Regulatory Roles of Noncoding RNA in Development and Disease

GAURAV KUMAR PANDEY
Abstract


Long noncoding RNAs (lncRNAs) are being realized as important players in gene regulation and their misregulation has been considered as one of the underlying causes for tumor initiation and progression in many human pathologies. In the current thesis, I have addressed the functional role of lncRNAs in development and disease model systems.

Genomic imprinting is an epigenetic phenomenon by which subset of genes are expressed in a parent of origin-specific manner. The Kcnq1 imprinted locus is epigenetically regulated by Kcnq1ot1 lncRNA. Deletion of an 890bp region at the 5’ end of Kcnq1ot1 in mouse resulted in the loss of silencing of neighboring ubiquitously imprinted genes (UIGs). In addition, we observed loss of DNA methylation at the UIG promoters. We have shown that Kcnq1ot1 RNA establishes CpG methylation by interacting with DNMT1. To explore the stability of lncRNA mediated silencing pathways, we have conditionally deleted Kcnq1ot1 in the mouse in a stage and tissue-specific manner. We have shown that Kcnq1ot1 is continuously required for maintaining the silencing of UIGs, whereas the silencing of the placental imprinted genes is maintained in an RNA independent manner.

To identify chromatin-associated lncRNA (CARs) on a genome-wide scale, we purified RNA from the sucrose gradient fractionated chromatin and subjected it to RNA sequencing. Our study has identified 141 intronic and 74 long intergenic CARs. Characterization of one of the CARs revealed that it regulates the expression of neighboring genes in cis by modulating the chromatin structure.

We have explored the functional role of lncRNA in tumor progression and initiation by using pediatric neuroblastoma. By transcriptional profiling of low- and high-risk tumors, we have identified several lncRNAs differentially expressed between these subtypes. We report an uncharacterized RNA NBAT-1, expressed at lower levels in high-risk tumors relative to low-risk tumors. Using neuroblastoma cell culture system, we demonstrated that NBAT-1 has anti-cell proliferative and anti-invasive properties. In addition, it promotes differentiation of neurons from undifferentiated neuroblastoma cell lines.

In summary, by employing mouse genetics, cell culture based model system and expression profiling in tumors, we have uncovered new roles of lncRNA in gene regulation.

Keywords: Noncoding RNA, Genomic Imprinting, Epigenetics, Neuroblastoma

Gaurav Kumar Pandey, Department of Immunology, Genetics and Pathology, Rudbeck laboratory, Uppsala University, SE-751 85 Uppsala, Sweden.

© Gaurav Kumar Pandey 2013

ISSN 1651-6206
urn:nbn:se:uu:diva-209596 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-209596)
Dedicated to my family
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I  Mohammad F, Mondal T, Guseva N, **Pandey GK**, Kanduri C. 


*-These authors contributed equally to the work.

Reprints were made with permission from the respective publishers.
Contents

Introduction ................................................................................................... 11
Complexity of the eukaryotic gene regulation ............................................. 11
Epigenetic regulators of gene expression ..................................................... 12
  DNA Methylation ................................................................................ 13
  Histone Modification ........................................................................... 14
  Noncoding RNA .................................................................................. 15
High-throughput strategies to identify ncRNAs on a genome wide scale .......................................................................................................... 16
  Array based approaches ....................................................................... 16
  Next generation sequencing methods .................................................. 17
  Immunoprecipitation based approaches ............................................... 19
LncRNAs as regulators of developmental processes .................................. 20
  Genomic Imprinting and ncRNA ......................................................... 20
  ncRNAs and Dosage compensation ..................................................... 24
  Other roles of LncRNAs in animal development ................................... 26
LncRNAs in disease ................................................................................. 28
  ANRIL LncRNA .................................................................................... 28
  HOTAIR LncRNA ................................................................................. 28
  MALATI LncRNA ................................................................................ 29
  LncRNAs in prostate cancer ................................................................. 29
  LncRNAs in neuroblastoma ................................................................... 30
Mechanisms of LncRNA mediated gene regulation .................................. 31
  1. Recruitment of Chromatin modifying enzymes ............................... 31
  2. Regulation of tumor suppressor or oncogene signaling ................... 33
  3. Sequestration of microRNA ............................................................ 33
  4. Regulators of mRNA processing, stability and translation .............. 33
  5. Decoys of transcription factors ........................................................ 34
Aims .............................................................................................................. 35
Results and Discussion ................................................................................. 36
  Paper I: Kcnqlot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1 ................................................................. 36
  Paper II: Long noncoding RNA mediated maintenance of DNA methylation and transcriptional gene silencing .................................................... 37
  Paper III: Characterization of the RNA content of chromatin ............. 39
PAPER IV: Transcriptomic profiling of primary neuroblastomas reveals a high-risk tumor associated long noncoding RNA *NBAT1*, with functional roles in cell proliferation and neuronal differentiation. 40

Acknowledgement .................................................................42

References.............................................................................45
Abbreviations

ncRNA  Non coding RNA  
IncRNA  Long noncoding RNA  
PRC2  Polycomb repressive complex 2  
PRE  Polycomb responsive element  
RdDM  RNA-directed DNA methylation  
RDR2  RNA-dependent RNA polymerase  
DCL3  DICER-like 3  
DRM2  DNA methyltransferase  
TGS  Transcriptional gene silencing  
ES  Embryonic stem cells  
iPSCs  Induced pluripotent stem cells  
CAR  Chromatin associated RNA  
LincRNA  Large intergenic noncoding RNA  
ICR  Imprinting control region  
DMR  Differentially methylated region  
UIG  Ubiquitously imprinted genes  
PIG  Placental imprinted genes  
XCI  X chromosome inactivation  
DCC  Dosage compensation complex  
HOX  Homeobox  
PCAT  Prostate cancer associated transcript
Introduction

The cellular complexity of higher eukaryotic organisms has generated a great deal of scientific interest since several years. Many theories and hypotheses have prevailed through the course of time explaining the intricate complexity, cellular diversity and tight gene regulation present in the highly evolved organisms like mammals. The initial hypothesis prevalent before the genomic revolution, were based on the assumptions that the number of genes linearly increase with the organismal size and developmental complexity. This increase was presumed mainly in terms of protein coding genes, as only they were envisaged to regulate the eukaryotic gene regulatory network. A dramatic change in this perception about the basis for the organismal complexity in higher organisms occurred after the genomic revolution in 90s and early 2000s. A very significant realization based on the results of high-throughput studies came in the form of the number of protein coding genes present in higher eukaryotes. The human genome contains around 20-25000 protein coding genes (far less than the earlier estimates) which is almost the same in number as present in nematode, *Caenorhabditis elegans* [1, 2]. This is termed as the ‘G-value’ paradox. It is very hard to visualize that similar number of protein coding genes could regulate, both a worm of 1000 cells and humans with diverse cell types and complex body plan. This contradiction in number of genes has made us to focus on the aspect of gene regulation, in order to understand and decipher the basis for complexities of the cellular diversity and development in higher organisms.

Complexity of the eukaryotic gene regulation

The conventional view regarding eukaryotic gene regulation has been largely protein centric. However, the evidences from the high-throughput studies have suggested that the eukaryotic genome is pervasively transcribed, and proteins coding RNAs form a very minor part (1.5%) of the transcriptional output. The major part of the transcribed genome is noncoding and is constituted of RNAs lacking coding potential, termed as noncoding RNAs (ncRNAs) [3-5]. High-throughput approaches have lead to identification of several classes of ncRNAs transcribed in the eukaryotic genome in diverse tissue types and developmental stages. The ncRNAs are mainly comprised of ribosomal RNAs (rRNA) and transfer RNAs (tRNAs) involved in house-
keeping functions, small RNAs such as microRNAs (miRNA), snoRNA, piwi-interacting RNAs (piRNA) and long noncoding RNAs (lncRNAs). The small RNAs are highly conserved and are involved in transcriptional and post-transcriptional gene silencing. In contrast, the lncRNAs (defined as RNAs greater than 200 nt in length) are poorly conserved and regulate gene expression through diverse mechanisms not yet fully understood[6-8]. However, lncRNAs have been reported in a myriad of biological processes such as genomic imprinting, RNA maturation and transport, cellular differentiation, pluripotency maintenance, X inactivation and signal transduction[5, 9-13]. With lncRNAs being implicated as key players in regulating gene expression in diverse cellular and developmental processes, it is widely accepted that lncRNAs form a major component of the eukaryotic gene regulatory network. Understanding their roles in greater detail and uncovering the mechanisms behind their functions would provide us with much improved understanding of the eukaryotic gene regulation and help us explain in part, the organismal and developmental complexity of higher organisms[14].

In the current thesis, I will begin with describing the different strategies that have been used to identify several different classes of lncRNAs. With increasing repertoire of lncRNAs being reported, it becomes relevant to understand their functional roles. Hence, I will discuss how lncRNAs are playing crucial roles in diverse developmental processes. Due to their significant roles in regulating these processes, it is highly probable that misregulation in expression of lncRNAs could be one of the underlying causes of disease such as cancer. Elaborating more on this point, I would state examples of various lncRNAs which have been reported to be misregulated in disease states and what kind of roles they are involved in disease progression. Finally, I will summarize the current known mechanisms of lncRNA mediated gene regulation.

Before, I start discussing the lncRNAs in detail, it is relevant to introduce the known epigenetic modifications and how in combination with these modifications, ncRNA bring about gene regulation.

**Epigenetic regulators of gene expression**

The diverse cell types in multicellular organisms have the same genetic composition, yet are distinctly different from each other in expression of genes. This differential expression can be understood, in part through the phenomenon of epigenetics. It is defined as the study of the inheritable modifications of the DNA or DNA associated proteins without any changes in the DNA sequences. Two well-known epigenetic modifications, DNA meth-
ylation and histone modification are known to act in conjugation with ncRNAs to bring about transcriptional regulation.

