstrong, D. and Beaudet, A. L. (2001). The SNRPN promoter is not required for genomic imprinting of the Prader-Willi/Angelman domain in mice. *Nat Genet* **28**, 232-40.

Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafreniere, R. G., Xing, Y., Lawrence, J. and Willard, H. F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**, 527-42.

Brown, S. W. and Nur, U. (1964). Heterochromatic Chromosomes in the Coccids. *Science* **145**, 130-6.

Caspary, T., Cleary, M. A., Baker, C. C., Guan, X. J. and Tilghman, S. M. (1998). Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster. *Mol Cell Biol* **18**, 3466-74.

Cattanach, B. M. and Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315**, 496-8.

Cerrato, F., Dean, W., Davies, K., Kagotani, K., Mitsuya, K., Okumura, K., Riccio, A. and Reik, W. (2003). Paternal imprints can be established on the maternal Igf2-H19 locus without altering replication timing of DNA. *Hum Mol Genet* **12**, 3123-32.

Chai, J. H., Locke, D. P., Ohta, T., Greally, J. M. and Nicholls, R. D. (2001). Retrotransposed genes such as Frat3 in the mouse Chromosome 7C Prader-Willi syndrome region acquire the imprinted status of their insertion site. *Mamm Genome* **12**, 813-21.

Chambeyron, S. and Bickmore, W. A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* **18**, 1119-30.

Charalambous, M., Smith, F. M., Bennett, W. R., Crew, T. E., Mackenzie, F. and Ward, A. (2003). Disruption of the imprinted Grb10 gene leads to disproportionate overgrowth by an Igf2-independent mechanism. *Proc Natl Acad Sci U S A* **100**, 8292-7.

Chaumeil, J., Le Baccon, P., Wutz, A. and Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev* **20**, 2223-37.

Chotalia, M., Smallwood, S. A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W. and Kelsey, G. (2009). Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev* **23**, 105-17.

Clark, L., Wei, M., Cattoretti, G., Mendelsohn, C. and Tycko, B. (2002). The Tnfrh1 (Tnfrsf23) gene is weakly imprinted in several organs and expressed at the trophoblast-decidua interface. *BMC Genet* **3**, 11.

Cohen, D. E., Davidow, L. S., Erwin, J. A., Xu, N., Warshawsky, D. and Lee, J. T. (2007). The DXPas34 repeat regulates random and imprinted X inactivation. *Dev Cell* **12**, 57-71.

Coombes, C., Arnaud, P., Gordon, E., Dean, W., Coar, E. A., Williamson, C. M., Feil, R., Peters, J. and Kelsey, G. (2003). Epigenetic properties and identification of an imprint mark in the Nesp-Gnasxl domain of the mouse Gnas imprinted locus. *Mol Cell Biol* **23**, 5475-88.

Costanzi, C. and Pehrson, J. R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599-601.

Courtier, B., Heard, E. and Avner, P. (1995). Xce haplotypes show modified methylation in a region of the active X chromosome lying 3' to Xist. *Proc Natl Acad Sci U S A* **92**, 3531-5.

Crouse, H. V. (1960). The Controlling Element in Sex Chromosome Behavior in *Sciara*. *Genetics* **45**, 1429-43.

Custodio, N., Carmo-Fonseca, M., Geraghty, F., Pereira, H. S., Grosveld, F. and Antoniou, M. (1999). Inefficient processing impairs release of RNA from the site of transcription. *Embo J* **18**, 2855-66.

da Rocha, S. T., Edwards, C. A., Ito, M., Ogata, T. and Ferguson-Smith, A. C. (2008). Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends Genet* **24**, 306-16.

Dao, D., Frank, D., Qian, N., O'Keefe, D., Vosatka, R. J., Walsh, C. P. and Tycko, B. (1998). IMPT1, an imprinted gene similar to polyspecific transporter and multi-drug resistance genes. *Hum Mol Genet* **7**, 597-608.

Davey, C. and Allan, J. (2003). Nucleosome positioning signals and potential H-DNA within the DNA sequence of the imprinting control region of the mouse Igf2r gene. *Biochim Biophys Acta* **1630**, 103-16.

Davies, K., Bowden, L., Smith, P., Dean, W., Hill, D., Furuumi, H., Sasaki, H., Cattanch, B. and Reik, W. (2002). Disruption of mesodermal enhancers for Igf2 in the minute mutant. *Development* **129**, 1657-68.

de Bustros, A., Nelkin, B. D., Silverman, A., Ehrlich, G., Poiesz, B. and Baylin, S. B. (1988). The short arm of chromosome 11 is a "hot spot" for hypermethylation in human neoplasia. *Proc Natl Acad Sci U S A* **85**, 5693-7.

de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M. et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* **7**, 663-76.

DeChiara, T. M., Efstratiadis, A. and Robertson, E. J. (1990). A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature* **345**, 78-80.

DeChiara, T. M., Robertson, E. J. and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64**, 849-59.

Dodge, J. E., Kang, Y. K., Beppu, H., Lei, H. and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. *Mol Cell Biol* **24**, 2478-86.

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. *Hum Mol Genet* **17**, 3021-9.

Engelstadter, J. (2008). Constraints on the evolution of asexual reproduction. *Bioessays* **30**, 1138-50.

Engemann, S., Stroedicke, M., Paulsen, M., Franck, O., Reinhardt, R., Lane, N., Reik, W. and Walter, J. (2000). Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. *Hum Mol Genet* **9**, 2691-706.

Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* **21**, 5427-40.

Fang, J., Chen, T., Chadwick, B., Li, E. and Zhang, Y. (2004). Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J Biol Chem* **279**, 52812-5.

Faust, C., Schumacher, A., Holdener, B. and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. *Development* **121**, 273-85.

Feil, R., Walter, J., Allen, N. D. and Reik, W. (1994). Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes. *Development* **120**, 2933-43.

Fitzpatrick, G. V., Soloway, P. D. and Higgins, M. J. (2002). Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat Genet* **32**, 426-31.

