UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 29

Long-range Control of Gene Expression by Imprinting Control Regions During Development and Neoplasia

NOOPUR THAKUR

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2005

ISSN 1651-6214
ISBN 91-554-6194-8
urn:nbn:se:uu:diva-5725

Dissertation presented at Uppsala University to be publicly examined in Lindahlsalen, Evolutionary Biology Centre, Uppsala, Monday, April 25, 2005 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Thakur, N. 2005. Long-range Control of Gene Expression by Imprinting Control Regions During Development and Neoplasia. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 29. 53 pp. Uppsala. ISBN 91-554-6194-8.

Genomic imprinting is an epigenetic phenomenon by which a subset of genes is expressed in a parent of origin specific manner. Most of the imprinted genes are located in clusters. Genetic evidences suggest that genes in imprinted clusters are regulated by Imprinting Control Regions (ICRs). To elucidate the mechanisms by which the imprinting is maintained in clusters, we have chosen a well characterized cluster at the distal end of mouse chromosome 7. This cluster contains 15 imprinted genes and they have been shown to be regulated by *H19* and *Kcnq1* ICRs.

The mouse *H19* ICR, which is shown to have a chromatin insulator function, is implicated in the regulation of *H19* and *Igf2* genes by interacting with the CTCF protein. It has been documented that CTCF is also involved in the maintenance of differential methylation at the ICR. In this investigation we demonstrated that CTCF maintained differential methylation is lost when we subjected the ICR containing episomal plasmids to de novo methylation machinery of the human choriocarcinoma cell line, JEG3, suggesting that the *H19* ICR looses its methylation privilege property under neoplastic conditions.

The *Kcnq1* ICR has been implicated in the regulation of 11 imprinted genes. The *Kcnq1* ICR is methylated on the active maternal allele but unmethylated on the inactive paternal allele and overlaps an oppositely oriented and paternally expressed gene known as *Kcnq1ot1*. In this investigation, we documented that the *Kcnq1* ICR controls the imprinting of neighboring genes by behaving as a bidirectional silencer and that this function is regulated by antisense RNA *Kcnq1ot1*. Furthermore, we have documented that duration of antisense transcription plays a critical role in the antisense RNA- mediated silencing.

In conclusion, this thesis provides more insights into the complex mechanistic aspects by which ICRs, control imprinting of genes in clusters during development and neoplasia.

Keywords: Genomic imprinting, Imprinting control region, Antisense RNA, de novo methylation

Noopur Thakur, Department of Physiology and Developmental Biology, Animal Development and Genetics, Norbyv 18 A, Uppsala University, SE-75236 Uppsala, Sweden

© Noopur Thakur 2005

ISSN 1651-6214
ISBN 91-554-6194-8
urn:nbn:se:uu:diva-5725 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-5725)

List of papers

I Kanduri C, Kanduri M, Liu L, **Thakur N**, Pfeifer S, Ohlsson R. The kinetics of deregulation of expression by de novo methylation of the *H19* imprinting control region in cancer cells. Cancer Res. 16, 4545-8. (2002).

II **Thakur N**, Kanduri M, Holmgren C, Mukhopadhyay R, Kanduri C. Bidirectional silencing and DNA methylation-sensitive methylation-spreading properties of the *Kcnq1* imprinting control region map to the same regions. J Biol Chem.11, 9514-9. (2003).

III **Thakur N**, Tiwari VK, Thomassin H, Pandey RR, Kanduri M, Gondor A, Grange T, Ohlsson R, Kanduri C. An antisense RNA regulates the bidirectional silencing property of the *Kcnq1* imprinting control region. Mol Cell Biol.18, 7855-62. (2004).

IV **Thakur N,** Kanduri C. Duration of antisense transcription: A critical component in antisense transcription-mediated bidirectional silencing. (Manuscript).

Papers I, II and III have been reproduced by the kind permission of the publishers.

Contents

Abbreviations

Inr	Initiator
CTCF	CCCTC binding factor
CpG	Cytosine paired with guanine dinucleotides
Dnmts	DNA methyltransferases
MeCPs	Methyl CpG binding proteins
MBD	Methyl binding domain
PGC	Primordial germ cell
HDAC	Histone deacetylase
HAT	Histone acetyl transferase
HMT	Histone methyl transferase
HP1	Heterochromatic protein 1
PWS	Prader-Willi syndrome
Igf2	Insulin like growth factor 2
BWS	Beckwith-Wiedemann syndrome
DMR	Differentially methylated region.
ICR	Imprinting control region
ChIP	Chromatin immunoprecipitation
Mb	Mega bases
DNase	Deoxyribonuclease
ORF	Open reading frame
Sir	Silent information regulatory protein
RITS	RNA induced transcriptional silencing
siRNAs	Small interfering RNAs
EICO	Expression-based Imprint Candidate Organizer.
Kcnq1ot1	*Kcnq1* overlapping transcript 1
Lit1	Long QT intronic transcript 1
ESTs	Expression sequence tags
XCI	X chromosome inactivation
Xa	Active X chromosome
Xi	Inactive X chromosome
Xist	X chromosome inactivation-specific transcript
Igf2r	Insulin-like growth factor 2 receptor
LMP1	Latent membrane protein1
RAR	Retinoic acid receptor
PML	Promyelocytic leukemia
PLZR	Promyelocytic leukemia zinc finger

Introduction

The presumed universality of the central dogma and the flow of genetic information was encapsulated by Jacques Monod's famous statement that ''***what was true for E. coli would also be true for the elephant***'' and ever since has dominated our conception of the nature of genetic information and the structure of genetic systems. Although Monod did suggest that RNA itself may have (other) functions, the prevailing orthodoxy has been that proteins not only constitute system, in both simple and complex organisms, but also constitute most of the regulatory control system, in both simple and complex organisms. The central dogma has therefore not only been taken to mean that most genes encode proteins, but also that proteins are sufficient in themselves to specify and organize the autopoietic programming of complex biological entities, an assumption that has pervaded molecular biology for decades. This assumption must now be reassessed, with the recent evidences implicating a greater role for RNA in complex genetic programming.

In recent years several lines of evidence document that sense and antisense RNAs control the gene activity by regulating the chromatin structure. The notable example for sense RNA-mediated gene regulation is *Xist* (X-chromosome inactivation specific transcript) controlled X-inactivation process. Although transcriptome analysis revealed more than 2500 antisense transcripts, surprisingly only a few of them have been functionally implicated. However, the mechanisms by which antisense transcription regulate gene expression is remain obscure. Hence, the present investigation by exploiting a relatively well characterized gene cluster at the distal end of the mouse chromosome 7 and its orthologous region on the human chromosome 11p15.5, we explored the mechanisms by which antisense expression regulate the parent of origin-specific expression of genes in this cluster. In addition, we investigated whether the molecular mechanisms that maintain the parent of origin-specific expression of genes in this cluster during normal development are maintained in neoplastic condition as well.

Functional organization of chromatin

Each eukaryotic chromosome contains one very long molecule of double stranded DNA, which is organized and compacted by proteins through a hierarchy of folded states, ultimately reaching a about 10,000-fold linear compaction of the DNA, prior to cell division. Much is known about the lowest level of chromosome organization, the nucleosome. Two molecules each of four different highly conserved proteins (known as histones), pack together to form a compact spool-like structure, which has the remarkable property of causing a short stretch of DNA (147 bp) to wrap tightly on its outer surface in $1^3/_4$ superhelical turns. The complex of the histone octamer (comprising of two molecules each of histones H2A, H2B, H3 and H4) with its wrapped DNA is known as a nucleosome core particle. This motif is repeated at approximately-conserved but irregular intervals, hundreds or thousands, or millions, of times along the length of the chromosomal DNA to form chromatin. Adjacent nucleosomes are connected by linker DNA, which is associated with another conserved protein histone H1 (*"Beads on a String")***.** This simple 'beads-on-a-string' arrangement is further folded into more condensed, ~30 nm thick fibers (Fig. 1).

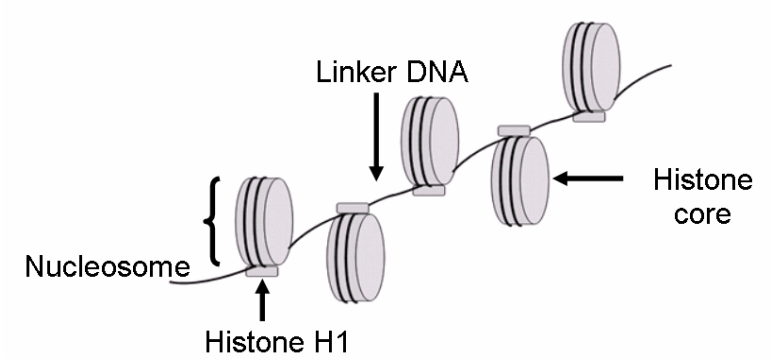

Figure 1. Organisation of nucleosomes in chromatin.

Such 30 nm fibers are then further condensed *in vivo* to form 100–400 nm thick interphase fibers or the more highly compacted metaphase chromosome structures (Grewal et al., 2003). This organization of DNA into chromatin fibers hinders its accessibility to proteins which 'read' and/or copy the nucleotide base sequence, and consequently such structures must be dynamic and capable of regulated unfolding–folding transitions. Hence, there is a tight link between gene expression and chromatin organization.

The appreciation of the link between chromatin compaction and gene expression dates back to cytological studies of the early 20th century. Using basic dyes to stain chromatin and visualize it under the microscope, Emil

Heitz in 1928 noted that chromatin of eukaryotic cells can be broadly distinguished into two forms: heterochromatin and euchromatin (Passarge et al., 1979).

Heterochromatin was defined as condensed regions in the nucleus that do not decondense during interphase whereas euchromatin was noted to readily decondense upon exit from mitosis. It was postulated that heterochromatin is the functionally inactive regions of the genome and euchromatin is the region where actual gene activity occurs. As we advance into 21st century, these chromatin domains are much better defined at the molecular level: heterochromatin regions are more "closed" in chromatin conformation; they contain few actively expressed genes, and replicate late in S-phase. In contrast, euchromatin is more "open" and accessible to nucleases, is rich in actively transcribed genes, and replicates early during S-phase (Cheung et al., 2005) (Fig. 2).

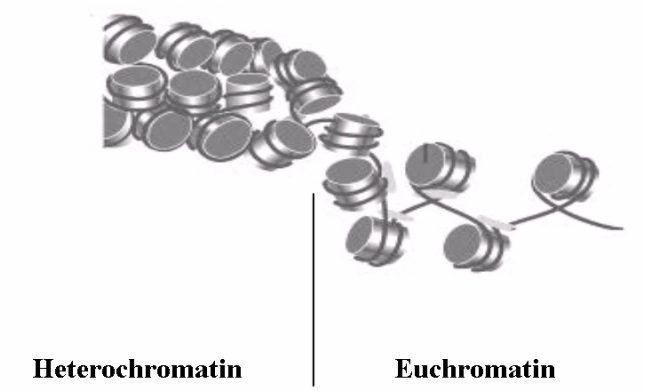

Figure 2. Functional and structural states of chromatin

These discrete functional domains are established by the concerted action of regulatory elements such as promoters, enhancers and silencers and individual domains are separated from one another by elements such as barriers and insulators.

Genetic regulation of gene expression

Decoding of genetic information stored in the DNA through a multistep process is called gene expression. According to the central dogma, DNA sequence of a gene is first transcribed into RNA, which is then translated into protein. In order to decode genetic information, a gene must have regulatory sequences associated with it. These are stretches of DNA which do not themselves code for protein but act as binding sites for RNA polymerase and

its accessory molecules as well as a variety of transcription factors. Every cell in mammals has the same DNA sequence; however, the genetic information decoded from the DNA sequence is different between tissues, suggesting that there is a regulation of gene activity in a tissue-specific manner. The selective expression of genes is one of the most fundamental control mechanisms governing a wide range of biological processes. It is now apparent that this selective expression is not only carried out at the genetic level but also at the epigenetic level.

The gene regulation at the genetic level is brought about by interplay between *cis*-regulatory sequences such as promoters, enhancers, insulators, silencers and *trans*-acting factors that bound to these elements.

Promoters

Eukaryotic promoters, like prokaryotic promoters, contain conserved sequences that are important for initiation of gene transcription. Eukaryotes, because of their added complexity, tend to have more conserved sequences in their promoters than do prokaryotes. Eukaryotic promoters reside immediately upstream, 30-100 bp, of the transcription initiation site of a gene, meaning that they act in *cis* in relation to the open reading frame (ORF). One important sequence in most eukaryotic promoters is found around –30 bp from the transcription start site, and has the sequence TATAAA (or something close to it). This promoter element is known as the **TATA Box**, which is analogous to the –10 bp element in prokaryotes. Another important element that is implicated in transcription initiation is Inr (initiator) element with a consensus sequence YYCAYYYYY (Smale et al., 1989). In addition, there are promoters with none of the previously described elements. These promoters contain a well-characterized conserved sequence GGCCAATCT at –80 bp, otherwise known as the **CCAAT Box.** The aforesaid conserved sequences in the core promoter region mainly help in organizing the basal transcription initiation machinery. Generally, this basal transcription machinery yields a low level of transcription from the core promoters in a tissue-nonspecific manner. However, a high level as well as tissue-specific transcription from the core promoter elements is controlled by another well-characterized *cis*-acting element, known as enhancers.

Enhancers

Enhancers are DNA sequences that after binding of activator proteins can activate transcription over distances as large as 80 kb (Jack et al., 1991). This activation is independent of enhancer orientation on DNA and their position upstream or downstream of the promoter. It has been proposed that in certain subset of genes, enhancer-promoter communications impart spatio-

temporal expression patterns during development and differentiation (Benerji et al., 1983).

Several models have been proposed to explain enhancer action over a distance (Fig. 3). All these models assume that activators work through a protein–protein interaction with some component of the transcription machinery (Bondarenko et al., 2003). The activator can either interact with a stable target (such as a protein already bound at the target promoter) or the target has to be created (enhancer-induced recruitment of a protein to the promoter).

According to the DNA looping model (Fig. 3A), the interaction of an enhancer-bound activator with a protein partner of the basal transcription machinery at the promoter is accompanied by looping of the intervening DNA. To find the promoter, the enhancer-bound activator can "scan" surrounding DNA regions by "hopping" and /or "scanning". A recent study on HNF-4α gene supports the loop-scanning model. By using chromatin immunopurification analysis over the HNF-4α regulatory region, the study has shown that the DNA-protein complex formed on the HNF-4α enhancer scans through the neighboring sequences until it forms a stable enhancer-promoter complex (Hatzis et al., 2002). The looping model is further supported by the observation that certain enhancers can activate transcription in trans when promoters and enhancers are located on separate, non-covalently linked plasmids. Alternatively, enhancer-bound activator can interact with the promoter bound protein without formation of intermediate complexes.

