Functional Analysis Of Second Intron Of Imprinted Mouse *Neuronatin* Gene

Thesis submitted for the degree of Doctor of Philosophy

To

The Department of Biochemistry,

School of Life Sciences,

University of Hyderabad

BY

DIVYA TEJ SOWPATI

Centre for DNA Fingerprinting and Diagnostics

Nampally,

Hyderabad – 500 001

2012

Reg. No. 06LBPH10

University of Hyderabad,
School of Life Sciences,
Department of Biochemistry,
Hyderabad – 500 046, INDIA

Declaration

The research work embodied in this thesis entitled "Functional Analysis of Second Intron of Imprinted Mouse *Neuronatin* Gene" has been carried out by me at the Centre for DNA Fingerprinting and Diagnostics, Hyderabad, under the guidance of Prof. Seyed E Hasnain and Dr. Sanjeev Khosla. I hereby declare that this work is original and has not been submitted in part or full for any other degree or diploma of any other university.

Divya Tej Sowpati
Centre for DNA Fingerprinting and Diagnostics
(CDFD), Hyderabad

University of Hyderabad,
School of Life Sciences,
Department of Biochemistry,
Hyderabad – 500 046, INDIA

Certificate

This is to certify that this thesis entitled "Functional Analysis of Second Intron of Imprinted Mouse *Neuronatin* Gene", submitted by Mr. Divya Tej Sowpati for the degree of Doctor of Philosophy to the University of Hyderabad is based on the work carried out by him at the Centre for DNA Fingerprinting and Diagnostics, Hyderabad. This work is original and has not been submitted in part or full for any other degree or diploma of any other university or institution.

Prof. Seyed E Hasnain
Thesis supervisor,
Indian Institute of Technology, Delhi

Dr. Sanjeev Khosla
Co-guide,
CDFD, Hyderabad

Head, Department of Biochemistry
University of Hyderabad,
Hyderabad

Dean, School of Life Sciences
University of Hyderabad,
Hyderabad

Table of Contents

ACKNOWLEDGEMENTS

I waited a long time for this day! Journey till here was not easy. In fact, it would have been impossible had it not been for so many people who helped me, and it feels great to have an opportunity to thank them.

To begin with, I would like to express my heartfelt gratitude to my PhD mentor, Dr. Sanjeev Khosla. Dr. Sanjeev is responsible for what I am today, and he has been most crucial in shaping my research attitude. He has taught me how to read, write, understand, talk, discuss and present in the scientific way. Moreover, he gave me liberty to think and work independently, and has been very encouraging when experiments did not work. The patience he had for me is incredible, and I could not have asked for a better PhD supervisor. I have imbibed substantially from him, and I feel fortunate to have been associated with him.

My earnest thanks go to Prof. Seyed E Hasnain for entertaining me as a special case and permitting me to be a part of CDFD. He has godfathered me throughout my doctoral career, allowed me to register at the University of Hyderabad as his student, and taken care of all formalities pertinent to my PhD that may have troubled me otherwise. On an extended note, I would like to thank Dr. J Gowrishankar for his continued support to work in CDFD, and also for permitting me to attend the IDIBELL conference at Barcelona apart from various other national conferences. I also take this chance to thank the Head and Dean, Department of Biochemistry, and Registrar, University of Hyderabad, for permitting me to register in their esteemed university.

I thank my doctoral committee members and various faculty of CDFD, in particular Drs. Gayatri Ramakrishna, Shekhar Mande, Mahalingam, Ranjan Sen, Murali Bashyam, Rashna Bhandari, Abhijeet Sardesai, Sunil Manna, Sangita Mukhopadhyay for the coursework classes and the valuable discussions and insights they provided. I am grateful to Dr. Rakesh Mishra, CCMB, for permitting me to work in the 'Fly Lab', and the invaluable conversations he has had with me.

I take this opportunity to thank Dwarakanath and Jyoti Lakshmi for their boundless help with animal work. I also thank the sequencing staff of NGTF, particularly Vartul and Ajay, for being very prompt in giving back the results, even when the work load was very high. I would like to extend my gratitude to the staff of Sandor Proteomics, especially Saroja, for her help with Real-Time experiments. I thank Suman for the countless times he aided me with mouse work. My thankfulness goes to Hina, Vasanthi, Hemant, Narendra and Pavan for not just helping me with Drosophila work, but for all the fun time I spent with them during my days in the Fly Lab. Many experiments would not have been possible without Shareef, Hassan, Gowrish, Nanci and Sudheer. The list is not complete without the administrative staff of CDFD, in particular Punitha and Santosh, who helped me many a time with various PhD related formalities.

I have been particularly lucky to have a great lab atmosphere during my stay here. Initially, Sudhish and Gokul made sure I did not feel out of place and helped me gel with the lab quickly. I will always treasure the outings we had, the places we visited, the restaurants we ate and the innumerable topics we discussed – scientific and otherwise. I thank Sudhish for the games of Table Tennis and Badminton, and Gokul for our shopping times, and the nights I spent at his place. I thank other present and past members of the lab, Usha, Malathi, Gautami, Mallikarjun, Neeraja, Garima, Amitava, Rachana, Bindu, Lalitha, Vaishnavo, Sneha, Imtiyaz and Thushara, who have constantly encouraged me, and helped me myriad times with various work. Even a day full of work did not feel stressful, thanks to the fun-filled atmosphere created by all of them. Special appreciation goes to Garima, who treated me like her own brother and for bringing me lunch almost every day, sometimes sacrificing hers and Amitava for the time I spent at CDFD hostel and for being a wonderful companion to eat at various restaurants. I am grateful to Lalitha for her help in protein work, and Vaishnavo for helping me with ChIP experiments and proof-reading my thesis. My sincere thanks to Ranjani who was my 'Bestie' for almost the entire duration of her stay at CDFD. I have been fortunate to meet many wonderful trainees during my course here and I have learnt something from each one of them. Rupa has greatly helped me with Drosophila work. I also thank Agila, Swathy, Shubha, Vinay and Madhulika for the refreshing times I had with them. Special thanks go to Lipsa, who is always there for me when I need her, and for her help in designing this thesis. She has been a constant source of encouragement and support, particularly towards the end of my PhD.

This doctorate has had a great impact on my personal life too, for I made some amazing friends during my course here. Devi has been much more than a lab mate, a true friend – always encouraging me and firmly believing in me. The support she provided during the times I was feeling low is unfathomable. Neelima has been a great companion, from the moments we had in the corridors of CDFD to the night stays at her place, every instance I spent with her was truly wonderful. Hari and Vidhi have been my 'hang-out' friends, particularly during weekends. In addition, they have aided me in taking some important decisions of my life.

All of them have relentlessly encouraged me, tolerated me when I was a pain and gave me emotional support when I needed it the most. They have been great advisors, both in my scientific and personal lives. They will always share a special place in my heart, and I am not exaggerating when I call them 'friends for life'.

I thank my batch mates Jisha, Ratheesh, Debashree, Sandeep, Gita, Kaiser, Yusuf, Tabrez, Ritu, Aisha and Jyoti for the great times we had during our coursework. I would like to extend my thanks to Nora for our 'dates', and Arvind for the *Chai*-time discussions. It was always good to be with Atul, Shubhada, Nagender, Santosh and Madhav. I thank the Cricket Team of CDFD for letting me be a part of it. On a similar note, I would like to convey my acknowledgements to each and every member of the CDFD family, who have helped me at some time in one way or the other.

This thesis in this shape, thanks to all the options and customizations possible in Microsoft Word. All the images were processed and collated using Illustrator and Photoshop by Adobe. Google, PubMed and Wikipedia were of particular help in finding suitable information contained in this thesis. All references were managed by EndNote. I thank the creators of all the above for providing such wonderful software and platforms.

I thank the God Almighty for making me what I am. I am also grateful to Him for blessing me with a wonderful family. My brother has been a great stress-buster, always helping me take my mind off work pressure with our mostly silly and sometimes intelligent conversations. I also thank him for the times he helped me with some basic problems related to Physics, Chemistry and Mathematics. Words cannot begin to describe the gratitude I have for my parents! The real reason I could achieve this feat is because of their unconditional love, support and belief in me. I thank them from the bottom of my heart for the struggle they have gone through to see me in this stage, and the patience they showed all these years. This thesis is dedicated to them.

TEJ

List of Abbreviations

µg	microgram
µl	microlitre
µM	micromolar
bp	Basepair
BSA	Bovine Serum Albumin
CBB	Coomassie Brilliant Blue
C-terminus	Carboxy terminus
dCTP	Deoxy cytosine triphosphate
DMR	Differentially methylated region
DMSO	Dimethyl Sulphoxide
DNA	Deoxy ribonucleic acid
Dnmt	DNA methyl transferase
dNTP	Deoxy Ribonucleotide Tri Phosphate
DTT	Dithiothreitol
E.coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetra Acetic acid
EMSA	Electrophoretic mobility shift assay
EtBr	Ethidium Bromide
GST	Glutathione S transferase
HCHO	Formaldehyde
HEPES	N-(2-Hydroxy ethyl) Peiperazine-N'-(2-ethanesulfonic acid)
ICR	Imprinting control region
IPTG	Isopropyl β-D thiogalactopyranoside
Kb	Kilo basepair
KCl	Potassium Chloride
KDa	Kilo Dalton
L	Litre
LB	Luria Broth
M	Molar
MALDI-TOF	Matrix assisted laser desorption/ionization – Time of Flight
Meg	Maternally expressed gene
MgCl2	Magnesium Chloride
min	Minute
ml	Millilitre
mM	Millimolar
N-terminus	Amino terminus
N2	Nitrogen
NaCl	Sodium Chloride
°C	degrees Celsius
OD	Optical Density
OD_{600}	Optical Density at 600 nm
PAGE	Poly Acrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction

Peg	Paternally expressed gene
PEG	Poly ethylene glycol
pmole	Picomole
PMSF	Phenyl methyl sulphonyl fluoride
PVDF	Polyvinylidene fluoride
RNA	Ribonucleic acid
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SDS	Sodium Dodecyl Sulphate
sec	Second
SSC	Sodium Saline Citrate
SSPE	Saline Sodium Phosphate –EDTA
TBE	Tris-borate EDTA
TE	Tris-EDTA
TEMED	N,N,N'N' Tetramethylethylenediamine
Tm	melting temperature
Tris	Tris-hydroxymethyl amino methane
UV	Ultra Violet

List of Figures and Tables

SYNOPSIS

The phenomenon of genomic imprinting in mammals is responsible for functional non-equivalence of parental alleles. This phenomenon represents a specialized mode of epigenetic transcriptional regulation for imprinted genes. Even two decades after the discovery of the first imprinted gene, molecular mechanisms governing genomic imprinting at various loci are not completely understood. The regulation of imprinted genes is very complex with several layers of control involving multiple *cis* and *trans* acting elements (Delaval and Feil, 2004). A majority of imprinted genes are found in clusters, often coordinately regulated by a single major *cis*-acting element known as an imprinting (or imprint) control region (ICR). Several mechanisms have been put forward to explain the regulation of imprinted expression at various loci. In some cases, the transcription of a non-coding RNA antisense to protein coding genes within the ICR results in their perturbed expression (Lewis and Reik, 2006). ICRs are also known to act as insulators that bind to insulator proteins and prevent the interaction of gene promoters with enhancers (Bell and Felsenfeld, 2000; Hark et al., 2000). In fact, gene silencing is the underlying mechanism for all the known ICRs. On the other hand, the mechanism by which the active allele of an imprinted gene is prevented from getting silenced remains largely unknown.

Neuronatin, a paternally expressed gene on chromosome 2 in mouse (chromosome 20 in human), is expressed in the central nervous system from mid-gestation through early postnatal development, correlating with the onset and termination of brain development in mice and humans (Joseph et al., 1994; Wijnholds et al., 1995). It has also been shown to be expressed in all cells of the developing pancreas in mice, becoming restricted to beta cells of pancreas in adulthood (Chu and Tsai, 2005). Several functions have been attributed to *Neuronatin*, including brain development, insulin secretion,

diabetes and obesity (Chu and Tsai, 2005; Joe et al., 2008; Mzhavia et al., 2008; Vrang et al., 2010; Wijnholds et al., 1995). The *Neuronatin* gene is located within the 8.5 kb intron of another gene, Bladder Cancer Associated Protein (*Bc10/Blcap*) (John et al., 2001), which encodes for a non-coding RNA. *Blcap* exhibits transcript-specific imprinting; *Blcapv1a* is preferentially transcribed from the maternal allele while *Blcapv2a* is expressed solely from the paternal allele and shares its promoter with *Neuronatin*.

DNA methylation analysis of the mouse *Neuronatin* locus showed exclusive methylation only on the maternal allele, and this region of differential methylation encompasses the whole *Neuronatin* gene, from its promoter to the last exon (Kagitani et al., 1997; Kikyo et al., 1997). Preliminary nuclease sensitivity assays previously done in the lab had identified a transcription-independent DNase I hypersensitive site, only on the unmethylated paternal allele. This hypersensitive site was mapped to the second intron of *Neuronatin.* The respective ICRs at other imprinted domains like *H19/Igf2, U2af1-rs1*, *Snrpn/Snurf, Kcnq1, Peg3* and *Gnas* showed mutual exclusiveness of DNA methylation and differential chromatin organisation on parental alleles (Coombes et al., 2003; Feil et al., 1997; Hark and Tilghman, 1998; Kim et al., 2003; Schweizer et al., 1999; Shibata et al., 1996; Thorvaldsen et al., 1998; Tremblay et al., 1997; Yatsuki et al., 2002). As the second intronic region of *Neuronatin* satisfied the biochemical criterion of mutual exclusiveness of DNA methylation and differential chromatin organization on parental alleles, it was hypothesized to be the putative ICR of this locus. This thesis describes our efforts in testing this hypothesis.

Chapter I presents a review of literature on epigenetic regulation of gene expression in eukaryotes with a special emphasis on the various *cis*-regulatory elements. This chapter

also discusses the phenomenon of genomic imprinting as it represents a very specialized way of transcriptional regulation adopted by mammals. Furthermore, information about ICRs – the *cis*-regulatory elements that control the expression within the imprinted domains has been elaborated in this chapter. Towards the end, Chapter I discusses the imprinted mouse *Neuronatin* locus in context of the aim of this thesis.

To examine the role of second intron of *Neuronatin* (NNI^2) in regulation of transcription, we performed a P-element reporter gene assay in *Drosophila.* Chapter II describes the analysis of constructs containing the 250 bp long NNI^2 upon injection into *Drosophila.* Flanked by *LoxP* sites, NNI^2 was introduced in both orientations into a vector containing *hsp70* promoter that drives the expression of *miniwhite* reporter gene. The vector also contained P-elements for facilitation of transposition into *Drosophila* genome. Both constructs were injected into white-eyed w^{1118} *Drosophila* and 13 unique transgenic lines were selected. For each transgenic line, the corresponding 'flipped-out' line where the intron was removed by Cre-mediated recombination was derived. Eye color of each transgenic line was compared to its flipped-out counterpart to assay for the effect of NNI^2 on transcription. 7 out of 13 transgenic lines showed a decrease in the intensity of red eye color when the NNI^2 was removed, irrespective of the orientation of intron, indicating that the NNI^2 was acting as a transcriptional activator. Further crosses of these transgenic lines with flies harboring mutations in chromatin modifying genes indicated the role of *Trx*, a H3K4 methyltransferase, and *Esc*, a H3K27 methyltransferase, in the mechanism by which NNI^2 modulates transcription.