Frank, D., Fortino, W., Clark, L., Musalo, R., Wang, W., Saxena, A., Li, C. M., Reik, W., Ludwig, T. and Tycko, B. (2002). Placental overgrowth in mice lacking the imprinted gene Ipl. *Proc Natl Acad Sci U S A* **99**, 7490-5.

Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E. and Busslinger, M. (2004). Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev* **18**, 411-22.

Gabory, A., Ripoche, M. A., Le Digarcher, A., Watrin, F., Ziyat, A., Forne, T., Jammes, H., Ainscough, J. F., Surani, M. A., Journot, L. et al. (2009). H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development* **136**, 3413-21.

Gardiner-Garden, M. and Frommer, M. (1987). CpG islands in vertebrate genomes. *J Mol Biol* **196**, 261-82.

Gingeras, T. R. (2007). Origin of phenotypes: genes and transcripts. *Genome Res* **17**, 682-90.

Goday, C. and Esteban, M. R. (2001). Chromosome elimination in sciarid flies. *Bioessays* **23**, 242-50.

Gould, T. D. and Pfeifer, K. (1998). Imprinting of mouse Kvlqt1 is developmentally regulated. *Hum Mol Genet* **7**, 483-7.

Greally, J. M., Starr, D. J., Hwang, S., Song, L., Jaarola, M. and Zemel, S. (1998). The mouse H19 locus mediates a transition between imprinted and non-imprinted DNA replication patterns. *Hum Mol Genet* **7**, 91-5.

Green, K., Lewis, A., Dawson, C., Dean, W., Reinhart, B., Chaillet, J. R. and Reik, W. (2007). A developmental window of opportunity for imprinted gene silencing mediated by DNA methylation and the Kcnq1ot1 noncoding RNA. *Mamm Genome* **18**, 32-42.

Gribnau, J., Hochedlinger, K., Hata, K., Li, E. and Jaenisch, R. (2003). Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* **17**, 759-73.

Guillemot, F., Caspary, T., Tilghman, S. M., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Anderson, D. J., Joyner, A. L., Rossant, J. and Nagy, A. (1995). Genomic imprinting of Mash2, a mouse gene required for trophoblast development. *Nat Genet* **9**, 235-42.

Gunaratne, P. H., Nakao, M., Ledbetter, D. H., Sutcliffe, J. S. and Chinault, A. C. (1995). Tissue-specific and allele-specific replication timing control in the imprinted human Prader-Willi syndrome region. *Genes Dev* **9**, 808-20.

Hagiwara, Y., Hirai, M., Nishiyama, K., Kanazawa, I., Ueda, T., Sakaki, Y. and Ito, T. (1997). Screening for imprinted genes by allelic message display: identification of a paternally expressed gene impact on mouse chromosome 18. *Proc Natl Acad Sci U S A* **94**, 9249-54.

Hanel, M. L. and Wevrick, R. (2001). Establishment and maintenance of DNA methylation patterns in mouse Ndn: implications for maintenance of imprinting in target genes of the imprinting center. *Mol Cell Biol* **21**, 2384-92.

Hansen, R. S., Canfield, T. K., Lamb, M. M., Gartler, S. M. and Laird, C. D. (1993). Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* **73**, 1403-9.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M. and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**, 486-9.

- Hata, K., Okano, M., Lei, H. and Li, E.** (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**, 1983-93.
- Helbig, R. and Fackelmayer, F. O.** (2003). Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. *Chromosoma* **112**, 173-82.
- Herman, H., Lu, M., Anggraini, M., Sikora, A., Chang, Y., Yoon, B. J. and Soloway, P. D.** (2003). Trans allele methylation and paramutation-like effects in mice. *Nat Genet* **34**, 199-202.
- Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R. and Sasaki, H.** (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev* **22**, 1607-16.
- Holmes, R., Williamson, C., Peters, J., Denny, P. and Wells, C.** (2003). A comprehensive transcript map of the mouse Gnas imprinted complex. *Genome Res* **13**, 1410-5.
- Hudson, Q. J., Kulinski, T. M., Huetter, S. P. and Barlow, D. P.** (2010). Genomic imprinting mechanisms in embryonic and extraembryonic mouse tissues. *Heredity in press*.
- Hurst, L. D. and McVean, G. T.** (1998). Do we understand the evolution of genomic imprinting? *Curr Opin Genet Dev* **8**, 701-8.
- Hutter, B., Helms, V. and Paulsen, M.** (2006). Tandem repeats in the CpG islands of imprinted genes. *Genomics* **88**, 323-32.
- Ideraabdullah, F. Y., Vigneau, S. and Bartolomei, M. S.** (2008). Genomic imprinting mechanisms in mammals. *Mutat Res* **647**, 77-85.
- Illingworth, R. S. and Bird, A. P.** (2009). CpG islands--'a rough guide'. *FEBS Lett* **583**, 1713-20.
- Itier, J. M., Tremp, G. L., Leonard, J. F., Multon, M. C., Ret, G., Schweighoffer, F., Tocque, B., Bluet-Pajot, M. T., Cormier, V. and Dautry, F.** (1998). Imprinted gene in postnatal growth role. *Nature* **393**, 125-6.
- Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. and Cheng, X.** (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* **449**, 248-51.
- Johnson, D. R.** (1974). Hairpin-tail: a case of post-reductional gene action in the mouse egg. *Genetics* **76**, 795-805.
- Johnson, D. R.** (1975). Further observations on the hairpin tail (Thp) mutation in the mouse. *Genet Res* **24**, 207-213.
- Kagotani, K., Takebayashi, S., Kohda, A., Taguchi, H., Paulsen, M., Walter, J., Reik, W. and Okumura, K.** (2002). Replication timing properties within the mouse distal chromosome 7 imprinting cluster. *Biosci Biotechnol Biochem* **66**, 1046-51.
- Kawahara, M., Wu, Q., Takahashi, N., Morita, S., Yamada, K., Ito, M., Ferguson-Smith, A. C. and Kono, T.** (2007). High-frequency generation of viable mice from engineered bi-maternal embryos. *Nat Biotechnol* **25**, 1045-50.
- Kelsey, G., Bodle, D., Miller, H. J., Beechey, C. V., Coombes, C., Peters, J. and Williamson, C. M.** (1999). Identification of imprinted loci by methylation-sensitive representational difference analysis: application to mouse distal chromosome 2. *Genomics* **62**, 129-38.
- Kernohan, K. D., Jiang, Y., Tremblay, D. C., Bonvissuto, A. C., Eubanks, J. H., Mann, M. R. and Berube, N. G.** (2010). ATRX Partners with Cohesin and MeCP2 and Contributes to Developmental Silencing of Imprinted Genes in the Brain. *Dev Cell* **18**, 191-202.
- Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D. J., Nicholls, R. D. and Cedar, H.** (1993). Allele-specific replication timing of imprinted gene regions. *Nature* **364**, 459-63.