The tracking model (Fig. 3B) suggests that the enhancer bound complex actively tracks the entire length of DNA in search of the promoter. The evidence for enhancer action by tracking mechanisms in eukaryotes is not experimentally proven. The spreading–looping model (Fig. 3C) suggests that binding of an activator to the enhancer induces cooperative binding–polymerization of a protein on the DNA (it could be an enhancer-binding protein or another activator). A series of relatively small DNA loops are formed during polymerization of an activator, and eventually the array of proteins reaches the target promoter. For example CHIP protein in Drosophila cannot bind to the DNA directly but can interact with numerous transcription factors and facilitates its action over a distance in vivo. It was proposed that CHIP is recruited by an activator protein bound at an enhancer; it works as a protein "bridge" between the activator bound at the enhancer and other proteins having multiple weak binding sites between the enhancer and promoter (Fig. 3C). As a result, a wave of small protein-stabilized chromatin loops is initiated at the enhancer and moves towards the promoter. The experimental evidence for this model is very weak.

Whatever the particular activation mechanism is, the activation signal has to be delivered from the enhancer to the promoter. The final result of enhancer-promoter communication is transcription activation through overcoming of a rate-limiting step that can be different on different promoters.

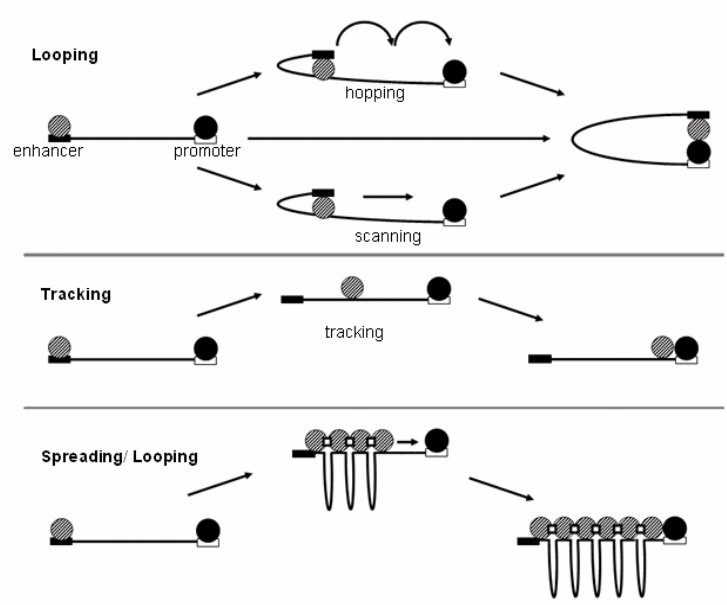

Figure 3. Various Models proposed for Enhancer- Promoter Interaction.

Insulators

Enhancers, apart from selectively activating the correct promoter over long distances, can also act promiscuously to activate transcription from the non specific promoters. As recent human genome sequence data suggest that most of the genes are organized in clusters and hence, there is a need to restrict enhancer actions without interfering with its actual function on its native locus. In the recent past several sequences have been characterized with a specific ability to interfere with communications between enhancer and promoter in a position-dependent manner without interfering with the activity of the enhancer. Such sequences are known as insulators (Burgess-Beusse et al., 2002). Elements with latter property have been found in Drosophila and vertebrates. In flies, the most studied insulator is *gypsy*, which when placed between an enhancer and promoter, specifically blocked the enhancer communication with the promoter (Gerasimova et al., 1996) (Fig. 4). Elements with enhancer blocking properties have also been characterized in vertebrates. The activity of the vertebrate insulators is primarily controlled by 11-zinc finger protein, CTCF. One of the vertebrate insulators, for example 5'-HS4 chicken β-globin insulator, apart from influencing the enhancer promoter communications has also been implicated in the barrier activity (Fig. 4). The barrier activity of the insulators confers position independent expression to transgenes stably integrated in the genome probably by pre-

venting the spreading of flanking heterochromatic regions (West et al., 2002). Recently, it has been documented that the enhancer blocking and barrier activities of the 5'-HS4 chicken β-globin insulator are carried out by distinct proteins (West et al., 2004). As a result of enhancer blocking and chromatin boundary activities, a genomic region flanked by insulators behaves as a functionally independent unit of gene regulation (a chromatin domain), isolated from both negative and positive affects of surrounding genomic regions.

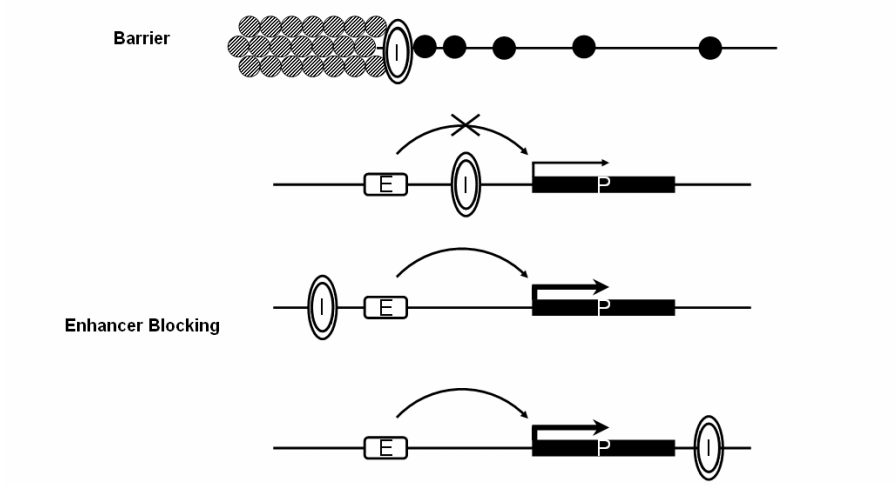

Figure 4. Barrier and Enhancer blocking functions of Insulators.

Silencers

Silencers are *cis*-acting DNA elements which establish the silenced state over the flanking promoters in an orientation and position-independent fashion. The mechanism of action of the silencers has not been investigated in detail. Based on recent investigations it has been suggested that silencer elements act by recruiting repressor complexes associated with histone deacetylases and methylases, thereby modifying the flanking chromatin into inactive state. This modified chromatin then forms a beacon for the recruitment of heterochromatic structural proteins which leads to the bi-directional spreading of the silenced chromatin (Fig. 5). For example, in *Saccharomyces cerevisiae* , the mating type genes, HML and HMR, are regulated by *cis*-acting elements, termed E and I silencers. These silencers work via recruitment of silent information regulatory proteins, Sir1-4. The recruitment of Sir Proteins by the E and I silencers result in the deacetylation of flanking regions thus creating a local heterochromatic region (Dhillon et al., 2002; Andrulis et al., 2003). The example for silencer elements in mammals in-

clude, regulation of human embryonic (epsilon) globin gene in both erythroid and non-erythroid cells by the silencer elements epsilonNRA-I and epsilonNRA-II, located 3 Kb downstream of the transcription start site. The complex and cooperative interactions of these negative regulatory elements have been suggested to play a critical role in the development-specific regulation of globin gene expression (Li et al., 1998).

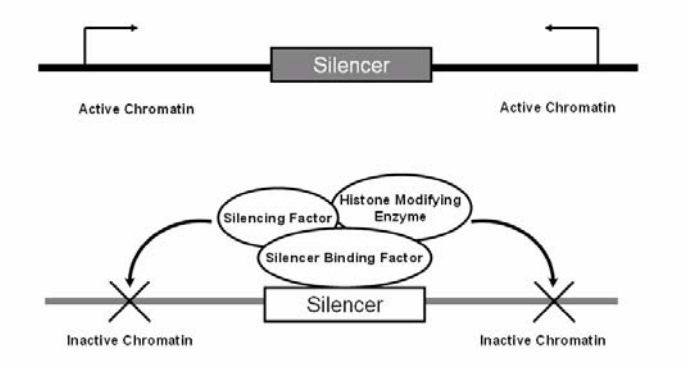

Figure 5. The mechanisms of silencer activity.

Epigenetic regulation of gene expression

As I described previously that every cell type in the higher organisms has similar DNA sequence but the genetic information decoded from the DNA sequence varies between the cell types. An important question that needs to be answered is, "how can different cell types with same DNA sequence achieve different genetic information"? The cell type-dependent decoding of information from the DNA sequence is possible only when the DNA sequence is layered with a mechanism that selectively filters the genetic information from the DNA sequence in a cell type–dependent manner. Thus epigenetic information plays a critical role in the tissue-dependent gene expression. "Epigenetics" can be defined as the study of heritable changes in gene expression that occur without a change in the DNA sequence. This far several epigenetic modifications have been characterized and among which DNA methylation, histone acetylation and methylation have been extensively investigated for their role in mammalian development and differentiation.

DNA Methylation

In eukaryotes, the most abundant covalent modification of DNA is methylation of cytosine residues at carbon 5 of the pyrimidine ring. This modification occurs primarily in the context of a simple sequence, 5´-CG-3´. The most striking property of DNA methylation, unlike other characterized epigenetic modifications this far, is its inheritance through cell divisions and its programmed alteration and fixation through development. Approximately 70% of the CpGs are methylated in the mammalian genome and CpGs are distributed between CpG islands and non-CpG islands. CpG islands are stretches of about 1.0 kilo base DNA with high GC content and CpG/GpC ratio of around 0.6 (Caiafa et al., 2005). There are about 45,000 CpG islands, most of which lie in the promoter regions of the genes. CpG islands are generally unmethylated except in the inactive X-chromosome and imprinted genes, where CpG islands are completely methylated (Antequera et al., 1993).

DNA Methylation Machinery

The cellular methylation patterns are established by a complex interplay of at least three DNA methyltransferases, DNA methyltransferase1 (DNMT1), DNMT3A and DNMT3B (Bestor, 2000) (Fig. 6). DNMT1 is the most abundant methyltransferase in mammalian cells which localizes to replication foci. DNMT1 is referred to as maintenance methyltransferase because it is believed to be responsible for copying the parental strand-specific methylation patterns to daughter strand after DNA replication. DNMT3A and 3B are thought to play a role in creating new methylation patterns independent of replication during early embryonic development. By genetic experiments it has been shown that all the three DNA methyltransferases are important for proper embryonic development. (Robertson 2002). In addition, two more proteins with a homology to DNA methyltransferases have been identified, DNMT2 and DNMT3L. Although DNMT2 has all the conserved methyltransferase motifs, no biological activity has been documented.

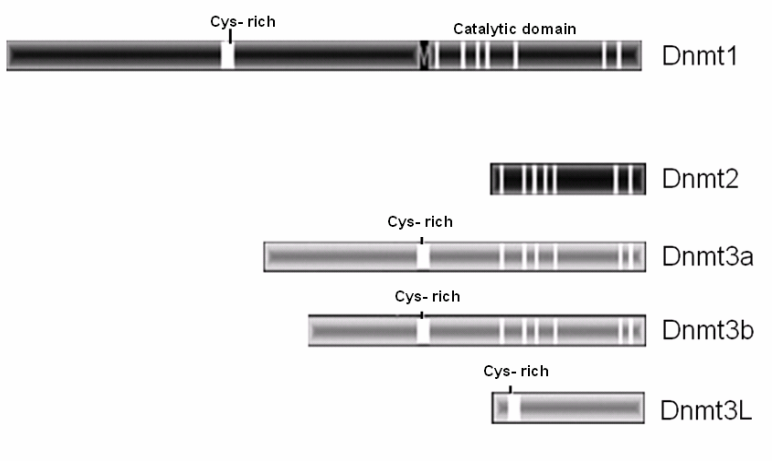

Figure 6. The family of DNA methyltransferases.

Moreover, the targeted deletion of DNMT2 in mouse has no phenotypic effect, suggesting that it does not play a critical role in methylation programming during mammalian development (Okano et al., 1998). *Dnmt3L* shares homology with DNA methyltransferases, *Dnmt3a* and *Dnmt3b*, but lacks enzymatic activity. Recently, it has been documented that Dnmt3L by cooperating with Dnmt3a is involved in the establishment of the parent of origin specific methylation marks.

DNA methylation and gene expression

DNA methylation has been implicated as one of the important mechanisms that regulate tissue-specific gene expression (Razin et al., 1980). Accumulating evidence over the last two decades suggests that the CpGs in the promoters of various genes are under-methylated in the tissue of expression and completely methylated in the non-expressing tissue. Two different mechanisms have been envisaged as to how DNA methylation controls gene expression: (I) DNA methylation in the pivotal *cis*-elements directly interferes with the binding of transcription factors, thereby preventing the transcription initiation (II) Alternatively, DNA methylation in the promoter regions attract methyl-specific CpG binding proteins such as MeCPs and MBDs associated with histone deacetylases and methyltransferases, thus creating heterochromatic-like region locally to inhibit gene expression (Scarano et al., 2005). Although direct effects of DNA methylation on transcription factor binding (such as E2F, CREB, cMYC, AP2 and NfkB) has been proven in some cases (Iguchi et al., 1989; Kovesdi et al., 1987; Prendergast et al., 1991), but several recent investigation document that methylation-dependent gene repression by methyl-specific CpG binding proteins is the most frequently occur-

ring mechanism. Several methyl-specific CpG binding proteins have been identified, which include MeCP2, MBD1, MBD2, MBD3 and MBD4 (Bird et al., 1999). All these proteins share a common methyl CpG binding domain. In addition, except MBD4, all of them are known to interact with HDACs, suggesting that methyl CpG binding proteins upon binding to methyl group attracts histone deacetylases thus creating a locally inactive chromatin, which then could be a target for heterochromatic machinery (Prokhortchouk et al., 2002). MBD4 has recently been characterized as a thymine DNA glycosylase that interacts with the mismatch repair protein MLH1. *In vivo*, MBD4 functions to reduce the mutability of methyl-CpG sites in the genome (Wu et al., 2003).

DNA methylation and Development

The importance of DNA methylation in genetic reprogramming during mammalian development and differentiation was elegantly substantiated by the finding that homozygous targeted deletion of the DNA methyltransferase1 gene in mouse is embryonic lethal, the embryo died before midgestation. In addition, DNMT3 targeted deletions in mouse also supported this view (Li et al., 1992). Methylation marks undergo genome-wide reprogramming during gametogenesis and during preimplantation development. During the development of primordial germ cells (PGC), methylation marks are erased by an unknown demethylation process (Lee et al., 2002). Studies in mice have shown that this erasure occurs immediately after the primordial germ cells enter into the gonads. The establishment of sex-specific patterns of methylation follows genome-wide demethylation during gametogenesis. In addition, these gamete-specific methylation marks undergo further reprogramming during the early embryonic stages. They are erased during early preimplantation development, methylation marks over the imprinted genes escape this genome wide reprogramming (Fig. 7.. The genome wide methylation patterns are again reestablished during the postimplantation development.