DNA Methylation

DNA methylation is one of the most well studied epigenetic mark present predominantly in the eukaryotic genome on the cytosine nucleotide (5-methyl cytosine) and, it is present as CpG dinucleotides across the genome[15].

The mammalian DNA methylation machinery consists of three families of DNA methyltransferases (DNMTs). DNMT3A and DNMT3B are the de novo methyltransferases involved primarily in the establishment of the methylation pattern whereas DNMT1 is the maintenance methyltransferase required for the faithful propagation of the methylation imprint over cellular divisions. Another member of the DNMT3 family, DNMT3L has been observed to be essential for establishment of methylation at most of the imprinted loci[16]. This DNMT lacks methyltransferase activity but can stimulate the activity of the de novo DNMTs like DNMT3A and DNMT3B[17, 18]. The classical view regarding methylation has been that once established, it is successfully perpetuated across generations by DNMT1 through its ability to specifically recognize the hemimethylated DNA. However, recent studies are providing evidences that other factors are required in addition to DNMT1, for maintenance of methylation. For example, UHRF1, a chromatin associated protein, which specifically binds to hemimethylated CpG nucleotides, has been shown to interact with DNMT1. In addition, loss of UHRF1 in embryonic stem (ES) cells leads to a drastic decrease in methylation indicating its role in DNA methylation[19]. It has been reported that UHRF1 is involved in maintenance of DNA methylation by regulating the stability of DNMT1, through its association with methylated histone H3 on lysine 9 residue[20].

The CpG dinucleotides are underrepresented in the eukaryotic genomes. It is due to the reason that, methylated cytosines are prone to deamination. They can get converted to uracil, which is removed by the base excision pathway or to thymine which makes it susceptible to mutagenesis. The majorities (50-70%) of the CpG dinucleotides are methylated in the vertebrate genome with an exception of hypomethylated regions called as CpG islands[21, 22]. These CpG islands are characterized as, the unmethylated regions present mostly near the gene promoters, which are around 1000 base pairs in length and show high GC content. The only exception to the unmethylated CpG islands, are the imprinted genes and genes undergoing X inactivation. NcRNAs, transcriptional events and transcription factors are known to be involved in methylation of these CpG islands present at imprinted genes and genes undergoing X inactivation [23, 24].
The unmethylated CpG islands are believed to keep the transcription initiation of the gene promoters going and methylation at these regions could lead to repression of the associated gene transcription. Indeed, these islands have been shown to be present over the transcription start site of the entire house keeping genes and on a significant number of tissue-specific or regulatory genes[25]. Also, a genome wide study in humans has found that the majority (72%) of the annotated gene promoters (Refseq database) are associated with high CpG content[26]. An interesting aspect associated with these islands is, how do CpG islands escape de novo methylation? It has been observed that CpG islands are enriched with active histone marks, histone H3 trimethyl lysine 4 (H3K4me3) suggesting that active chromatin at these islands could be crucial for their unmethylated state. Indeed, it has been reported that H3K4me3 prevents the access of DNMT3L, which can directly interact with unmethylated tail of histone H3[27]. In addition, the protein Cfp1 has been found to be present over the unmethylated CpG islands, whose depletion leads to reduced levels of H3K4me3 indicating its involvement in maintaining the unmethylated state of the CpG islands by regulating the local active chromatin structure[28].

Histone Modification

In eukaryotes the DNA is packed with proteins known as histones and forms the basic structural components of the eukaryotic chromatin. The dimers of core histone proteins H2A, H2B, H3 and H4 constitute the core octameric nucleosomal molecule with 147bp of DNA wrapped around it. Along with these four core histone proteins there is histone protein H1 which forms an important component of higher order chromatin organization. The N-terminal tails of the histones undergo post translational modifications (PTM). These PTM include modifications such as methylation, acetylation, ubiquitylation, phosphorylation and sumoylation. Most of these modifications occur mainly on the lysine (K) and arginine (R) residues. These modifications affect the accessibility of various regulatory factors and complexes to chromatin and thus play a vital role in regulating gene expression. H3K4 di or tri methylation (me) are mainly associated with active promoters, H3K36me3 and H3K79me2 are associated with the transcribed regions whereas H3K27me3, H3K9me2 and H3K9me3 are present mainly over repressed genes[29]. The distal regulatory elements such as enhancers are enriched mainly with H3K4 monomethylation and H3K27 acetylation. The active mark (H3K4me3) and repressive mark (H3K27me3) can be present together in ES cells as a bivalent chromatin[30]. This bivalent chromatin feature is thought to provide genes a flexibility to commit itself to an active or inactive state in specific lineages during subsequent differentiation. These histone modifications are catalyzed by specific enzymes such as histone methyltransferases (HMTs), histone acetylases (HATs) and histone deacetyl-
lases (HDACs). Polycomb repressive complex 1 (PRC1) and 2 (PRC2) are involved in regulation of repressive chromatin mark, H3K27me3 [31]. The PRC2 complex is composed of three major proteins, SUZ12, EED and EZH2 which is a HMT. These PRC2 members are known to be recruited by polycomb repressive elements (PRE) in *Drosophila* for PRC2 mediated gene repression[32]. However not much is known about presence of such elements in mammals. Another interesting aspect associated with PRC2 members is, how does the faithful propagation of the repressive marks take place. Biochemical and structural studies have revealed that the repressive mark H3K27me3 promotes its own propagation on the newly incorporated nucleosomes at the end of mitosis[33]. This is done by the allosteric activation of the PRC2 activity by the c-terminal domain of the PRC2 member EED[34].

**Noncoding RNA**

NcRNAs are emerging as novel epigenetic regulators which can work in conjugation with both DNA methylation and histone modifications machineries to bring about gene regulation. The role of small RNAs in transcriptional gene silencing (TGS) and post-transcriptional gene silencing is very well studied. We will discuss here the examples of mainly small RNAs and how they drive TGS via DNA methylation or histone modifications. Also, we will try to correlate it with the known information about their longer counterparts (lncRNAs) in TGS.

*(A) RNA and DNA methylation*

Role of ncRNAs in regulating gene expression via DNA methylation is reported in both plants and animals. RNA-directed DNA methylation (RdDM) was first reported in plants, following the observation that upon introduction of viroid cDNAs in tobacco genome, only the viroid sequences became methylated whereas the neighboring genomic DNA remained unaltered [35]. Since then, mechanisms of RdDM have been well studied in *Arabidopsis*. In *Arabidopsis*, single stranded long transcripts are transcribed from the repetitive heterochromatic regions by a specialized polymerase, RNA Pol IV. The single strands of these transcripts are used to produce the complementary strand by a RNA-dependent RNA polymerase (RDR2), through its association with RNA Pol IV. Once double stranded (ds) RNAs are produced, DICER-like 3 (DCL3) enzymes cleave these ds transcripts to generate 24 nt, small interfering RNAs (siRNAs). These siRNAs once loaded with the Argonaute 4 protein (AGO4) during their nuclear-cytoplasmic-nuclear transport, target the *denovo* DNA methyltransferase (DRM2) to methylate and silence the target genes[7, 36].

Along with small RNAs, examples of IncRNA are also known to be involved in regulating DNA methylation are also known. *Kcnq1ot1* IncRNA has been shown to interact with DNMT1 to regulate the differential methyla-
tion at the somatic DMRs of Cdkn1c and Slc22a18 genes[37]. In fact, using a conditional knockout approach, it was shown that continuous expression of RNA is required to maintain the differential methylation of these imprinted genes across cell generations during mouse development[38]. Similarly, Tsix lncRNA is reported to silence the Xist promoter by associating with DNMT3A[39, 40]. In another example, DNMT3b has been observed to associate with triplex formed by promoter RNAs and ribosomal DNA (rDNA) promoter to methylate and silence rDNA[41].

(B) RNA and histone modifications- In yeast, small RNAs have been shown to be critical for heterochromatin initiation and maintenance at the centromeric repeats. Removal of the components of RNAi machinery affected the silent state of the transgenes integrated at the centromeres. Loss of repressive histone mark H3K9 methylation coupled with increase in H3K4me3 and transcriptional activation of the repeat sequence was also observed. According to the current model of TGS in yeast, dsRNA transcribed from these repetitive regions are processed into siRNAs by Dcr1. These siRNAs guide the RNA-induced transcriptional silencing complex (RITSC) to pericentromeric regions along with Chp1 (chromodomain protein). Once localized, it silences the corresponding loci through recruitment of the H3K9me3 methyltransferase, Clr4[7, 36]. Similar to small RNAs, lncRNAs such as Kcnq1ot1, Air have been shown to interact with mammalian H3K9me3 HMT, G9a [42, 43]. Additionally, lncRNAs (RepA, Kcnq1ot1, HOTAIR, ANRIL) have also been observed to associate with other repressive modifiers such as PRC2 complex[42, 44-46].

High-throughput strategies to identify ncRNAs on a genome wide scale

Research groups have employed various strategies to globally identify and catalogue the lncRNAs in mammalian transcriptome. We would discuss some of these approaches which have evolved through the years to identify lncRNAs and the current methods or criteria which have been used to efficiently annotate them.

Array based approaches

At the beginning of 21st century, the sequencing of the human genome was nearly complete and the next challenge ahead was to decode the hidden biological information in these sequences. To interrogate these sequences at a global scale, high-throughput approaches were needed to be employed and the advent of microarray technology proved to be a very useful tool at that
time. The initial glimpse into presence of noncoding transcribed regions in mammalian genome was provided with use of microarrays. Kapranov et al used an oligonucleotide array to map transcribed region of human chromosome 22, in 11 different cell lines and observed that many unannotated regions are transcribed in comparison to annotated regions [47]. Similar results were described in another study using polyA(+) placental RNA on chromosome 22 where, unannotated transcribed regions (putative noncoding regions) were observed in equal number to annotated protein coding genes [48].