- Knoll, J. H., Cheng, S. D. and Lalonde, M.** (1994). Allele specificity of DNA replication timing in the Angelman/Prader-Willi syndrome imprinted chromosomal region. *Nat Genet* **6**, 41-6.
- Koerner, M. V.** (2006). Mapping the limits of replication asynchrony around the imprinted *Igf2r*/Air gene cluster. In *Fakultät für Lebenswissenschaften*, vol. Diploma thesis (ed., pp. 96. Vienna: Universität Wien.
- Koerner, M. V., Pauler, F. M., Huang, R. and Barlow, D. P.** (2009). The function of non-coding RNAs in genomic imprinting. *Development* **136**, 1771-83.
- Kouzarides, T.** (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.
- Kuzmin, A., Han, Z., Golding, M. C., Mann, M. R., Latham, K. E. and Varmuza, S.** (2008). The PcG gene *Sfmbt2* is paternally expressed in extraembryonic tissues. *Gene Expr Patterns* **8**, 107-16.
- Landers, M., Bancescu, D. L., Le Meur, E., Rougeulle, C., Glatt-Deeley, H., Brannan, C., Muscatelli, F. and Lalonde, M.** (2004). Regulation of the large (approximately 1000 kb) imprinted murine *Ube3a* antisense transcript by alternative exons upstream of *Snurf/Snrpn*. *Nucleic Acids Res* **32**, 3480-92.
- Larsen, F., Gundersen, G., Lopez, R. and Prydz, H.** (1992). CpG islands as gene markers in the human genome. *Genomics* **13**, 1095-107.
- 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* **136**, 437-48.
- Leeb, M. and Wutz, A.** (2007). Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *J Cell Biol* **178**, 219-29.
- Lefebvre, L., Viville, S., Barton, S. C., Ishino, F., Keverne, E. B. and Surani, M. A.** (1998). Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene *Mest*. *Nat Genet* **20**, 163-9.
- Leff, S. E., Brannan, C. I., Reed, M. L., Ozcelik, T., Francke, U., Copeland, N. G. and Jenkins, N. A.** (1992). Maternal imprinting of the mouse *Snrpn* gene and conserved linkage homology with the human Prader-Willi syndrome region. *Nat Genet* **2**, 259-64.
- Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R. and Li, E.** (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**, 3195-205.
- Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S. M.** (1995). Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* **375**, 34-9.
- 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. *Mech Dev* **61**, 141-9.
- Lewis, A., Green, K., Dawson, C., Redrup, L., Huynh, K. D., Lee, J. T., Hemberger, M. and Reik, W.** (2006). Epigenetic dynamics of the *Kcnq1* imprinted domain in the early embryo. *Development* **133**, 4203-10.
- Lewis, A., Mitsuya, K., Constancia, M. and Reik, W.** (2004a). Tandem repeat hypothesis in imprinting: deletion of a conserved direct repeat element upstream of H19 has no effect on imprinting in the *Igf2*-H19 region. *Mol Cell Biol* **24**, 5650-6.
- Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., Higgins, M., Feil, R. and Reik, W.** (2004b). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet* **36**, 1291-5.
- Li, E., Beard, C. and Jaenisch, R.** (1993). Role for DNA methylation in genomic imprinting. *Nature* **366**, 362-5.
- Li, E., Bestor, T. H. and Jaenisch, R.** (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-26.

- Li, L., Keverne, E. B., Aparicio, S. A., Ishino, F., Barton, S. C. and Surani, M. A.** (1999). Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science* **284**, 330-3.
- Li, T., Vu, T. H., Zeng, Z. L., Nguyen, B. T., Hayward, B. E., Bonthron, D. T., Hu, J. F. and Hoffman, A. R.** (2000). Tissue-specific expression of antisense and sense transcripts at the imprinted Gnas locus. *Genomics* **69**, 295-304.
- Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P. and Ferguson-Smith, A. C.** (2008). A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell* **15**, 547-57.
- Lin, S. P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J. and Ferguson-Smith, A. C.** (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. *Nat Genet* **35**, 97-102.
- Liu, J., Yu, S., Litman, D., Chen, W. and Weinstein, L. S.** (2000). Identification of a methylation imprint mark within the mouse Gnas locus. *Mol Cell Biol* **20**, 5808-17.
- 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. *Nat Genet* **25**, 19-21.
- Lyon, M. F. and Glenister, P. H.** (1977). Factors affecting the observed number of young resulting from adjacent-2 disjunction in mice carrying a translocation. *Genet Res* **29**, 83-92.
- Macleod, D., Charlton, J., Mullins, J. and Bird, A. P.** (1994). Sp1 sites in the mouse apt gene promoter are required to prevent methylation of the CpG island. *Genes Dev* **8**, 2282-92.
- Mager, J., Montgomery, N. D., de Villena, F. P. and Magnuson, T.** (2003). Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat Genet* **33**, 502-7.
- Mancini-DiNardo, D., Steele, S. J., Ingram, R. S. and Tilghman, S. M.** (2003). A differentially methylated region within the gene Kcnq1 functions as an imprinted promoter and silencer. *Hum Mol Genet* **12**, 283-94.
- Mancini-Dinardo, D., Steele, S. J., 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 Dev* **20**, 1268-82.
- Marin, M., Karis, A., Visser, P., Grosveld, F. and Philipson, S.** (1997). Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. *Cell* **89**, 619-28.
- Marks, P. A., Richon, V. M., Breslow, R. and Rifkind, R. A.** (2001). Histone deacetylase inhibitors as new cancer drugs. *Curr Opin Oncol* **13**, 477-83.
- Mattick, J. S. and Makunin, I. V.** (2006). Non-coding RNA. *Hum Mol Genet* **15 Spec No 1**, R17-29.
- McGrath, J. and Solter, D.** (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37**, 179-83.
- Mermoud, J. E., Costanzi, C., Pehrson, J. R. and Brockdorff, N.** (1999). Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J Cell Biol* **147**, 1399-408.
- Mietton, F., Sengupta, A. K., Molla, A., Picchi, G., Barral, S., Heliot, L., Grange, T., Wutz, A. and Dimitrov, S.** (2009). Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. *Mol Cell Biol* **29**, 150-6.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P. et al.** (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-60.
- Miri, K. and Varmuza, S.** (2009). Imprinting and extraembryonic tissues-mom takes control. *Int Rev Cell Mol Biol* **276**, 215-62.