The third chapter elucidates the generation and characterization of mice deleted for second intron of *Neuronatin* ($NN\Delta I^2$) from the endogenous locus. The aim of this experiment was to examine the role of NNI^2 in establishment and maintenance of

imprints at mouse *Neuronatin* locus. Analysis of progeny from crosses where NNΔI^2 was inherited from either the mother or father showed no expression of *Neuronatin* in brain of paternal heterozygous and homozygous NNΔI^2 mice, whereas the expression in brain of maternal heterozygous NNΔI^2 mice was found to be equal to that of wild type mice. DNA methylation analysis of *Neuronatin* promoter revealed the hypermethylation of promoter in brain of paternal heterozygous and homozygous NNΔI^2 mice, concomitant with the loss of expression. This indicated that as observed in *Drosophila*, NNI^2 acts as a transcriptional activator, even in mice. The methylation status of *Neuronatin* promoter of maternal heterozygous NNΔI^2 mice was similar to that of wild type. Phenotypic analysis revealed obesity in paternal heterozygous and homozygous NNΔI^2 mice. Surprisingly, maternal heterozygous NNΔI^2 mice, which did not show any change in *Neuronatin* expression, had lesser body weight compared to wild type mice. Moreover, all NNΔI^2 mice showed loss of control on their body weight with age. Further analysis revealed abnormal expression of both *Neuronatin* and *Blcap*, in a transcript-, tissue-, and parent-of-origin specific manner. This indicated the role of NNI^2 not only in maintaining the transcription of *Neuronatin*, but also that of *Blcap*. However, no other genes in 250 kb vicinity of *Neuronatin* were affected by this deletion.

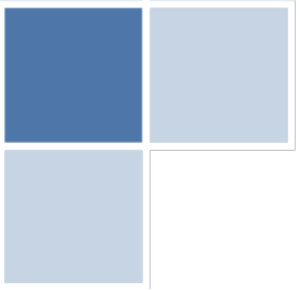
Previous studies from our laboratory had indicated the possibility that the differential chromatin organization observed for two alleles of *Neuronatin* could be due to the exclusive binding of non-histone protein(s) to one of the parental alleles in a methylation-dependent manner. Similar observations have been made in other studies on the imprinted genes like *H19/Igf2, Peg3* and *Gnas* that had shown the differential binding of proteins like CTCF and YY1 to the two alleles within the imprinting control region (Bell and Felsenfeld, 2000; Kim, 2008; Szabo et al., 2004). Yeast Mono Hybrid

and MALDI-TOF experiments previously conducted in our lab had identified several possible proteins that could bind to a 38 bp GC rich region within the NNI^2, referred to as GC motif. In Chapter IV, we describe the testing of proteins Cbx5 and hnRNPK for their ability to bind the GC motif in a methylation-dependent manner. EMSA (Electrophoretic Mobility Shift Assay) analyses revealed that Cbx5 preferentially binds to the unmethylated GC motif while methylated GC motif was preferred by hnRNPK.

In this thesis, we report the identification of the imprinting control region (NNI^2 ICR) for the mouse *Neuronatin/Blcap* locus. NNI^2 ICR was found to be a transcriptional activator in a reporter gene assay in *Drosophila*. This observation was strengthened by Knock-Out studies where NNΔI^2 mice showed loss of imprinted expression for both the *Neuronatin* and *Blcap* genes. The significance of these findings in context of various phenotypic consequences observed in NNΔI^2 mice has been discussed in chapter V.

CHAPTER I

CIS-REGULATORY ELEMENTS IN EUKARYOTIC GENOMES



Different phenotypes from the same genotype arise in a cell due to differential expression of genes. This regulation of gene expression is achieved by the combination of various *trans*-factors interacting with *cis*-regulatory regions. This interplay defines the transcriptome or expression profile of a cell. While certain combinations ensure the expression of 'house-keeping' genes in all cells, tissue-specific transcription of certain genes in response to external stimuli is achieved by a different set of interactions.

While regulation of transcription is observed in prokaryotes, it is much simpler than what is seen in single-cell eukaryotes and higher vertebrates (Levine and Tjian, 2003). Some of the complexities in eukaryotic gene expression arise due to mechanisms such as splicing, post-translational modifications, chromatin structure and modifications, etc. Moreover, regulation can be achieved at different levels including gene structure, transcription initiation and elongation, RNA processing, mRNA transport, stability and translation etc. Though regulation at each level contributes equally in modulating the phenotypic response of a given genotype, the focus of this chapter is on the regulatory elements and factors that provide the interphase between gene structure and transcription initiation. The key to regulation at transcription initiation is the RNA Polymerase. The process of transcription initiation by RNA Pol II (responsible for mRNA synthesis in eukaryotes) includes the assembly of various transcription factors into transcription machinery, which facilitates the positioning of RNA Pol II at the promoter. The activity of RNA Pol II is also dependent on several auxiliary factors including DNA elements other than promoter, histone and non-histone proteins, epigenetic modifications etc. Presented below is a review of literature on epigenetic regulation of gene expression in eukaryotes with a special emphasis on the *cis*-regulatory elements.

1.1 CIS-REGULATORY ELEMENTS IN EUKARYOTIC GENOMES:

Almost all genes in a eukaryotic cell can be constitutively expressed at a minimal rate. It is the regulatory elements that are responsible for their appropriate expression levels. The initial school of thought was that the complexity of an organism is directly proportional to the number of genes it has. However, comparative genome-wide analyses proved otherwise: *Drosophila*, with roughly 14,000 genes (Adams et al., 2000) has many more cell types and tissues than *Caenorhabditis elegans* which has approximately 20,000 genes (Ruvkun and Hobert, 1998). However, the size of the genome is proportional to the complexity, indicating that it is not the coding sequences, but the regulatory regions present around them that confer the physiological and behavioural complexity to an organism. It is proposed that more the number of regulatory elements in a genome, the diverse the gene expression profiles will be, and hence the more complex the organism can be (Levine and Tjian, 2003). Apart from the motifs that are responsible for the positioning of RNA Pol II at the promoter, several eukaryotic *cis*-elements that can positively or negatively influence the transcription rates have been identified. As described below, these regulatory elements are known to influence transcription even from several kilobases away and have been categorized as enhancers, silencers and boundary elements based on the underlying mechanism.

1.1.1 ENHANCERS:

Enhancers, the first regulatory elements to be identified, are specific DNA motifs that up-regulate gene expression by sequestering transcription factors and promoting the formation of pre-initiation complex at the core promoter of their target genes. The assembly of general transcription factors (GTF) at the core promoter can be sufficient for a successful initiation of transcription. However, the rate of transcription from such

a promoter could be minimal. This rate can be modified by the interaction of the transcription factors with the enhancer elements. Initially it was thought that the enhancers worked by direct interaction with RNA Polymerase II, the proof for which came from mutational studies on the C-Terminal Domain (CTD) of RNA Polymerase II. The deletion of CTD resulted in abolishment of transcription regulation by all studied enhancers, as the mediator assembly failed in binding to RNA Pol II (Roberts et al., 1993; Stringer et al., 1990). However, now the general belief is that the effect of enhancers is indirect, with intermediary factors messengering their signals to transcription machinery (Bjorklund and Kim, 1996).

A well-studied example of enhancer proximal to gene promoter is the CCAAT box, usually present 75 bp upstream to the transcription initiation site of several eukaryotic genes (Jones et al., 1985). The 9 bp consensus sequence (GGCCAATCT) of the CCAAT box is recognized by the transcription factor CT1/NF1 which in turn recruits various transcription factor assemblies (Santoro et al., 1988). The 6 bp GC box (GGGCGG) is another proximal enhancer located within the 100 bp region upstream of the promoter. It is recognized and bound by the transcription factor Sp1 (Gidoni et al., 1984). Both CCAAT and GC boxes have the potential to strongly upregulate their target genes, particularly housekeeping genes.

Enhancer elements for several genes involved in response to external stimuli are situated further upstream to these boxes. For example, in response to the steroid hormone family of glucocorticoids, the homo-dimers of glucocorticoid receptors (GR) bind to specific DNA sequences known as glucocorticoid response elements or GREs (Tsai et al., 1988; Wrange et al., 1989). A GRE has a consensus sequence of 5'GGTACANNNTGTTCT 3' (Zilliacus et al., 1995), and binding of the homo-dimeric GR to this sequence recruits chromatin modifiers, transcription machinery and co-factors,

including Brg1, the histone acetylase CBP/p300 and RNA Pol II (Chakravarti et al., 1996; Fryer and Archer, 1998).

While enhancers were originally thought to be present upstream and close to promoters, studies in *Drosophila* led to the discovery of enhancer elements that can act from a distance of 10 kb (Latchman, 2008). In mammals, enhancers situated at a distance greater than 100 kb have also been identified (Birney et al., 2007). One such example is the limb-specific enhancer of the human *SHH* (Sonic HedgeHog) gene, called ZRS (Lettice et al., 2003; Visel et al., 2009). The ZRS enhancer is located 1 megabase away from the *SHH* gene, in the intron of the neighbouring gene, *LMBR1* (Clark et al., 2000). The homologous sequences of both the gene and the enhancer have also been identified in mice (Sagai et al., 2005). Moreover, many enhancers have been shown to work irrespective of their position and orientation with respect to the target promoters. For example, the HS2 enhancer of the human β-globin gene upregulated the transcription of reporter CAT (chloramphenicol acetyltransferase) gene irrespective of whether the enhancer was placed i) in the genomic or reverse genomic orientation, ii) in a position 5' or 3' to the gene, and iii) at various distances up to 6 kb from the gene (Kong et al., 1997). In addition to indicating the long range and position-independent potential of enhancer elements, these observations also imply the role of nucleosome condensation and chromatin looping in bringing the enhancers close to intermediary factors and target promoters.

1.1.2 SILENCERS:

Silencers are elements that negatively regulate the transcription of certain genes and were initially defined as elements that repress the activity of a native promoter in a position- and orientation-independent manner (Brand et al., 1985). However, they are now generally considered as specific sequences, usually short, located upstream to gene

promoters and capable of recruiting specific transcription factors, which in turn carry out specific functions. Silencing can be achieved by mechanisms which can negatively affect transcription either through modification of chromatin structure, or cytoplasmic retention of transcription factors, or intron splicing or hindering of general transcription factor assembly.

Silencers can be broadly classified into two groups: classical silencers and passive silencers or negative regulatory elements (NREs). Classical silencers are short, active and position-independent motifs that work by disrupting the pre-initiation complexes formed at their target promoters. One such classical silencer element was found near human *THYROTROPIN* (*hTSHβ*) gene (Kim et al., 1996). This element mapped to an AT-rich sequence present 140 bp upstream to the promoter. AT-rich sequences have been indicated in the process of nucleotide-unpairing and unzipping of DNA helix during interphase (Bode et al., 1992). Moreover, a protein called Oct-1 is known to recognize and bind this region, which in turn brings about silencing by interfering with TFIIH mediated helicase activity at the locus (Kim et al., 1996). Another example of a classical silencer is the region present upstream to the promoter of chicken ovalbumin gene. This silencer region was mapped to the nucleotides -239 to -220, and it repressed the activity of heterologous thymidine kinase promoter in a reporter-gene transient transfection assay (Haecker et al., 1995).

Passive silencers or negative regulatory elements (NREs) are position-dependent and work by inhibiting the physical interaction of transcription factors with their specific DNA-binding sites. They may also work by affecting other transcriptional events such as splicing, polyadenylation signalling or transcriptional elongation. Unlike classical silencers which usually are located upstream of the target promoters, NREs are

found at various places of the gene body including UTRs, exons, introns or downstream to the ORF.

The 5' UTR/exon 1 of the human α*1-CHIMAERIN* gene is a good example of an NRE. Deletion of a 70 bp region containing this element resulted in a 5-fold increase in the promoter activity (Dong and Lim, 1996). Furthermore, the NRE was also shown to repress the activity of heterologous TK promoter in a position-dependent manner. This silencer works by affecting transcriptional elongation, either by premature termination or by pausing RNA Pol II (Dong and Lim, 1996). Another example of an NRE is the silencer capable of repressing the promoter of *Osteocalcin* gene in proliferating osteoblasts in rats. The NRE was mapped to nucleotides +39 to +104 of *Osteocalcin* and has been shown to repress the activity of heterologous SV40 early promoter only when present 3' to the reporter gene (Frenkel et al., 1993).

Along with their interaction with transcription factors, silencers interact with Polycomb group (PcG) of proteins that show no sequence-specificity and are indicated to play a major role in stable and heritable silencing of several homeotic genes in *Drosophila* during development (Kingston et al., 1996; Muller, 1995; Simon, 1995). Several PcG proteins have been shown to have repressor activity when tethered to DNA-binding domains, suggesting that their mechanism is indirect and they interact with their target sequences via other pre-bound proteins (Satijn et al., 1997). Homologues of PcG have been found in cell species from yeast (Rme1p), *Xenopus* (XPc) to humans (HPH1) (Alkema et al., 1997; Reijnen et al., 1995; Shimizu et al., 1997).

1.1.3 INSULATORS AND BOUNDARY ELEMENTS:

The third transcriptional regulatory mechanism that has been identified blocks the interaction of enhancers with specific promoters. Certain DNA sequences, when placed between a promoter and enhancer, block the interaction of the two. Such DNA

elements are known as insulator elements. Insulators were first identified by their ability to protect a gene from external influences, thereby preventing incorrect activation or silencing (Bushey et al., 2008). Insulator sequences are recognized and bound by specific insulator proteins to achieve the insulating effect (Kim et al., 2007). Several different insulator elements, defined by the proteins that bind them, have been identified in *Drosophila*. These insulator proteins include Boundary Element Associated Factors (BEAF), GAGA Binding Factor (GAF), Suppressor of hairy wing (Su(Hw)), and more recently, CTCF (Valenzuela and Kamakaka, 2006). Sequence motifs to which these insulator proteins bind have also been delineated. For example, a sequence called *scs* (specific chromatin structures) in *Drosophila*, which contains direct repeats of 5' GCTGNG 3' is bound by the insulator protein Zw5 (Gaszner et al., 1999). Similarly, the sequences containing multiple repeats of 5' CGATA 3', also known as *scs'*, are bound by BEAF (boundary element associated factor) (Jiang et al., 2009). The vertebrate CTCF protein recognizes and binds the sequence 5' CCCTC 3' (Dunn and Davie, 2003; Kim et al., 2007). Insulator elements can work in combination, either by interaction with other insulators, or with themselves. For example, the insulator pair *scs*/*scs'*, which flank the either sides of the *Drosophila* 87A7 heat shock puff of polytene chromosomes, are bound by the proteins Zw5 and BEAF respectively. Both the insulator elements interact, forming a chromatin loop that insulates this region from external regulatory elements (Blanton et al., 2003; Farkas and Udvardy, 1992). Similarly, analysis of *Drosophila* lines that carry two insertions of the *gypsy* retrotransposon (bound by the insulator protein Su(Hw)) indicated that the chromosomal insertion sites colocalize within the nucleus in a Su(Hw)-dependent manner (Gerasimova et al., 2000). CTCF protein was initially identified as a transcriptional repressor at the chicken *myc* and *lysozyme* genes. Genome-wide high throughput approaches have since identified a number of CTCF

binding sites, thereby establishing the role of CTCF in gene expression, insulation, nucleosome positioning, and dosage compensation.