With technological advances, arrays have become one of most commonly used methods to identify or screen IncRNAs associated to a specific biological question. There are several reports which have used microarrays or tiling arrays to identify IncRNAs. Using a customized microarray, Dinger et al identified 945 IncRNAs expressed during differentiation of mouse ES cells to embryoid bodies [49]. Out of these, 174 IncRNAs were observed to be differentially expressed and correlated with pluripotency or differentiation events [50]. In additional studies, Mercer et al reported systematic analysis of IncRNA expression during neuronal and oligodendrocyte lineage specification, neuronal-glial fate switching and neural lineage maturation [51]. Another elegant approach has used information generated from genome-wide chromatin maps of H3K4me3-H3K36me3 (active marks at promoter and gene body respectively) domains across four mouse cell types, to generate a tilling array with 2.1 million oligonucleotide probes representing the identified 350 K4-K36 domains. Around 1,600 long intervening noncoding RNAs (lincRNAs) were identified from the analyzed four mouse cell types [52, 53]. Extending the same tilling array, around 3,300 human lincRNAs were identified mapping to K4-K36 chromatin signatures. Using this catalog of identified lincRNAs, Loewer et al designed a microarray of 900 lincRNA and hybridized it to RNA from fibroblast cells, derived iPSCs and ES cells. They detected 133 and 104 lincRNAs as induced or repressed across all iPSCs and ES cells as compared to fibroblast cells [54].

Next generation sequencing methods

RNA sequencing or famously known as RNA-seq is a powerful tool which provides in depth sequencing of transcripts at a single nucleotide resolution allowing identification, detection and quantification of known as well as novel IncRNAs. The steps followed in a standard RNA-seq procedure are presented below.
RNA-seq provides selective advantage over microarrays in identification of novel transcripts and has proved to be a handy tool in unbiased identification of diverse classes of noncoding transcripts across cell types. We would discuss some of these classes of ncRNAs whose genome-wide view/presence has been possible through use of these next generation sequencing (NGS) approaches.

(A) **Antisense transcripts**- These classes of ncRNAs are transcribed from the complementary strand of the sense transcript and overlap in part with it. Large scale cDNA sequencing in the FANTOM project have revealed the presence of widespread antisense transcription in mammalian genome. It was observed that at least 25-40% of mammalian protein coding genes have overlapping transcription[55]. The function of majority of these antisense transcripts remains unkown, however there are many very well studied examples of antisense IncRNAs of protein coding genes or noncoding RNAs in the mammalian genome such as Kcnq1/Kcnq1ot1, Igf2r/Air, Xist/Tsix[10].

(B) **Long intronic RNAs**- These classes of noncoding transcripts are present within the introns of either protein coding genes or noncoding RNAs. Mainly introns have been known to harbor small RNAs such as snoRNAs or microRNAs. However, recent large scale transcriptomic and computational analysis have reported that many intronic IncRNAs are present in mammalian genome[56, 57]. An example of such RNA is COLDAIR, an intronic transcript implicated in vernalization in plants. It is involved in epigenetic suppression of
FLOWERING LOCUS (FLC) through its interaction with PRC2 [58].

(C) **Bidirectional promoter transcription**— With high-throughput approaches such as NGS, it has been realized that initiation of RNA polymerase (RNA Pol II) transcripts can be bidirectional. Cryptic unstable transcripts (CUTs) are products of such a bidirectional transcription reported in unicellular eukaryotes[59]. Two studies using NGS or tiling arrays, have observed that CUTs are well defined transcription units and are rapidly subjected to degradation by exosomes. In addition, divergent transcription from known promoters can give rise to stable unannotated transcripts (SUTs). Similar bidirectional transcription has also been observed in higher eukaryotes[60, 61]. Exosome depletion and subsequent tiling array analysis in human fibroblasts has revealed a new class of short, poly-adenylated and unstable RNAs; promoter upstream transcripts or PROMPTs[62].

(D) **Enhancer RNAs**— Enhancers are regulatory DNA elements which can activate gene expression independent of their proximity or orientation to the target gene. NGS studies have revealed that several enhancers express lncRNAs, resulting into a new class of lncRNAs; enhancer RNAs or eRNAs. For example, based on ChIP profiles of p300/CBP binding, H3K4me1 and RNA PolII occupancy, Kim et al identified 2000 enhancers which encode enhancer RNAs[63]. A similar approach was employed in human macrophages to identify several transcribed enhancers[64].

**Immunoprecipitation based approaches**

LncRNAs like Xist, Kncq1ot1 and many others have been reported to associate with epigenetic regulators like EZH2 to regulate gene expression. Immunoprecipitation followed by microarray or sequencing has been used to identify such chromatin modifying enzymes associated RNAs. Using such an approach Zhao et al identified > 9000 PRC2 interacting RNAs[65]. Another modified method called as CLIP (cross-linking and immunoprecipitation) uses UV crosslinking followed by RNase digestion to retain only the protein–interacting RNA region, which is then immunoprecipitated and purified fragments are subjected to sequencing. CLIP has been used to identify intronic lncRNAs directly associated with PRC2 complex[66]. Many of the lncRNAs such as Kcnq1ot1, Xist, and HOTAIR use chromatin as a substrate to execute their biological functions. Thus, it can be postulated that there could be many other lncRNAs regulating gene expression through their chromatin association property. Based on this presumption, around 200 chromatin associated RNAs (CARs) were identified upon sequencing of RNAs from the chromatin fraction of human fibroblast cells. Functional analysis of one of the intergenic CARs, intergenic 10 showed that, it epige-
netically regulates the expression of neighboring genes[67]. With the number of identified lncRNAs constantly increasing, it becomes necessary to explore their plausible biological functions.

LncRNAs as regulators of developmental processes

A number of lncRNAs have been reported in developmental processes of multicellular organisms. In the following section, I will discuss some of these developmental processes which have served as model system for the understanding of lncRNAs mediated gene regulation.

Genomic Imprinting and ncRNA

Genomic imprinting is an epigenetic process that causes mono-allelic expression of genes in a parent of origin-specific manner. Such genes are referred as ‘imprinted’. Some of these genes are expressed only from the paternal allele, and the others are expressed from the maternal allele. Genomic imprinting was observed by researchers in plants, insects and mammals around 50 years ago. Historically, the term ‘imprinting’ was first coined by Helen Crouse in 1960s during her studies in *Sciara coprophila*[68]. Based on her observation that there is a selective loss of paternally inherited, one or two X chromosomes in females and males respectively, She mentions in her article that—

"the ‘imprint’ a chromosome bears is unrelated to the genic constitution of the chromosome and is determined only by the sex of the germ line through which the chromosome has been inherited."

Since then, subsequent genetic studies performed in mouse over the years have established genomic imprinting as a paradigm for epigenetic research. Currently, there are 150 imprinted genes reported in mouse and humans. The characteristic feature associated with imprinted genes is that they are mostly present in clusters. The presence of these genes in cluster indicates towards presence of coordinated regulation of imprinted genes in a given imprinted locus. Indeed, these imprinted clusters are under control of a *cis*-acting control element known as the imprint control region (ICR). The ICR is methylated on one of the parental alleles, whereas the other paternal allele remains unmethylated and hence the term differentially methylated region or DMR. This DNA methylation ‘imprint’ is established by *denovo* methyltransferase DNMT3A/3L during germ cell development at these ICRs and subsequently maintained by maintenance methyltransferase DNMT1 in diploid cells. The methylation imprint at ICRs is erased in primordial germ cells in the devel-
oping embryo and a new methylation mark is re-established based on the sex of the organism.

A distinct feature associated with the majority of imprinted gene clusters is, the presence of one or more lncRNAs. In this respect, imprinted clusters provide us with an excellent model system for understanding the mechanistic aspects of lncRNA mediated gene regulation. However, mono-allelic silencing occurs by different mechanisms amongst these clusters. ICR can acts as an insulator and control the access of imprinted genes to the regulatory elements ($H19/Igf2$ locus) or, act as a promoter for a lncRNA which is involved in silencing of the imprinted genes ($Kcnq1& Igf2r$ locus). We will discuss in detail about some well-studied imprinted clusters and how nicely they have contributed in our understanding of functional roles played by lncRNAs in regulating gene expression—

$H19/Igf2$ locus: The $H19/Igf2$ locus resides on mouse chromosome 7, and the orthologous region is present on chromosome 11 in humans. Insulin growth factor 2 ($Igf2$) and $H19$ were the initial identified imprinted genes in mammals during early 90s. $Igf2$ is expressed from the paternal allele and is present 90kb upstream of the maternally expressed $H19$ gene. Both $H19$ and $Igf2$ share common tissue-specific enhancers located downstream of the $H19$ gene. The ICR regulating mono-allelic expression of both these genes is present upstream of the $H19$ gene. The ICR is maternally unmethylated, and contains multiple binding sites for the insulator protein CTCF. Binding of CTCF at the ICR blocks the $Igf2$ promoter from accessing the downstream enhancers and thus keeps $Igf2$ maternally repressed. This, in turn, allows the maternally expressed $H19$ to access the downstream enhancers. On the paternal allele, the ICR is methylated and prevents the binding of CTCF. In the absence of insulator protein CTCF, $Igf2$ can access the downstream enhancers and express from the paternally derived allele[69, 70]. Later studies, using 3C approach have validated the presence of higher order chromatin interactions (looping) between the promoters and enhancers which are influenced by the presence of parent specific differential methylation and CTCF binding[71].
Figure 2: Model depicting the insulator based regulation at the H19/Igf2 locus. The ICR is unmethylated on the maternal allele and allows binding of CTCF, preventing access of enhancer to Igf2. On the paternal allele, binding of CTCF is prevented due to methylation of ICR and therefore allows Igf2 to interact with downstream enhancers.