- Moldon, A., Malapeira, J., Gabrielli, N., Gogol, M., Gomez-Escoda, B., Ivanova, T., Seidel, C. and Ayte, J.** (2008). Promoter-driven splicing regulation in fission yeast. *Nature* **455**, 997-1000.
- Montgomery, N. D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S. J., Otte, A. P. and Magnuson, T.** (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol* **15**, 942-7.
- Moon, Y. S., Smas, C. M., Lee, K., Villena, J. A., Kim, K. H., Yun, E. J. and Sul, H. S.** (2002). Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. *Mol Cell Biol* **22**, 5585-92.
- Moore, T. and Haig, D.** (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* **7**, 45-9.
- Mooslehner, K. A. and Allen, N. D.** (1999). Cloning of the mouse organic cation transporter 2 gene, Slc22a2, from an enhancer-trap transgene integration locus. *Mamm Genome* **10**, 218-24.
- Mummaneni, P., Yates, P., Simpson, J., Rose, J. and Turker, M. S.** (1998). The primary function of a redundant Sp1 binding site in the mouse aprt gene promoter is to block epigenetic gene inactivation. *Nucleic Acids Res* **26**, 5163-9.
- 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. *Nat Genet* **36**, 889-93.
- 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**, 1717-20.
- Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R.** (2003). Manipulating the Mouse Embryo. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Neumann, B., Kubicka, P. and Barlow, D. P.** (1995). Characteristics of imprinted genes. *Nat Genet* **9**, 12-3.
- Okamoto, I., Otte, A. P., Allis, C. D., Reinberg, D. and Heard, E.** (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303**, 644-9.
- Okamura, K. and Ito, T.** (2006). Lessons from comparative analysis of species-specific imprinted genes. *Cytogenet Genome Res* **113**, 159-64.
- Okamura, K., Wintle, R. F. and Scherer, S. W.** (2008). Characterization of the differentially methylated region of the Impact gene that exhibits Glres-specific imprinting. *Genome Biol* **9**, R160.
- Ono, R., Nakamura, K., Inoue, K., Naruse, M., Usami, T., Wakisaka-Saito, N., Hino, T., Suzuki-Migishima, R., Ogonuki, N., Miki, H. et al.** (2006). Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat Genet* **38**, 101-6.
- Pachnis, V., Belayew, A. and Tilghman, S. M.** (1984). Locus unlinked to alpha-fetoprotein under the control of the murine raf and Rif genes. *Proc Natl Acad Sci U S A* **81**, 5523-7.
- Paldi, A., Gyapay, G. and Jami, J.** (1995). Imprinted chromosomal regions of the human genome display sex-specific meiotic recombination frequencies. *Curr Biol* **5**, 1030-5.
- Pandey, R. R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D. and Kanduri, C.** (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* **32**, 232-46.
- Pannetier, M., Julien, E., Schotta, G., Tardat, M., Sardet, C., Jenuwein, T. and Feil, R.** (2008). PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. *EMBO Rep* **9**, 998-1005.