Initially, insulators were considered to work by simply stalling the RNA Pol II complex, causing termination of transcription elongation (Kong et al., 1997; Tuan et al., 1992). This model, known as physical barrier model, was based on the observation that insertion of lacO/R complex between an enhancer and a promoter resulted in transcription termination (Ling et al., 2004). Further support for this model came from experiments where RNA Pol II accumulated at the 5' enhancer of chicken β-globin gene instead of its promoter (Zhao and Dean, 2004). However, some intron-localized insulator elements are capable of insulating the promoters from downstream enhancers without resulting in truncation of gene product, indicating that at least in some cases RNA Pol II can transcribe 'through' certain insulators (Geyer and Corces, 1992). Therefore, an alternative hypothesis called the loop domain model was proposed, which predicts that insulators interact with each other or other nuclear structures through formation of chromatin loops (Bushey et al., 2008). Evidence for chromatin loop formation in regulation of endogenous loci in mice came from studying the *H19* Imprinting Control Region (ICR), also known as *H19*-DMD. Two differentially methylated regions (DMRs) named DMR1 (located upstream to the promoter of *Igf2*) and DMR2 (within exon 6 of *Igf2*), interact with *H19*-DMD on maternal and paternal chromosomes respectively thereby determining whether or not *Igf2* is placed in a silent chromatin region and its subsequent interaction with the downstream enhancers (Murrell et al., 2004). The specific interaction of DMRs with the ICR is determined by whether or not CTCF occupies its target sites on the *H19*-DMD (Bell and Felsenfeld, 2000).

1.2 EPIGENETIC MODIFICATIONS AND THEIR ROLE IN GENE REGULATION:

Variation in gene expression is brought about by a complex interplay of various factors, their intermediates and target sequences. This interplay is controlled by modifications of the DNA itself, and its organization in the chromatin. These modifications are known as epigenetic modifications. Two well-studied epigenetic modifications, methylation of DNA and post-translational modifications of histones, have been implicated in regulation of gene expression.

1.2.1 DNA METHYLATION:

DNA methylation was first identified in the nucleic acid of *M. tuberculosis*, which was known at that time as tuberculinic acid (Wyatt, 1951). The function of DNA methylation was unclear for a long time, till it was proposed in prokaryotes as a mechanism to identify foreign DNA, so that it can be efficiently cleaved and disposed (Noyer-Weidner and Trautner, 1993). However, over the years, various functions have been attributed to methylation of DNA like silencing of foreign DNA, suppression of transcription of repetitive elements, long-term cellular memory, genomic imprinting, dosage compensation, gene regulation, and carcinogenesis (Jones and Takai, 2001; Miller and Sweatt, 2007; Singer-Sam and Riggs, 1993; Wainfan and Poirier, 1992). Though DNA methylation occurs in many eukaryotic species, its importance in gene regulation has been extensively studied in mammals. DNA methylation in mammals predominantly occurs on the 5^{th} carbon of cytosines present in the context CpG. The CpG dinucleotide is greatly under-represented in mammalian genomes due to a high spontaneous deamination rate of 5-methylcytosine into thymine (Antequera and Bird, 1993). For example, in humans, the number of CpG dinucleotides is only 21% of its

expected frequency (Lander et al., 2001). Importantly, 70 to 80% of these CpG dinucleotides are methylated (Bird, 1995; Tucker, 2001). These methylated regions typically map to repeats and introns of the genes. However, in contrast to the rest of the genome, smaller regions known as CpG islands show the expected CpG dinucleotide frequency. A CpG island is defined as a region with GC content higher than 60 % and has a CpG to GpC ratio of greater than 0.6. Several views exist on the length of these islands; while many consider 100bp as the minimum length, some put the limit as 500bp. These islands most frequently are located at 5' regulatory regions of genes, and are maintained in an unmethylated state, which is required though insufficient for the active transcription of a gene (Suzuki and Bird, 2008).

Methylation of DNA in a non-CpG context is most prominently observed in embryonic stem cells (Haines et al., 2001), where 25% of all methylation is in a non-CpG context (Lister et al., 2009). This indicates that the ES cells may use different DNA methylation mechanisms to regulate gene expression. Moreover, non-CpG methylation predominantly mapped to gene bodies and was depleted at protein-binding regions and enhancers. It was also observed that non-CpG methylation was lost upon differentiation and was restored in induced pluripotent stem (iPS) cells (Lister et al., 2009). Prevalent non-CpG methylation has also been recently observed in mouse frontal cortex, where the levels are comparable to those observed in ES cells and is negatively correlated with gene transcription (Xie et al., 2012).

In mammals, DNA methylation is carried out by a set of enzymes known as DNA methyltransferases or DNMTs. DNMTs are categorized into three groups: Dnmt1 family, known as maintenance methyltransferases that recognize, bind to and methylate hemimethylated DNA strands after cell replication (Yoder et al., 1997), Dnmt3 family,

known as *de novo* methyltransferases that can bind and act on unmethylated DNA and methylate cytosines on both strands of DNA (Okano et al., 1999), Dnmt2, though structurally similar to other DNA methyltransferases, has no known DNA methylation activity but has been shown to act as a tRNA methyltransferase and has a role in RNA processing during stress (Goll et al., 2006; Thiagarajan et al., 2011).

Removal of methyl group from cytosine can be achieved by an active or passive mechanism. The passive method involves removal of methylation with each round of DNA replication in absence of a DNA methyltransferase activity (Laird, 1997). Active demethylation is achieved by a set of different proteins. TET-domain containing proteins allow conversion of 5'methyl cytosine to hydroxyl-, formyl- and carboxyl-cytosine (Ito et al., 2011), which are replaced with cytosine by the APOBEC family of enzymes (Bhutani et al., 2011).

Role of DNA methylation in gene regulation is well established. The ability of DNA methylation to silence genes was initially shown for the genes on the inactive X chromosome (Singer-Sam and Riggs, 1993). It was postulated that DNA methylation caused chromatin condensation, which in turn reduced the accessibility of DNA to transcription factors (Kass et al., 1997). Even methylation of one CpG dinucleotide in the promoter of genes can regulate transcription. The regulation at such genes is achieved by DNA methylation-dependent binding of transcription factors, for example AP-2 (Comb and Goodman, 1990). In addition, methylated DNA attracts methyl DNA binding proteins like MeCP2 that are known components of repressive chromatin and can recruit histone deacetylases (Hendrich and Bird, 1998). DNA methylation can also interfere with the binding of transcription factors to the target sequences by obscuring

the binding sites (Nan et al., 1998). Furthermore, methylation can up-regulate gene expression by antagonizing the activity of silencers (Jones and Takai, 2001).

1.2.2 CHROMATIN ORGANIZATION AND HISTONE MODIFICATIONS:

Genomic DNA is compacted in the eukaryotic nucleus through its interaction with positively charged histone proteins into chromatin. While the role of chromatin was initially thought to be a structural one and restricted to compacting DNA, many studies recently have shown the importance of different levels of chromatin compaction in regulation of transcription, nuclear organization, DNA repair, etc. (Green and Almouzni, 2002). The chromatin organization is achieved by the wrapping of DNA into a nucleosome formed by an octet of the 4 core histone proteins – two each of H2A, H2B, H3 and H4. Approximately 146 bp of DNA is wrapped around histone core in each nucleosome. In the nucleosomal conformation chromatin is the least compact. With the help of other histone and non-histone chromatin proteins, chromatin can be further compacted into a solenoid where the nucleosomal arrays are organized into a helical structure. Higher order compaction can be achieved by organizing chromatin into looped arrangement.

Chromatin organization influences gene expression in two ways: (i) As a structural impediment to transcription, and (ii) Through the histone code. Association of DNA with nucleosomes can inhibit its transcription by sterically hindering the DNA-transcription factor interaction (Narlikar et al., 2002). To transcribe a gene, chromatin is reorganized to facilitate the accessibility of promoter to transcription factors and RNA polymerase. Though structural impediment to transcription is the most obvious way by which chromatin can play a role in gene regulation, it is the histone code that has been shown to be the major player in regulating transcription. Histone code, as was

proposed by Thomas Jenuwein and David Allis (Jenuwein and Allis, 2001), is the post-translational modifications of N-termini of core histones such as methylation, acetylation, phosphorylation and ubiquitination at specific residues. (Clayton and Mahadevan, 2003; Rea et al., 2000; Turner, 2000; Zhang, 2003). The complexity of the histone code is not only due to the number of N-terminal amino acids that can be modified but also due to the different combinations that can be modified (Fischle et al., 2003). This complexity is further enriched by the specificity of enzymes that recognize and modify each residue (Strahl and Allis, 2000). In addition, each modification brings with it potential to interact with specific protein or RNA factors. For example, methylated and acetylated lysines are bound by proteins containing chromo and bromodomains respectively (Kouzarides, 2007). Several specific modifications have been well studied and are associated either with activation or repression of the gene they are present at. For example, the methylation of lysine present at 4^{th} position of histone H3 is known to activate a gene (Turner, 2000) whereas methylation of lysines at positions 9 and 27 of histone H3 leads to repression of the gene (Rea et al., 2000). Few histone modifications and their related function are listed in Table 1.1.

Modification/ Lysine	Mono-methylation	Di-methylation	Tri-methylation	Acetylation	References
H3K4	Activation		Activation		Barski et al., 2007; Benevolenskaya, 2007; Jenuwein and Allis, 2001; Rosenfeld et al., 2009; Steger et al., 2008; Wang et al., 2008
H3K9	Activation	Repression	Repression	Activation	
H3K14				Activation	
H3K27	Activation	Repression	Repression		
H3K79	Activation	Activation	Both		
H4K20	Activation				
H2BK5	Activation		Repression		

Table 1.1: Histone Code. Role of various histone modifications in activation/repression of associated genes. H3, H4 and H2B denote Histones. K# denotes the lysine at the particular amino acid position which is modified

1.3 GENOMIC IMPRINTING: A SPECIAL PHENOMENON OF TRANSCRIPTIONAL CONTROL:

Somatic cells of diploid organisms contain two copies of all autosomal genes, one copy inherited from each parent. Most of the genes can be expressed from either copy. However, a subset of genes, known as imprinted genes, is expressed only from one allele and the expressing allele is chosen based on which parent the allele was inherited from (Kelsey, 2007). This parent-of-origin specific expression and the epigenetic mechanisms that underlie the non-equivalence of parental alleles constitute the phenomenon of genomic imprinting. Imprinted genes are functionally hemizygous and known to play important roles in embryonic development, growth, behavior, development of various lineages and also some human diseases (Charalambous et al., 2007; Kiefer, 2007; Mikkelsen et al., 2007; Ubeda and Wilkins, 2008; Wood and Oakey, 2006).

The phenomenon of genomic imprinting was first identified in the insects Sciara (Crouse, 1960) and coccids (Nelson-Rees, 1962) in early 1960s. In mammals, genomic imprinting was first reported by the laboratories of Solter and Surani, who showed through nuclear transplantation experiments that embryos formed by two male (androgenetic) or female (gynogenetic) pronuclei failed to develop normally and died early during embryogenesis (McGrath and Solter, 1984; Surani et al., 1984). These observations suggested that the contributions to the embryo from both parents are unequal and only diploid progeny inheriting a copy from each parent was capable of complete development. Cattanach *et al* (1985) while working with maternal and paternal disomic mice arrived at similar results. The results of his experiments also

indicated that only specific regions/genes within the mammalian genome are involved in parent-of-origin specific effects (Cattanach and Kirk, 1985).

Most imprinted genes identified so far are found in clusters. Organization of imprinted genes in clusters indicates that genes within these clusters are regulated co-ordinately, often by shared *cis*-regulatory elements (Edwards and Ferguson-Smith, 2007). This could also suggest that the mechanisms that regulate genomic imprinting are not gene-specific and can work over long distances. For many of the imprinted loci, regions have been identified that are correlated with setting up of allele-specific imprints. They are known to acquire parent-of-origin dependent, allele-specific methylation and chromatin organization. These regions can be categorized as differentially methylated regions (DMRs) and imprinting (or imprint) control regions (ICRs).

1.3.1 DIFFERENTIALLY METHYLATED REGIONS:

DMRs are regions within the mammalian genome that acquire parent-of-origin specific methylation (Kelsey, 2007). Some of these DMRs are known as germ line or gametic DMRs (gDMRs) because they acquire differential methylation during gametogenesis in only one gamete (either egg or sperm) and this methylation is maintained on the same allele throughout development (Bartolomei and Ferguson-Smith, 2011). The second class of DMRs are known as somatic DMRs (sDMRs) as they acquire allele-specific methylation only in a post-implantation embryo (John and Lefebvre, 2011). Almost all imprinted domains have been found to harbour at least one germ line DMR, and one to several somatic DMRs. Though both types of DMRs are believed to be important for the imprinted expression of a gene, the gDMRs have a particular relevance to the phenomenon of genomic imprinting. gDMRs are strongly correlated with establishment

Locus	Chromosomal location	Genomic Location of DMR*	DMR (ICR) location with respect to gene	Methylated allele	Other genes in the locus	References
Zdbf2	Cent 1	63296857 – 63327099	10 kb upstream of *Zdbf2* Exon 1	Paternal	*Gpr1*	(Hiura et al., 2010; Kobayashi et al., 2009)
Mcts2	Dist 2	152512010 – 152512663 152533500 – 152533900	Within Exon 1 of *Mcts2* 21 kb downstream of *Mcts2* Exon 1	Maternal	*H13*	(Wood et al., 2007; Xie et al., 2012)
Gnas†	Dist 2	174109010 – 174113395 174118404 – 174125287 174152611 – 174153503	9 kb upstream of *Gnasxl* promoter Within Promoter of *Gnasxl* (ICR) Within *Gnas1A* promoter	Paternal Maternal Maternal	*Gnasxl*, *Nesp*, *Nesp-as*, *GnasExon 1A*	(Coombes et al., 2003; Kelsey et al., 1999; Peters et al., 1999; Williamson et al., 2004)
Peg10	Prox 6	4696303 – 4699370	Within Promoter of *Peg10*	Maternal	*Sgce*, *Ppp1r9a*, *Pon2*, *Pon3*,	(Ono et al., 2003)

Peg10					*Asb4*, *Calcr*, *Tpfi2*	
Mest	Prox 6	30685840 – 30689965	Promoter/Exon 1 of *Mest*	Maternal	*Klf14*, *Copg2*, *Copg2-as*, *MirN335*	(Lefebvre et al., 1997; Lucifero et al., 2002)
Nap1L5	Prox 6	58857395 – 58857788	*Nap1L5* promoter/Exon 1	Maternal		(Smith et al., 2003; Wood et al., 2007)
Peg3	Prox 7	6680067 – 6685920	*Peg3* promoter/Exon 1	Maternal	*Zim1*, *Zim2*, *Zim3*, *Usp29*, *Zfp264*	(Li et al., 2000; Lucifero et al., 2002)
Snrpn	Cent 7	67148026 – 67150169 67285188 – 67285250 69493100 – 69493181	*Snrpn* promoter/exon 1 (ICR) Within *Snrpn* U exon *Ndn* promoter/ exon 1	Maternal Maternal Maternal	*Atp10a*, *Ube3a*, *Snurf*, *Ndn*,	(Chai et al., 2001; Dindot et al., 2009; Hanel and Wevrick, 2001; Hershko et al.,