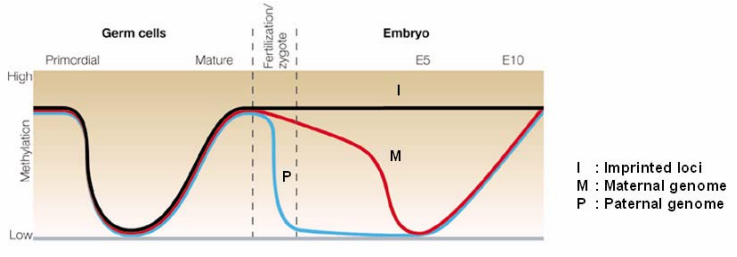

Figure 7. Genome wide reprogramming of methylation marks

It is suggested that reprogramming of methylation marks during gametogenesis helps in resetting the imprinting marks, whereas methylation reprogramming during early embryonic development, would help in ensuring proper embryonic development, probably by regulating the spatiotemporal expression of early embryonic-specific genes (Fig. 7).

Histone modifications

Each of the core histones has a related globular domain that mediates histone–histone interactions within the octamer, and that organizes the two wraps of nucleosomal DNA. Each histone also harbors an amino-terminal tail domain of 20–35 residues that is rich in basic amino acids and extends from the surface of the nucleosome. These histone 'tails' do not contribute significantly to the structure of individual nucleosomes nor to their stability, but they do play an essential role in controlling the folding of nucleosomal arrays into higher-order structures. Histones are subject to an enormous number of post-translational modifications, including acetylation and methylation of lysines(K) and arginines(R), phosphorylation of serines(S) and threonines(T), ubiquitylation and sumoylation of lysines, as well as ribosylation (Peterson et al., 2004).

Adding to the complexity is the fact that each lysine residue can accept one, two or even three methyl groups, and an arginine can be either mono- or di-methylated. The majority of these post-translational marks occur on the amino-terminal tail of histones. Given the number of new modification sites that are identified each year, it seems likely that nearly every histone residue that is accessible may be a target for post-translational modification.

Why should histones be the target for so much enzymatic activity? Given that chromatin is the physiological template for all DNA-mediated processes, histone modifications are likely to control the structure and/or function of the chromatin fiber, with different modifications yielding distinct functional consequences. Indeed, recent studies have shown that site-specific combinations of histone modifications correlate well with particular biological functions. For instance, the combination of H4 K8 acetylation, H3 K14 acetylation, and H3 S10 phosphorylation is often associated with transcription. Conversely, tri-methylation of H3 K9 and the lack of H3 and H4 acetylation correlate with transcriptional repression in higher eukaryotes (Lachner et al., 2002). The three best understood epigenetic modifications that mark heterochromatin are histone hypoacetylation, histone H3 lysine9 methylation and DNA methylation. Histone H3 and H4 acetylation and Histone H3 lysine 4 methylation mark the euchromatic regions. Histone H3 lysine 4 methylation inhibits histone H3 lysine 9 methylation and vice versa, thus determining the active or inactive state.

The crosstalk between these epigenetic modifications can lead to the formation of heterochromatin (Ben-Porath et al., 2001) (Fig. 8). Several self-reinforcing loops and feedback mechanisms drive the formation, maintenance and spreading of heterochromatin (Richards et al., 2002).

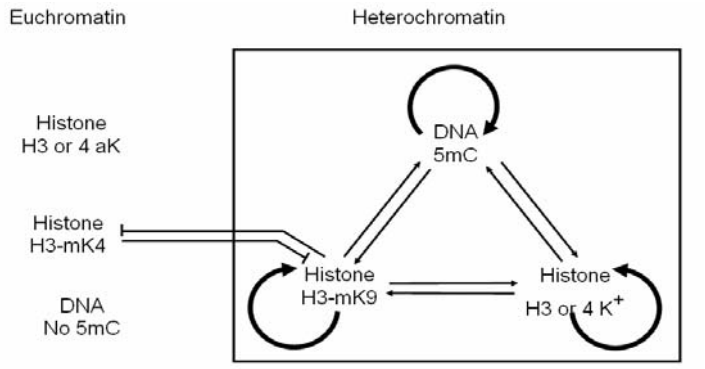

Figure 8. Cooperative interactions between various epigenetic modifications.

Although histone modifications have been studied for over 30 years, the identification of the histone modifying enzymes themselves remained elusive until the first nuclear histone acetyltransferase (HAT), a *Tetrahymena* homolog of yeast Gcn5, was identified in 1996. *In vivo* studies in yeast had previously characterized Gcn5 as a transcriptional co-activator protein, and thus its identification as a HAT solidified the view that histone modifications directly regulate transcription (Brownell et al., 1996). Subsequently, a variety of other transcriptional co-activators, such as CBP/p300 were found to have intrinsic HAT activity (Chan et al., 2001), and many co-repressors, such as Rpd3, were found to have histone deacetylase (HDAC) activity (Rundlett et al., 1996). Histone modifying enzymes are now organized into large HAT, HDAC, histone methyltransferase (HMT) and histone kinase families.

A quite different strategy is used by small noncoding RNAs to target histone H3 K9 methylation to chromatin surrounding mammalian and fission yeast centromeres. These centromeric regions are characterized by repetitive DNA sequences that are transcribed at low levels by opposing promoters. The double-stranded RNAs (dsRNA) formed due to opposing transcription trigger RNA interference pathway, which in turn processes the dsRNA into small, 21–23 nucleotide RNAs. Recent studies have shown that an intact RNAi pathway is essential for targeting H3 K9 methylation to centromeric chromatin, and that these small RNAs actually associate with RNA induced transcriptional silencing complex (RITS). The resulting novel ribonucleoprotein complex ultimately targets the Clr4 histone methyl transferase to cen-

tromeric repeats, via either RNA–RNA (nascent centromeric transcripts) or RNA–DNA homologous pairing (Hall et al., 2002). Subsequent histone methylation leads to the recruitment of proteins such as heterochromatin protein 1 (HP1), which directs formation of highly condensed, heterochromatin structures required for centromere function (Dillon, 2004).

There is now a wealth of examples where specific histone modifications control the binding of nonhistone proteins to the chromatin fiber. These nonhistone proteins then elicit the function that is associated with a particular histone mark. A hallmark of many proteins that bind to histone tails is the presence of small histone binding modules. For example, some chromodomains bind to methylated lysines, whereas bromodomains specify binding to acetylated lysines. Furthermore, these modules often bind to only a particular modified histone residue. For example, the chromodomain within HP1 interacts specifically with a methylated K9 of histone H3, whereas the chromodomain of the Polycomb protein binds to a methylated K27 of histone H3. In contrast, the binding of bromodomains to different acetylated lysines does not show much specificity. For instance, acetylation of K8 within histone H4 can promote the recruitment of the ATP-dependent chromatin remodeling enzyme, human SWI/SNF – via a bromodomain within the Brg1 subunit – but a similar bromodomain within the Swi2 subunit of the yeast SWI/SNF complex interacts with a broader range of acetylated H3 and H4 tails (Craig, 2005).

In addition to canonical nucleosomes, *in vivo* chromatin arrays also contain histones variants. For instance, nucleosomes assembled at yeast and mammalian centromeres contain a histone H3 variant, Cse4/CENP-A, which is essential for centromere function or assembly. Another histone H3 variant, H3.3, replaces canonical histone H3 during transcription, generating a mark of the transcription event (McKittrick et al., 2004). Several variants of histone H2A have also been identified. The macro-H2A variant is restricted to metazoans and functions in X chromosome inactivation (Ladurner et al., 2003).

Histone variants, distinct patterns of posttranslational modifications of histones, and histone tail binding proteins all contribute to establishment of various ‘open’ or ‘closed’ chromatin domains that have specialized folding properties and biological functions. Some of these domains can be propagated through DNA replication and mitosis, guaranteeing the inheritance of chromatin states to progeny (Henikoff et al., 2002). Histone H3 lysine9 methylation may play a central role in the stability of these chromatin states, as to date no enzymes are known that catalyze H3 lysine9 demethylation. Only one histone demethylase has been identified to date and is specific for mono and di methylated H3 lysine 4 (Shi et al., 2004). Furthermore, several nonhistone proteins, such as HP1 or the PRC1 Polycomb complex, apart from binding to methylated histone lysines, they are also implicated in recruiting the histone methylases, thus providing a means for templating new histone

methylation events; for example, spreading the inactive chromatin to adjacent nucleosomes.

How 'open' chromatin states are propagated through cell divisions is not clear, especially as histone lysine acetylation or serine phosphorylation can be rapidly reversed by HDACs or histone phosphatases. Future studies will no doubt continue to identify the functional and biochemical properties of new chromatin domains as well as to elucidate the principles that govern their maintenance and propagation.

Cooperative interactions between different epigenetic modifiers

There are now several examples suggesting a functional interaction between H3-K9 methylation and DNA methylation. In the filamentous fungus *Neurospora crassa,* the H3-K9 methyltransferase dim-5 is necessary for normal genomic DNA methylation. In addition, when lysine 9 of H3 was replaced with non-methylatable amino acids, DNA methylation was lost, suggesting a dependence of DNA methylation on H3-K9 methylation (Tamaru et al., 2001).

This view has been further expanded by experiments in Arabidopsis. It was shown that the DNA methyltransferase CMT3 interacts with the heterochromatin protein 1 (HP1) that binds to K9-methylated H3 (Jackson et al., 2002). Intriguingly, loss of DNA methylation resulted in a reduction of H3-K9 methylation in Arabidopsis (Soppe et al., 2002). These results strongly suggest a mutual relationship between DNA methylation and H3-K9 methylation. Cooperative interactions between epigenetic modifications have also been observed in organisms with mutations in the RNAi pathway. In the fission yeast *Schizosaccharomyces pombe*, deletion of various RNAi-associated genes resulted in aberrant modification and derepression of centromeric sequences (Volpe et al., 2002).

Because epigenetic deregulation was accompanied by an accumulation of complementary transcripts from the centromeres, this result provided a strong indication that small interfering RNAs (siRNAs) are involved in regulating epigenetic modifications. A more recent study in Arabidopsis also revealed a connection between RNAi and DNA methylation. In this study, it has been demonstrated that mutation in the ARGONAUTE4 gene affects both DNA methylation and histone H3-K9 methylation patterns (Zilberman et al., 2003). The mechanistic basis of the interaction between siRNAs and other epigenetic factors has yet to be determined. It is possible that siRNAs are involved in targeting DNA and histone methyltransferases to epigenetically regulated loci. In particular, interactions between DNA methylation and histone modifications can be established, maintained and reinforced by defined adaptor molecules, like methyl-DNA binding proteins and siRNAs (Fig. 9). In this scenario, feedback mechanisms become possible at various levels. For example, siRNAs might target DNA methyltransferases

(DNMTs) and histone modifying enzymes (HMEs) to epigenetic control elements (Schramke et al., 2004). DNA methylation recruits methyl-DNA binding proteins that in turn recruit histone deacetylases and histone methyltransferases (Ben-Porath et al., 2001). Many types of interactions are conceivable that would reinforce the stability of epigenetic chromatin (Fig. 9). In addition, cooperative interactions would also provide a mechanism for the spreading of epigenetic states over large chromosomal domains (Fig. 9).

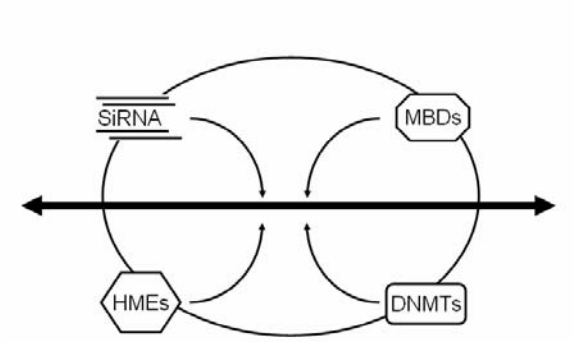

Figure 9. Cooperative interactions between epigenetic modifiers.

Genomic imprinting

While most mammalian autosomal genes are expressed from both parental chromosomes, a few are expressed from one of the two alleles in a parent of origin-specific manner. Such genes are said to be imprinted. When an imprinted gene passes through the germline, it acquires a ´mark´ that identifies the parental origin of that allele. Some imprinted genes are expressed from the maternally inherited chromosomes and others from the paternally inherited chromosomes, suggesting that the maternal and paternal genomes are not functionally equivalent and that both maternal and paternal genomes are required for normal mammalian development. Genomic imprinting has also been described in plants where the process is believed to have evolved independently from that in mammals. Genomic imprinting was identified for the first time in the early 1980s through classical manipulation experiments on mouse embryos. Pronuclear transplantation was used to generate 'Gynogenetic' or 'Androgenetic' conceptuses, which, respectively, have two sets of maternal chromosomes and no paternal contribution or vice versa (Hoffman et al., 2000; Surani et al., 1984.). These embryos fail to develop properly and die before term despite being diploid. Furthermore, the defects presented in androgenones and gynogenones are strikingly different. Gynogenetic embryos die before or at mid-gestation and are growth retarded with poor de-

velopment of the extra-embryonic tissues. In contrast, androgenetic conceptuses have a more restricted developmental potential. Their embryonic components develop poorly while extra-embryonic tissues are better formed with an overgrown trophoblast.

These embryo reconstitution experiments suggested that maternal and paternal contributions to the developing mammalian embryo are different. It was proposed that a specific 'imprinting' of the paternal and maternal genomes occurs during the development of the egg and the sperm, resulting in the requirement of both genomes after fertilization for normal full-term development. An apparent functional opposition between paternally expressed growth promoting genes and maternally expressed growth suppressing genes provided evidence for the hypothesis that imprinting evolved from a conflict for resources between the maternal and paternally inherited genomes in the offspring. This theory suggests that fathers are driven to extract maximal resources for their offspring from the mother, while the maternal resources are best protected for future pregnancies, if she is able to limit this demand to some extent (Wilkins et al., 2003). The conflict hypothesis fits for most, though not all, imprinted genes. To date more than 70 imprinted genes have been identified. The database, EICO (Expression-based Imprint Candidate Organizer), found 2101 candidate imprinted genes derived by comparing mRNA expression profiles between parthenogenotes and androgenotes using RIKEN cDNA microarrays. Of the 2101 candidate imprinted genes, 1403 showed maternal expression and 698 showed paternal expression (Nikaido et al., 2004).