Kcnq1 locus: The Kcnq1 imprinted locus is present on chromosome 7 in mouse and, the orthologous region is present on chromosome 11 in humans. It is constituted of 8-10 maternally expressed protein coding genes and an lncRNA, Kcnq1ot1. The ICR at this locus is located in the intron 10 of the Kcnq1 gene where it serves as promoter for the expression of lncRNA. Kcnq1ot1 RNA is 91Kb, nuclear localized, unspliced and expressed only from the paternal allele as its promoter is unmethylated, whereas it is repressed on the maternal allele due to methylation of its promoter[42, 72]. The imprinted genes at the Kcnq1 locus show varied gene repression in different tissue types and developmental stages. Based on the extent of gene repression, the imprinted genes at this locus can be categorized into three distinct classes; ubiquitously imprinted genes (UIGs) which are repressed in both embryonic and extra-embryonic lineages, placental imprinted genes (PIGs) which show repression only in the placental tissue, and non-imprinted genes (NIGs) which escape the RNA mediated paternal silencing at the Kcnq1 locus.

The paternal expression of Kcnq1ot1 RNA and silencing of imprinted genes on the paternal allele indicate towards a role of Kcnq1ot1 in silencing of its neighboring genes in cis. Deletion, truncation and destabilization studies on the Kcnq1ot1 lncRNA in mouse and cell based systems have shown that Kcnq1ot1 RNA is required for the allelic silencing of the protein coding genes in cis[42, 73, 74]. Furthermore, using an episomal system, a 890bp functional sequence, crucial for its silencing has been reported which underscores the role of RNA in gene silencing[75].

The Kcnq1ot1 adopts distinct mechanisms to bring about silencing of the imprinted genes at the Kcnq1 locus[9]. Silencing of PIGs has been observed to be regulated primarily by histone modifications. Mouse lacking histone modifiers such as G9a, Ezh2 and Eed have an effect on the silencing of PIGs. Furthermore, the Kcnq1ot1 lncRNA has been demonstrated to interact
with chromatin modifying enzymes such as G9a and Ezh2 and recruit them to the gene promoters of the imprinted genes in cis in a lineage specific manner. Unlike PIGs, DNA methylation is reported to be involved in silencing of UIGs. Moreover, the lncRNA Kcnq1ot1 has been observed to interact with maintenance methyltransferase DNMT1, and recruit it to somatic DMRs of UIGs, Cdkn1c and Slc22a18[37].

Taken together it has been suggested that, the lncRNA Kcnq1ot1 through its association with DNA and chromatin modifying enzymes is involved in regulating multilayered silencing mechanisms operational at the Kcnq1 imprinted locus.

**Figure 3:** Schematic representation of the imprinted loci (A) Kcnq1 locus and (B) Igf2r locus. The Kcnq1 locus is constituted of paternally expressed RNA, Kcnq1ot1 which is involved in silencing of the neighboring genes. Genes colored in red are imprinted in all tissues(UIGs), blue in only placental tissue (PIGs) and green are non-imprinted genes(NIGs). (B) The Igf2r locus contains a paternally expressed lncRNA, Air. The genes colored in red represent the imprinted genes whereas genes colored in green represent non-imprinted genes.

**The Igf2r locus-** The Igf2r locus is present on chromosome 17 in mouse. An unspliced, polyadenylated, 108Kb lncRNA, Air is transcribed from this locus. Similar to the Kcnq1 locus, the ICR is unmethylated on the paternal allele and serves as a promoter for the expression of Air lncRNA. The paternal expression of Air lncRNA correlates with the silencing of its three neighboring genes Igf2r, Sl22a2 and Slc22a3.

Studies using deletion of the ICR or the paternally inherited truncated allele of the Air transcript have directly implicated the RNA or its transcription in gene repression over the Igf2r locus[76, 77]. Furthermore, Air lncRNA has been shown to interact with H3K9 methyltransferase G9a and specifical-
ly target it to the promoter of Slc22a3 gene in a lineage specific manner[43]. Therefore, similar to Kcnq1ot1 RNA, Air too is involved in epigenetic silencing by targeting repressive histone machinery to specific gene promoters in a lineage specific manner.

ncRNAs and Dosage compensation

Sex of an organism is determined by the composition of the sex chromosome, such as XO or XY in males and XX in females. Distinct regulatory mechanisms are present, to balance the dosage of X-linked gene products in cells of either sex. This process of dosage compensation is well studied in flies (Drosophila) and mammals (mouse) and serves as an excellent example, where significant insights into regulatory roles of lncRNAs have been elucidated. We will discuss these roles in greater detail in the following section.

ncRNAs in mammalian X -inactivation— In mammals, there are two X chromosomes (XX) in females as compared to one in males (XY). Therefore, X –chromosome inactivation (XCI) takes place in females to equalize the dosage of X-linked gene products between both the sexes. During female development, XCI occurs in two forms: imprinted XCI, observed in pre-implantation embryo where paternally inherited X chromosome (Xp) is preferentially silenced, and random XCI, observed in post-implantation epiblast lineage where either X chromosome is silenced randomly.

Studies on X chromosome have postulated the presence of a X inactivation center (Xic), which is essential to trigger the XCI. This region is around 500Kb in size and harbors five lncRNAs known to regulate X inactivation. The most well-known lncRNA transcribed from the Xic is Xist. It is a 17kb, nuclear localized, spliced transcript and expressed exclusively from the inactive X chromosome (Xi)[78]. It coats the entire Xi chromosome in cis- from which it is expressed and its upregulation is necessary for the onset of XCI. The Xist lncRNA contains several repeats of which repeat A has been shown to be crucial for its silencing activity. Interestingly, a shorter RNA, ‘RepA’ (Repeat A RNA), around 1.6Kb in size is transcribed from the repeat A region of Xist and is believed to interact and recruit the PRC2 complex to Xi[44]. The repressive epigenetic memory once initiated by Xist expression, can induce several changes on the Xi chromatin by recruitment of factors such as PcG proteins, Ash2I, hnRNP U (SAF-A), macroH2a and epigenetic modifications such as histoneH4 hypoacetylation and DNA methylation, which are believed to contribute in faithful maintenance of the repressed state of Xi[40, 79-81].

An intriguing question in relation to XCI is, whether Xist lncRNA is also required for maintenance of the repressed state of Xi during development or its role is limited only in establishment of XCI? Using mouse/human somatic cell hybrids, Brown et al, observed that, X-linked genes remained silent in
somatic hybrids lacking the Xic consisting of Xist RNA [82]. Later, Csankovszki et al noted that, conditionally deleting Xist after XCI in MEFs did not reactivate the X-linked genes[83]. These observations suggest that, Xist expression is dispensable for maintenance of XCI. However, the continued expression and association of Xist RNA with Xi in somatic cells of females during entire lifespan make us to revisit its role during maintenance. In fact, the lack of Xist causes a decrease in levels of histone macroH2A1.2, characteristic of Xi in conditionally mutant Xi cells. Reactivation of two genes; GFP (transgene) and Hprt (endogenous) due to destabilization of inactive state of X chromosome was also reported in conditional allele lacking Xist[84]. Furthermore, it resulted in an increase in the active chromatin marks; histone H4 acetylation and H3K4me3 throughout Xi[85, 86]. In a recent study in mouse hematopoietic stem cells, the conditional deletion of Xist at E10.5 (after establishment of XCI) resulted in upregulation of 86 X-linked genes. Collectively, it can be summarized that, Xist is indispensable for the establishment of XCI whereas it could be involved in maintaining specific features of the Xi chromatin.

The Xic harbors another nuclear localized lncRNA, Tsix, with antagonistic function to Xist. It is transcribed in the antisense orientation to Xist and its deletion in the mouse ES cells results in activation of Xist in cis. Therefore, in combination with Xist, Tsix is involved in regulating XCI by maintaining the repression of Xist on the active X chromosome. Another ncRNA, Xite is located upstream of the Tsix promoter and is believed to positively regulate Tsix[86].

Another important and unique aspect in XCI is positive and negative regulation of the Xist by lncRNAs. For example, Tsix functions as a repressor of Xist on the Xa, another lncRNA upstream of the Xist, Jpx acts as a trans-activator of the Xist expression on the Xi. A recent investigation has demonstrated that Jpx regulates Xist activation by binding and titrating away the CTCF protein from the Xist promoter[87].

**roX RNAs and dosage compensation in Drosophila melanogaster** — Unlike mammals, dosage compensation in Drosophila is achieved by hyperactivation of the single X chromosome in males. Although the ways to achieve dosage compensation are exactly opposite to the mammals, nevertheless lncRNAs act as a common theme regulating dosage compensation in mammals and Drosophila. Two lncRNAs; RNA on X1 and 2(roX1 and roX2) are known to be transcribed from the X chromosome. These two RNAs along with male specific proteins (male specific lethal proteins-MSL1, MSL2 and MSL3; maleless helicase-MLE; and histone acetyltransferase males absent on the first-MOF) form the dosage compensation complex (DCC) which is spread across the X chromosome. This coating of the DCC complex acetylates the H4K16, resulting in an open chromatin and thus brings about an increased transcriptional output of the single X-chromosome in males[88].
How these two lncRNAs interact and function with the MSL complex to bring about the X chromosome activation has remained enigmatic? Two recent studies have shed light on the contributions of these two lncRNAs. Using CLIP or RNA affinity purification strategy, critical binding sites for MLE helicase on roX RNA have been reported. The MLE helicase once bound to the conserved motifs on roX RNAs leads to a conformational change in the interacting roX RNAs. This leads to an unmasking of binding sites for MSL2 subunit and thus triggers DCC assembly[89, 90]. Taken together, roX RNAs present an interesting RNA based mechanism of gene regulation.