Snrpn		69521307 – 69522167 69564012 – 69565740 69608471 – 69609019	Within *Magel2* promoter *Mkrn3* promoter/exon 1 Within *Peg12* promoter	Maternal Maternal Maternal	*Magel2*, *Mkrn3*, *Pec2*, *Pec3*, *Peg12*, *Ube3a-as*, *Zfp127as*, *SnoRNAs*	1999; Sharp et al., 2010; Shemer et al., 1997; Xie et al., 2012)
Inpp5f	Dist 7	135831638 – 135831747	*Inpp5f*-V2 promoter/exon1	Maternal	*Inpp5f-V2*, *Inpp5f-V3*	(Choi et al., 2005)
Igf2	Dist 7	149763483 – 149765230 149765791 – 149767931 NA NA	Within *H19* promoter 2 kb upstream of *H19* exon 1 (ICR) Within *Igf2* promoter 1 Within *Igf2* exon 6	Paternal Paternal	*H19*, *Ins2*	(Bartolomei et al., 1993; Ferguson-Smith et al., 1993; Thorvaldsen et al., 1998; Tremblay et al., 1997; Ueda et al., 2000)
Kcnq1	Dist 7	150480736 – 150482006	Within *Kcnq1-ot1* promoter (ICR)	Maternal	*Ascl2*,	(Bhogal et al., 2004;

Kcnq1		150645240 – 150647381 150649567 – 150649883	*Cdkn1c* promoter/exon 1 2 kb upstream of *Cdkn1c* exon 1	Paternal Paternal	*Tspan32*, *Cd81*, *Tssc4*, *Kcnq1-ot1*, *Cdkn1c*, *Slc22a18*, *Phlda2*, *Nap1l4*, *Osdp15*,	Engemann et al., 2000; Fitzpatrick et al., 2002; Lewis et al., 2004; Yatsuki et al., 2002)
Rasgrf1	Dist 9	89771945 – 89780072	30 kb upstream of *Rasgrf1* exon 1 (ICR)	Paternal	*A19*, *Mir184*	(Yoon et al., 2002)
Plagl1	Prox 10	12809928 – 12812145	*Plagl1* promoter/exon 1	Maternal		(Smith et al., 2002)
Grb10	Prox 11	11923325 – 11926800 11935418 – 11935544	Within *Grb10* promoter (ICR) 9 kb upstream of *Grb10* promoter	Maternal Maternal	*Ddc*, *Cobl*, *Grb10-as*	(Arnaud et al., 2003; Hikichi et al., 2003; Xie et al., 2012)
U2afl-rs1	Prox 11	22871545 – 22874145	*U2afl-rs1* promoter/exon 1	Maternal	*Commd1*	(Shibata et al., 1997)

Dlk1	Dist 12	110697919 – 110700243 110763965 – 110768609 110777813 – 110781249	Within Last exon (exon 6) of *Dlk1* 13 kb upstream *Gtl2* (ICR) *Gtl2* promoter/exon 1	Paternal Paternal Paternal	*Begain*, *Dio3*, *Rtl1*, *Gtl2*, *Rian*, *Mirg*, *miRNAs*	(Lin et al., 2003; Takada et al., 2002)
Ksnk9	Dist 15	72632245 – 72641614	Within *Peg13* promoter	Maternal	*Peg13*	(Smith et al., 2003)
Slc38a4	Dist 15	96884431 – 96886172	Within *Slc38a4* promoter	Maternal		(Chotalia et al., 2009; Smith et al., 2003)
Igf2r	Prox 17	12934626 – 12935815 12962643 – 12962696	Within *Airn* promoter (ICR) Within *Igf2r* promoter	Maternal Paternal	*Airn*, *Slc22a2*, *Slc22a3*	(Stoger et al., 1993; Wutz et al., 1997)
Impact	Cent 18	13131356 – 13133257	*Impact* promoter/exon 1	Maternal		(Okamura et al., 2000)

Table 1.2: Mouse germ line and somatic DMRs associated with imprinted loci. DMRs methylated in either **maternal** or **paternal** germ line are indicated. Genes in red are **maternally expressed**. Genes in blue are **paternally expressed**. * - genomic locations of DMRs are taken from Xie et al (2012). † - *Gnas* is biallelically expressed in most tissues, but shows biased maternal expression in tissues like brown fat and kidney (Yu et al., 1998)

of imprints for several imprinted genes of the locus. Of the 21 known gDMRs implied in regulation of imprinting, most of the gDMRs (17) acquire methylation in growing oocytes while only 4 of them inherit methylation from the paternal germ line (Chotalia et al., 2009; Hiura et al., 2006; Kobayashi et al., 2009; Lucifero et al., 2004; Ueda et al., 1992). Table 1.2 lists some important and well-studied germ line and somatic DMRs.

1.3.2 IMPRINTING CONTROL REGIONS:

ICRs are termed so because they constitute a cis-acting regulatory region that controls the imprinting status of the genes at an imprinting locus. The ICRs usually overlap with the gDMR for a particular locus. There can be several DMRs in an imprinted locus and a DMR can encompass the whole gene body but the ICRs are usually limited to only a sub region within one of them (Bartolomei and Ferguson-Smith, 2011). Apart from allele-specific DNA methylation, an important biochemical criterion for an ICR is allele-specific chromatin organization. All the ICRs discovered till date have been found to have allele-specific chromatin organization (Kato and Sasaki, 2005; Khosla et al., 1999). Moreover, genetically deleting an ICR leads to loss of imprinted expression of genes in the locus (Lin et al., 2003; Sleutels et al., 2002). Provided below are examples of some of the known ICRs:

*1.3.2a H19/Igf2 ICR***:**

The *H19*/*Igf2* locus was the first imprinted locus to be identified(Bartolomei et al., 1991). It is present on distal end of mouse chromosome 7 and contains three imprinted genes, two paternally expressed protein coding genes, *Igf2* and *Ins*, and one maternally expressed ncRNA, *H19* (DeChiara et al., 1991; Giddings et al., 1994). Both *H19* and *Igf2* are regulated by endodermal and mesodermal enhancers present 9 kb and 11 kb downstream to the *H19* gene (Drewell et al., 2002; Leighton et al., 1995). There are 4 DMRs identified at this locus (3 paternally and 1 maternally methylated, Table 1.2

and Figure 1.1). The *H19*/*Igf2* ICR is located 2 kb upstream to the TSS of *H19* gene. It is referred to as *H19*-DMD (differentially methylated domain), and is part of the paternally methylated *H19*-DMR (Bartolomei et al., 1993; Feil et al., 1994; Ferguson-Smith et al., 1993; Moore et al., 1997). Deletion of this region results in upregulation of *H19* gene and repression of *Igf2* expression upon paternal transmission. A reciprocal effect (down regulation of *H19* and upregulation of *Igf2*) is seen when the same deletion is inherited from the mother (Thorvaldsen et al., 1998). The *H19*-DMD has 4 CpG-rich repeats which are recognized and bound by the CCCTC-binding factor, CTCF (Bell and Felsenfeld, 2000). As discussed earlier in this chapter, CTCF occupies these sites on the maternal chromosome thereby insulating the interaction between downstream enhancers and the promoter of *Igf2*. However, this does not happen on the paternal chromosome as the binding of CTCF is prevented by the methylation of its target sites (Murrell et al., 2004).

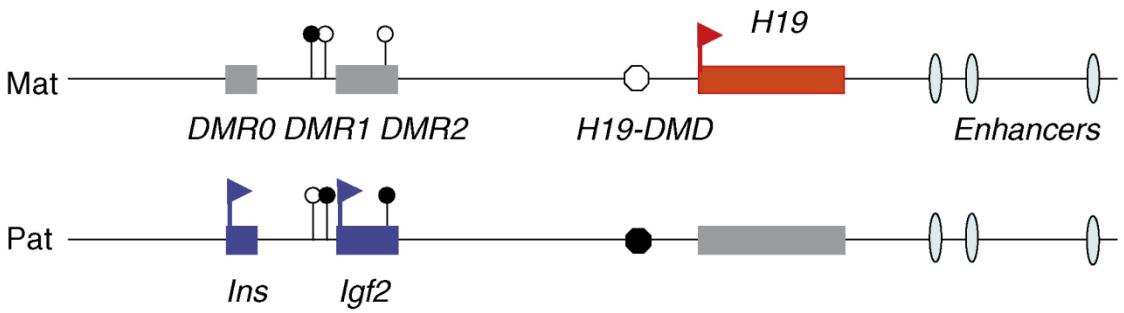


Figure 1.1: *H19/Igf2* locus. Genes are indicated as horizontal rectangles. **Maternally expressed allele is shown in red**, **paternally expressed allele is in blue**. **Grey rectangles indicate repressed alleles**. Raised arrows indicate transcription and its orientation. Mat – maternal allele, Pat – paternal allele. *H19*-DMD is indicated as octagon, lollipops indicate DMRs. Filled and unfilled octagons and lollipops indicate their methylated or unmethylated state respectively. *H19*/*Igf2* enhancers are shown as ellipses.

1.3.2b Igf2r/Air ICR:

Located at the proximal end of chromosome 17 in mice, the *Igf2r*/*Air* region contains three maternally expressed protein coding genes (*Igf2r*, *Slc22a2*, *Slc22a3*) and one paternally expressed antisense ncRNA, *Air* (Barlow et al., 1991; Zwart et al., 2001).

Slc22a1, another gene present at the locus, is biallelically expressed. The ICR of this locus encompasses the *Air* promoter and is located within the intron of *Igf2r* (Fig. 1.2). Deletion of a 3.7 kb region that leads to the truncation of *Air* transcript results in biallelic expression of all maternally expressed genes, indicating that *Air* has the potential to repress the paternally inherited alleles (Sleutels et al., 2002).

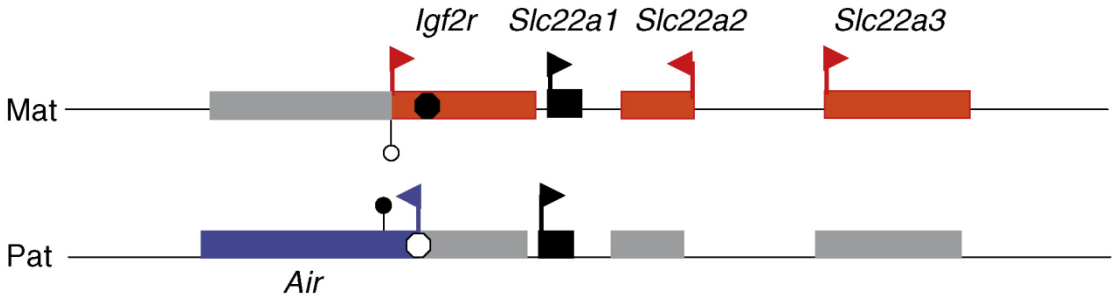


Figure 1.2: *Igf2r/Air* locus. Genes are indicated as horizontal rectangles. **Maternally expressed allele is shown in red**, **paternally expressed allele is in blue**. **Grey rectangles indicate repressed alleles**. Raised arrows indicate transcription and its orientation. Mat – maternal allele, Pat – paternal allele. The ICR for the locus is indicated as octagon, lollipops indicate secondary DMRs. Filled and unfilled octagons and lollipops indicate their methylated or unmethylated state respectively.

1.3.2c Kcnq1 ICR:

The mechanism of imprinting regulation observed at the *Kcnq1* locus present on mouse chromosome 7 is similar to that of *Igf2/Air* locus. Just as the *Igf2r/Air* locus, the *Kcnq1* cluster contains several protein coding genes (*Kcnq1*, *Ascl2*, *Cdnk1c*, *Phlda2* etc.) which are maternally expressed, and a single non coding RNA, *Kcnqot1*, which is paternally expressed and transcribed antisense to the other genes (Smilinich et al., 1999). The *Kcnq1* ICR includes the promoter of *Kcnq1ot1* and its truncation results in loss of imprinted expression of all the genes at that locus, including those that show placenta-specific imprinting (Mancini-Dinardo et al., 2006). However, it is not clearly understood whether the ncRNA *Kcnqot1* regulates the imprinted expression of other genes or whether the presence of an antisense transcript affects the process of transcription.

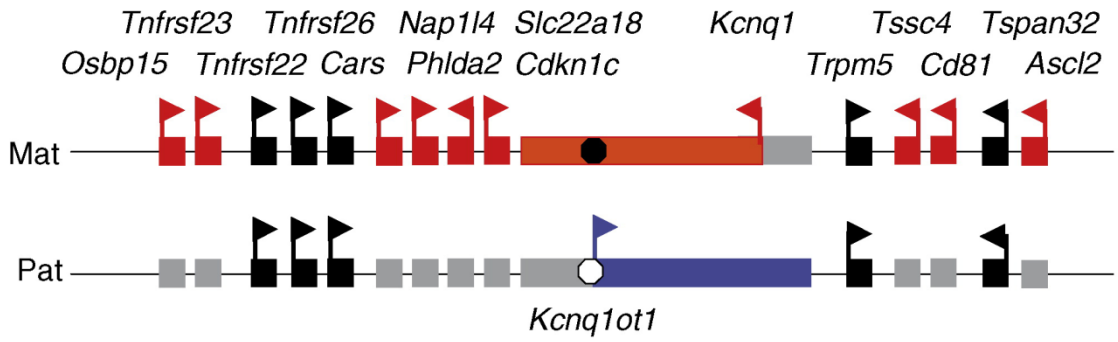


Figure 1.3: *Kcnq1* locus. Genes are indicated as horizontal rectangles. **Maternally expressed allele is shown in red**, **paternally expressed allele is in blue**. **Grey rectangles indicate repressed alleles**. Raised arrows indicate transcription and its orientation. Mat – maternal allele, Pat – paternal allele. *Kcnq1* ICR is indicated as an octagon. Filled and unfilled octagon indicates its methylated or unmethylated state respectively.

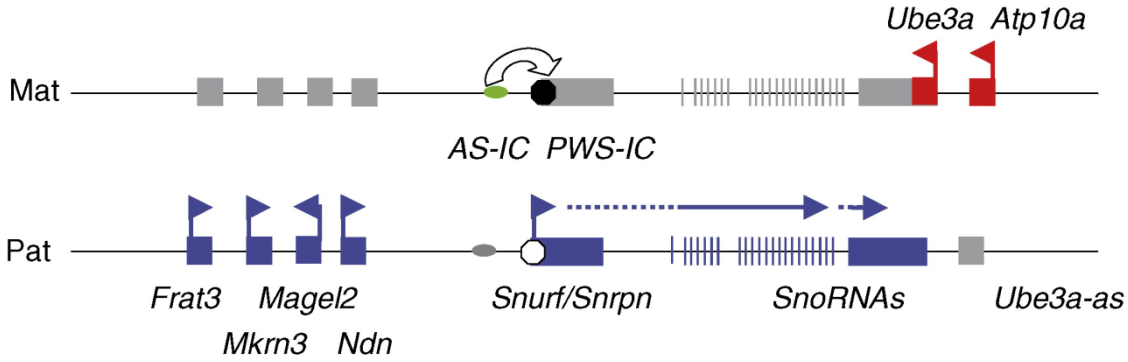


Figure 1.4: *Snrpn* locus. Genes are indicated as horizontal rectangles. **Maternally expressed allele is shown in red**, **paternally expressed allele is in blue**. **Grey rectangles indicate repressed alleles**. Raised arrows indicate transcription and its orientation. Mat – maternal allele, Pat – paternal allele. *Snrpn* ICR is indicated as octagon. Filled and unfilled octagon indicates its methylated or unmethylated state respectively.