Mechanisms that regulate genomic imprinting

Many of imprinted genes identified this far are organized in clusters, which can be as large as 2 Mb, as in the case of the PWS/AS domain (Kantor et al., 2004). The reasons for the cluster formation are not known but it has been suggested that clustering allows sharing of common *cis*-regulatory elements. This suggestion is supported by the fact that differentially methylated imprinting control regions (ICRs or DMRs) associated with imprinted genes play a critical role in the regulation of parent of origin-specific expression patterns of genes in the cluster, as targeted deletion of these regions result in large-scale disturbances in the expression patterns in the cluster (Fitzpatrick et al., 2002; Thorvaldsen et al., 1998). The better understanding of role of these imprinting control regions in the maintenance of imprinting in clusters was possible due to efforts of several laboratories in the recent past.

How does an imprinting control region (ICR) regulate parent of origin-specific expression of neighboring genes? Before going into the details of how imprinting control regions regulate gene expression in the clusters, I would like to discuss some of the characteristic features that are common among the imprinting control regions. One of the most distinguishing fea-

tures of the imprinting control regions is that they are differentially methylated on the parental alleles. These differential methylation marks at the ICRs are established during early gametogenesis. Although not much is known about how differential methylation is established in the early gametogenesis and maintained thereupon in the development, recent studies implicated a few transcription factors (such as CTCF and NF-Y) at least in the maintenance of differential methylation during development (Pant et al., 2003; Pandey et al., 2004). Establishment of differential methylation during gametogenesis could be due to interplay between the availability of transcription factors and/or DNA methyltransferases. It has been shown that DNMT3L is essential for the establishment of maternal methylation imprints and appropriate expression of maternally imprinted genes, as defect in DNMT3L resulted in the disruption of maternal methylation imprints in homozygous oocytes (Hata et al., 2002). DNMT3L interacts with DNMT3A and DNMT3B and co-localizes with these enzymes in the nuclei of transfected cells, suggesting that DNMT3L may regulate maternal-specific imprints via the DNMT3 family of enzymes (Suetake et al., 2004).

Another interesting feature of the ICRs is that they are often embedded with CpG islands and highly repeated motifs (Reik et al., 2001). The role of repeat motifs in the ICR function remains unclear, but they could be involved in attracting methyltransferases to ICRs. In addition to DNA methylation, ICRs also harbor other epigenetic modifications such as histone acetylation and methylation (Delaval et al., 2004). Although DNA methyltransferase knockouts revealed relaxation in the imprinting of genes (Caspary et al., 1998), suggesting a causal role for DNA methylation in the ICR function, it is quite possible that other epigenetic modification such as histone acetylation and methylation could be involved in setting up of differential chromatin structures at the ICRs initially during early gametogenesis which then could form a target for DNA methylation.

The critical role of ICRs in the maintenance of imprinting primarily came from the studies on relatively well-characterized imprinted cluster at the distal end of the mouse chromosome 7 and its orthologous region on human 11p15.5 (Paulsen et al., 2000). This cluster spans about 1.0 Mb in mouse and harbors many imprinted genes like *Osbp15*, *Phld2a*, *Slc22a1l*, *Kcnq1*, *Tssc4*, *Cd81*, *Ascl2, Ins2, Igf2* and *H19*. This cluster is divided into two subdomains (Caspary et al., 1998; Fitzpatrick et al., 2002; Thorvaldsen et al., 1998). The domain 1 consists of only three genes *H19*, insulin-like growth factor 2 (*Igf2*) and *Ins2*, where as the domain2 consists of 11 imprinted genes. *Igf2*, which encodes an embryonic mitogen, and *H19* which produces a noncoding RNA of unknown function are located ~80 kb apart (Tilghman et al., 1995). These two genes are reciprocally imprinted; that is *H19* is expressed from the maternal allele whereas *Igf2* is expressed from the paternal allele (Tilghman et al., 1995). Coordinated expression of these genes in tissues of mesoderm and definitive endoderm origin is due to the sharing of a

set of enhancers located downstream of *H19* (Leighton et al., 1995). The monoallelic expression of *H19* and *Igf2* genes is regulated by a differentially methylated imprinting control region located –2 to –4 kb upstream of the *H19* promoter. This region is methylated on the paternal chromosome but unmethylated on the maternal chromosome (Thorvaldsen et al., 2002; Tremblay et al., 1995). The maternal inheritance of a deletion encompassing ICR resulted in the activation of silent maternal *Igf2* allele, suggesting that this region play a critical role in the imprinting of *H19* and *Igf2* genes (Thorvaldsen et al., 2002). Chromatin analysis of the *H19* ICR revealed four DNaseI hypersensitive sites and surprisingly all four hypersensitive sites mapped to 21 bp repeats. Functional dissection of these 21 bp revealed binding sites for a 11-zinc finger protein, CTCF (Hark et al., 1998; Kanduri, Holmgren et al., 2000; Kanduri, Pant et al., 2000; Szabo et al., 2000) (Fig. 10). *In vitro* transfection experiments using an episomal-based system carrying wild type and modified ICR carrying mutations at the CTCF target sites revealed that the *H19* ICR behaves as a chromatin insulator and that the insulator activity is regulated by CTCF (Kanduri, Pant et al., 2000). More importantly, maternal inheritance of the *H19* ICR carrying mutations at the CTCF target sites leads to reactivation of silent maternal *Igf2* gene, suggesting that the *H19* ICR on the maternal chromosome mediates silencing of *Igf2* in *cis* through chromatin insulator function. In addition, these studies have implicated CTCF in the maintenance of differential methylation at the ICR, as the loss of occupancy of CTCF at the H19ICR on the maternal chromosome resulted in methylation of the ICR (Pant et al., 2003; Schoenherr et al., 2003). However, whether CTCF is involved in the establishment of differential methylation during gametogenesis, remains unclear.

In addition to the *H19* ICR, several other elements have also been implicated in the *H19* and *Igf2* imprinting. For example, *Igf2* DMR1 located upstream of *Igf2* promoter 1, and DMR2 located within exon 6 of *Igf2*, behaves as a methylation-sensitive silencer and a methylation-dependent activator, respectively (Constancia et al., 2000; Murrell et al., 2001). All these DMRs are methylated on the paternal chromosome. Interestingly, a recent study documented that the unmethylated maternal H19 ICR is required to protect DMR1 and DMR2 from methylation, and similarly, DMR1 protects DMR2 from methylation. Since this coordination is not due to linear spreading of methylation, it may be caused by long-range chromatin interactions. The differentially methylated regions (DMRs) containing insulators, silencers and activators, were shown to have physical contacts between them and these interactions were changeable depending on their epigenetic state, presumably enabling *Igf2* to move between an active and a silent chromatin domain (Kato et al., 2005).

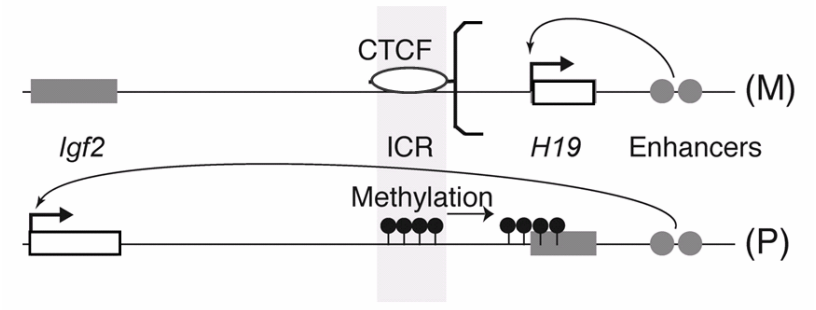

Figure 10. The mechanism of imprinting at the *H19* and *Igf2* locus.

Deletion of the *H19* ICR affects imprinting of *Igf2* and *Ins2*, while the imprinting of genes in domain2 remains unaffected suggesting that an additional imprinting control element in domain2 may be involved in controlling the imprinting of genes in this domain. Based on chromosomal break points and methylation changes in BWS patients, a differentially methylated CpG island identified in the intron 10 of the mouse *Kcnq1* gene has been proposed to be involved in the regulation of imprinting (Smilinich et al., 1999). This region is methylated on the active maternal allele, but unmethylated on the inactive paternal allele of *Kcnq1* and overlaps with an oppositely oriented and paternally expressed gene known as *Kcnq1ot1* or *Lit1* (Mitsuya et al., 1999). (Fig. 11). Targeted deletion of this region in the human paternal chromosome 11 propagated in the chicken DT40 cell line resulted in the activation of the normally silent paternal alleles of *KCNQ1* and *CDNK1C*. The suggestions that this region, termed *Kcnq1* imprinting control region (ICR), has a pivotal role in the maintenance of imprinting of neighboring genes has recently been confirmed by targeted deletion experiments in the mouse (Fitzpatrick et al., 2002). Bi-directional activation of genes in the *Kcnq1* ICR deletion suggests that the ICR executes different mechanisms to that of *H19* ICR. In that scenario, it is important to know about the mechanisms that are employed by the *Kcnq1* ICR in order to control the imprinting of neighboring genes in the domain2. Since actively transcribed antisense transcript is associated with paternally unmethylated *Kcnq1* ICR, it is possible that bi-directional influence of the *Kcnq1* ICR could be just due to an antisense RNA production. If antisense RNA has a role, then how does it execute its influence not only at the overlapping but also at the non-overlapping side. Therefore the main objective of this thesis is to understand the molecular details underlying the *Kcnq1* ICR function and to check whether these molecular details involve antisense RNA.

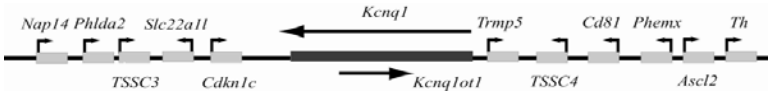

Figure 11. Physical map of the imprinted *Kcnq1* domain.

It is clear from the aforesaid details that ICRs associated with the imprinted clusters play a critical role in the maintenance of imprinting and that they employ diverse mechanisms to achieve this.

Antisense transcription and gene expression

During much of the age of molecular biology, attention was focused primarily on the protein-coding genetic units in the genome. The perspective began to change with the discovery of non-coding RNAs, as they have been implicated in the roles of catalysis and gene regulation. The functional significance of many of these transcripts is currently not known, but the class of non-coding "antisense" transcripts has quickly arisen as an important regulator of epigenetic information (Kelley et al., 2000). The term "antisense" implies a gene that overlaps in sequence with a second genetic unit but is transcribed in the opposite orientation. Before the sequence of the human genome became available, a few human naturally occurring antisense transcripts (NATs) had been identified by groups studying specific genomic loci (Giovanni et al., 2004). However, the first evidence that antisense transcription was a common feature of eukaryotic genomes came from the analysis of reverse complementarity between all available human mRNA sequences (Lavorgna et al., 2004). This study identified 87 genomic loci encoding natural antisense transcripts, and predicted that more than 800 noncoding antisense transcripts would exist in the human genome (Lehner et al., 2002). Subsequent studies both confirmed and extended these observations using databases of mRNAs and expressed sequence tags (ESTs). In particular, Yelin *et al.* have identified 2667 human NATs of which more than 1600 are predicted to be true NATs. More recently, analysis of many fully sequenced mouse cDNAs has predicted the existence of as many as 2500 mammalian *cis*-NATs (Yelin et al., 2003). As some antisense transcripts have been shown to regulate gene expression, it is possible that antisense transcription might be a common mechanism of regulating gene expression in eukaryotic cells.

In a diploid organism, this type of genetic arrangement presents an opportunity for *cis*-regulation and direct interaction between the products of "sense" and "antisense" transcripts. Generally antisense RNA transcription leads to the silencing of the overlapping gene, as reported in case of an indi-

vidual with an inherited form of anemia (alpha thalassemia) who has a deletion that results in a truncated, widely expressed LUC7L transcription unit becoming juxtaposed to a structurally normal alpha globin gene HBA2 in an antisense orientation. In these patients, the silencing of HBA2 gene and methylation of flanking CpG island is observed and interestingly, this feature is linked to antisense transcription (Tufarelli et al., 2003).This finding identified a new mechanism underlying human genetic disease. Antisense genes are also known to occur at the X inactivation center and at many, if not all, imprinted domains (Table 1) (Rougeulle et al., 2002; Sleutels et al., 2002), where silencing effects of antisense RNA occurs over large regions spanning more than 100Kb. In particular, antisense transcription from the imprinted domains not only silences overlapping regions but also the non-overlapping regions by an unknown mechanism.

Table 1. *Sense and Antisense RNAs in X-chromosome inactivation and imprinting.*

Region	Sense		Antisense		
Xic	*Xist*	(P)	*Tsix*	(M)	
Igf2r locus	*Igf2r*	(M)	*Air*	(P)	
PWS/AS region	*Ube3a*	(M)	*Snurf1/Snprn/Ipw/ SnoRNAs/Ube3a-as* (multicistronic)	(P)	
PWS/AS region	*Zfp127*	(P)	*Zfp127-as*	(P)	
BWS region	*Kcnq1*	(M)	*Kcnq1-as*	(P)	
Gnas locus	*Gnas*	(M)	*Gnas-as*		(P)
H19/Igf2 locus	*Igf2*	(P)	*Igf2-as*	(P)	
Peg1 locus	*Copg2*	(M)	*Copg2-as*	(P)	

M: Maternal allele, P: Paternal allele

Antisense RNA and X-chromosome inactivation

X-chromosome inactivation (XCI) is an epigenetically regulated process in female mammals by which, one of the two X chromosomes is silenced in early embryogenesis to compensate for the differences in X-linked gene dosage between males (XY) and females (XX) (Lyon,1961). In mice, XCI is found in two forms – "random" and "imprinted". In random XCI, XX cells can select either the maternal or the paternal X for inactivation. Random XCI is observed in all placental mammals (eutherians) examined so far, including humans and mice (Boumil et al., 2001). In contrast, the paternally inherited X chromosome is exclusively inactivated in imprinted XCI (Takagi et al., 1975). This is observed today in extant marsupials (metatherians). Interestingly, in some placental mammals such as the mouse, both forms can be

observed with random XCI taking place in epiblast lineages (embryo proper) and imprinted XCI occurring in the extraembryonic tissues (e.g., yolk sac, trophectoderm).

Both imprinted and random XCI are controlled by a *cis*-acting master switch, known as the "X inactivation center" (Xic) (Brown et al., 1991), which is required and sufficient to direct X chromosome counting, choice, and the initiation of silencing. Oddly, all known regulatory components of the Xic involve non-coding transcripts – *Xist* (Brown et al., 1991a) and *Tsix* (Lee et al., 1999). The induction of XCI is controlled by the "ying and yang" relationship between *Xist* and *Tsix*. The *Xist* gene product is a 17-kb *cis*-acting RNA that spreads along the X chromosome destined to become inactive chromosome and "paints" that X chromosome exclusively. This painting action of *Xist* RNA initiates the silencing step presumptively by recruiting a complex of silencing proteins, as it propagates along the chromosome (Cohen et al., 2002). Prior to the onset of XCI, *Xist* RNA is actually made from both X's but is not stable at this time. At the onset of XCI, *Xist* becomes repressed on the future active X (Xa) and becomes markedly upregulated presumably stabilized on the future inactive X (Xi). This switch in *Xist* expression is regulated by its antisense partner, *Tsix*, a gene initiating 12 kb downstream of *Xist* and transcribed across the entire *Xist* locus on the opposite DNA strand. *Tsix* is expressed on all X's prior to the onset of XCI. At the onset of XCI, *Tsix* repression on the future Xi is required for the upregulation of *Xist* RNA and *Tsix* continued expression is necessary to maintain the active state of the future Xa (Fig 12). As with many antisense genes in autosomally imprinted domains, *Tsix* also initiates within a putative imprinting control element (ICE) that controls imprinted XCI (Sado et al., 2001). Curiously, this element also appears to control stochastic choice in the random form of XCI.