Other roles of lncRNAs in animal development

In addition to classically studied functions of lncRNAs in genomic imprinting and dosage compensation, several examples of lncRNAs and their roles in body pattern formation and cellular differentiation have come into light in recent years. I would discuss briefly the regulatory roles of lncRNAs in these developmental processes.

lncRNAs in Homeobox (Hox) gene regulation: Hox genes are a group of related genes organized in clusters and are crucial for early embryonic development in animals. There are 39 Hox genes in vertebrates, clustered into four chromosomal loci (HoxA-HoxD) that are involved in regulating the anterio-posterior axis formation during embryogenesis. The differential expression of Hox genes is under regulation of epigenetic modifiers such as Trithorax (Trx) and PcG group of proteins. In addition, several lncRNAs have been reported at these loci, some of which have been functionally implicated in repressing and activating target Hox genes. The first example is, HOTAIR, a 2.2Kb transcript which is expressed from the HoxC cluster is involved in gene repression of the HoxD locus in trans- in mammalian cells. It has been shown that, the HOTAIR lncRNA interacts with PRC2 complex at its 5’ end and with LSD1, a H3K4 demethylase at its 3’ end. Therefore, the HOTAIR lncRNA serves as a scaffold for assembly of both these complexes and this property allows the lncRNA to regulate the targeting of these complexes to silence target HoxD genes[45, 91].

Other lncRNAs at the Hox locus, HOTTIP and Mistral have been reported to activate transcription of Hox genes[92, 93]. HOTTIP is a 3.7 Kb transcript and expressed from the anatomical distal region ie 5’ tip of the HOXA locus. HOTTIP has been shown to activate the transcription of HoxA genes by associating with the WDR5-MLL complex and targeting it to these genes. In mouse ES cells, another lncRNA, Mistral is involved in epigenetic activation of homeotic genes Hoxa6 and Hoxa7. This unspliced and polyadenylated 798 nt RNA interacts with WDR5-MLL1 complex resulting in a changed chromosomal conformation and hence activation of its target genes.
LncRNA in pluripotency maintenance and cellular differentiation: ES cells are pluripotent cells with unique ability to differentiate into all the lineages derived from any of the three primary germ layer cells (ectoderm, mesoderm and endoderm). This pluripotent state of ES cells is tightly regulated and maintained by key transcription factors, such as Oct4, Sox2 and Nanog. In addition to these well-known transcription factors, several lncRNAs have been implicated in maintenance of pluripotency. Using a loss-of-function approach, Guttman et al have tested the role of 147 lncRNAs expressed in the mouse ES cells[94]. They found 26 lncRNA with roles in the maintenance of pluripotency, whereas expression of 30 lncRNAs was crucial for the repression of lineage specific programs associated with the differentiation. A significant number of these RNAs are under the regulation of ES cell specific transcription factors such as Oct4, Sox2 and Nanog, and many of them associate with diverse chromatin modifiers, potentially revealing their role in regulating gene expression programs by forming specific RNA-protein complexes.

In addition to maintenance of pluripotent state of ES cells, many lncRNAs have been reported to be crucial for differentiation into specific lineages. Ng et al have identified 4 lncRNAs crucial for neuronal differentiation, through their association with proteins such as SUZ12 and REST[95]. One of these RNAs, RMST has been reported to interact with SOX2 and this RMST-SOX2 complex regulates many downstream genes during neurogenesis[96]. On a similar line, a 590nt lncRNA, Braveheart (Bvht) was found to be crucial for mouse ES cells differentiation towards cardiovascular lineage. Bvht epigenetically regulates cardiovascular lineage commitment through its association with PRC2 member, SUZ12[97]. All these mentioned lncRNAs having roles during lineage commitment act as molecular scaffold for protein complexes, however other examples of lncRNAs employing different regulatory mechanisms during differentiation have also been characterized. A cytoplasmic lncRNA, linc-MD1, controls muscle differentiation in mouse and human cells by acting as a sponge for micro RNAs miR-133 and miR-135. Overexpression of linc-MD1 up regulates transcription factors, MAML1 and MEF2C, known to activate genes related to muscle differentiation. These transcription factors are targets of miR-133 and miR-135 and thus linc-MD1 acts as a decoy for these microRNAs and contributes in regulation of muscle differentiation[98]. Another lncRNA, TINCR (terminally differentiation-induced ncRNA) is essential for human epidermal differentiation. Analyzing TINCR RNA- and protein interactome revealed that TINCR interacts with mRNAs of genes involved in differentiation (through a 25 nucleotide ‘TINCR motif’) and staufen (STAU1) protein. It is proposed that, TINCR stabilizes expression of mRNAs essential for epidermal differentiation, through its association with STAU1[49].
With all these examples, it is evident that lncRNAs are an integral component of regulatory circuit during embryonic development and cellular differentiation.

LncRNAs in disease

With several examples of lncRNAs functions discussed in earlier section, it is clear that these molecules play a significant role in mammalian developmental processes. In this section I would discuss the examples of lncRNAs which have been reported to be misregulated in disease conditions. Also, I would discuss how understanding these diseases in the context of lncRNA expression are contributing to our understanding of lncRNA mediated gene regulation.

**ANRIL lncRNA**

This lncRNA was first reported from genetic analysis of familial melanoma-neural system tumors with large germ line deletion (42Kb) encompassing INK4 locus\[99\]. This locus harbors three tumors suppressor genes; *p15/CDKN2B/INK4b*, *p16/CDKN2A/INK4a*, *p14/ARF* along with antisense lncRNA ANRIL (antisense RNA in INK4 locus). Altered expression of these genes has been observed in many tumors. Additionally, the regions encompassing ANRIL gene contains disease associated SNPs\[100, 101\].

**ANRIL** lncRNA is transcribed in an antisense direction to INK4b. It has been reported that ANRIL interacts with chromobox 7 (CBX7) protein, a member of PRC1 complex, and controls the transcriptional repression of the INK4a tumor suppressor genes. Increased levels of both ANRIL and CBX7 have been reported in prostate cancer correlating with reduced levels of INK4a. Another study observed that, ANRIL interacts with SUZ12 and recruits this PRC2 member to repress INK4b expression. This suggests that ANRIL is involved in regulating epigenetic repression of INK4 locus through its association with both PRC1 and PRC2 members. These observations also indicate that, altered expression of ANRIL is involved in tumor initiation via silencing the three tumor suppressor genes. However the detailed role of ANRIL in INK4 locus regulation and mechanisms by which it initiates cancer progression need to be fully elucidated\[46, 102\].

**HOTAIR lncRNA**

*HOTAIR* is a very well-known example of lncRNA involved in cancer progression. Gupta et al first reported that increased expression of HOTAIR was associated with primary breast cancer and metastasis. Over expression of HOTAIR lncRNA resulted in an increased invasiveness and metastasis in
epithelial cancer cells. This increased expression of *HOTAIR* lncRNA leads to modified occupancy of PRC2 complex to many metastasis suppressor genes resulting in their aggressive phenotype. Conversely, depletion of *HOTAIR* decreased the invasive potential of the MCF7 cell line. In addition to breast cancer, increased expression of *HOTAIR* has been reported in several other cancers such as adenocarcinoma, melanoma, squamous cell carcinoma, hepatocellular carcinoma etc. It is therefore believed that *HOTAIR* contributes to cancer progression and metastasis by remodeling the chromatin landscape through its interaction with the PRC2 complex\[103\].

**MALAT1** lncRNA

*MALAT1* was initially reported as a prognostic parameter associated with patient survival and metastasis in stage I non-small cell lung cancers\[104\]. Since its identification several studies have reported elevated levels of *MALAT1* in cancers of liver, breast, cervix and bladder and linked its differential expression to cancer recurrence and metastasis. In addition, overexpression experiments observed that increased levels of *MALAT1* resulted in increased cell proliferation, migration and tumor formation in nude mice and conversely, depleted levels of *MALAT1* resulted in impaired migration and reduced tumor formation in mouse xenografts\[105\]. The mechanisms behind these effects of *MALAT1* and its role in cancer can be partly understood by its regulatory functions.

It is a highly expressed, conserved nuclear RNA and is localized to nuclear speckles, although it is not crucial for the integrity of these nuclear speckles. *MALAT1* has been shown to interact with pre-splicing factors (SR proteins) and regulate their level and distribution in nuclear speckles domains, known to be enriched for pre-mRNA splicing factors\[106\]. In extension to this observation, Tripathi *et al* have recently shown that *MALAT1* promotes cell proliferation by modulating levels of cell-cycle regulated oncogenic transcription factors such as *B-MYB*\[107\]. It is therefore proposed that *MALAT1* acts as molecular sponge and impairs the levels of pre-mRNA splicing factors leading to misregulation of cell cycle progression genes and thus contributes in tumor progression.

LncRNAs in prostate cancer

*PCA3* is one of the initially reported lncRNA observed to be highly expressed in primary prostate cancers. Recently, FDA has approved the use of *PCA3* for prostate cancer detection\[100, 108\]. After its identification many other lncRNAs in prostate cancer have been reported. A study using transcriptomic profiling approach has sequenced RNA from a cohort of 102 prostate tissue and cell lines, and identified 121 unannotated prostate cancer associated transcripts (*PCATs*). Functional characterization of one of these
LncRNAs, PCAT-1 revealed that this lncRNA is selectively unregulated in prostate cancer[109]. It is involved in repression of genes involved in mitosis and cell division such as BRCA2, a known tumor suppressor gene. Another recent study has identified another lncRNA, PCAT-114 or also named as second chromosome locus associated with prostate-1 (SchLAP1) as a strong predictor of clinical outcomes in aggressive prostate cancer. Over expression and depletion experiments suggests that, SchLAP1 contributes to cancer cells invasiveness and metastasis. It was observed that SchLAP1 antagonizes SWI/SNF complex and in turn impairs its genome wide occupany and regulatory functions. These observations indicate that SCHLAP1 promotes aggressive cancer in part by regulating downstream targets of SWI/SNF complex[110]. In addition, two lncRNAs PRNCR1 and PCGEM1 have been recently reported to be highly expressed in aggressive prostate cancer. Both these lncRNAs bind to androgen receptors (AR) and enhance AR mediated gene expression programs resulting in an increased proliferation and metastasis in prostate cancers[111].