1.3.2d Snrpn ICR:

The *Snrpn* locus is one of the largest imprinted domains located on mouse chromosome 7, and it includes several paternally expressed protein coding genes such as *Mkrn3*, *Magel2*, *Ndn*, *Snrpn* (Nicholls and Knepper, 2001). The region also encodes several paternally expressed ncRNAs, snoRNAs, and *Ube3a-as*, all of which are derived from a single long transcript regulated by alternative transcription from various promoters present upstream to that of *Snrpn* (Landers et al., 2004; Runte et al., 2001).

Ube3a and *Atp10a* are two maternally expressed genes of the locus (Albrecht et al., 1997; Kashiwagi et al., 2003). Failure of proper imprinting at this region results in two distinct neurological syndromes in humans, Prader-Willi Syndrome (PWS) and Angelman Syndrome (AS). Analysis of patients with these syndromes revealed microdeletions at this region and therefore helped map the ICR of the locus. As represented in Figure 1.4, the ICR of this region is a bipartite element, comprising of Prader-Willi Syndrome-Imprinting Centre (PWS-IC) and Angelman Syndrome-Imprinting Centre (AS-IC). The PWS-IC is considered a positive element that activates the gene expression on the paternal allele, and is located 4.3 kb upstream of *SNRPN* exon 1 (Ohta et al., 1999). The PWS-IC is methylated in oocytes in both mouse (Lucifero et al., 2002) and humans (Geuns et al., 2003), and deletion of this region in mice results in repression of all paternally expressed genes and activation of *Ube3a* (Yang et al., 1998). The AS-IC, a negative element, acts to silence the PWS-IC during oogenesis, thereby repressing the paternally expressed genes on the future maternal allele (Shemer et al., 2000). The AS-IC is an 880 bp region located 35 kb upstream to *SNRPN* and contains two alternative exons of *SNRPN* (Buiting et al., 1999). The AS-IC in mice has been recently identified and mapped to the region consisting of U1-U3 alternative exons of *Snrpn* (Smith et al., 2011).

1.3.2e Gnas ICR:

The *Gnas* cluster of imprinted genes located at distal end of chromosome 2 in mice contains 4 different genes (Fig. 1.5). The 3' portion of the cluster contains the *Gnas* gene which transcribes from three different promoters and first exons to form various tissue-specific transcripts. The protein coding *Gnas* shows biallelic expression in most tissues while showing preferential expression from the maternal allele in tissues such as brown fat and kidney (Yu et al., 1998). A transcript variant of *Gnas*, known as *Exon1A*,

is paternally expressed and encodes a ncRNA. Located upstream to the *Gnas* gene are the maternally expressed protein-coding *Nesp* (Peters et al., 1999) and a paternally expressed non-coding RNA transcribed antisense, called *Nespas* (Wroe et al., 2000). Two maternally methylated gDMRs have been identified within this locus. The deletion of the DMR located within the promoter of *Exon1A* results in biallelic expression of *Gnas* in tissues where it normally shows biased maternal expression, but does not affect any other genes of the cluster (Liu et al., 2005; Williamson et al., 2006). The second germ line DMR, which is rather large in size and encompasses the promoter of *Nespas* is perhaps the ICR of this locus, (Coombes et al., 2003) as paternal transmission of a 1.6 kb deletion within this DMR resulted in loss of imprinting of all the genes of the cluster and the offspring died 2 days after birth. However, maternal transmission (unmethylated allele) of the deletion did not show any effect and the offspring were viable and fertile (Williamson et al., 2006).

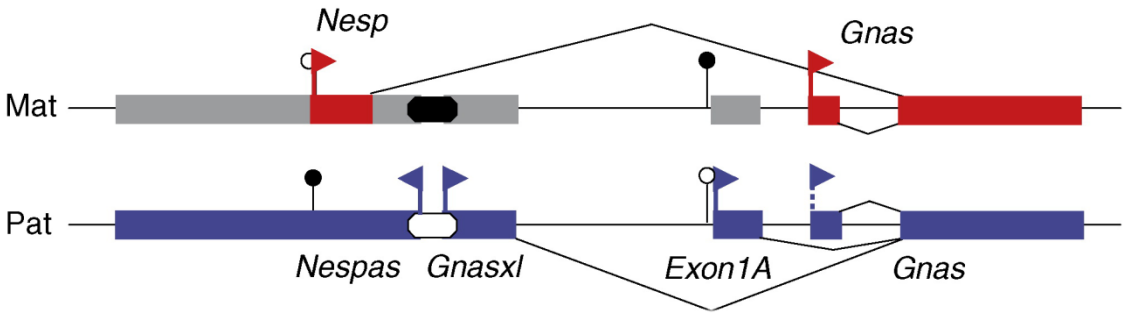


Figure 1.5: *Gnas* locus Genes are indicated as horizontal rectangles. **Maternally expressed allele is shown in red**, **paternally expressed allele is in blue**. **Grey rectangles indicate repressed alleles**. Raised arrows indicate transcription and its orientation. Mat – maternal allele, Pat – paternal allele. Germ line DMRs are indicated as octagons, lollipops indicate secondary DMRs. Filled and unfilled octagons and lollipops indicate their methylated or unmethylated state respectively.

1.3.2f Dlk1/Dio3 ICR:

This is another example of a paternally methylated ICR. *Dlk1-Dio3* region, present on mouse chromosome 12, contains three protein-coding genes that are paternally expressed – *Dlk1*, *Dio3* and *Rtl1* (Takada et al., 2000). Several ncRNAs like

Gtl2, *antiRtl1*, *sno*RNAs and microRNAs are maternally expressed (Cavaille et al., 2002; Seitz et al., 2004; Seitz et al., 2003). However, it is hypothesized that all the ncRNAs on the maternal chromosome arise from a single transcript regulated by *Gtl2* associated elements (Schuster-Gossler et al., 1998). Four different DMRs are identified for the locus. The intergenic DMR (IG-DMR) located 10-15 kb upstream of *Gtl2* acts as the ICR of the locus (Lin et al., 2007). The IG-DMR acquires methylation in paternal germ line. Transmission of deletion of IG-DMR from female germ line results in a switch of the expression profile with biallelic expression of protein-coding genes, but the paternal transmission of the IG-DMR deletion does not result in changed expression of any gene (Lin et al., 2003). IG-DMR is similar to *H19*-DMD as they both are located upstream of ncRNA promoters and are paternally methylated (Takada et al., 2002). However, the underlying mechanisms seem different because the IG-DMR does not contain any CTCF sites and lacks the insulator activity exhibited by the *H19*-DMD (Paulsen et al., 2001). Moreover, the paternal transmission of deletion of IG-DMR does not affect the imprinting status of the locus.

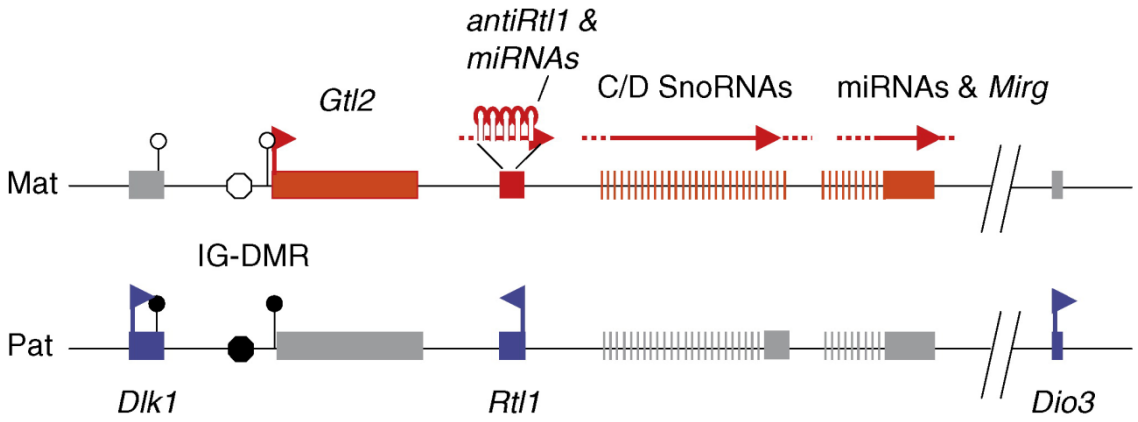


Figure 1.6: *Dlk1/Dio3* locus. Genes are indicated as horizontal rectangles. **Maternally expressed allele is shown in red**, **paternally expressed allele is in blue**. **Grey rectangles indicate repressed alleles**. Raised arrows indicate transcription and its orientation. Mat – maternal allele, Pat – paternal allele. The ICR for this locus is indicated as an octagon, lollipops indicate secondary DMRs. Filled and unfilled octagons and lollipops indicate their methylated or unmethylated state respectively.

1.4 AIM OF THE THESIS:

1.4.1 IMPRINTED MOUSE *NEURONATIN* GENE:

Neuronatin was first identified as a gene selectively expressed during mouse post-natal brain development, in a screen looking for genes differentially expressed during neural development (Joseph et al., 1994). The approximately 1.2 kb long RNA encoded by the *Neuronatin* gene was shown to produce two variants of Neuronatin protein, alpha and beta (Joseph et al., 1995). Both the transcript variants have been shown to code for putative transmembrane proteins with hydrophobic N-terminus and hydrophilic C-terminal end. *Neuronatin* has been predicted to be involved in signal transduction, cell to cell communication and cell adhesion (Joseph et al., 1995; Wijnholds et al., 1995). *Neuronatin* exhibits tissue specific expression – though it is expressed in tissues like Brain, Adrenal Gland, Pancreas etc, it remains silent in tissues such as Liver. Even in brain it has been shown to exhibit segment specific expression (Wijnholds et al., 1995). Though initially the only attributed role of *Neuronatin* was in hindbrain development, it has been shown recently to play crucial roles in insulin secretion, diabetes and obesity, and is a known interacting factor of Leptin (Joe et al., 2008; Mzhavia et al., 2008; Suh et al., 2005; Vrang et al., 2010).

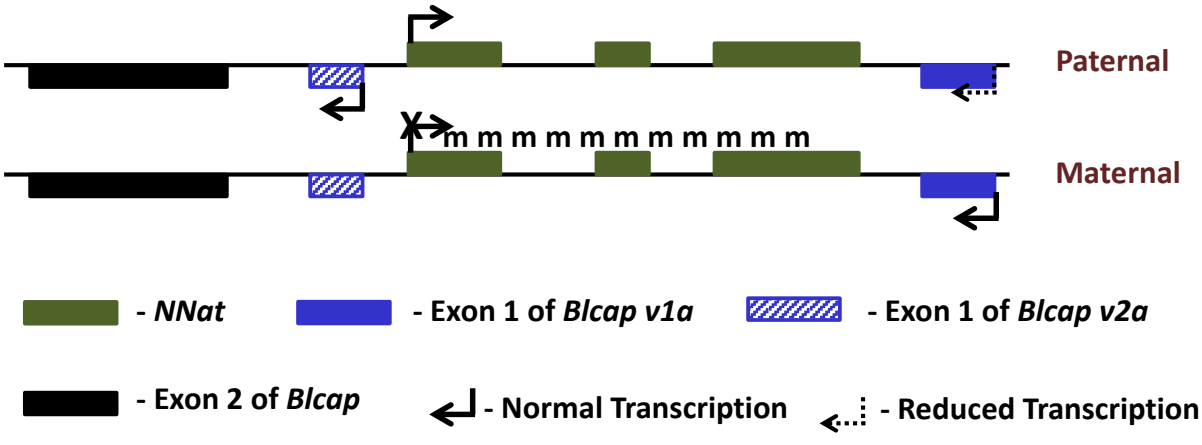


Figure 1.7: *Neuronatin* locus. 'm' indicates methylation.

At the genomic level, the *Neuronatin* gene is approximately 2.5 kb long, with three exons and two introns (Fig. 1.7). The alpha isoform has all the three exons while splicing out of middle exon results in the beta isoform. *Neuronatin* is conserved in rat, mouse and human. Two separate studies identified the imprinted expression of *Neuronatin.* Using subtractive hybridization, Kaginati et al (1997) showed loss of *Neuronatin* expression in E8.5 parthenogenetic embryos. The same was shown using differential display by Kikyo et al (1997). Mouse *Neuronatin* is located at distal end of chromosome 2 in mice (Kagitani et al., 1997), between the T26H and T2Wa breakpoints (Kikyo et al., 1997), whereas it maps to chromosome 20q11.2 in humans (Evans et al., 2001). Careful analysis of genomic organization at *Neuronatin* locus indicated that the entire gene is present within the single intron of another gene, *Bc10/Blcap*, which encodes for a non-coding RNA (John et al., 2001). Though initially *Bc10* was found to be expressed bi-allelically and thought to be a non-imprinted gene, it was recently shown to have alternate splicing and transcript- and tissue-specific imprinting (Schulz et al., 2009). As *Bc10* encodes for a non-coding RNA and is transcribed antisense to *Neuronatin*, it was initially thought to play a role in regulation of imprinted expression of *Neuronatin*. However, studies on a 30 kb transgene that contained *Neuronatin* but excluded the promoter and first exon of *Bc10* showed an imprinted expression of *Neuronatin* from the transgene, indicating that *Bc10* does not play role in regulation of imprinting at this locus (John et al., 2001). Therefore, unlike for other clustered imprinted loci whose enhancers are shown to act over long distances (Edwards and Ferguson-Smith, 2007; Lewis and Reik, 2006), it was suggested the imprinting control region for *Neuronatin* locus must be small and located within the 8.5 kb long intron of *Bc10* (John et al., 2001; Sowpati et al., 2008).

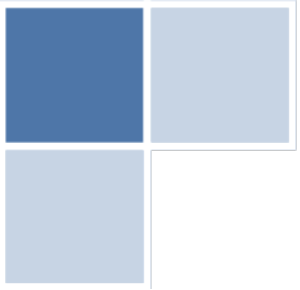
The transcriptionally silent maternal allele of *Neuronatin* is shown to be methylated (Kikyo et al., 1997) and the differentially methylated region (DMR) of *Neuronatin* locus extends from the promoter to the end of third exon of the gene (John et al., 2001). DMRs of imprinted loci are known to span across several kilobases and encompass whole genes (Thorvaldsen et al., 1998; Williamson et al., 2006), In contrast, ICRs for imprinted loci are usually limited to a small region within a DMR. Importantly, ICRs are the only regions within the DMRs that show a differential chromatin organization on parental alleles. This has been shown either as regions of differential nuclease sensitivity or of differential histone modifications (Delaval and Feil, 2004; Kato and Sasaki, 2005; Khosla et al., 1999; Lewis et al., 2006). The nuclease hypersensitive sites were always found on the unmethylated allele, suggesting a mutual exclusiveness of DNA methylation and nuclease sensitivity on the two parental alleles. It has been suggested previously that this mutual exclusiveness indicates a possible mechanism by which ICRs are protected from methylation, as it is unclear whether differential histone profile is a cause or consequence of mono-allelic expression (Sowpati et al., 2008).