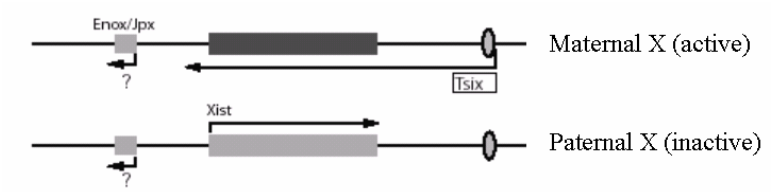

Figure 12. Antisense RNA in X-inactivation.

Antisense RNA and Genomic imprinting

A considerable proportion of imprinted genes are associated with antisense transcripts (at present 15%). Surprisingly, all antisense transcripts discovered so far in the imprinted clusters are imprinted and in most of the cases they are expressed from paternally inherited chromosome. Among all the an-

tisense transcripts that have been characterized in various imprinted clusters, only one antisense transcript, *Air* from *Igf2r* locus has been relatively well investigated. The promoter of *Air* transcript maps to differentially methylated imprinting control region (ICR) in the intron 2 of the *Igf2r* gene. *Air* is actively expressed from the unmethylated paternal ICR whereas it is silenced on the maternal methylated ICR. The *Air* transcript overlaps with entire *Igf2r* transcription unit as well as its promoter in opposite orientation (Rougeulle et al., 2002). (Fig 13A) It has recently been documented that the *Igf2r* ICR encompassing *Air* promoter controls the parent of origin-specific expression of neighboring genes and that this property is regulated by *Air* transcript (Sleutels et al., 2002). However, the mechanism by which *Air* transcript controls the imprinting of neighboring genes remains unclear. As I mentioned earlier, the promoter of another antisense transcript (*Kcnq1ot1*) has recently been mapped to intron 10 of the *Kcnq1* gene in a relatively well-investigated imprinted cluster at the distal mouse chromosome 7 and its orthologous region in the human chromosome 11p15.5 (Fig. 13B). The present investigation is mainly focused on whether or not antisense RNA has any link to the imprinting mechanism in this region.

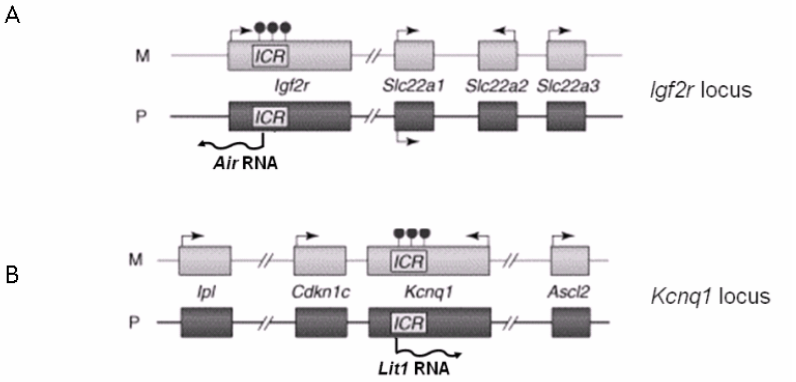

Figure 13. Antisense RNA in imprinting.

Epigenetics and cancer

In recent years, several lines of evidence have accumulated to suggest that epigenetic dysregulation is one of the main causes of cancer development and progression. This dysregulation includes hypomethylation leading to oncogene activation and chromosomal instability, hypermethylation leading to tumor suppressor gene silencing (Fig. 14). Loss of DNA methylation at CpG dinucleotides was the first epigenetic abnormality to be identified in cancer cells (Feinberg et al., 2004).

Genome-wide hypomethylation of DNA has several mechanistic implications. Recently, it has been found that some of the CpG islands are normally methylated in somatic tissues. These methylated islands can become hypomethylated in cancer thus activating the nearby genes. Examples of genes that are affected by hypomethylation include oncogenes such as *HRAS* and the 'CT' genes — those that are expressed normally in the testis and aberrantly in tumours (Feinberg et al., 1983).

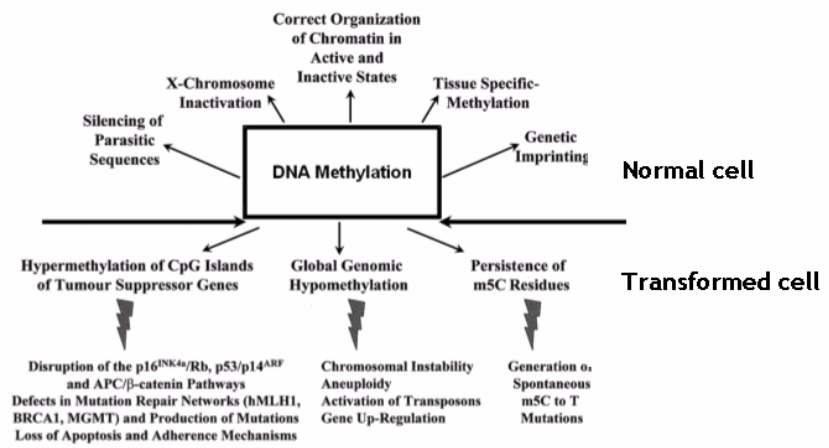

Figure 14. The role of DNA methylation in cancer

Several studies have proposed that hypomethylation in cancer is linked to chromosomal instability. Consequences of hypomethylation are particularly severe in pericentromeric satellite sequences as unbalanced chromosomal translocations with breakpoints in the pericentromeric DNA of chromosomes 1 and 16 has been detected in several cancers (Wilm´s tumor, ovarian and breast cancers). Potential connection between hypomethylation and chromosomal instability is further suggested by the fact that hypomethylation of L1 retrotransposons in colorectal cancers, which might promote chromosomal rearrangement (Suter et al., 2004). In addition to gene amplification, hypomethylation of the multi-drug resistance gene *MDR1* correlates with increased expression and drug resistance in acute myelogenous leukemia (Nakayama et al., 1998).

More intriguingly, despite decrease in genome-wide methylation level, it is frequently observed that there are focal increases in the methylation level, particularly at the CpG islands that are associated with promoters of tumor suppressor and cancer related genes. Hypermethylation of a tumor-suppressor gene as a potential epigenetic inactivating mechanism was first described for the *RB* gene in sporadic cases of human retinoblastoma. Aberrant DNA hypermethylation of *RB* gene affects numerous pathways, includ-

ing cell cycle, DNA repair, and hormonal responses, and therefore alters the perfect epigenetic equilibrium of the cell. Although the mechanism(s) by which this hypermethylation occurs over certain selective promoters remain largely unknown but the recent studies in this direction provide insights into this novel phenomenon. For example, in nasopharyngeal carcinoma, Latent membrane protein 1 (LMP1) of Epstein-Barr virus is implicated in the repression of the *E-cadherin* promoter thus enhancing the invasive capacity of cancer cells. It has been documented that LMP1 represses the *E-cadherin* promoter by increasing the levels of expression of DNMT1, DNMT3A and DNMT3B, which in turn hypermethylated the *E-cadherin* promoter (Tsai et al., 2002). Other examples for the selective promoter methylation are the two translocation events involving the retinoic acid receptor-α, PML-RARα and PLZR-RARα. The leukemia-promoting PML-RARα (promyelocytic leukemia) fusion protein induces hypermethylation and silencing of RARα target genes via recruitment of HDACs and DNMTase activity. The promyelocytic leukemia zinc finger (PLZR) protein also associates with transcriptional repressors such as HDACs and PcG proteins and the PLZR-RARα fusion protein likewise represses genes normally activated by RAR (Lund et al., 2004).

The fidelity of the epigenetic states at differentially methylated imprinting control regions (ICRs) is critical in maintaining the parent of origin-specific expression of genes in clusters. The epigenetic lesions at ICRs in cancer cause widespread deregulation in gene expression patterns in clusters causing loss of imprinting (LOI). The parent of origin-specific expression of most of the genes in the cluster at the human chromosome 11p15.5 is disrupted in several types of cancers. Several lines of evidence suggest that the loss of expression of various genes in this cluster is mostly due to epigenetic lesions at ICRs. The epigenetic lesions at ICRs could be as a result of gain of methylation or loss of methylation on the parental alleles. For example, the *H19* ICR is biallelically methylated in the Wilm´s tumor patients, allowing the activation of normally silent paternal *IGF2* allele (Cui et al., 2001). However, in several colorectal cancers the *H19* ICR is hypomethylated, suggesting that maintenance of differentially methylation at the ICRs is quite crucial for the normal development. Loss of imprinting of *IGF2, KCNQ1OT1* and *CDKN1C* is also associated with Beckwith Wiedemann Syndrome (BWS) patients. These BWS patients are at an increased risk of developing certain type of cancers (Khatib et al., 2004; Fukuzawa et al., 2003; Bliek et al., 2004). The most common epigenetic alteration associated with BWS (50% of cases) is the loss of methylation at KvDMR (Cerrato et al., 2002), associated with loss of imprinting of *KCNQ1OT1* and *Cdkn1c* (Diaz-Meyer et al., 2003). Expression of the normally silent maternal allele of *IGF2* occurs in 25–50% of BWS cases and for most of these cases the cause, whether genetic or epigenetic, is not known. A few of them are associated with hypermethylation of the *H19* promoter leading to biallelic ex-

pression of *IGF2*, which is referred to as *H19*-dependent loss of imprinting of *IGF2*. However, most cases of loss of imprinting of the *IGF2* gene are associated with normal monoallelic maternal expression of the *H19* gene. This is referred to as H19-independent loss of imprinting of *IGF2* (Weksberg et al., 2003). It is unclear how hypo and hyper methylation patterns are established at the ICRs in cancers, but I presume that it could be due to dysregulation of functions at the ICR or alternatively dysregulation of DNA methylation machinery. In this investigation, by employing JEG-3 cell line with a high de novo methylation capacity, we provide insights as to how *H19* ICR could be hypermethylated in cancer cells.

Given that all the epigenetic modifications currently known, affect gene expression and that deregulation of certain genes involved in cell proliferation and differentiation result in uncontrolled cell growth, it is not surprising that all types of epigenetic modifications mentioned above have been linked to cancer. For example, promoter CpG methylation-mediated tumor suppressor gene silencing has been observed in many types of cancers. Likewise, a number of histone acetyltransferases have been found to be rearranged or mutated in cancer. In addition, the Aurora/Ipl kinases, which phosphorylate histone H3, have been found to be over-expressed in cancer. Finally, inhibitors of histone deacetylase have entered clinical trials for certain leukemia. Like that of histone acetylation, histone methylation has attracted great attention recently due to its function in the control of gene expression. Interestingly, many of the identified histone methyltransferases are in one way or another linked to cancer. The connection between histone methylation and cancer is well exemplified in prostate cancers that express higher levels of EZH2 showed a poorer prognosis, suggesting that EZH2 levels can be a potential biomarker for the prediction of the risk factor of individual prostate cancer patients. Up-regulation of EZH2 is not limited to prostate cancer. In fact, recent studies indicate that the EZH2 levels also directly correlate with breast cancer aggressiveness. Similar to the finding in metastatic prostate cancer, EZH2-mediated breast cell invasion requires an intact SET domain. EZH2 encodes a histone methyltransferase that methylates lysine 27 of histone H3, a modification linked to transcriptional repression (Robertson, 2002).

In summary, given that epigenetic modifications play a crucial role in controlling the important developmental decisions by regulating the chromatin structure and transcription, it is no surprise that dysregulation of mechanisms that maintain the fidelity of epigenetic states can lead to cancer development and progression.

Aims

Paper I

Wilm´s tumors with LOI show aberrant methylation of the maternal *H19* ICR, specifically involving CTCF binding sites, which in all likelihood abrogates CTCF binding and hence, allows *IGF2* activation. However, the mechanisms by which hypermethylation at the *H19* ICR occurs are remain unknown. Since point mutations of CTCF target sites within the ICR lead to loss of methylation protection during mouse development, mechanisms that interfere with CTCF–DNA interactions are likely to be culprits in maintaining the differential methylation levels at the ICR. We envisage two alternative scenarios by which aberrant methylation at the ICR could occur. (I) Mutations in CTCF zinc fingers that lead to loss of its ability to interact with the ICR are likely to lead to aberrant methylation. (II) Over-expression of methyltransferases may by brute force overcome the methylation protection of the maternal *H19* ICR allele. Since CTCF mutations are very rare in Wilm´s tumors, we favor the latter scenario as plausible explanation for the aberrant methylation at the ICR. Previously, we have shown that the JEG-3 (choriocarcinoma) cell line has high *de novo* methylation property on the transfected episomes. In this paper, we were interested to address the epigenetic stability of the *H19* ICR under the neoplastic situation by subjecting the *H19* ICR under the challenge of *de novo* methylation machinery of the JEG-3 cell line. We also attempted to understand the kinetics of loss of methylation protection property of the human *H19* ICR and its effect on the expression of *H19* reporter gene.

Paper II

The *Kcnq1* ICR (3.6Kb) has been implicated in the long range control of imprinted gene expression at the distal end of the mouse chromosome 7, as targeted deletion of this region on the paternal chromosome results in the activation of neighboring imprinted genes spread over several hundreds of kilo bases. Since the *Kcnq1* ICR controls the parent of origin-specific expression of the flanking genes spread over several hundreds of kilo bases, in this investigation we were interested to know the mechanisms by which this action is achieved. Additionally, we attempted to fine-map the regions responsible for various novel functions of the ICR.

Paper III

We have documented in paper II that the *Kcnq1* ICR harbors bi-directional silencing and that this property maps to region1. Earlier studies have shown that the paternally unmethylated *Kcnq1* ICR contains a promoter for actively transcribed antisense transcript, *Kcnq1ot1*. We were therefore interested in addressing whether the bi-directional silencing activity of the ICR involves antisense RNA. Following are the main aims of this paper; (I) to fine-mapping of the antisense promoter using genomic footprinting, (II) to investigate if transcription from an antisense promoter forms a molecular basis for the bi-directional silencing property of the *Kcnq1* ICR? (III) to investigate the potential role of dsRNA, formed as a result of opposing transcription between antisense promoter and sense promoter, mediated RNA interference in the bi-directional silencing property of the *Kcnq1* ICR.