LncRNAs in neuroblastoma

Neuroblastoma is a disease of sympathetic nervous system affecting children below the age of 1.5 years. These tumors occur generally in adrenal medulla but can occur in other organs such as chest, abdomen and pelvis. Neuroblastomas are highly heterogeneous tumors with variable clinical outcomes and hence can be classified into various subtypes. However, in general these tumors can be divided into two broad categories- Low-risk tumors and high-risk tumors. Low-risk tumors are characterized by whole chromosomal changes, favorable prognosis and have the propensity to undergo regression. On the other hand high-risk tumors have characteristic chromosomal aberrations such as deletion of 11q and 1p arm, gain of 17q and amplification of MYCN oncogene, unfavorable prognosis and are highly aggressive. Many studies have reported mutations, epigenetic alterations and expression signatures for efficient stratification of these tumors[112].

Few lncRNAs have also been reported to be associated with neuroblastoma. Transcribed ultra-conserved regions or T-UCRs are lncRNAs, transcribed from 481 regions, which were found to be highly conserved between human, mouse and rat and have been found to be dysregulated in tumors such as hepatocellular carcinoma and leukemia[113]. In a study, expression of all 481 T-UCRs was investigated in primary neuroblastomas. 7 T-UCRs were found to be overexpressed in MYCN amplified tumors. In addition, three of these lncRNAs were found to be upregulated upon MYCN induction, pointing towards their MYCN dependent regulation[114]. Another lncRNA, ncRAN (noncoding RNA expressed in aggressive neuroblastomas) has been reported in the aggressive neuroblastoma associated 17q region[115]. Depletion of ncRAN in neuroblastoma cell lines inhibited cell growth. Also, over-
expression of \textit{ncRAN} in bladder cancer cell line has been reported to result in increased cell proliferation and invasiveness. However, the mechanisms and exact roles of these \textit{T-UCRs} and \textit{ncRAN} in neuroblastoma progression are largely unknown. Therefore, identifying and functionally characterization lncRNA signatures in neuroblastoma would lead to a much improved understanding of the disease.

\textbf{Mechanisms of LncRNA mediated gene regulation}

In the current thesis, I have discussed so far various examples of lncRNA in different aspects of development and maintenance of cellular states. In addition, I have provided examples of misregulated lncRNAs in diseases. In this section of the thesis, I will discuss various mechanisms adopted by lncRNA for regulating gene expression.

1. \textbf{Recruitment of Chromatin modifying enzymes}

One of the most well-known mechanisms adopted by lncRNAs is through their ability to interact with chromatin modifying enzymes and recruit these epigenetic modifiers to target gene loci in \textit{cis} or guide these modifiers to distant loci in \textit{trans}. \textit{Cis} regulatory RNAs such as \textit{Xist/RepA}, \textit{Kcnq1ot1}, and \textit{ANRIL} have been observed to bind to repressive complex PRC2 to regulate gene expression in a localized manner. The recruitment of these complexes could be through putative protein binding domains/motifs in the RNA. A GC rich double stem loop structure (28 nt) in \textit{Rep A} lncRNA directly interacts with EZH2\cite{44}. Such structural motifs or functional domains in the RNA could serve as scaffolds for assembly of protein complexes and facilitate them to efficiently target the repressive gene machinery. For example, the \textit{Kcnq1ot1} RNA is reported to interact with both PRC2 & G9a as well as with the maintenance methyltrasferase, DNMT1 and regulate lineage specific gene repression via recruiting these complexes to specific gene promoters and somatic DMRs respectively\cite{37, 42}. Germ line deletion of 890 bp silencing domain, present at its 5’ end significantly reduces the chromatin association and interaction of the \textit{Kcnq1ot1} RNA with DNMT1, resulting in loss of gene repression. The structural conformations/ binding motifs which recruit and the mechanisms through which lncRNAs stabilize the assembly of these protein complexes are yet to be fully characterized in mammals. In relation to this aspect of lncRNA mechanism, an interesting example of lncRNA recruitment and protein assembly has been recently reported in Drosophila, where binding of one protein factor to conserved motif in \textit{roX RNA} leads to change in structural confirmation of the interacting RNA and triggers assembly of other proteins\cite{89, 90}.
In contrast to the local gene repression by cis regulating RNAs, the trans regulatory RNAs such as \textit{HOTAIR} and \textit{linc-p21} might adopt additional mechanism to regulate distant gene repression. It is proposed that, they can target the repressive gene machinery in trans by binding to the specific genes as RNA: DNA duplexes or RNA: DNA: triplexes\cite{116}. For example, recruitment of Dnmt3b by promoter RNA (pRNAs) generated from the intergenic spacer regions (IGS) of the rDNA repeats is proposed to be dependent on the DNA: RNA triplex between pRNA and rDNA promoter\cite{41}.

\textbf{Figure 4:} Schematic representing mechanisms of lncRNA mediated gene regulation. LncRNA can act as Scaffold for chromatin modifying enzymes(A), regulators of signaling pathways of tumor suppressor genes(B), sponges for microRNAs (C), decoys of transcription factors(D) and can regulate mRNA processing, translation and stability(E).
2. Regulation of tumor suppressor or oncogene signaling

LncRNAs can play critical regulatory roles in disease states through regulating expression of tumor suppressor genes or their downstream signaling. For example, gene repression of tumor suppressor genes, p15 and p16 is observed in many cancers and is associated with ANRIL mediated targeting of the repressive chromatin machinery to these target tumor suppressors [102, 117]. Another lncRNA, linc-p21 has emerged as a crucial p53 response gene. In mouse lung, sarcoma and lymphoma tumors, linc-p21 was observed to be induced upon activation of p53 signaling and transcriptionally repress p53 target genes through its association with heterogeneous ribonucleoprotein K (hnRNP-K)[118]. Similarly, an lncRNA PANDA has been reported to be overexpressed in metastatic ductal carcinoma and is activated in a p53 dependent manner. It limits expression of pro-apoptotic genes by sequestering away the transcription factor NF-Y from promoters of these genes[119].

3. Sequestration of microRNA

Several lncRNAs have micro-RNA binding sites and thus can serve as ‘sponges’ of these molecules and regulate their levels during cellular processes. An interesting ‘competing endogenous RNA (ceRNA)’ hypothesis has been proposed, where an intricate network of cross talk between coding and noncoding RNAs is present, and changes in levels of one microRNAs can lead to concomitant changes in expression of other member of the network[120]. For example, H19 lncRNA has binding sites for let-7 microRNA and regulates muscle differentiation by regulating levels of let-7[121]. Similarly, linc-MD1 lncRNA regulates levels of transcription factors, MAML1 and MEF2C involved in muscle differentiation by acting as a sponge for micro-RNAs, mir-133 and mir-135[98]. Along with lncRNAs, pseudogenes such as PTENP1 can also act as sponges for microRNAs and in turn regulate the expression of their target genes[122].

4. Regulators of mRNA processing, stability and translation

LncRNAs can also exert their regulatory control through their roles at the level of mRNA processing, stability or translational control. As discussed earlier, lncRNA MALAT1 controls mRNA splicing by regulating the levels of pre-mRNA splicing (SR) factor. Similarly, a relatively new class of lncRNA, sno-lncRNAs is transcribed from the intronic region of 15q11-q13 (Prader-willy syndrome region) locus. This too can influence splicing events through its association with the splicing factor FOXC2 [123]. In addition, lncRNAs can influence stability of transcripts through their interaction with target mRNAs. It has been observed that Alu-repeat containing cytoplasmic lncRNAs can interact with Alu elements in the 3’ UTR of a mRNA and
these double stranded structures formed through RNA-RNA interaction can create binding sites for Staufen-mediated decay (SMD) and in turn activate the mRNA decay pathway[124]. Another important aspect of lncRNA mediated regulation is through translational control. Antisense lncRNAs like Uch1As have been observed to upregulate the translation of its sense mRNA, Uch1 through a repeat SINEB2 element. Several other sense–antisense lncRNAs with SINEB2 elements have been identified and translation control could be employed by such RNAs[125].

5. Decoys of transcription factors
An interesting theme in lncRNA mediated gene regulation is through acting as ‘molecular decoys’ of transcription factors or chromatin modifiers or other protein factors. For example, PANDA lncRNA regulates the expression of pro-apoptotic gene through titrating away transcription factor NF-Y [119]. Similarly, MALAT1 acts as a decoy for the SR proteins and regulates splicing.
Aims

The main objective of the current thesis was to understand the role of lncRNA in development and disease. Using Kcnq1 imprinted locus as a model system, we have investigated the role of Kcnq1ot1 RNA in establishment and maintenance of the epigenetic memory. Next, using genome wide approaches and pediatric tumors as disease model, we have identified regulatory RNAs in normal and disease condition. The specific aims of the thesis are as follows-

Paper I
1. To functionally characterize the critical sequences within Kcnq1ot1 RNA required for epigenetic silencing of neighboring genes, in vivo.

Paper II
1. To study the role of Kcnq1ot1 RNA in the maintenance of transcriptional gene silencing at different stages of mouse embryonic development.
2. To investigate the plausible mechanism by which non-imprinted genes escape noncoding RNA mediated silencing.

Paper III
1. Identification and characterization of Chromatin associated RNAs (CARs) from human fibroblast cells.

Paper IV
1. Identification of differentially expressed lncRNAs between low- and high-risk neuroblastomas.
2. To functionally characterize the regulatory roles of these identified lncRNAs in different subtypes of neuroblastomas.
Results and Discussion

Paper I: *Kcnqlot1* noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1.