Previous analysis of chromatin organization at *Neuronatin* locus revealed the presence of two DNase I hypersensitive sites – one of which mapped to its promoter and the other to its second intronic region. The hypersensitive site at the promoter region was only found in tissues where *Neuronatin* was expressed, but the hypersensitive site that mapped to the second intron was present in all tissues tested, irrespective of the transcription status of *Neuronatin*. Interestingly, the hypersensitive site of the second intron was only present on the unmethylated paternal allele, from which *Neuronatin* is transcribed. As the second intronic region satisfied the biochemical criterion of mutual exclusiveness of DNA methylation and specialized chromatin organization on the parental alleles, it was proposed to be the putative ICR of this locus.

In this thesis, we have tested this hypothesis. Firstly, we examined the role of DNA sequences present within the second intron of *Neuronatin* in transcriptional regulation of a reporter gene in *Drosophila* P-element reporter gene assay. In addition, we analysed the mice deleted for second intron of *Neuronatin* to examine whether it has a role in initiation, establishment and maintenance of imprinting at *Neuronatin* locus in its normal genomic context. Finally, we tested two proteins for their ability to bind within the second intron in a methylation-dependent manner. Our results indicate that the second intronic region of *Neuronatin* acts as the imprinting control region for this locus.

CHAPTER II

FUNCTIONAL ANALYSIS OF NEURONATIN'S SECOND INTRON IN DROSOPHILA



2.1 INTRODUCTION:

Biochemical characterization previously done in our laboratory revealed that the two parental alleles of the mouse *Neuronatin/Blcap* locus are organised in different chromatin conformations. As described in Chapter I, the transcribed paternal allele of *Neuronatin* showed presence of two DNase I hypersensitive (HS) sites, which were not present on the methylated maternal allele. The HS site that localised to the second intron of *Neuronatin* was found to be independent of its transcriptional status. This suggested that the second intron could have a regulatory role in the expression of *Neuronatin*. Moreover, this mutual exclusiveness between nuclease sensitivity and DNA methylation has been defined as a biochemical characteristic of imprinting control regions (ICRs) (Coombes et al., 2003; Khosla et al., 1999; Schweizer et al., 1999). To analyse whether this indeed is true, we decided to characterize the second intron of *Neuronatin* for its capabilities in regulating transcription, and more importantly in controlling the imprinting status of the *Neuronatin/Blcap* locus.

This chapter describes our efforts in examining the role of DNA sequences present within the second intron of *Neuronatin* (NNI^2) in transcriptional regulation of a reporter gene in *Drosophila* P-element reporter gene assay. *Drosophila* was chosen, as it is an excellent model organism for studying putative regulatory regions with well-defined genetic protocols. In addition, there is negligible CpG methylation reported in *Drosophila* (Krauss and Reuter, 2011). Therefore, the intron when introduced will remain unmethylated, thereby mimicking the paternal allele on which the transcription-independent HS site was observed.

P-elements, one of the most studied transposons, are present specifically in *Drosophila* (Engels, 1992). An autonomous P-element is 2907 bp long, featuring a 31 bp

terminal inverted repeat and an 11 bp sub-terminal inverted repeat (O'Hare and Rubin, 1983). These repeats are important but not completely sufficient for transposition (Mullins et al., 1989). The donor site is excised and reinserted at the recipient site, thereby resulting in a direct duplication of 8 bp (O'Hare and Rubin, 1983). The middle sequence encodes for the transposase, an 87 kDa protein, that is required in *trans* for the mobility and integration of P-elements (Kaufman et al., 1989). The P-elements in wild type flies give rise to a variety of phenotypes collectively known as hybrid disgenesis (Kidwell et al., 1977). P-element transposition has become one of the indispensible tools of *Drosophila* genetics because the transposition will occur provided that the terminal repeats and an externally supplied transposase are present. This allows replacing the internal transposase with a gene of interest, most often a reporter or a marker. Though the P element insertions have been found at thousands of genomic loci, not all sites are equally susceptible to the event. However, euchromatic regions are often more vulnerable than heterochromatic sites, with the frequency ranging from 10^{-2} to 10^{-6} (Berg and Spradling, 1991).

2.2 MATERIALS AND METHODS:

2.2.1 PREPARATION OF CONSTRUCTS:

2.2.1a Choice of vectors:

The vector pLML was chosen for the initial cloning primarily because of the *LoxP* sites present on either sides of its MCS region. The region of interest now flanked by *LoxP* sites was sub-cloned into the pCaSpeR vector, one of the most used vectors for P-element transposition. This vector contains the required terminal inverted P repeats, and the original transposase gene is replaced by the reporter *miniwhite* gene (results in

red eye color) that is driven by the *hsp70* promoter. The MCS is located just upstream of the promoter region. The presence of *LoxP* sites facilitates the removal or flipping-out of the region of interest by introducing Cre recombinase, hence allowing the possibility to validate the effect seen on the expression of reporter gene. The strategy for construction of reporter gene cassettes is outlined in Figure 2.1.

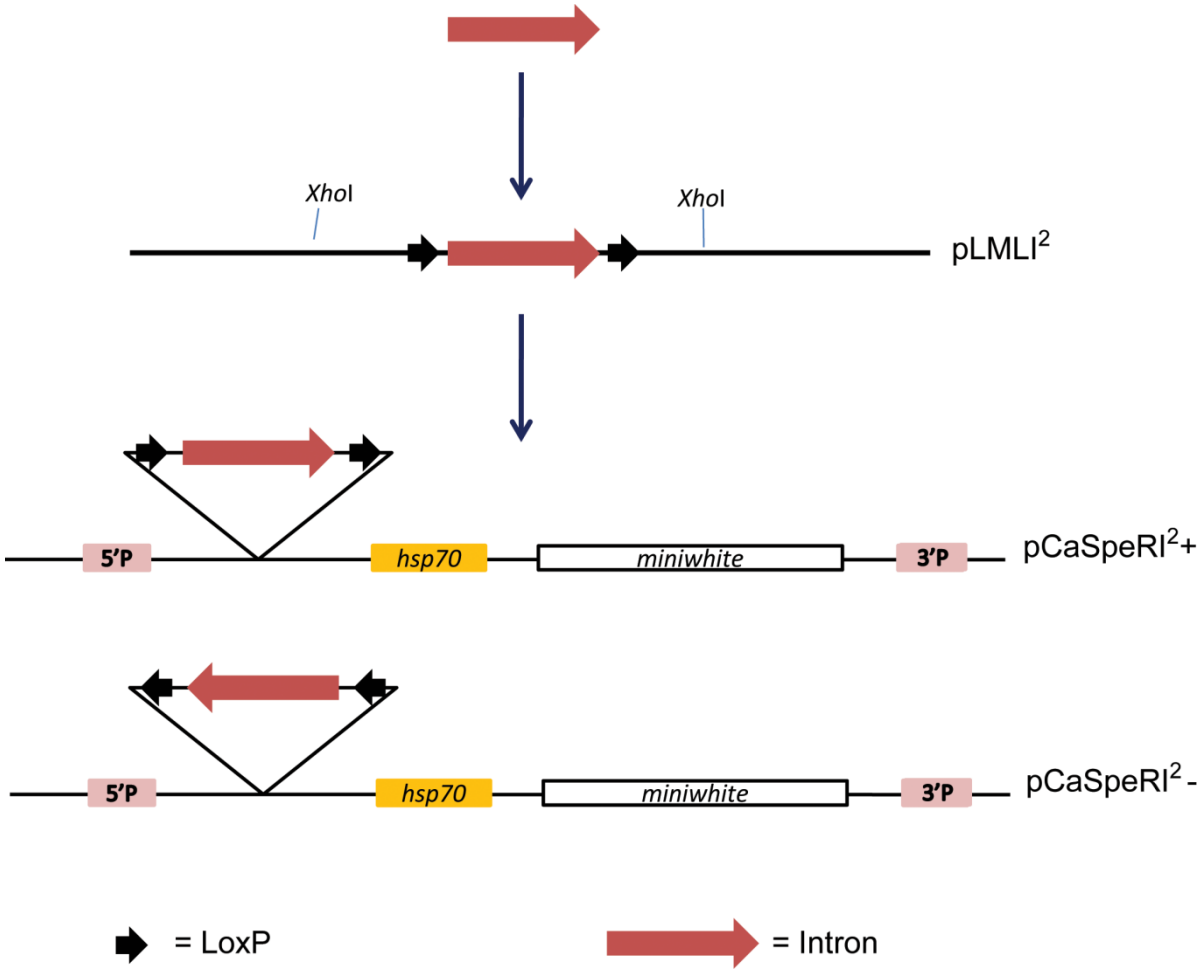


Figure 2.1: Strategy for reporter gene cassette construction. The vector contained *miniwhite* reporter gene under the control of *hsp70* promoter. The 250 bp NNI^2 flanked by *LoxP* sites was introduced upstream to *hsp70* promoter in both orientations (indicated by arrows). 5'P and 3'P are P elements present on either side of the reporter cassette.

2.2.1b Construction of pLMLI^2:

The intron was first cloned into the vector pLML. To achieve this, the second intron of *Neuronatin* was amplified using PCR. The PCR product was purified and

ligated with pLML/*Sma*I. The ligation mix was transformed into *E.coli* (DH10B) and positive clones were screened by digestion with *EcoR*I and *Hind*III. Positive clones were sequenced using insert specific primers to analyse the orientation of the insert. The clone with intron in positive orientation (same as in mouse genomic context) was named pLMLI^{2}+ and the clone with intron in negative orientation (the reverse complement) was named pLMLI^{2}-. The primers used for amplification of NNI^{2} are:

Forward: TGAGGTATACTTAAGTTGTGGGTCC

Reverse: CACCTGCGGTGAGAGACCCAGGAC

2.2.1c Construction of pCaSpeRI^{2}+ and I^{2}-:

The intron with *LoxP* sites on either side was sub-cloned into the vector pCaSpeR. For this, the plasmid pLMLI^{2}+ was isolated and digested with *Xho*I to release the fragment consisting of intron surrounded by *LoxP* repeats on either side. Simultaneously, pCaSpeR vector was digested with *Xho*I that cuts upstream of the *hsp70* promoter. The vector and insert were ligated and transformed into *E.coli* (DH10B). The positive clones were screened by digestion with *Xho*I and the orientation of the insert was analysed by sequencing. The clone with intron in positive orientation with respect to the promoter was named pCaSpeRI^{2}+ and its reverse complement was named pCaSpeRI^{2}-. Both the constructs were isolated and purified using QIAGEN MidiPrep kit (Cat. # 12143).

2.2.2 MICROINJECTION:

2.2.2a Collection of embryos:

100-200 white eyed mutants of *Drosophila* (w^{1118}) were transferred per cage, (cages made by taping an egg-laying plate to the bottom of a plastic beaker with air

holes punched with a needle (20 gauge or smaller)). The flies were moved to day-for-night schedule (exposed to light during night and kept in dark during day time) 2 days before collections to improve the number of eggs laid during day time. Before the first round of injection, a fresh egg laying plate was placed in the cage for 1 hour to induce the flies to lay any overdeveloped eggs. After that, eggs were collected every 30 – 60 min on a new egg-laying plate and injected within the next 45 min, before cellularization takes place. Once the eggs were collected, they were brushed from agar onto an egg strainer using a soft, wet paint brush.

2.2.2b Treatment of embryos:

The eggs were washed with water or PBS and excess wash was blotted onto a paper towel. The eggs were dechorionated by placing the egg strainer in a petridish containing 50% bleach in water for 3 min with occasional swirling. The eggs were then rinsed thoroughly with water and transferred to a thin strip of agar using a fine paintbrush. The eggs were then aligned to their poles in a straight line with a needle and transferred onto the edge of a coverslip that was previously brushed with tape glue or rubber glue such that the animal poles of the embryos faced outwards. After 40 – 50 embryos were lined up, the coverslip was placed in a desiccator for 7 min. After desiccation, the coverslip was attached to a micro-slide such that the end of the coverslip with the embryos hangs 2 – 3 mm over the edge of the slide. The embryos were then covered with a thin layer of halocarbon oil to prevent further desiccation and to provide moisture.

2.2.2c Injection of DNA into embryos:

The construct DNA was precipitated and dissolved in injection buffer solution (contains DNA and transposase plasmid in 3:1 ratio). The injector needle was attached to the micro-injector and the output volume was calibrated. The needle was inserted

gently into the animal pole of the embryo with the tip precisely orthogonal to the membrane at the site of insertion. The DNA solution was injected into the embryo using a pneumatic pump and the needle was gradually removed to avoid plasma leakage. The slide/coverslip sandwiches with injected embryos were placed in a humidified chamber at 18 – 25°C. The edges of coverslips with oil and embryos were slightly tilted downwards (2°) to prevent the oil from spreading over the entire coverslip and exposure of the embryos to the air.

2.2.3 CROSSES AND BALANCING:

2.2.3a Fly nomenclature:

The entire chapter follows the traditional way of listing flies and their genotypes, as explained below. The genotype is always written in the order of X/Y, 2, 3 and 4 chromosomes, and the genotype of each chromosome pair is separated with a ';' (e.g. *f*; *CyO*; *Tm2* – represents the genotype of one each of the first, second and third chromosomes). The genotype of the chromosome is only listed when it differs from wild type (e.g. *Fm3*; *Tm6* – indicates the balancers present in the fly for X and third chromosomes). When representing the genotype of the same pair of chromosomes, the two chromosomes are separated using a '/' (e.g. *CyO*/*Pin*; *Tm2*/*Tm6* – indicates the genotype of flies with four balancer chromosomes, two each for second and third chromosomes respectively). Two mutant loci on the same chromosome are written in continuation with a space in between (e.g. *Tm6 Tb*). Genotypes and mutants are always italicized; small letters indicate a recessive phenotype while dominant alleles are written either in complete capitals or only the first letter is capitalized. The gene names in *Drosophila* often indicate their respective mutant phenotypes (e.g. *white* gene gives red eye color while a mutant *white* gene results in a white eye).

2.2.3b Balancers:

Balancers are chromosomes in which the gross sequence is scrambled due to multiple rearrangements induced by radiation. These rearrangements predominantly include inversions, but translocations, deletions, duplications etc. are also used. The scrambled sequence renders the balancers incapable of pairing or recombining with their normal homologues during meiosis. In addition, they carry dominant marker mutations to track their presence in the fly, as well as recessive markers which are recessive lethals in most cases, to prevent any fly from receiving two copies of the same balancer. The presence of a balancer negatively impacts the health of the stock. Fly stocks carrying multiple balancer chromosomes (e.g. double balancer flies, having four balancer chromosomes for second and third chromosomes) are often pretty weak and need to be maintained at higher numbers than wild type stocks.

Many different balancers exist for first (X), second and third chromosomes. The fourth chromosome does not need a balancer as it shows negligible amount of recombination. The balancers are often named following a well-established convention; the first letter of the name indicates the chromosome pair it balances (F, S and T used for first, second and third chromosomes respectively), the second letter is usually an M which stands for "multiply inverted" while the third number is used to name a series of balancers for the same chromosome. It is also conventional to write the dominant marker mutation after the name (e.g. *TM6 Tb* – a balancer for third chromosome that carries the *Tubby* marker that results in shorter pupae compared to wild type pupae). Sometimes, the balancer is denoted using just the dominant marker mutations, e.g. *CyO* denotes a balancer for the second chromosome that carries the *Curly* gene. All the balancers used in this study are listed in Table 2.1.