Paper IV

In the papers II and III, we have documented that the *Kcnq1* ICR harbors bi-directional silencing property and that this feature is controlled by an antisense transcription. In this paper, we wished to address the mechanisms by which antisense transcription executes bi-directional silencing of overlapping and non-overlapping genes.

Results and discussion

Paper I

Rapid methylation of H19 minigene subsequently leads to its repression

By exploiting an episomal model system and the *de novo* methylation capacity of the JEG3 cell line, we have shown that *H19* minigene rapidly becomes methylated and silenced, mimicking the inactivation of the maternal *H19* allele in a range of cancers. After 15 days of selection with hygromycin from post-transfection, the cells were harvested after every four days to extract DNA and RNA. To perform methylation analysis, DNA was restricted by methylation sensitive enzymes *HhaI* and *HpaII*, and subjected to southern blot hybridization using the probe specific for *H19* coding region as well as promoter. The results showed that both *H19* coding region as well as promoter becomes progressively methylated and complete methylation can be seen at day 23 of post-transfection. RNase protection assay was performed by using RNA obtained from different passages to assess the expression levels of *H19* reporter gene. No linear correlation was found between the *de novo* methylation of the mouse *H19* promoter and its silencing. This observation can be explained in two possible ways. Firstly, the transcriptional activity of *H19* gene tolerated a certain degree of methylation in the early stages of transfection and secondly, methylation acts as a template for the repressive chromatin conformation and this process is time dependent.

Methylation of the mouse H19 ICR occurred slowly and correlated with the loss of DNase1 nuclease hypersensitive sites

The maternal copy of the *H19* ICR escapes genome-wide *de novo* methylation during early embryonic development and therefore remains unmethylated during somatic cell propagation. Considering the high *de novo* methylation property of the JEG3 cell line, we asked whether the *H19* ICR escapes from being heavily methylated by the *de novo* methylation activity of JEG3 and maintains its methylation privilege or becomes methylated. Methylation analysis was performed by southern blot hybridization using the probe specific for the *H19* ICR. Interestingly, the *H19* ICR was not methylated in the initial passages but gradually loses its methylation protection property and becomes methylated 27 days after transfection. Thus the JEG3 cell line takes more time to inflict its *de novo* methylation activity on the *H19* ICR than on the *H19* minigene. Bisulphite sequencing of *H19* ICR revealed that the CpG

sites flanking the CTCF target sites are *de novo* methylated. In addition, DNaseI hypersensitivity studies indicated that CTCF protein occupancy is gradually lost in relation to *de novo* methylation. Collectively, these observations suggest that CTCF protein occupancy does not offer protection against the *de novo* methylation machinery of the JEG-3 cell line and that this situation in the JEG3 cell line reflects an aberrant methylation of the *H19* ICR in neoplastic condition.

De novo methylation of the human H19 ICR results in the loss of its insulator function

Methylation analysis was performed on the human *H19* ICR and the results showed that the human *H19* ICR became more rapidly methylated than the mouse *H19* ICR suggesting inherent differences in chromatin conformation between the human and mouse ICRs. Given the high affinity human *H19* ICR for *de novo* methylation in JEG-3 cells, we wished to investigate its Insulator function under methylation conditions. The results show that the loss of insulator function under conditions of methylation, suggesting that rapid *de novo* methylation of the ICR would have probably disrupted the CTCF-protein complex-mediated insulator function and that rapid *de novo* methylation of the human *H19* ICR as compared to the mouse *H19* ICR, indicates participation of *cis*-acting elements other than CTCF in methylation protection in mouse. In sum, it can be inferred from this study that aberrant epigenetic activity is one of the main characteristic features of cancer cells and that it might also be the effective contributor to the development of neoplasia.

Paper II

The Kcnq1 ICR is a bidirectional silencer in JEG3 cells

An earlier study from our lab suggests that the *Kcnq1* ICR acts as an insulator in the Hep-3B cell line. Because the imprinting of *Kcnq1* and other neighboring genes is tissue-specifically regulated and that the observed insulator function in the Hep-3B cell line cannot be explained fully the bidirectional action of the ICR in regulating the parent of origin-specific expression patterns of neighboring genes in all tissues, we were interested in knowing whether the *Kcnq1* ICR function depends on the cell type used. To address this issue in detail, we have employed trophectodermally derived cell line, JEG3. We have inserted the 3.6 kb ICR fragment into an episomal-based system, pREP4H19 in both orientations in relation to the mouse *H19* promoter both at the insulating (between enhancer and promoter) and silencing positions (upstream of the SV40 enhancer) to address whether the 3.6 kb ICR fragment harbors insulator activity, as has been documented or silencer activity that influences expression of genes that are located on both sides of

the ICR fragment. We have transfected these plasmids transiently into the JEG-3 cell line. *H19* reporter gene expression was analyzed by RNase protection assay on RNA extracted from the transient transfections. The results from the latter experiment indicated that the 3.6 kb *Kcnq1* ICR fragment could silence the *H19* reporter gene in both orientations although to different magnitudes i.e., in PS4 orientation (where the antisense promoter of the *Kcnq1* ICR faces the *H19* promoter) the activity is reduced by 10 fold whereas in NS11 (the antisense promoter of the *Kcnq1* ICR faces away from the *H19* promoter), the activity is reduced by 7 fold. We have also found silencing of *H19* reporter gene in the silencing position but only in one orientation. These results indicate that the 3.6 kb *Kcnq1* ICR behaves as a silencer in a position independent manner. To asses the bi-directional action of the ICR at the insulating position, we have analyzed the *hygromycin* gene activity by counting the hygromycin resistant colonies after selection with hygromycin. The results from these experiments indicated that the 3.6 kb *Kcnq1* ICR silences the *hygromycin* gene at the insulating position in an orientation-independent manner, suggesting that the *Kcnq1* ICR harbors bi-directional silencing activity and that this bi-directional silencing property mimics the *in vivo* situation, where the paternal deletion of the *Kcnq1* ICR leads to the activation of the neighboring genes located on both sides of ICR.

The Kcnq1 ICR spreads DNA methylation over the H19 reporter gene in an orientation-dependent and methylation-sensitive manner

Since the above experiments suggest that the *Kcnq1* ICR is a bi-directional silencer, we were interested in understanding the molecular mechanisms by which it silences neighboring reporter genes. As the evidence accumulated over the past two decades suggests that DNA methylation plays a critical in gene silencing, we wished to know whether bi-directional silencing action of the ICR involves DNA methylation. For this purpose, we have analyzed the DNA methylation status over the neighboring *H19* promoter. For DNA methylation analysis, DNA collected at various time points from the JEG3 cells transiently transfected with various episomal constructs, were subjected to southern blot hybridization after restriction digestion of the DNA with *HpaII* and *HhaI* restriction enzymes using the probe specific for *H19* gene. The results indicated that the entire *H19* gene along with promoter was methylated in the *Kcnq1* ICR-dependent fashion, suggesting that the *Kcnq1* ICR-mediated bi-directional silencing involves DNA methylation. More importantly, studies on the kinetics of DNA methylation and silencing suggest that the silencing precedes DNA methylation. Surprisingly, methylation occurred over the *H19* promoter by the *Kcnq1* ICR only in PS4 orientation but not in NS11. An orientation-dependent methylation of the *H19* promoter suggests that *de novo* methylation of the *H19* promoter occurs only during efficient silencing conditions. More importantly, we have shown that the property of methylation spreading is unique to the *Kcnq1* ICR, as the *H19* ICR in the

same episomal constructs could not spread methylation over *H19* promoter. Also, these observations suggest that methylation of the *H19* promoter is due to the *Kcnq1* ICR but not due to the *de novo* methylation property of the JEG-3 cell line.

Fine-mapping of the regions in the 3.6 kb Kcnq1 ICR responsible for silencing as well as methylation spreading

To fine-map the regions responsible for silencing and methylation spreading, we generated serial deletion within the 3.6kb *Kcnq1* ICR based on the information obtained from the DNaseI hypersensitive sites. We cloned all these deletion fragments into pREPH19 episomal-based system and carried out bi-directional silencing assay and methylation-spreading analysis, as described above. The results from these experiments indicated that both the bi-directional silencing as well as methylation spreading properties map to two regions that we name them as region 1 (R1) and region 2 (R2). Removal of 300 bp (R1) or 600 bp (R2) resulted in loss of silencing as well as methylation spreading, suggesting that both R1 and R2 are critical for achieving 100% silencing and loss of any one of these regions resulted in loss of silencing. However, it is not clear how these two fine-mapped regions help the *Kcnq1* ICR to execute these two novel functions but we have proposed two alternative explanations that could provide some understanding; (I) an antisense transcript originating from R1 could interact with R2 to form a repressive chromatin structure, which can later spread bi-directionally to silence neighboring genes (II) double stranded RNA (dsRNA) resulted from opposing transcription could trigger RNA interference that in turn could recruit heterochromatic machinery, as has been described in *S.cerevisiae* .

The Kcnq1 ICR is methylated in an orientation specific manner

We next addressed whether or not the *cis*-acting elements in the *Kcnq1* ICR would protect ICR from *de novo* methylation spreading, as the ICRs have been shown to be equipped with *cis*-acting elements that offer protection against *de novo* methylation. Methylation analysis of the *Kcnq1* ICR was performed in both PS4 and NS11 episomes that carry the ICR in two different orientations. Results from these analyses indicated a few surprising observations. We found that the *Kcnq1* ICR remains unmethylated and spreads methylation over the neighboring *H19* promoter in PS4 orientation, whereas the ICR becomes methylated in NS11 orientation and hence methylation spreading was not detected over *H19* promoter. More over, methylation-sensitive methylation spreading of the *Kcnq1* ICR was also detected when the *Kcnq1* ICR was *in vitro* methylated, where *in vitro* methylated *Kcnq1* ICR lacked the methylation spreading property, suggesting that methylation spreading property of the *Kcnq1* ICR is specific to unmethylated ICR but not to the methylated ICR. Bisulphite-sequencing analysis of the CpGs in the R1 region of the *Kcnq1* ICR was performed on the DNA extracted from PS4 and

NS11 episomal plasmids stably propagated for 42 days in the Hep-3B cell line. Bisulphite sequencing analysis further confirmed the observation that the *Kcnq1* ICR is methylated in an orientation-dependent manner. The methylation-sensitive methylation spreading of *the Kcnq1* ICR can very well explain the loss of methylation of the maternal *Kcnq1* ICR in Beckwith-Wiedemann syndrome resulting in the spreading of methylation to the neighboring genes and thus leading to the pathological inactivation of those genes.

Paper III

Identification of crucial cis-acting elements within the Kcnq1ot1 promoter by genomic footprinting

In the paper II, we have documented that the *Kcnq1* ICR harbors bi-directional silencing and that the region1, which encompasses majority of DNaseI hypersensitive sites of the ICR, plays a critical role in the bi-directional silencing property of the *Kcnq1* ICR. Since the paternal unmethylated *Kcnq1* ICR harbors a promoter for an actively transcribed antisense transcript, and an independent study fine-mapped an antisense promoter to region1, we were therefore interested to know whether the bi-directional silencing property of the *Kcnq1* ICR can be related to the production of antisense RNA, *Kcnq1ot1*. To address this issue in detail, first we have to fine-map the promoter of antisense transcript. For that purpose we have chosen genomic footprinting approach to fine map antisense promoter. Genomic footprint approach was undertaken on episomal plasmids stably propagated in cultured cells containing the *Kcnq1* ICR in PS4 orientation. The results from DNase1 genomic footprints indicated the presence of a DNaseI footprint spanning 300-350 bp. Bioinformatic analysis of this region suggested the presence of potential binding sites for the following factors: OCT1, CCAAT, YY1, GATA, CTCF and CREB. The footprints obtained using an episomal-based system, were similar to the DNaseI footprints obtained from the day 14.5 embryonic liver. These results clearly indicated that the *Kcnq1* ICR recapitulated all of the features of *in vivo* chromatin conformation. Most of the *cis*-acting elements characterized by DNase1 genomic footprint bind to transcription factors that support basal transcription and it is most likely that DNase1 footprint in region1 could be a promoter for antisense transcript.

Targeted deletion of the DNaseI footprint in the Kcnq1ot1 promoter results in bi-directional activation of reporter genes

After having identified the transcription factor binding sites that mostly support basal transcription, we next aimed at understanding the functional significance of these transcription factor binding sites in the bidirectional silencing activity of the *Kcnq1* ICR. To this end, we introduced the unique *Age1* restriction sites within the *Kcnq1* ICR using site-directed mutagenesis approach which allowed selective deletion of the DNaseI footprint. The reporter gene assays, as described previously, were carried out on episomal plasmids, containing the wild type and modified ICRs, transfected transiently into the JEG-3 cell line. The results showed that the loss of the whole DNase1 footprinted region results in the loss of the antisense, *Kcnq1ot1*, transcription as well as parallel loss of silencing of *H19* and *hygromycin* reporter genes, suggesting that this region is crucial for both antisense transcription and silencing activity. To explore the potential role of CTCF target site in the antisense transcription and bi-directional silencing, we selectively deleted the footprinted region spanning the CTCF target site. The results indicated that the deletion of this region neither effects the antisense (*Kcnq1ot1*) transcription nor the bidirectional silencing activity. By using this deletion approach we were able to find out the crucial elements responsible for both the antisense transcription and the bi-directional silencing.

Truncation of the Kcnq1ot1 RNA results in loss of bidirectional silencing

By selectively deleting the region encompassing the antisense promoter, we have shown that antisense transcription plays an important role in silencing. However, this does not rule out the possibility that some of the *cis*-acting elements in the deleted region could independently silence the flanking reporter genes without interfering with antisense transcription. To address this possibility and to check the direct role of the antisense RNA in the bidirectional silencing property of the *Kcnq1* ICR, we truncated antisense RNA (*Kcnq1ot1*) by inserting SV40 polyadenylation sequence at 0.15, 0.75 and 1.7 Kb downstream of the antisense transcription start site and tested their effects on antisense transcription and silencing using approaches, as described previously. Insertion of polyadenylation sequence did not interfere with production of antisense RNA, as the levels of antisense RNA remain similar between the wild type and PS4 episome inserted with polyadenylation sequences at 0.75 kb and 1.7 kb. However, the bi-directional silencing property was lost in these mutants, suggesting that the antisense promoter activity *per se* is not important but the production of antisense RNA is important. Absence of silencing in the mutant that carries a polyadenylation sequence at 1.7 kb downstream of the antisense transcription start site (which has complete antisense transcription unit of 3.6 kb ICR), suggests that antisense transcription beyond the polyadenylation sequence is important. Based on these observations, we have proposed that an increase in the

duration of antisense transcription would simply increases the presence antisense RNA at the site of transcription, which would in turn help in antisense RNA-mediated recruitment of the heterochromatic complexes.

dsRNA is not involved in the bidirectional silencing of the Kcnq1 ICR

As the *Kcnq1ot1* and *H19* promoters face each other in PS4 episomal construct, there is a potential possibility for the formation of dsRNA due to convergent transcription. The formation of dsRNA could trigger RNA interference silencing pathway. To understand the functional role of the dsRNA in the silencing, first we looked for the *Kncq1ot1* RNA over *H19* reporter in various constructs having the *Kcnq1 ICR* at both the insulating and silencing positions by RNase protection assay as well as RT-PCR. We found that at the insulator position, *Kcnq1ot1* expression was detected over the *H19* gene (PS4) but at the silencing position (PC3), no antisense RNA was detected over *H19* reporter gene even though it leads to silencing suggesting that dsRNA has no role to play in the bi-directional silencing of the *Kcnq1* ICR. To further confirm this observation, we deleted *H19* promoter and most of the coding region of *H19* in the construct having *Kcnq1* ICR at the insulator position (PS4). Deletion of the *H19* gene has no effect on the expression of *hygromycin* gene, as it was still repressed, confirming that the bi-directional silencing property of the *Kcnq1* ICR does not involve production of dsRNA.