- Pant, V., Mariano, P., Kanduri, C., Mattsson, A., Lobanenkov, V., Heuchel, R. and Ohlsson, R.** (2003). The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the H19 imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains. *Genes Dev* **17**, 586-90.
- Paoloni-Giacobino, A., D'Aiuto, L., Cirio, M. C., Reinhart, B. and Chaillet, J. R.** (2007). Conserved features of imprinted differentially methylated domains. *Gene* **399**, 33-45.
- Pauler, F. M., Koerner, M. V. and Barlow, D. P.** (2007). Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet* **23**, 284-92.
- Pauler, F. M., Sloane, M. A., Huang, R., Regha, K., Koerner, M. V., Tamir, I., Sommer, A., Aszodi, A., Jenuwein, T. and Barlow, D. P.** (2009). H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res* **19**, 221-33.
- Pauler, F. M., Stricker, S. H., Warczok, K. E. and Barlow, D. P.** (2005). Long-range DNase I hypersensitivity mapping reveals the imprinted *Igf2r* and *Air* promoters share cis-regulatory elements. *Genome Res* **15**, 1379-87.
- Peters, J. and Williamson, C. M.** (2008). Control of imprinting at the *Gnas* cluster. *Adv Exp Med Biol* **626**, 16-26.
- Plath, K., Talbot, D., Hamer, K. M., Otte, A. P., Yang, T. P., Jaenisch, R. and Panning, B.** (2004). Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J Cell Biol* **167**, 1025-35.
- Prissette, M., El-Maarri, O., Arnaud, D., Walter, J. and Avner, P.** (2001). Methylation profiles of DXPas34 during the onset of X-inactivation. *Hum Mol Genet* **10**, 31-8.
- Pullirsch, D., Hartel, R., Kishimoto, H., Leeb, M., Steiner, G. and Wutz, A.** (2010). The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. *Development* **137**, 935-43.
- Qian, N., Frank, D., O'Keefe, D., Dao, D., Zhao, L., Yuan, L., Wang, Q., Keating, M., Walsh, C. and Tycko, B.** (1997). The IPL gene on chromosome 11p15.5 is imprinted in humans and mice and is similar to TDAG51, implicated in Fas expression and apoptosis. *Hum Mol Genet* **6**, 2021-9.
- Rasmussen, T. P., Mastrangelo, M. A., Eden, A., Pehrson, J. R. and Jaenisch, R.** (2000). Dynamic relocalization of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation. *J Cell Biol* **150**, 1189-98.
- Reed, M. R., Riggs, A. D. and Mann, J. R.** (2001). Deletion of a direct repeat element has no effect on *Igf2* and H19 imprinting. *Mamm Genome* **12**, 873-6.
- Regha, K., Latos, P. A. and Spahn, L.** (2006). The imprinted mouse *Igf2r/Air* cluster--a model maternal imprinting system. *Cytogenet Genome Res* **113**, 165-77.
- Regha, K., Sloane, M. A., Huang, R., Pauler, F. M., Warczok, K. E., Melikant, B., Radolf, M., Martens, J. H., Schotta, G., Jenuwein, T. et al.** (2007). Active and repressive chromatin are interspersed without spreading in an imprinted gene cluster in the mammalian genome. *Mol Cell* **27**, 353-66.
- Reinhart, B., Eljanne, M. and Chaillet, J. R.** (2002). Shared role for differentially methylated domains of imprinted genes. *Mol Cell Biol* **22**, 2089-98.
- Reinhart, B., Paoloni-Giacobino, A. and Chaillet, J. R.** (2006). Specific differentially methylated domain sequences direct the maintenance of methylation at imprinted genes. *Mol Cell Biol* **26**, 8347-56.
- Renfree, M. B., Hore, T. A., Shaw, G., Graves, J. A. and Pask, A. J.** (2009). Evolution of genomic imprinting: insights from marsupials and monotremes. *Annu Rev Genomics Hum Genet* **10**, 241-62.

- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, 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**, 1311-23.
- Robinson, W. P. and Lalande, M. (1995). Sex-specific meiotic recombination in the Prader-Willi/Angelman syndrome imprinted region. *Hum Mol Genet* **4**, 801-6.
- Rougeulle, C., Glatt, H. and Lalande, M. (1997). The Angelman syndrome candidate gene, UBE3A/E6-AP, is imprinted in brain. *Nat Genet* **17**, 14-5.
- Rougeulle, C. and Heard, E. (2002). Antisense RNA in imprinting: spreading silence through Air. *Trends Genet* **18**, 434-7.
- Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-86.
- Ryu, W. S. and Mertz, J. E. (1989). Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. *J Virol* **63**, 4386-94.
- Sambrook, J. and Russel, D. W. (2001). *Molecular Cloning - A Laboratory Manual* (Third Edition). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Sandelin, A., Carninci, P., Lenhard, B., Ponjavic, J., Hayashizaki, Y. and Hume, D. A. (2007). Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet* **8**, 424-36.
- Sandhu, K. S., Shi, C., Sjolinder, M., Zhao, Z., Gondor, A., Liu, L., Tiwari, V. K., Guibert, S., Emilsson, L., Imreh, M. P. et al. (2009). Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. *Genes Dev* **23**, 2598-603.
- Schmidt, J. V., Levorse, J. M. and Tilghman, S. M. (1999). Enhancer competition between H19 and Igf2 does not mediate their imprinting. *Proc Natl Acad Sci U S A* **96**, 9733-8.
- Schoeftner, S., Sengupta, A. K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T. and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *Embo J* **25**, 3110-22.
- Seidl, C. I. (2006). Transcriptional characterization of the Air ncRNA and its role in regulation of imprinted expression in the Igf2r cluster in visceral yolk sac. In *Fakultaet fuer Lebenswissenschaften*, vol. Dr. rer. nat. (ed., pp. 97. Vienna: University of Vienna.
- Seidl, C. I., Stricker, S. H. and Barlow, D. P. (2006). The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *Embo J* **25**, 3565-75.
- Seitz, H., Royo, H., Bortolin, M. L., Lin, S. P., Ferguson-Smith, A. C. and Cavaille, J. (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* **14**, 1741-8.
- Selig, S., Okumura, K., Ward, D. C. and Cedar, H. (1992). Delineation of DNA replication time zones by fluorescence in situ hybridization. *Embo J* **11**, 1217-25.
- Sharman, G. B. (1971). Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* **230**, 231-2.
- Shearwin, K. E., Callen, B. P. and Egan, J. B. (2005). Transcriptional interference--a crash course. *Trends Genet* **21**, 339-45.
- Shemer, R., Birger, Y., Riggs, A. D. and Razin, A. (1997). Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. *Proc Natl Acad Sci U S A* **94**, 10267-72.
- Shi, W., Lefebvre, L., Yu, Y., Otto, S., Krella, A., Orth, A. and Fundele, R. (2004). Loss-of-imprinting of Peg1 in mouse interspecies hybrids is correlated with altered growth. *Genesis* **39**, 65-72.
- Shin, J. Y., Fitzpatrick, G. V. and Higgins, M. J. (2008). Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *Embo J* **27**, 168-78.