Name	Chromosome	Marker	Phenotype
*Fm3**	I	*Ultra-Bar*	Eye is extremely bar shaped
Fm7a	I	*Bar*	Eye shape is bar
Pin	II	*Pin*	Body bristles are blunt
CyO	II	*Curly*	Wings are curled
Tm2	III	*Tm2*	Adult flies have bristles on halteres
Tm6	III	*Tm6, Tb*	More than 2 bristles on shoulders Pupae are shorter

Table 2.1: Balancing markers used in this study. The column 'chromosome' represents the chromosomal location of the marker. Marker – the phenotypic marker. The phenotype of each marker is mentioned in the last column. '*'- *Fm3* results in red eye color and hence cannot be used in this study for balancing.

2.2.3c Screening of transgenic lines:

Most larvae emerged from the injected embryos two to three days after incubation. The larvae were transferred to new food vials (~15 – 20 larvae per vial) and incubated at 25°C till they hatched into adult flies (approximately 10 days). These flies, named G0 flies, were crossed back to w^{1118} flies to assay for the eye color differences, and to get rid of any remnant transposase added before injection. The progeny of the cross were screened for eye color and every fly with colored eyes was considered an individual line.

2.2.3d Determination of site of integration:

Colored flies from each individual line were crossed to w^{1118} flies to expand the line. From the F1 progeny, a colored male was taken and crossed to w^{1118} virgins. The site of integration was considered to be on the first chromosome if all the male progeny of F2 cross were white-eyed and all female progeny had colored eyes (Fig. 2.2). All the

lines with integration on first chromosome were separated and balanced. Males from remaining lines were crossed to double balancer virgins (genotype *CyO/Pin; TM2/TM6*), which are referred to as DB virgins hereafter. Male F1 progeny with colored eyes and a marker each for second and third chromosomes was again crossed to DB virgins. The site of integration was then decided as explained in Figure 2.3.

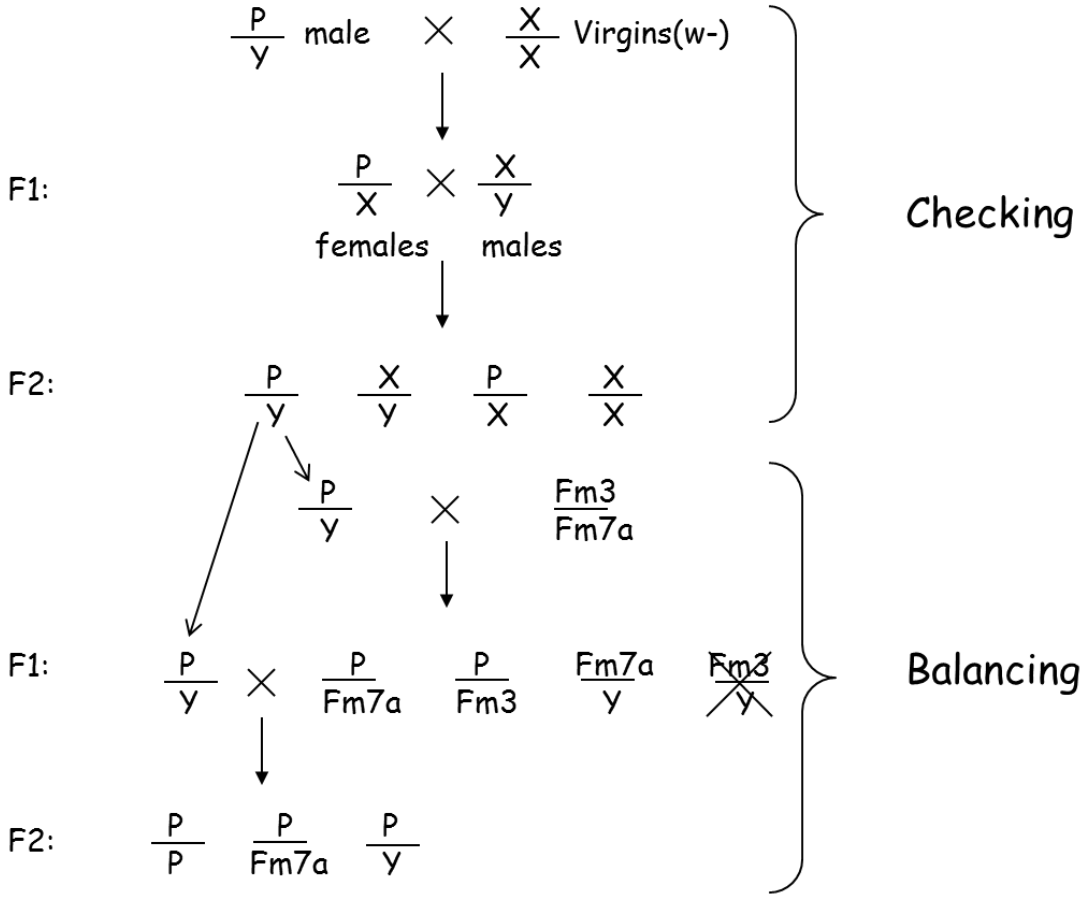


Figure 2.2: Strategy for creating balanced stock of transgenic *Drosophila* on first (X) chromosome. Genotypes of only the first chromosome are represented. P indicates X chromosome containing reporter transgene. X and Y indicate wild type X and Y chromosomes (applicable only to males) respectively. A cross is indicated with a 'X'. *Fm3* and *Fm7a* are balancing markers used for the first chromosome. F1 and F2 indicate first and second generations of progeny respectively. A crossed genotype indicates its lethality.

2.2.3e Balancing:

All the lines with integration on first chromosome were balanced using *Fm3*/*Fm7a* flies. A transgenic male was crossed to *Fm3*/*Fm7a* virgins. *Fm7a* females with colored eyes from F1 progeny were back-crossed to transgenic males to achieve a balanced homozygous stock (Fig. 2.2). For lines with integration on second and third chromosomes, *CyO* and *Tm6* balancers were chosen respectively. After the second cross with DB virgins, progeny with colored eyes and either *CyO* or *Tm6* (depending on chromosome of integration) were taken and self-crossed to achieve a balanced homozygous stock (Fig. 2.3).

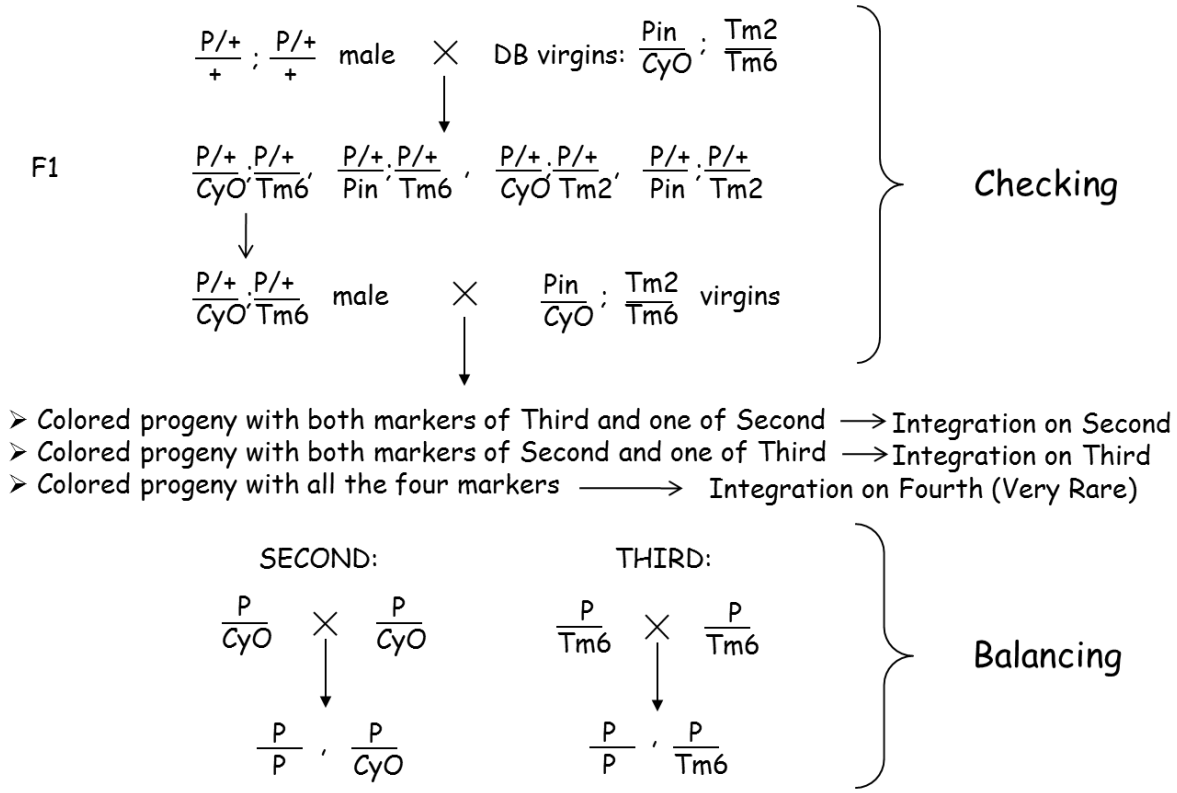


Figure 2.3: Strategy for creating balanced stock of transgenic *Drosophila* on second and third chromosomes. Genotypes of only the second and third chromosomes are represented. P indicates chromosome containing reporter transgene. + indicates wild type second or third chromosomes. A cross is indicated with a 'X'. *Pin* and *CyO* are balancing markers for the second chromosome, while *Tm2* and *Tm6* are third chromosome balancing markers. F1 indicates first generation of progeny.

2.2.3f Elimination of redundant lines:

After all the lines were balanced, eye color of both heterozygous and homozygous flies was analysed to remove any possible redundant lines. Two lines were considered redundant if they originated from the same G0 fly, have the same chromosome of integration, and the eye color of the flies is the same. All the redundant flies were discarded and the unique lines were used for further analysis.

2.2.4 FLIPPING OUT:

2.2.4a Cross with Cre flies:

Males from each unique transgenic line were taken and crossed with *Cre*/*CyO* virgins. Males from the F1 progeny were taken and crossed again to *Cre*/*CyO* virgins. Two to three males from each F2 cross were taken and considered as individual lines and crossed to w^{1118} virgins to expand the line.

2.2.4b Confirmation with PCR:

From the expanded lines, progeny with colored eyes were taken and processed for DNA isolation. A single fly per line was used as the starting material. The fly was taken into a 1.5 ml tube and stored at -20°C for 30 min. 100 µl of extraction buffer (25 mM NaCl, 10 mM Tris pH 8.0 and 1 mM EDTA) was added to each tube and incubated at 95°C for 10 min. The fly was crushed by passing it through a 200 µl pipette tip. To the crushed mixture, Proteinase K was added to a final concentration of 200 µg/ml and was incubated at 55°C overnight. The lysate was incubated at 95°C for 15 min to inactivate Proteinase K and centrifuged at 13000 rpm for 15 min. The supernatant was collected into a fresh 1.5 ml tube and was stored at -20°C till further use. 1 µl of this supernatant was used as template for PCR with primers targeting the pCaSpeR vector.

2.2.4c Cross with balancers:

Once the flies were confirmed for flip-out, one line for each original transgenic line was selected for further experiments. As the chromosome of integration is already known, they were directly crossed to suitable balancer flies. From the progeny, flies with suitable markers (*Fm7a*, *CyO* and *Tm6* for first, second and third respectively) were self-crossed to achieve balanced homozygous stocks.

2.2.4d Extraction of eye-pigment:

20 flies of same age and sex were selected from each line (original transgenic and the flipped-out counterpart) and etherized to their death. The heads were chopped using a fresh scalpel and collected into 1.5 ml micro-centrifuge tubes. The heads were homogenized in 200 µl of extraction buffer (0.1% NH_4OH and Chloroform, 1:1) using a micro-pestle. The extract was vortexed vigorously for 15 secs and left at RT for 15 min. The mix was centrifuged at maximum speed for 10 min and supernatant was collected into a fresh 1.5 ml tube. The OD of the solution was measured at 485 nm.

2.3 RESULTS:

2.3.1 DERIVATION OF NNI^2-SPECIFIC *Drosophila* TRANSGENIC LINES:

To examine the possibility that the second intron of *Neuronatin* (NNI^2) is able to influence transcription, reporter gene assay in *Drosophila* was conducted. The following strategy was used to introduce this intronic sequence in a vector containing the *miniwhite* gene driven by *hsp70* promoter. As described in Materials and Methods and shown in Figure 2.1, the intron was inserted upstream of the *hsp70* promoter in both orientations. Each of these constructs was injected into approximately 1000 embryos of w^{1118} flies.

LINE	CHROM	MARKER(P)	P/+ male	P/+ female	ΔP/+ male	ΔP/+ female
A 33.1	I	*Fm7a*	-	Dark Red	-	Dark orange
A33.8.1	I	*Fm7a*	-	Light Orange	-	Light Orange
A34.2.2	III	*Tm2*	Red	Red	Orange	Orange
A39.4.3	III	*Tm2*	Dark Red	Dark Red	Dark Red	Dark Red
A44.1.1	II	*CyO*	Orange	Orange	Dark Yellow	Dark yellow
A44.2.1	I	*Fm7a*	-	Dark Red	-	Dark orange
A86.1.1	II	*CyO*	Dark orange	Orange	Yellow	Yellow
A91.1	II	*CyO*	Dark red	Dark red	Dark red	Dark red
B55.1.1	II	*CyO*	Red	Red	Red	Red
B55.5.1	II	*CyO*	Yellow, variegation		Variegation goes down	
B88.1.1	III	*Tm2*	Orange	Orange	Orange	Orange
B88.7.2	I	*Fm7a*	-	Yellow	-	Yellow
B88.9.1	I	*Fm7a*	Dark orange, variegation		-	Yellow

Table 2.2: Eye color comparison for heterozygous transgenic lines. Eye color was assayed on 3 day-old flies. Lines highlighted in yellow show a color difference upon flip-out of NNI[2]. - indicates that the genotype is not possible

LINE	CHROM	MARKER(P)	P/P male	P/P female	ΔP/ΔP male	ΔP/ΔP female
A 33.1	I	*Fm7a*	Dark Red	Dark Red	Dark Orange	Orange
A33.8.1	I	*Fm7a*	Orange	Orange	Light Orange	Light Orange
A34.2.2	III	*Tm2*	Dark Red	Dark Red	Orange	Light Orange
A39.4.3	III	*Tm2*	Dark Red	Dark Red	Dark Red	Dark Red
A44.1.1	II	*CyO*	Orange	Orange	Yellow	Yellow
A44.2.1	I	*Fm7a*	Dark Red	Dark Red	Dark Orange	Orange
A86.1.1	II	*CyO*	Dark orange	Orange	Yellow	Yellow
A91.1	II	*CyO*	Dark red	Dark red	Dark Red	Dark Red
B55.1.1	II	*CyO*	Red	Red	Red	Red
B55.5.1	II	*CyO*	Yellow, variegation		Pale Yellow with Variegation	
B88.1.1	III	*Tm2*	Red	Red	Red	Red
B88.7.2	I	*Fm7a*	Yellow	Yellow	Yellow	Yellow
B88.9.1	I	*Fm7a*	Dark orange, variegation		Pale Yellow with Variegation	

Table 2.3: Eye color comparison for homozygous transgenic lines. Eye color was assayed on 3 day-old flies. Lines highlighted in yellow show a color difference upon flip-out of NNI[2].