*Kcnqlot1* is an IncRNA paternally expressed from the intron 10 of the *Kcnql* gene. Earlier studies, using cell based episomal system and mouse knockouts have demonstrated that *Kcnqlot1* is involved in bidirectional, epigenetic silencing at the *Kcnql* locus, in a lineage specific manner. Further attempts have been made to identify and characterize the crucial sequences in the *Kcnqlot1* RNA, necessary for its silencing activity. A cell based episomal system was used to search for functional sequences within the noncoding RNA. The episomal system comprises of two reporter genes Hygromycin (*Hyg*) and *H19*. In this construct the *Hyg* gene is located upstream to the promoter of *Kcnqlot1* whereas the *H19* gene is present downstream of the *Kcnqlot1* transcription start site. The *Kcnqlot1* RNA upon transcription is involved in bidirectional silencing of the two reporter genes. During identification of the functional sequences within the RNA, it was found that deletion of 890bp from the RNA resulted in the loss of bidirectional silencing of both the reporter genes within the episomal system. Further characterization of the 890bp sequence demonstrated that, it could silence the reporter genes in an orientation dependent and position independent manner.

In the current study, we wanted to investigate the role of 890bp present at the 5’ end of the *Kcnqlot1* on the silencing status of imprinted genes at the *Kcnql* locus, in an *in vivo* condition. Towards this end, we have generated a knockout mouse lacking the critical 890 base pairs through homologous recombination. Homozygous mice lacking 890bp (Δ890 -/-) were normal and fertile. In addition, no effect on the stability, localization, expression and promoter methylation of *Kcnqlot1* RNA was observed in Δ890 -/-. In order to analyze the effect of 890 bp deletion on imprinted genes, we performed allele-specific imprinting assays on the RNAs collected from placenta (extra-embryonic) and liver (embryonic) tissues of mouse embryos. We observed a loss of silencing of the ubiquitously imprinted genes (UIGs-*Kcnql*, *Slc22a18*, *Phlda2*, *Cdkn1c*) in the paternally derived *Kcnqlot1* allele lacking 890bp. However, the UIGs showed variations in their loss of imprinting in a lineage specific manner. *Kcnql* showed a loss of imprinting in both placenta
and liver, Phlda2 and Cdkn1c showed loss of imprinting only in placental tissue whereas Slc22a18 showed loss of imprinting only in liver tissue. In addition, we did not observe any changes on the imprinting status of the placental imprinted genes (PIGs-Ascl2, Tssc4, Cd81 and Osbpl5). Taken together, our results suggest that 890bp of the Kcnq1ot1 RNA are essential for the proper imprinting of UIGs.

The UIGs are known to be regulated via DNA methylation. Mouse lacking the maintenance methyltransferase, DNMT1 shows a loss of silencing of UIGs. In order to investigate the reasons behind loss of silencing of UIGs in ∆890 mice, we analyzed the methylation status of somatic DMRs of Cdkn1c and Slc22a18 genes. We observed a loss of methylation at the DMRs of both the genes in placenta as well as liver tissues indicating that methylation could be important for the silencing at these UIGs. To further investigate, the role of RNA in regulating the methylation, we performed RNA immunoprecipitation (RIP) with DNMT1 antibody. We found that RNA interacts with the maintenance methyltrasferase, DNMT1 and this interaction was significantly reduced in mice lacking 890bp, explaining the loss of methylation observed at the somatic DMRs of Cdkn1c and Slc22a18 genes. We also found significant enrichment of DNMT1 at the somatic DMRs of Cdkn1c and Slc22a18, which was reduced in ∆890 mice.

Kcnq1ot1 is known to associate with chromatin and it mediates its transcriptional silencing by organizing a three dimensional, contracted repressive compartment. We analyzed the effect of deletion of 890bp on these properties of the noncoding RNA using Chromatin RNA immunoprecipitation (ChRIP) and 3C. We observed a reduced chromatin association of the RNA and release in the contracted state of the repressive compartment in mice lacking 890bp. This effect on the higher order chromatin structure in ∆890 mice could explain the loss of silencing of Kcnq1 and Phlda2 gene which are not known to have any DMRs.

In summary, the current investigation elucidates that, 890bp of the Kcnq1ot1 RNA is required for RNA mediated transcriptional silencing of the UIGs, through its association with the DNA methylation machinery.

Paper II: Long noncoding RNA mediated maintenance of DNA methylation and transcriptional gene silencing.

In Paper I we have investigated the role of RNA in establishment of epigenetic silencing. The Kcnq1ot1 IncRNA starts expressing from 2 cell stage onwards, however the gene silencing at the Kcnq1 locus occurs in a lineage and developmental stage specific manner. The imprinting of UIGs is established at the blastocyst stage (E3.5) whereas PIGs start showing imprinted expression from the post-implantation stages (E5.5). In addition, there are
non-imprinted genes (NIGs) in this locus which remain biallelically expressed in all developmental stages. The Kcnq1ot1 RNA has been shown to associate with DNA and chromatin modifiers and establish multilayered silencing pathways at the Kcnq1 locus. However, the role of IncRNA in maintenance of epigenetic memory once established has remained mainly unknown. Another intriguing question is how certain genes escape transcriptional gene silencing by noncoding RNAs?

In the current investigation, using conditional knockout mouse, we have attempted to answer these questions. We have generated a transgenic mouse where the Kcnq1 ICR was flanked with two loxP sequences (ICR^{2lox}). We have deleted the Kcnq1ot1 RNA at specific embryonic developmental stages by crossing the ICR^{2lox} mice with two different mouse lines expressing, the cre recombinase enzyme. Firstly, we crossed the male ICR^{2lox} mice with Meox cre/+ females, that expressed cre recombinase only in the epiblast lineages which give rise to embryonic tissue. Conditional deletion of the Kcnq1ot1 at E 5.5 resulted in a loss of silencing of UIGs in the embryonic lineages (liver, brain and heart) suggesting that RNA is required to maintain the imprinting of UIGs. In line with earlier observation, we observed reduced weight in mice lacking Kcnq1ot1 RNA explained the increased expression of Cdkn1c. Furthermore, we wanted to investigate whether RNA is also required to maintain silencing at later stages of development or there are developmental stage specific differences in the requirement of RNA in maintenance of epigenetic memory. Therefore, we crossed the male ICR^{2lox} mice with CAG cre/+ female mice and fed the pregnant female with 5mg of tamoxifen at E8.5d to conditionally delete the RNA in both embryonic and extra-embryonic lineages. Conditionally deleting the RNA at later stages of development resulted in loss of UIGs, however the silencing of PIGs remained unchanged.

As ubiquitously imprinted genes such as Cdkn1c and Slc2218 are known to have somatic DMRs, we analyzed the status of methylation over these genes. A loss of methylation at these DMRs was observed correlating with the loss of silencing of these UIGs at two different developmental stages. We also observed a decreased enrichment of DNMT1 at these DMRs explaining the decreased methylation. Taken together, we observed that Kcnq1ot1 RNA is required to maintain gene silencing at the UIGs. Moreover, continuous presence of RNA over the two somatic DMRs is essential to stably maintain the DNA methylation across cell generations, contrary to the general perception that DNA methylation is a stable mark. We also investigated the repressive histone modification at the Kcnq1 locus, upon deletion of RNA at E8.5. The repressive histone modifications remained unchanged indicating that silencing of the PIGs is maintained in a RNA independent manner.

Finally, we have investigated the plausible mechanisms by which NIGs escape the silencing mediated by Kcnq1ot1 ncRNA. The repressive modifications H3K27me3 were found to be absent over the NIGs, Nap1l4 and
Cars. As the H3K27me3 modifications and the active histone modifications like H3K27act are known to be present in a mutually exclusive fashion, we profiled the enhancer specific marks H3K27act and H3K4me1 along the one mega base Kcnq1 domain using ChIP on chip. The non-imprinted genes were found to be enriched with enhancer specific modifications, H3K27act and H3K4me1. These results indicate that, enhancer specific active histone marks over the NIGs could be involved in evading RNA mediated silencing.

Paper III: Characterization of the RNA content of chromatin.

With the advent of high throughput analysis it has been well accepted that majority of the genome is transcribed. This has indicated towards the presence of large number of noncoding RNAs as compared to protein coding genes. There have been several evidences from studies which have implicated noncoding RNAs in epigenetic regulation. For example, Kcnq1ot1 and Airn noncoding RNAs at their respective imprinted locus are involved in epigenetic silencing by recruiting repressive machinery to target genes by associating with the chromatin. Similarly Xist has been implicated in recruiting histone modifiers during X chromosome inactivation. These studies have pointed towards the presence of many more noncoding RNAs in the genome which might be epigenetically regulating diverse functions through their association with the chromatin.

In line with this assumption, we have identified chromatin associated RNAs (CARs) on a genome wide scale. Using sucrose gradient purified chromatin, we have isolated RNA from the soluble chromatin fractions in human fibroblast cells and sequenced it using Illumina platform. Analysis of the sequenced reads revealed that, majority of the reads (71%) mapped to intergenic and intronic regions whereas only 29% mapped to exons of protein coding genes. By further analyzing the intergenic and intronic reads, we identified 141 CARs mapping to intronic regions and 74 CARs mapping to intergenic regions. We observed that 52% of the intronic CARs overlapped with the known totally intronic (TIN) and partially intronic (PIN) transcripts suggesting that CARs are independent transcripts. In addition, we observed that 15 of the intergenic CARs were part of identified lincRNAs of which 9 have been shown to associate with PRC2 members.

The identified CARs were found to be evolutionary conserved across placental mammals. We validated the chromatin association of selected CARs using complementary technique, Chromatin RNA immunoprecipitation (ChRIP). One of the intergenic CARs, Intergenic 10 was functionally characterized. We show that, Intergenic 10 regulates the expression of its two
neighboring genes by modulating the active chromatin structure over their gene promoters.