- Shuster, M., Dhar, M. S., Olins, A. L., Olins, D. E., Howell, C. Y., Gollin, S. M. and Chaillet, J. R.** (1998). Parental alleles of an imprinted mouse transgene replicate synchronously. *Dev Genet* **23**, 275-84.
- Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P. and Brockdorff, N.** (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* **4**, 481-95.
- Simon, I., Tenzen, T., Reubinoff, B. E., Hillman, D., McCarrey, J. R. and Cedar, H.** (1999). Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* **401**, 929-32.
- 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. *Embo J* **22**, 3696-704.
- Sleutels, F., Zwart, R. and Barlow, D. P.** (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**, 810-3.
- Smilnich, N. J., Day, C. D., Fitzpatrick, G. V., Caldwell, G. M., Lossie, A. C., Cooper, P. R., Smallwood, A. C., Joyce, J. A., Schofield, P. N., Reik, W. et al.** (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc Natl Acad Sci U S A* **96**, 8064-9.
- Smrzka, O. W., Fae, I., Stoger, R., Kurzbauer, R., Fischer, G. F., Henn, T., Weith, A. and Barlow, D. P.** (1995). Conservation of a maternal-specific methylation signal at the human IGF2R locus. *Hum Mol Genet* **4**, 1945-52.
- Solter, D.** (1988). Differential imprinting and expression of maternal and paternal genomes. *Annu Rev Genet* **22**, 127-46.
- Stoger, 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.
- Stricker, S. H., Steenpass, L., Pauler, F. M., Santoro, F., Latos, P. A., Huang, R., Koerner, M. V., Sloane, M. A., Warczok, K. E. and Barlow, D. P.** (2008). Silencing and transcriptional properties of the imprinted Airn ncRNA are independent of the endogenous promoter. *Embo J* **27**, 3116-28.
- Sunahara, S., Nakamura, K., Nakao, K., Gondo, Y., Nagata, Y. and Katsuki, M.** (2000). The oocyte-specific methylated region of the U2afbp-rs/U2af1-rs1 gene is dispensable for its imprinted methylation. *Biochem Biophys Res Commun* **268**, 590-5.
- 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**, 548-50.
- Suske, G.** (1999). The Sp-family of transcription factors. *Gene* **238**, 291-300.
- Sutcliffe, J. S., Nakao, M., Christian, S., Orstavik, K. H., Tommerup, N., Ledbetter, D. H. and Beaudet, A. L.** (1994). Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nat Genet* **8**, 52-8.
- Swain, J. L., Stewart, T. A. and Leder, P.** (1987). Parental legacy determines methylation and expression of an autosomal transgene: a molecular mechanism for parental imprinting. *Cell* **50**, 719-27.
- Szabo, P. E. and Mann, J. R.** (1995). Allele-specific expression and total expression levels of imprinted genes during early mouse development: implications for imprinting mechanisms. *Genes Dev* **9**, 3097-108.
- Takada, S., Paulsen, M., Tevendale, M., Tsai, C. E., Kelsey, G., Cattanach, B. M. and Ferguson-Smith, A. C.** (2002). Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. *Hum Mol Genet* **11**, 77-86.
- Takagi, N. and Sasaki, M.** (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256**, 640-2.

Terranova, R., Yokobayashi, S., Stadler, M. B., Otte, A. P., van Lohuizen, M., Orkin, S. H. and Peters, A. H. (2008). Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell* **15**, 668-79.

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 Dev* **12**, 3693-702.

Thorvaldsen, J. L., Mann, M. R., Nwoko, O., Duran, K. L. and Bartolomei, M. S. (2002). Analysis of sequence upstream of the endogenous H19 gene reveals elements both essential and dispensable for imprinting. *Mol Cell Biol* **22**, 2450-62.

Tierling, S., Dalbert, S., Schoppenhorst, S., Tsai, C. E., Oliger, S., Ferguson-Smith, A. C., Paulsen, M. and Walter, J. (2006). High-resolution map and imprinting analysis of the Gtl2-Dnchc1 domain on mouse chromosome 12. *Genomics* **87**, 225-35.

Todaro, G. J. and Green, H. (1963). Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J Cell Biol* **17**, 299-313.

Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M. and Bartolomei, M. S. (1995). A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. *Nat Genet* **9**, 407-13.

Tucker, K. L., Beard, C., Dausmann, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E. and Jaenisch, R. (1996). Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev* **10**, 1008-20.

Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y. and Feil, R. (2004). Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**, 1296-300.

Varmuza, S. and Mann, M. (1994). Genomic imprinting--defusing the ovarian time bomb. *Trends Genet* **10**, 118-23.

Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Akin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A. et al. (2006). Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. *Dev Cell* **11**, 711-22.

Verhaagh, S., Barlow, D. P. and Zwart, R. (2001). The extraneuronal monoamine transporter Slc22a3/Orct3 co-localizes with the Maa0 metabolizing enzyme in mouse placenta. *Mech Dev* **100**, 127-30.

Verona, R. I., Mann, M. R. and Bartolomei, M. S. (2003). Genomic imprinting: intricacies of epigenetic regulation in clusters. *Annu Rev Cell Dev Biol* **19**, 237-59.

Vigneau, S., Augui, S., Navarro, P., Avner, P. and Clerc, P. (2006). An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. *Proc Natl Acad Sci U S A* **103**, 7390-5.

Vrana, P. B., Guan, X. J., Ingram, R. S. and Tilghman, S. M. (1998). Genomic imprinting is disrupted in interspecific *Peromyscus* hybrids. *Nat Genet* **20**, 362-5.

Wagschal, A. and Feil, R. (2006). Genomic imprinting in the placenta. *Cytogenet Genome Res* **113**, 90-8.

Wagschal, A., Sutherland, H. G., Woodfine, K., Henckel, A., Chebli, K., Schulz, R., Oakey, R. J., Bickmore, W. A. and Feil, R. (2008). G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Mol Cell Biol* **28**, 1104-13.

Wang, Z. Q., Fung, M. R., Barlow, D. P. and Wagner, E. F. (1994). Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature* **372**, 464-7.

Wassenegger, M. (2005). The role of the RNAi machinery in heterochromatin formation. *Cell* **122**, 13-6.