Absence of methylation spreading in the truncated Kcnq1ot1 RNA and Kcnq1ot1 promoter mutant

In paper II, We have documented that the *Kcnq1 ICR* has a methylation sensitive methylation spreading property and that this methylation occurs as a consequence of silencing rather than acting as a cause for silencing. To check whether the methylation spreading property of the *Kcnq1* ICR is *Kncq1ot1* dependent or independent, we carried out methylation analysis over the *H19* reporter gene in the episomal constructs carrying selective antisense promoter deletions or polyadenylation sequence (PS4 polyA 0.75 and PS4 polyA 1.7). We observed loss of CpG methylation over the *H19* gene when the length of the *Kcnq1ot1* was reduced by insertion of polyadenylation sequence and also when the *Kcnq1ot1* promoter was selectively deleted. However, the deletion of the CTCF site did not affect the methylation spreading property of the *Kcnq1* ICR. These results explicitly document that the methylation-spreading property of the *Kcnq1* ICR is tightly linked to antisense RNA production from the ICR and that the antisense RNA-mediated bi-directional silencing could be due to the recruitment of heterochromatin machinery associated with HDACs and HMTs, thus modifying flanking chromatin into an inactive state, which later forms a target for DNA methylation.

Paper IV

Duration of antisense transcription is a crucial regulator of bi-directional silencing

In the paper III, we have proposed that degree of silencing probably depends on the length of antisense RNA; this would allow its association for longer time with the site of transcription. In turn, this helps to increase the recruitment of heterochromatic machinery. To address the functional significance of duration of antisense transcription in the silencing process, we truncated antisense RNA by inserting the SV40 polyadenylation sequence at 5.5 (PS4 polyA 5.5) and 9.2 (PS4 polyA 9.2) kb downstream of the antisense transcription start site. These modified episomal constructs were transfected into the JEG3 cell line for eight days and silencing assay were carried out, as described previously. To check whether the insertion of polyA truncated antisense transcript, we initially carried out RT-PCRs by using primers that map to prior and after the polyadenylation sequence insertion. The results from the RT-PCR analysis suggested that the polyadenylation sequence indeed truncated antisense transcript, as no antisense transcript was detected after the polyadenylation sequence insertion. We next analyzed the activity of *H19* and *hygromycin* genes in these polyA constructs. We observed pronounced silencing of both reporter genes in PS4 polyA 9.2, as compared to PS4 polyA 5.5, indicating that an increase in the length of antisense RNA increases the degree of silencing. One of the questions that rise due to the latter observation is that how an increase in the length of antisense RNA increases the degree of silencing. We can answer the latter contention by considering two possible explanations; (I) if length of RNA is longer, it takes longer time by the transcriptional process to synthesize it, thus allowing the RNA to stay for more time at the site of transcription, this in turn helps in increased recruitment of the heterochromatic complexes to the site of transcription, or (II) alternatively, an interaction between trans-acting factors that associate with unidentified *cis*-acting elements in the ICR (probably the *cis*-acting elements in the Region 2, for details refer paper II) and the basal transcription machinery of antisense RNA resulting into basal transcription machinery that spread heterochromatin as it transcribes through the sequence in a manner that is analogous to RNA polymerase-dependent spreading of heterochromatin structures at the RDN1 locus of *S.cerevisiae*. In this context, intensity of silencing increases with increase in the length of antisense transcript.

Antisense RNA-mediated bi-directional silencing does not involve dsRNA

In the paper III, we have ruled out the possible involvement of dsRNA by selectively deleting the one of the opposing promoters, *H19* promoter, in PS4 episome. Although this observation rules out the possible dsRNA formation due to opposing transcription between *H19* and *Kcnq1ot1* promoter

but does not exclude the involvement of other promoters, such as EBNA, which is present 3.0 kb upstream of *H19* promoter. To conclusively rule out the involvement of dsRNA, we deleted the *H19* transcription unit from PS4 polyA 9.2, as the polyadenylation sequence in this construct truncate antisense RNA prior to reaching the EBNA transcription unit. Since we cannot analyze the *H19* gene activity due to its deletion from the construct, we have assayed the activity of *hygromycin* gene, which showed significant repression. These results suggest the dsRNA-mediated silencing pathway is not a component of *Kcnq1* ICR-mediated bidirectional silencing process.

Bi-directional silencing occurs irrespective of identity of the sequences through which Kcnq1ot1 transcription occurs

To address whether the identity of the sequences through which antisense RNA passes has any role to play in the silencing process, we added an extra 3.2 Kb native sequence from the *Kcnq1* locus to already existing 1.7kb sequence to make the total length of the native *Kcnq1* gene sequence downstream of antisense promoter to 4.9kb. We inserted the SV40 polyadenylation sequence at the end of 4.9kb. We mapped the antisense transcript by RT-PCR prior and after the polyadenylation sequence insertion and found that it lack antisense transcription beyond the polyadenylation sequence insertion site, suggesting that polyadenylation indeed truncated antisense transcript. Analysis of the reporter genes activity in this construct showed significant silencing. These results suggest that the identity of the sequences, through which *Kcnq1ot1* transcription occurs, does not play a role in the bidirectional silencing.

Antisense transcription silences overlapping gene prior to non-overlapping genes.

Since *H19* and *hygromycin* genes mimic in spatial alignment the *in vivo* situation in relation to the mouse *Kcnq1* ICR in PS4 episome, we wanted to address the kinetics of silencing of the overlapping *H19* and the non-overlapping *hygromycin* gene. In order to understand the latter issue in detail, we have transiently transfected PS4, PH19 (formerly referred as pREPH19 in paper II) and PS4 polyA 9.2 into the JEG-3 cell line for 2, 4 and 8 days and assayed the activity of the overlapping *H19* and the non-overlapping *hygromycin* by RNase protection assay, as described previously. The latter analysis indicated that the overlapping *H19* reporter gene is silenced much earlier when compared to the non-overlapping *hygromycin* gene. One of the plausible explanations for the quick silencing of the overlapping gene could be due to antisense transcription-mediated recruitment of heterochromatic machinery forms heterochromatin nucleation site at the overlapping side due to its proximity and then spreading of heterochromatin from the nucleation site to the non overlapping side occurs gradually.

Spreading of epigenetic modifications over the overlapping H19 reporter gene promoter is tightly linked to antisense transcription

We have earlier suggested that the DNA methylation occurs as a consequence of silencing rather than a cause for the silencing. We were interested in understanding what other epigenetic modifications that occur prior to DNA methylation and play a causal role in the silencing process mediated by the *Kcnq1* ICR. To this end, we performed chromatin immunopurification assays (ChIP) using antibodies to tri-methylated lysine 9 of H3 (mK9), di-methylated lysine 4 of H3 (mK4), Acetyl lysine 9 H3 (aK9), methyl CpG binding proteins such as MBD2, MeCP2 and Polycomb group proteins (PcG) complex member EED, on the crosslinked chromatin obtained from cells transfected with PS4 and PS4CAT3 (*Kcnq1 ICR* carrying mutations at the NF-Y binding sites and lack antisense transcription) for 8 days. The results from these experiments showed that the inactive chromatin modifications like mK9, methyl CpG binding proteins MBD2 and MeCP2 and EED were associated with the overlapping *H19* promoter region in relation to antisense transcription (PS4), while the active chromatin modifications such as methylated K4 and acetylated K9 were associated with the promoter in the absence of antisense transcription (PS4CAT3). The appearance of these epigenetic modifications that lead to the formation of inactive chromatin structures over *H19* promoter in the presence of antisense transcription, suggests that the antisense transcription mediates silencing by recruiting the heterochromatin machinery.

Conclusions

The efforts from several labs suggested that differentially methylated imprinting control regions (ICRs) play a critical role in the maintenance of parent of origin-specific expression patterns in imprinted clusters. In this thesis, by using a relatively well-investigated imprinted cluster located at the distal end of the mouse chromosome 7, we document that ICRs employ diverse methylation-sensitive mechanisms to control the parent of origin-specific expression of genes.

By exploiting the human and mouse *H19* ICRs and the JEG-3 cell line with high *de novo* methylation property, we demonstrate that the loss of methylation privilege status of the *H19* ICR in the JEG-3 cell line is reminiscent to aberrant methylation of the ICR in neoplastic conditions, implying the fact that aberrant epigenetic activity is one of the main features of cancer cells.

More importantly, the present investigations on the *Kcnq1* ICR unleash new perspectives into antisense RNA mediated bi-directional silencing process. In addition, the studies indicate that ICRs employ divergent mechanisms to control the parent of origin-specific expression of genes in clusters.

Acknowledgements

First of all, I dedicate all my academic achievements so far, to the members of my family, specially my parents. They have been, and will continue to be the sole driving force for me in all walks of life.
My mentor, Dr. Chandrasekhar Kanduri deserves special thanks for helping me settle in this new place and shaping up my scientific abilities during the Ph.D. program.
I sincerely thank the Head of our Department, Prof. Rolf Ohlsson, without whose help my Ph.D. would have been both scientifically and logistically impossible.
I am deeply indebted to Prof. Reinald H. Fundele for his moral support, and constructive criticisms, whenever needed.
I am also very thankful to my master supervisors Prof. S.C.Lakhotia, Prof. Rajeev Raman, Prof. Mercy Raman, and others in BHU who opened a new world of science to me.
Mrs. Helena Malmikumpu (for being so patient to all our demands and being so helpful), Mrs. Anita Mattson (for a nice time in the lab. working together), Mrs. Rosemarie Löfberg (for her administrative support), and Carina Ostman (for her skiing lessons and other helps), deserve gratefulness next to none.
My numerous colleagues, Rolf Ericsson (for moments of humour while troubleshooting computers), Piero and Claes (for making up the league of extraordinary gentleman), Rosita (for her multifaceted helps), Anita and Wen Qiang Yu (for giving medical advices), Sha and Irina (for their nice advices and helpful behaviour), Taras (for his Russian jokes and chatting), Clara (for a nice company in the office), Pandey, Magda and Sreenivasulu (for their cooperation).
Uppsala has brought to me wonderful friends like Yang & Wei, HongYu, Tong & Wei, Tiger, Shinshin & WenYu, Reza & Nazila, Sonchita & Bhoopi. They have all been very special and helpful to me. They all deserve more than just thanks.
Above all, I am thankful to Uppsala for bringing to me my other half, Umashankar.
May the almighty bless us all!

****

References

Andrulis, E. D., D. C. Zappulla, K. Alexieva-Botcheva, C. Evangelista, and R. Sternglanz. 2004. One-hybrid screens at the Saccharomyces cerevisiae HMR locus identify novel transcriptional silencing factors. Genetics **166:**631-5.

Antequera, F., and A. Bird. 1993. Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci U S A **90:**11995-9.

Banerji J, Olson L, Schaffner W. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **3**:729-40.

Ben Lehner, Gary Williams, R. Duncan Campbell and Christopher M. Sanderson. 2002. Antisense transcripts in the human genome. Trends Genet **18**: 63–65.

Ben-Porath I, Cedar H. 2001. Epigenetic crosstalk. Mol Cell **5**:933-5.

Bestor, T. H. 2000. The DNA methyltransferases of mammals. Hum Mol Genet **9:**2395-402.

Bird, A. P., and A. P. Wolffe. 1999. Methylation-induced repression--belts, braces, and chromatin. Cell **99:**451-454.

Bliek J, Gicquel C, Maas S, Gaston V, Le Bouc Y, Mannens M. 2004. Epigenotyping as a tool for the prediction of tumor risk and tumor type in patients with Beckwith-Wiedemann syndrome (BWS). J Pediatr **6**:796-9.

Bondarenko, V. A., Y. V. Liu, Y. I. Jiang, and V. M. Studitsky. 2003. Communication over a large distance: enhancers and insulators. Biochem Cell Biol **81:**241-51.

Boumil, R. M., and J. T. Lee. 2001. Forty years of decoding the silence in X-chromosome inactivation. Hum Mol Genet **10:**2225-32.

Brown, C., A. Ballabio, J. Rupert, R. Lafreniere, M. Grompe, R. Tonlorenzo, and W. HF. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature **349:**38-42.

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell **6**:843-51.

Burgess-Beusse, B., C. Farrell, M. Gaszner, M. Litt, V. Mutskov, F. Recillas-Targa, M. Simpson, A. West, and G. Felsenfeld. 2002. The insulation of genes from external enhancers and silencing chromatin. Proc Natl Acad Sci U S A **99 Suppl 4:**16433-7. Epub 2002 Aug 1.

Caiafa P, Zampieri M. 2005. DNA methylation and chromatin structure: the puzzling CpG islands. J Cell Biochem **2**:257-65.

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

Cerrato F, Vernucci M, Pedone PV, Chiariotti L, Sebastio G, Bruni CB, Riccio A.2002 The 5' end of the KCNQ1OT1 gene is hypomethylated in the Beckwith-Wiedemann syndrome. Hum Genet **1**:105-7. Epub 2002 Jun 14.

Chan HM, La Thangue NB. 2001. p300/CBP proteins: HATs for transcriptional bridges and scaffolds **13**:2363-73.

Cheung, P., and P. Lau. 2005. Epigenetic regulation by histone methylation and histone variants. Mol Endocrinol **19:**563-73. Epub 2005 Jan 27.

Cohen, D. E., and J. T. Lee. 2002. X-chromosome inactivation and the search for chromosome-wide silencers. Curr Opin Genet Dev **12:**219-24.

Constancia, M., W. Dean, S. Lopes, T. Moore, G. Kelsey, and W. Reik. 2000. Deletion of a silencer element in *Igf2* results in loss of imprinting independent of *H19*. Nature Genet **26:**203-6.