PAPER IV: Transcriptomic profiling of primary neuroblastomas reveals a high-risk tumor associated long noncoding RNA NBAT1, with functional roles in cell proliferation and neuronal differentiation.

LncRNAs are being reported in regulatory roles in many biological pathways. Therefore, misregulation of lncRNA mediated functions could have an adverse effect on the cell and might lead to diseases such as cancer. Indeed, many examples of misregulated lncRNAs in diseases are being reported. Identifying such lncRNAs could serve as important information for better understanding of the disease conditions. In the current investigation, we have aimed to identify lncRNAs with probable roles in neuroblastoma tumors and characterize them in order to better understand lncRNA mediated gene regulation.

Neuroblastoma is a tumor of sympathetic nervous system affecting mostly children below the age group of 1.5 years. These tumors are highly heterogeneous in nature and can be categorized into distinct clinical subtypes based on their clinical outcomes. However, broadly these tumors can be classified as either aggressive, high-risk tumors (characteristic chromosomal aberrations such as 1p, 11q- deletion or MYCN amplification) or regressive, low-risk tumors (only copy number chromosomal variations). With an aim to identify differentially expressed lncRNAs between low- and high-risk tumors, we have sequenced the transcriptomes of tumors representing, low-risk and two high-risk subtypes (11q- deletion and MYCN amplified). We obtained around 584, 756 and 70 differentially expressed genes (both protein coding genes and lncRNAs) between the low-risk vs. 11q-, low-risk vs. MNA (MYCN amplified) and low-risk vs. high-risk tumors respectively. Furthermore, we obtained 24 (out of 70 differentially expressed genes) novel lncRNAs as differentially expressed between low-and high-risk subtypes. We have validated the expression of some of these lncRNAs using qRT-PCR in an independent cohort of neuroblastoma tumors.

Many characteristic aberrations are associated with aggressive neuroblastomas such as amplification on chromosome 2, deletions on chromosome 1p arm and 11q arm. These regions are believed to have putative tumor suppressor genes or oncogenes which could contribute towards neuroblastoma aggressiveness. Therefore, we wanted to investigate if any of the differentially expressed lncRNA between these subtypes mapped to these characteristic nonrandom chromosomal locations. We have found lncRNAs mapping to MNA amplified region, 11q arm and 1p arm as differentially expressed
between the neuroblastoma subtypes. These RNAs could be involved in tumor suppressor or oncogenic functions in these neuroblastoma subtypes.

We observed an uncharacterized lncRNA, which we named as \textit{NBAT-1} (Neuroblastoma associated transcript-1), as one of the top ranked candidate to be differentially expressed between low-risk and high-risk neuroblastoma tumors. \textit{NBAT-1} has a relatively lower expression in high-risk tumors as compared to low-risk tumors. We have validated its expression in an independent cohort of 37 primary neuroblastoma tumors using qRT-PCR. Interestingly, an aggressive neuroblastoma associated SNP is known to reside in the intron of \textit{NBAT-1}. Our analysis on SNP and its effect on \textit{NBAT-1} expression revealed that risk allele status at the SNP correlated with lowered expression of \textit{NBAT-1} in high-risk tumors. In order to further understand the role of \textit{NBAT-1} in neuroblastoma progression, we have depleted the RNA using sh-RNAs against \textit{NBAT-1} in SH-SY5Y cells. Knocking down of the RNA resulted in an increased proliferation and invasiveness in the SH-SY5Y cells. We have sequenced the RNA obtained from \textit{NBAT-1} depleted as well as control cells. We found an enrichment of genes involved in cell proliferation, migration and neuronal differentiation. The change in expression of genes in cell proliferation and migration upon \textit{NBAT-1} knockdown indicates towards its tumor suppressor role and explains the observed phenotype of increased proliferation and invasiveness. As we also observed neuronal differentiation as another important biological process to be affected upon \textit{NBAT-1} knockdown, we investigated the role of \textit{NBAT-1} in neuronal differentiation in SH-SY5Y cells. We observed that \textit{NBAT-1} is significantly up-regulated during neuronal differentiation and depletion of the RNA resulted in perturbed differentiation process.

Taken together, in the current investigation, we have identified several lncRNAs which could serve as potential biomarkers in neuroblastoma prognosis. Moreover, we report an lncRNA, \textit{NBAT-1} as differentially expressed between low- and high-risk tumors. Our functional studies on \textit{NBAT-1} indicate that it could promote neuronal differentiation along with inhibiting increased cell proliferation and invasiveness in primary neuroblastomas.
Acknowledgement

The work presented here in this thesis was carried out at Department of Immunology, genetics and pathology (IGP), Uppsala University, as well as in Department of Medical and Clinical genetics, Gothenburg University.

Firstly, I would like to sincerely thank my supervisor, Prof Chandrasekhar Kanduri for all his guidance, support, and motivation during this work. I am also grateful to him for letting me work on some very interesting projects in his laboratory. I will always be thankful to him for all the discussions and his valuable suggestions, which have immensely helped me to improve as a student.

Thanks to my Co-supervisor, Prof. Claes Wadelius for organizing chromatin journal club. It was a great source of keeping up to date on recent and interesting scientific studies.

Thanks to Dr. Fredrik Öberg, Dr Lena Åslund and other members of the chromatin journal club for the interesting articles and discussions. Thanks to Prof. Helena Jerenberg Wiklund for suggestions whenever needed.

I would like to thank Dr. Fredrik Hedborg with his help and support with the tumor project.

I would like to thank, head of the department, Prof. Lena Claesson-Welsh for the constructive research environment at the IGP. Thanks to all the past and present members of administration for their help whenever needed. Special thanks to Christina for all her help with my applications and forms. I would also like to thank all the members of animal facility for taking good care of the mouse.

I am thankful to my senior lab colleague, Pandey Ji for his professional help during my initial days in the lab. It was also a pleasure to listen to your cool ideas. Thanks to Pratima Bhabhi for the good food at your place.

Many thanks to Faizaan bhai for teaching me fine details of mouse work. I am thankful to you for your valuable suggestion and tips which were of great
use in my experiments. You also have been an inspiration for cooking good food. Thanks to Fatima Bhabhi for all the nice parties at your home.

Thanks to Natalia, for teaching me ice-skating, for all the nice parties and organizing many wonderful trips.

I would like to thank, Tanmoy (Dada) for always being a very nice and dear friend. Thanks to you for your moral support and valuable suggestions throughout my work. I have always admired your scientific intellect.

Thanks to Matthieu for always being ready to help whenever required. It has been very nice to have you as a colleague in the lab and know your thoughts on various topics. Thanks for coming to my wedding.

Thanks to Kankadeb (KK) for his unconditional help in and outside the lab. Thanks for being very colorful and jovial member of the lab. It has been really great to have you around.

Thanks to Sanhita for her timely help in my experiments.

I would like to sincerely thank Santhilal for your constant support with the data analysis and patiently answering all my questions. Thanks to Sirisha for her help whenever required. Thanks to the past members of the lab, Roshan (Bhaiya) and Abiarchana for their cooperation in my experiments. Thanks to Arif for his suggestions in the lab.

I would like to thank Meena for all the good food and parties at your home.

I would like to thank all the members of the third floor at Rudbeck lab for nice chats and discussion during lunch and tea time. Thanks to Nicola, Mi, Lesley, Larry, Annika, Tobias, Akhtar, Amar, Demet, Jelena, Sara, Naren da, Sneganshu Dada, Soumi di, Geeta, and Lothar for making my stay at IGP wonderful. It has been very nice to know you all.

Thanks to the members of our neighboring lab at Uppsala. Thanks to Charlotte, Antonia, Pernilla, Inger for being very nice to us and for all your help whenever needed. Special thanks to Prasoon (Baba) for your eternal knowledge on all the subjects. Thanks for your cooperation with all the things at Uppsala.

Thanks to members of the Wadelius lab- Mehdi, Ola and Madhu for the discussions during journal club.
I would like to thank, Prof. Per Elias and Prof. Claes Gustafsson for their support and cooperation during shifting to our new lab in Gothenburg. Thanks to Dr. Erik Larsson for all his ideas and discussions with the data analysis.

Thanks to Sashidhar for all his help with the mouse work.

I want to thank Prof. Saumitra Das for providing me with the opportunity to work in his lab at IISc, Bangalore at the initial stages of my research carrier. Thanks to all the lab members at IISc, Sharath, Rammana garu, Nithya, and Bhupi singh for your nice company. Special thanks to Bhupi singh for your support and cooperation during my stay at IISc. I want to specially thank Deb dada (Janab) for always being there for me whenever needed and always suggesting me constructively.

I want to thank all my MSc teachers at KUD Dharwad for their encouragement and belief in me. I want to thanks all my MSc batch mates and friends for the very nice and memorable days spent with them at Dharwad. Thanks to Dr. Narahari for his guidance after my MSc.

I want to thank the Indian community at Uppsala for making my stay at Uppsala, very memorable. The years spent with Kali bhai, Jeetu mama, Puneet ji and Dada at Doblensgatan have been one of the best and memorable times of my life. Thanks to all of you for the great time. I will always cherish these memories. Thanks to Kali and Sulena, Nimesh and Smitha, Sandeep and Navya for all the get together and social gatherings. I also want to specially thank, Umash ji and Noopur Ji for all the nice dinners and interesting chats at Rudbeck. It has been wonderful to known you all.

I also want to thank, my Indian friends in Gothenburg, Pankaj and Prisha, Rakesh and Pooja.

I want to thank all my family members for their support and encouragement. I want to specially thank my parents and my younger brother Saurabh, for their unconditional love, affection, encouragement and moral support without which nothing would have been possible.

I want to thank my wife, Nalinee for all her love, support and understanding during my research studies. And lastly, I want to thank my little daughter, Nushi for filling my life with lots of joy and happiness. You are the cutest baby in the world.
References


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 940

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, 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 Medicine.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-209596