- Watanabe, T., Yoshimura, A., Mishima, Y., Endo, Y., Shiroishi, T., Koide, T., Sasaki, H., Asakura, H. and Kominami, R.** (2000). Differential chromatin packaging of genomic imprinted regions between expressed and non-expressed alleles. *Hum Mol Genet* **9**, 3029-35.
- Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T. et al.** (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796-801.
- Wilkie, A. O.** (1994). The molecular basis of genetic dominance. *J Med Genet* **31**, 89-98.
- Williamson, C. M., Turner, M. D., Ball, S. T., Nottingham, W. T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G. et al.** (2006). Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. *Nat Genet* **38**, 350-5.
- Windham, C. Q. and Jones, P. A.** (1997). Expression of H19 does not influence the timing of replication of the Igf2/H19 imprinted region. *Dev Genet* **20**, 29-35.
- Wood, A. J., Schulz, R., Woodfine, K., Koltowska, K., Beechey, C. V., Peters, J., Bourc'his, D. and Oakey, R. J.** (2008). Regulation of alternative polyadenylation by genomic imprinting. *Genes Dev* **22**, 1141-6.
- Wroe, S. F., Kelsey, G., Skinner, J. A., Bodle, D., Ball, S. T., Beechey, C. V., Peters, J. and Williamson, C. M.** (2000). An imprinted transcript, antisense to Nesp, adds complexity to the cluster of imprinted genes at the mouse Gnas locus. *Proc Natl Acad Sci U S A* **97**, 3342-6.
- Wutz, A. and Gribnau, J.** (2007). X inactivation Xplained. *Curr Opin Genet Dev* **17**, 387-93.
- Wutz, A., Rasmussen, T. P. and Jaenisch, R.** (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* **30**, 167-74.
- 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**, 745-9.
- 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**, 1881-7.
- Yamasaki, Y., Kayashima, T., Soejima, H., Kinoshita, A., Yoshiura, K., Matsumoto, N., Ohta, T., Urano, T., Masuzaki, H., Ishimaru, T. et al.** (2005). Neuron-specific relaxation of Igf2r imprinting is associated with neuron-specific histone modifications and lack of its antisense transcript Air. *Hum Mol Genet* **14**, 2511-20.
- Yokota, H., Singer, M. J., van den Engh, G. J. and Trask, B. J.** (1997). Regional differences in the compaction of chromatin in human G0/G1 interphase nuclei. *Chromosome Res* **5**, 157-66.
- Yoon, B. J., Herman, H., Sikora, A., Smith, L. T., Plass, C. and Soloway, P. D.** (2002). Regulation of DNA methylation of Rasgrf1. *Nat Genet* **30**, 92-6.
- Yoshida, M., Hoshikawa, Y., Koseki, K., Mori, K. and Beppu, T.** (1990). Structural specificity for biological activity of trichostatin A, a specific inhibitor of mammalian cell cycle with potent differentiation-inducing activity in Friend leukemia cells. *J Antibiot (Tokyo)* **43**, 1101-6.
- Yu, S., Gavrillova, O., Chen, H., Lee, R., Liu, J., Pacak, K., Parlow, A. F., Quon, M. J., Reitman, M. L. and Weinstein, L. S.** (2000). Paternal versus maternal transmission of a stimulatory G-protein alpha subunit knockout produces opposite effects on energy metabolism. *J Clin Invest* **105**, 615-23.
- Zakharova, I. S., Shevchenko, A. I. and Zakian, S. M.** (2009). Monoallelic gene expression in mammals. *Chromosoma* **118**, 279-90.

Zupkovitz, G., Tischler, J., Posch, M., Sadzak, I., Ramsauer, K., Egger, G., Grausenburger, R., Schweifer, N., Chiocca, S., Decker, T. et al. (2006). Negative and positive regulation of gene expression by mouse histone deacetylase 1. *Mol Cell Biol* **26**, 7913-28.

Zwart, R., Sleutels, F., Wutz, A., Schinkel, A. H. and Barlow, D. P. (2001). Bidirectional action of the *Igf2r* imprint control element on upstream and downstream imprinted genes. *Genes Dev* **15**, 2361-6.

6. CURRICULUM VITAE

6.1 Personal data

Name	Martha Veronika Körner
Date of birth	13 th of February, 1982
Place of birth	Schärding, Austria
Parents	Veronika and Harald Körner
Marital status	single

6.2 Education

09/1988 – 07/1992	Elementary school in Esternberg (Austria)
09/1992 – 06/2000	Gymnasium in Schärding (Austria) 21 st of June 2000 Matura exam passed with distinction
10/2000 – 10/2002	Study of Biology (A437) at the University of Vienna
10/2002 – 02/2006	Study of Microbiology and Genetics (A441) at the University of Vienna, passed with distinction
12/2004 – 01/2006	Diploma thesis 'Mapping the limits of replication asynchrony around the imprinted <i>Igf2r/Air</i> gene cluster' in the research group of Denise P. Barlow, PhD CeMM - Center for Molecular Medicine of the Austrian Academy of Sciences c/o Department of Microbiology, Immunobiology and Genetics Dr. Bohrg. 9/4 1030 Wien
1 st of February 2006	Diploma exam, passed with distinction
since 02/2006	Dissertation 'Dissecting the multiple roles of the <i>Igf2r</i> imprint control element' in the research group of Denise P. Barlow, PhD CeMM - Center for Molecular Medicine of the Austrian Academy of Sciences c/o Department of Microbiology, Immunobiology and Genetics Dr. Bohrg. 9/4 1030 Wien