Craig, J. M. 2005. Heterochromatin--many flavours, common themes. Bioessays **27:**17-28.

Cui H, Niemitz EL, Ravenel JD, Onyango P, Brandenburg SA, Lobanenkov VV, Feinberg AP. 2001. Loss of imprinting of insulin-like growth factor-II in Wilms' tumor commonly involves altered methylation but not mutations of CTCF or its binding site Cancer Res **13**:4947-50.

Delaval, K., and R. Feil. 2004. Epigenetic regulation of mammalian genomic imprinting. Curr Opin Genet Dev **14:**188-95.

Dhillon, N., and R. T. Kamakaka. 2002. Breaking through to the other side: silencers and barriers. Curr Opin Genet Dev **12:**188-92

Diaz-Meyer N, Day CD, Khatod K, Maher ER, Cooper W, Reik W, Junien C, Graham G, Algar E, Der Kaloustian VM, Higgins MJ. 2003. Silencing of CDKN1C (p57KIP2) is associated with hypomethylation at KvDMR1 in Beckwith-Wiedemann syndrome. J Med Genet **40**:797-801.

Dillon, N. 2004. Heterochromatin structure and function. Biol Cell **96:**631-7.

Feinberg AP, Tycko B. 2004 .The history of cancer epigenetics. Nat Rev Cancer **4**:143-53.

Feinberg AP, Vogelstein B. 1983. Hypomethylation of ras oncogenes in primary human cancers. Biochem Biophys Res Commun **111**:4754.

Fitzpatrick GV, S. P., Higgins MJ. 2002. Reginal loss of imprinting and growth deficiency in mice with targeted deletion of KvDMR1. Nat Genet **32:**426-431.

Fitzpatrick, G. V., P. D. Soloway, and M. J. Higgins. 2002. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. Nat Genet **32:**426-31. Epub 2002 Sep 9.

Fukuzawa R, Hata J, Hayashi Y, Ikeda H, Reeve AE. 2003. Beckwith-Wiedemann syndrome-associated hepatoblastoma: wnt signal activation occurs later in tumorigenesis in patients with 11p15.5 uniparental disomy. Pediatr Dev Pathol **6**:299-306.

Gerasimova, T., and V. Corces. 1996. Boundary and insulator elements in chromosomes. Curr Opin Genet Dev**:**185-92.

Giovanni Lavorgna, Dvir Dahary, Ben Lehner, Rotem Sorek, Christopher M. Sanderson and Giorgio Casari. 2004. In search of antisense. Trends in Biochemical Sciences **29**: 88-94.

Grewal, S. I., and D. Moazed. 2003. Heterochromatin and epigenetic control of gene expression. Science **301:**798-802.

Hall IM, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, Grewal SI. 2002. Establishment and maintenance of a heterochromatin domain. Science **5590**:2232-7. Epub 2002 Sep 05.

Hark, A. T., and S. M. Tilghman. 1998. Chromatin conformation of the *H19* epigenetic mark. Hum Mol Genet. **7:**1979-85.

Hata, K., M. Okano, H. Lei, and E. Li. 2002. Dnmt3L cooperates with the Dnmt3 family of *de novo* DNA methyltransferases to establish maternal imprints in mice. Development **129:**1983-93.

Hatzis, P., and I. Talianidis. 2002. Dynamics of enhancer-promoter communication during differentiation-induced gene activation. Mol Cell **10:**1467-77.

Henikoff, S., T. Furuyama, and K. Ahmad. 2004. Histone variants, nucleosome assembly and epigenetic inheritance. Trends Genet **20:**320-6.

Hoffman, A. R., T. H. Vu, and J. Hu. 2000. Mechanisms of genomic imprinting. Growth Horm IGF Res **10 Suppl A:**S18-9.

Iguchi-Ariga, S. M., and W. Schaffner. 1989. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev **3:**612-9.

Jack J, Dorsett D, Delotto Y, Liu S. 1991. Expression of the cut locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer. Development **3**:735-47.

Jackson, J. P., A. M. Lindroth, X. Cao, and S. E. Jacobsen. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature **416:**556-60. Epub 2002 Mar 17.

Kanduri, C., C. Holmgren, G. Franklin, M. Pilartz, E. Ullerås, M. Kanduri, L. Liu, V. Ginjala, E. Ulleras, R. Mattsson, and R. Ohlsson. 2000. The 5'-flank of the murine *H19* gene in an unusual chromatin conformation unidirectionally blocks enhancer-promoter communication. Curr Biol **10:**449-457.

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

Kantor B, Makedonski K, Green-Finberg Y, Shemer R, Razin A. 2004. Control elements within the PWS/AS imprinting box and their function in the imprinting process. Hum Mol Genet **7**:751-62.

Kato, Y., and H. Sasaki. 2005. Imprinting and looping: epigenetic marks control interactions between regulatory elements. Bioessays **27:**1-4.

Kelley, R. L., and M. I. Kuroda. 2000. Noncoding RNA genes in dosage compensation and imprinting. Cell **103:**9-12.

Khatib Z, Levi A, Pefkarou A, Escalon E. 2004. Acute lymphocytic leukemia in a child with Beckwith-Wiedemann syndrome. JPediatrHematol Oncol **26**:45-7.

Kovesdi, I., R. Reichel, and J. R. Nevins. 1987. Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. Proc Natl Acad Sci U S A **84:**2180-4.

Lachner M, Jenuwein T. 2002. The many faces of histone lysine methylation. Curr Opin Cell Biol **3**:286-98.

Ladurner, A. G. 2003. Inactivating chromosomes: a macro domain that minimizes transcription. Mol Cell **12:**1-3.

Lee, J. T., L. S. Davidow, and D. Warshawsky. 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet **21:**400-4.

Lee, J., K. Inoue, R. Ono, N. Ogonuki, T. Kohda, T. Kaneko-Ishino, A. Ogura, and F. Ishino. 2002. Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. Development **129:**1807-17.

Leighton, P. A., J. R. Saam, R. S. Ingram, C. L. Stewart, and S. M. Tilghman. 1995. An enhancer deletion affects both H19 and Igf2 expression. Genes Dev. **9:**2079-2089.

Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **6**:915-26.

Li, J., C. T. Noguchi, W. Miller, R. Hardison, and A. N. Schechter. 1998. Multiple regulatory elements in the 5'-flanking sequence of the human epsilon-globin gene. J Biol Chem **273:**10202-9.

Lund AH, van Lohuizen M. 2004. Epigenetics and cancer.Genes Dev **19**:2315-35.

Lyon, M. F. 1961. Gene action in the X chromosome of the mouse (*Mus musculus* L.). Nature **190:**373.

McKittrick, E., P. R. Gafken, K. Ahmad, and S. Henikoff. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proc Natl Acad Sci U S A **101:**1525-30. Epub 2004 Jan 19.

Mitsuya, K., M. Meguro, M. P. Lee, M. Katoh, T. C. Schulz, H. Kugoh, M. A. Yoshida, N. Niikawa, A. P. Feinberg, and M. Oshimura. 1999. LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. Hum Mol Genet **8:**1209-17.

Murrell, A., S. Heeson, L. Bowden, M. Constancia, W. Dean, G. Kelsey, and W. Reik. 2001. An intragenic methylated region in the imprinted Igf2 gene augments transcription. EMBO Rep **2:**1101-6. Epub 2001 Nov 21.

Nakayama M, Wada M, Harada T, Nagayama J, Kusaba H, Ohshima K, Kozuru M, Komatsu H, Ueda R, Kuwano M. 1998. Hypomethylation status of CpG sites at the promoter region and overexpression of the human MDR1 gene in acute myeloid leukemias. Blood **11**:4296-307.

Nikaido, I., C. Saito, A. Wakamoto, Y. Tomaru, T. Arakawa, Y. Hayashizaki, and Y. Okazaki. 2004. EICO (Expression-based Imprint Candidate Organizer): finding disease-related imprinted genes. Nucleic Acids Res **32:**D548-51.

Okano M, Xie S, Li E. 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res **11**:2536-40.

Pandey, R. R., M. Ceribelli, P. B. Singh, J. Ericsson, R. Mantovani, and C. Kanduri. 2004. NF-Y regulates the antisense promoter, bidirectional silencing, and differential epigenetic marks of the Kcnq1 imprinting control region. J Biol Chem **279:**52685-93. Epub 2004 Sep 29.

Pant, V., P. Mariano, C. Kanduri, A. Mattsson, V. Lobanenkov, R. Heuchel, and R. Ohlsson. 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:**586-590.

Passarge E. 1979. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. Am J Hum Genet **2**:106-15.

Paulsen, M., K. Davies, L. Bowden, A. Villar, O. Franck, M. Fuermann, W. Dean, T. Moore, N. Rodrigues, K. Davies, R. Hu, A. Feinberg, E. Maher, W. Reik, and J. Walter. 2000. Syntenic organization of the mouse distal chromosome 7 imprinting cluster and the Beckwith-Wiedemann syndrome region in chromosome 11p15.5. Hum Mol Genet **7:**1149-59.

Peterson, C. L., and M. A. Laniel. 2004. Histones and histone modifications. Curr Biol **14:**R546-51.

Prendergast, G. C., and E. B. Ziff. 1991. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. Science **251:**186-9.

Prokhortchouk E, Hendrich B. 2002. Methyl-CpG binding proteins and cancer. Are MeCpGs more important than MBDs? Oncogene **21**:5394-5399.

Razin, A., and A. D. Riggs. 1980. DNA methylation and gene function. Science **210:**604-10.

Reik, W., and W. Dean. 2001. DNA methylation and mammalian epigenetics. Electrophoresis **22:**2838-43.

Richards, E. J., and S. C. Elgin. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell **108:**489-500.

Robertson KD. 2002 . DNA methylation and chromatin - unraveling the tangled web. Oncogene **35**:5361-79.

Robertson, K. D. 2002. DNA methylation and chromatin - unraveling the tangled web. Oncogene **21:**5361-79.

Rougeulle, C., and E. Heard. 2002. Antisense RNA in imprinting: spreading silence through Air. Trends Genet **18:**434-7.

Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M. 1996. HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. Proc Natl Acad Sci U S A **25**:14503-8.

Sado, T., Z. Wang, H. Sasaki, and E. Li. 2001. Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development **128:**1275-86.

Scarano, M. I., M. Strazzullo, M. R. Matarazzo, and M. D'Esposito. 2005. DNA methylation 40 years later: Its role in human health and disease. J Cell Physiol **12:**12.

Schoenherr, C., J. Levorse, and S. Tilghman. 2003. CTCF maintains differential methylation at the *Igf2/H19* locus. Nat Genet **33:**66-9.

Schramke V, Allshire R. 2004. Those interfering little RNAs! Silencing and eliminating chromatin. Curr Opin Genet Dev **2**:174-80.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell **7**:941-53.

Sleutels, F., R. Zwart, and D. P. Barlow. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature **415:**810-3.

Smale, S.T., and and D.Baltimore. 1989. The initiator as a transcription control element. Cell **57**:103-13.

Smilinich, N. J., C. D. Day, G. V. Fitzpatrick, G. M. Caldwell, A. C. Lossie, P. R. Cooper, A. C. Smallwood, J. A. Joyce, P. N. Schofield, W. Reik, R. D. Nicholls, R. Weksberg, D. J. Driscoll, E. R. Maher, T. B. Shows, and M. J. Higgins. 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.

Soppe, W. J., Z. Jasencakova, A. Houben, T. Kakutani, A. Meister, M. S. Huang, S. E. Jacobsen, I. Schubert, and P. F. Fransz. 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. Embo J **21:**6549-59.

Suetake, I., F. Shinozaki, J. Miyagawa, H. Takeshima, and S. Tajima. 2004. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem **279:**27816-23. Epub 2004 Apr 21.

Surani, M., S. Barton, and M. Norris. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature **308:**548-550.

Suter CM, Martin DI, Ward RL. 2004. Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue. Int J Colorectal **2**:95-101. Epub 2003 Oct 08.

Szabó, P., S.-H. Tang, A. Rentsendorj, G. Pfeifer, and J. Mann. 2000. Maternal-specific footprints at putative CTCF sites in the *H19* imprinting control region give evidence for insulator function. Curr Biol **10:**607-610.

Takagi, N., and M. Sasaki. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature **256:**640-2.

Tamaru, H., and E. U. Selker. 2001. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature **414:**277-83.

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

Tilghman SM, B. M., Webber AL, Brunkow ME, Saam J, Leighton PA and Pfeifer K. 1995. Genomic imprinting of the H19 and Igf2 genes in the mouse. Cambridge University Press.

Tremblay, K. D., J. R. Saam, R. S. Ingram, S. M. Tilghman, and M. S. Bartolomei. 1995. A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. Nature Genetics **9:**407-413.

Tsai CN, Tsai CL, Tse KP, Chang HY, Chang YS. 2002. The Epstein-Barr virus oncogene product, latent membrane protein 1, induces the downregulation of E-cadherin gene expression via activation of DNA methyltransferases. Proc Natl Acad Sci U S A **15**:10084-9.

Tufarelli, C., J. A. Stanley, D. Garrick, J. A. Sharpe, H. Ayyub, W. G. Wood, and D. R. Higgs. 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. Nat Genet **34:**157-65.

Weksberg R, Smith AC, Squire J, Sadowski P. 2003. Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. Hum Mol Genet **12**: 61-8.

West, A. G., M. Gaszner, and G. Felsenfeld. 2002. Insulators: many functions, many mechanisms. Genes Dev **16:**271-88.

West, A. G., S. Huang, M. Gaszner, M. D. Litt, and G. Felsenfeld. 2004. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. Mol Cell **16:**453-63.

Wilkins, J.F., and D.Haig. 2003. What good is genomic imprinting: the function of parent-specific gene expression. Nat Rev Genet **4**: 359-68.

Volpe, T. A., C. Kidner, I. M. Hall, G. Teng, S. I. Grewal, and R. A. Martienssen. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science **297:**1833-7. Epub 2002 Aug 22.

Wu P, Qiu C, Sohail A, Zhang X, Bhagwat AS, Cheng X. 2003. Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. J Biol Chem **7**:5285-91.

Yelin, R., D. Dahary, R. Sorek, E. Y. Levanon, O. Goldstein, A. Shoshan, A. Diber, S. Biton, Y. Tamir, R. Khosravi, S. Nemzer, E. Pinner, S. Walach, J. Bernstein, K. Savitsky, and G. Rotman. 2003. Widespread occurrence of antisense transcription in the human genome. Nat Biotechnol **21:**379-86. Epub 2003 Mar 17.

Zilberman, D., X. Cao, and S. E. Jacobsen. 2003. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science **299:**716-9. Epub 2003 Jan 9.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 29

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-5725

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2005