6.3 Conferences

- 11.4.05 – 14.4.05 'Genomic imprinting, development and disease Conference', Oxford (United Kingdom)
poster: 'Analysis of replication asynchrony around the imprinted mouse *Igf2r* domain by fluorescence *in situ* hybridisation'
Martha V. Koerner, Tuncay Baubec and Denise P. Barlow
awarded with 1st poster price
- 27.6.06 – 30.6.06 'NoE Annual Meeting 2006', Naples (Italy)
poster: 'Replication asynchrony along the imprinted mouse *Igf2r* cluster extends over 3Mb and is independent of *Air* and the ICE'
Martha V. Koerner, Tuncay Baubec and Denise P. Barlow
- 28.10.06 – 31.10.06 'HEROIC Annual Meeting 2006', Uppsala (Sweden)
talk: 'Replication asynchrony at the 500kb imprinted *Igf2r* cluster extends over 3Mbp and is independent of the ICE and the *Air* ncRNA'
- 18.12.06 – 19.12.06 '25 years of embryonic stem cells in Cambridge', Cambridge (United Kingdom)
poster: 'An ES cell differentiation model system to study the developmental onset of imprinted *Igf2r* expression'
Laura Spahn, Paulina Latos, Florian M. Pauler, Stefan H. Stricker, Martha V. Koerner and Denise P. Barlow
- 5.8.07 – 11.8.07 'Gordon Research Conference on Epigenetics', Holderness, NH (USA)
poster: 'Asynchronous behaviour of a large chromatin domain containing the imprinted *Igf2r* gene cluster'
Martha V. Koerner, Tuncay Baubec and Denise P. Barlow
- 10.10.07 – 13.10.07 'HEROIC Annual Meeting 2007', Nijmegen (Netherlands)
- 29.6.08 – 4.7.08 'Gordon Research Conference on the biology of post-transcriptional gene regulation', Waterville, ME (USA)
poster: 'Characterisation of co- and post-transcriptional features of the imprinted *Air* ncRNA and identification of functional elements with its transcriptional unit'
Martha V. Koerner, Stefan H. Stricker, Christine IM. Seidl, Florian M. Pauler and Denise P. Barlow
- 6.10.08 – 8.10.08 'HEROIC Annual Meeting 2008', Vienna (Austria)
talk: 'Parental-specific nuclear organisation of the imprinted *Igf2r* cluster'

- 29.8.09 – 01.9.09 'EMBO meeting 2009', Amsterdam (Netherlands)
- 3.9.09 – 6.9.09 'From imprinting to the epigenome in 25 years', Cambridge (United Kingdom)
- 4.10.09 – 7.10.09 'HEROIC Annual Meeting 2009', Paris (France)
talk: 'Repeats within imprinted macro ncRNAs – functional or ICEing on the cake?'

6.4 Publications

Pauler FM*, Koerner MV*, Barlow DP. Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet.* 2007 Jun;23(6):284-92.

Stricker SH, Steenpass L, Pauler FM, Santoro F, Latos PA, Huang R, Koerner MV, Sloane MA, Warczok KE, Barlow DP. Silencing and transcriptional properties of the imprinted Airn ncRNA are independent of the endogenous promoter. *EMBO J.* 2008 Dec 3;27(23):3116-28.

Pauler FM*, Sloane MA*, Huang R, Regha K, Koerner MV, Tamir I, Sommer A, Aszodi A, Jenuwein T, Barlow DP. H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res.* 2009 Feb;19(2):221-33.

Latos PA*, Stricker SH*, Steenpass L*, Pauler FM, Huang R, Senergin BH, Regha K, Koerner MV, Warczok KE, Unger C, Barlow DP. An in vitro ES cell imprinting model shows that imprinted expression of the Igf2r gene arises from an allele-specific expression bias. *Development.* 2009 Feb;136(3):437-48.

Koerner MV*, Pauler FM*, Huang R, Barlow DP. The function of non-coding RNAs in genomic imprinting. *Development.* 2009 Jun;136(11):1771-83.

Koerner MV and Barlow DP. Genomic imprinting – an epigenetic gene-regulatory model. *Curr Opin Genet Dev.* 2010 Feb. [Epub ahead of print]

7. ACKNOWLEDGEMENTS

I thank all past and present members of the Barlow lab – without you, working here would not have been the same. I thank Denise Barlow for giving me the opportunity to do my thesis in her lab, for her support, her care and her mentoring. I thank Florian Pauler who is the best co-worker one can imagine, discussed projects, problems and weird ideas with me, who was a source of inspirations during common projects and writing of reviews and made my working life so much fun. I thank Paulina Latos for her friendship, all the great times I spent with her and teaching me how to work with ES cells. Thanks to Stefan Stricker for his constant encouragement and intellectual input. I thank Yvonne Schichl and Stefan Stricker for starting off the repeat deletion project. Furthermore, I want to thank Yvonne Schichl for her faithful friendship and her constant support. I thank Federica Santoro for her help to provide me with free time during my weekends and for always being in a good mood. Thanks to Huang Ru for all her Chinese wisdom. I thank Katarzyna Warczok for teaching me how to work with mice and for help with Southern blotting. Thanks to Simon Hör for his calmness and advices. Thanks to Romana Bittner and Ruth Fuchs for making working so much easier by providing lab stocks and keeping the lab in order. I thank Angelika Eisner and Xenia Weissmann for organising all the bureaucratic things. Thanks to Quanah Hudson and Mathew Sloane for bringing the end of the world closer to me. I also want to thank Stephan Hütter, Philipp Günzl, Tomasz Kulinski, Irena Vlatkovic, Lucas Rudigier, Basak Senegin, Regha Kakkad, Laura Steenpass, Christine Seidl and Iveta Yotova for a great lab membership. Furthermore, I thank Denise Barlow, Florian Pauler and Stefan Stricker for careful reading of this thesis.

I thank Renée Schröder and Anton Wutz for their time and their support as my committee members.

I thank Karin Aumayr, Gabriele Stengl and Pawel Pasierbek for help with the 3D FISH and the S phase fractionation. I thank Marisa Bartolomei and Karl Pfeifer for providing me with pMEFs with deletions in the *Igf2* cluster. I thank Anton Wutz,

Arabella Meixner, Erwin Wagner and Maria Sibilina for providing me with advice and materials for ES cell culture.

Ein großes Dankeschön geht auch an meine Familie, vor allem an meine Eltern Veronika und Harald. Ich bin euch unendlich dankbar für eure Unterstützung. Ihr habt es mir ermöglicht, meine Zukunft so zu gestalten, wie ich sie mir vorstelle.