The survival rate was approximately 10% with 113 larvae hatching for I^2+ and 84 larvae hatching for I^2-. Larvae were grown in groups of 15 - 20 per vial till eclosion, and each fly was back-crossed to w^{1118} flies. The F1 progeny included flies with varying shades of red eye color, and also a high percentage of white-eyed flies. All the white-eyed flies were discarded and each fly with colored eye was given a number with prefix 'A' or 'B' for I^2+ and I^2- respectively. These flies were considered G0 (generation zero) and each G0 fly was analysed for the site of P-element integration and crossed with appropriate balancers (see Materials and Methods). For the ease of scoring, all lines were balanced with *Fm7a*, *CyO* and *Tm6* for integrations on first, second and third chromosomes respectively. Once the homozygous stocks were made, all redundant lines were discarded (see Materials and Methods). Out of the initial 86 lines (46 for I^2+ and 42 for I^2-), only 13 were considered unique: 8 for I^2+ and 5 for I^2-. The intron was flipped-out from each of the unique lines and crossed to appropriate balancers to achieve balanced homozygous stocks (described in Materials and Methods).

2.3.2 EYE-COLOR ANALYSIS:

Eye color of heterozygous and homozygous flies of transgenic lines and their flipped-out counterparts was compared. The eye colors varied from light yellow to dark red, with few lines showing a variegated phenotype. An approximate shade of the eye color of each line and its flipped-out counterpart is outlined in Tables 2.2 and 2.3. As observed in Figures 2.4 and 2.5, 7 out of 13 lines showed a difference in eye color when NNI^2 was flipped-out, while the remaining 6 lines showed no difference. Out of these 7 lines, 5 had intron in positive orientation (I^2+), and 2 in negative orientation (I^2-). Interestingly, all the 7 lines showed a lighter shade of red color when the second intron was flipped out as compared to their normal counterparts.

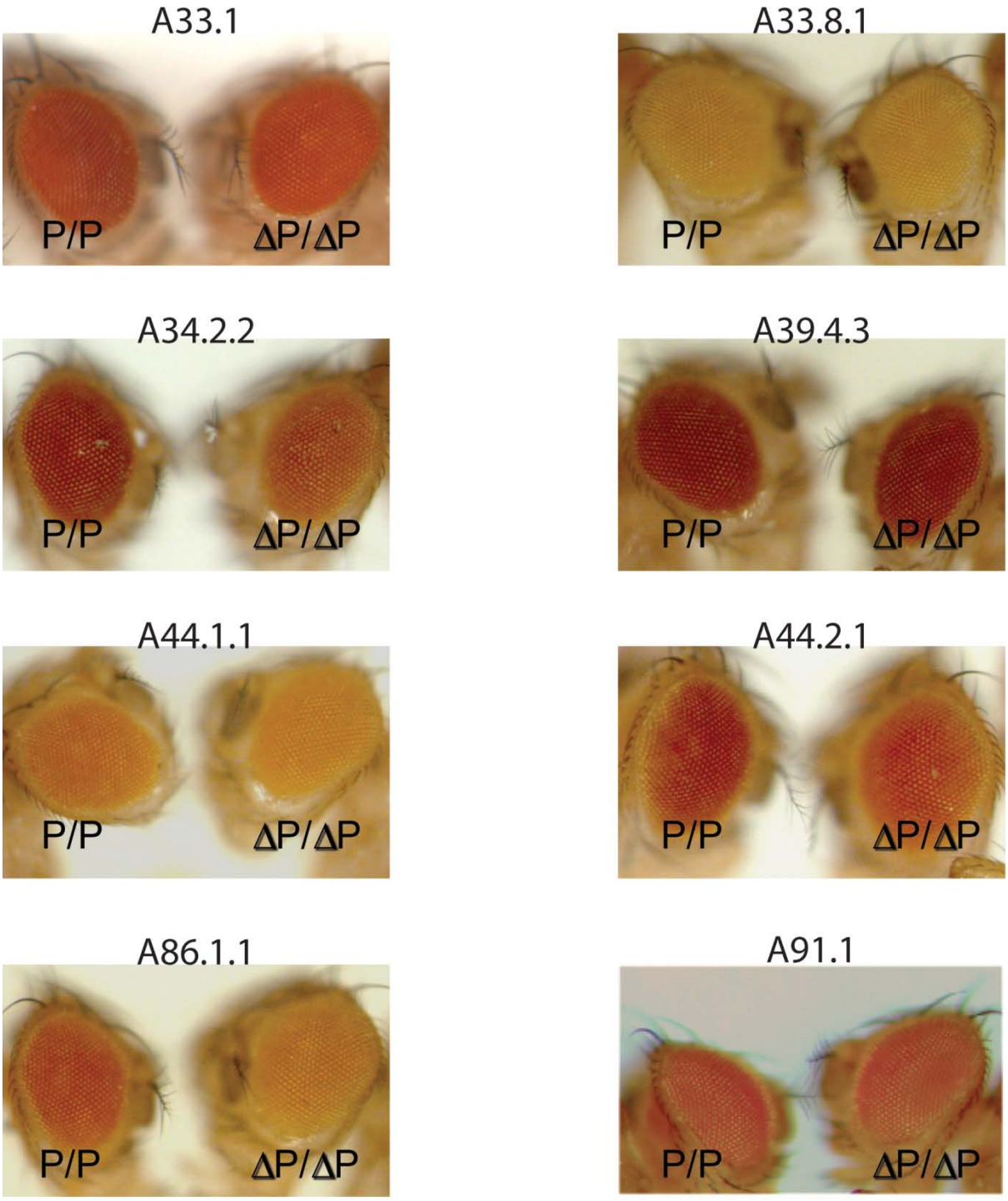


Fig. 2.4: Eye color comparison for the various homozygous NNI^{2}+ transgenic lines. Names of transgenic lines are indicated on top of each panel. P/P – homozygous transgenic line, ΔP/ΔP – their respective flipped-out counterparts. All pictures are for 3 day-old males.

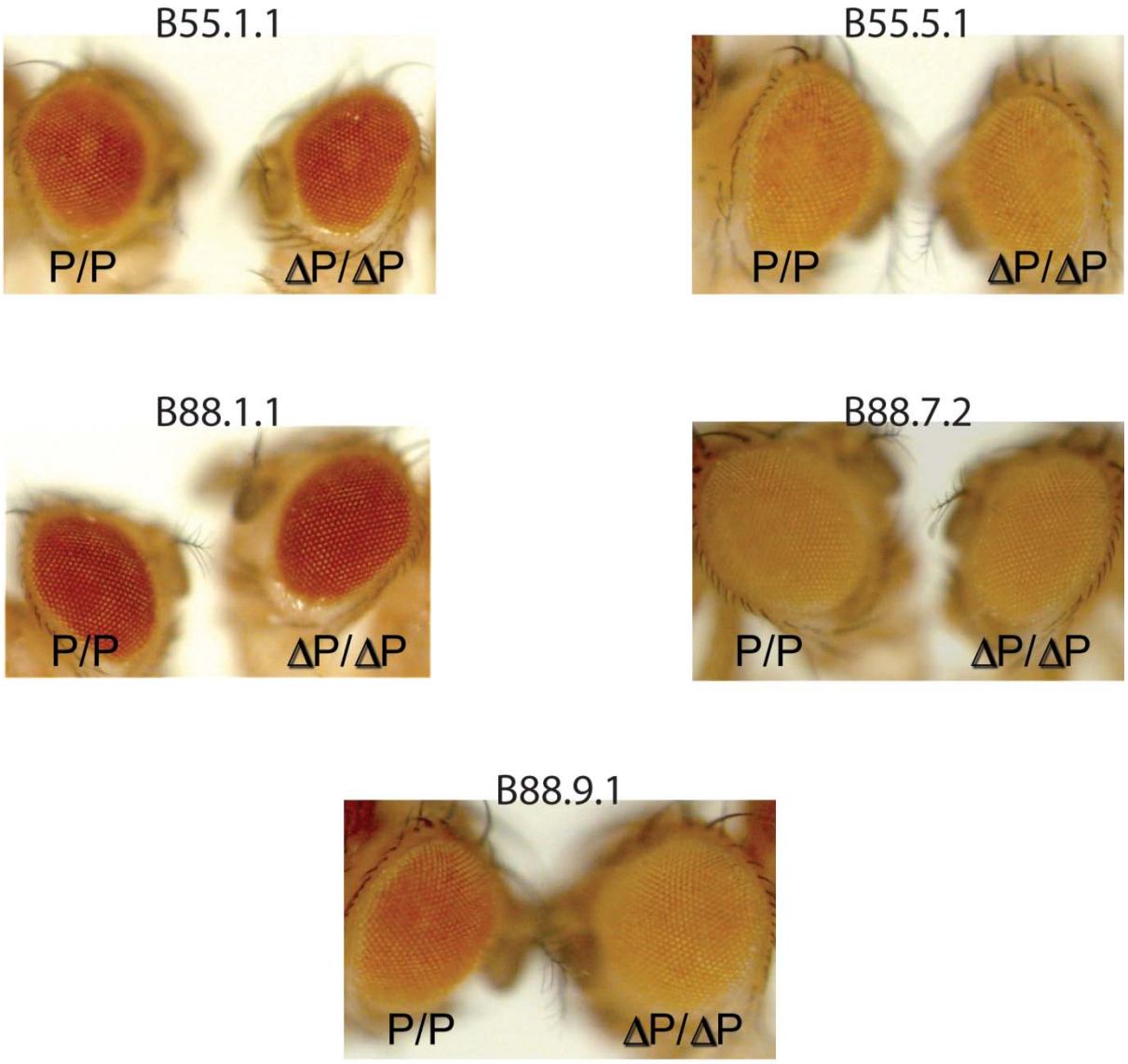


Figure 2.5: Eye color comparison for the various homozygous NNI^2- transgenic lines. Names of transgenic lines are indicated on top of each panel. P/P – homozygous transgenic line, ΔP/ΔP – their respective flipped-out counterparts. All pictures are for 3 day-old males.

To confirm these findings, quantitative estimation of total red pigment from the eyes of each NNI^2 line was performed. The values confirmed the difference observed visually. The difference observed was statistically significant for all the transgenic lines tested ($p < 0.05$, Student's t-test) (Fig. 2.6). This suggested that the putative ICR of *Neuronatin* could be acting as an activator of the reporter transgene.

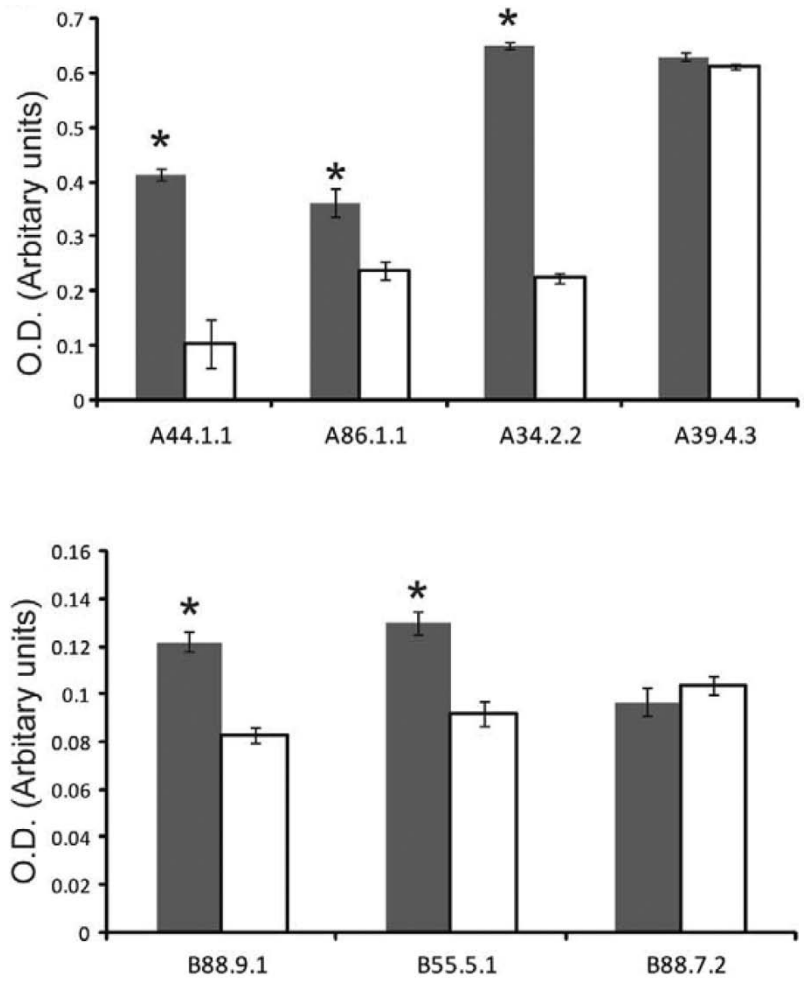


Figure 2.6: Eye pigment analysis for NNI[2] *Drosophila* transgenic lines. Columns shaded in grey indicate transgenic lines while the non-shaded columns represent their 'flipped-out' counterparts. Data is shown for representative lines only. * indicates significant difference (Student's t test, $p < 0.05$)

2.3.3 CROSSES WITH MUTANTS:

To examine if the effect seen on eye color is because of the interaction of *Neuronatin's* putative ICR with any known homologs of chromatin/histone modifying proteins in *Drosophila*, the 7 lines that were showing difference in eye color were crossed to various stock flies. Each stock carried a mutation of several known regulators of gene expression. The various stocks and their respective mutations are detailed in Table 2.4. Two lines (A39.4.3 and B88.7.2) that did not show a difference in eye color upon flip-out of intron were also crossed to the mutants as controls.

Mutation	Chrom	Balancer	Function of the protein	TG/PcG
Su(var)3-91	III	*Tm6*	H3K9 methyltransferase	PcG
Su(var)2-102	II	*CyO*	DNA binding	PcG
Psc	II	*CyO*	Zinc finger, chromatin remodelling	PcG
Pc1	III	*Tm6*	Gene silencing	PcG
Esc	II	*CyO*	H3K27 methyltransferase	PcG
Trx	III	*Tm6*	H3K4 methylation H3 acetylation	TG
Brm	III	*Tm2*	DNA binding and helicase activity, contains bromodomain	TG
Ash1	III	*Tm2*	H3K4 methylation	TG
TrlR85	III	*Tm6*	Zinc finger, chromatin binding	TG

Table 2.4: List of chromatin modulator mutants used in this study. Column 'Chrom' indicates the chromosome on which the gene is present. 'Balancer' indicates the balancing marker present on the chromosome homologue. The last column indicates the whether the protein belongs to Trithorax (TG) or Polycomb (PcG) group of chromatin proteins.

Most of the lines that were analysed did not show any difference when crossed to various mutants. However, the lines A34.2.2 and A86.1.1 showed a difference in the eye color when crossed with flies mutant for *Trx* and *Esc* respectively (Fig. 2.7). As expected, in case of line A34.2.2, a decrease in the intensity of red eye color was observed in *Trx* mutant background as compared to the flies with w- background. The intensity of red eye color increased when flies of line A86.1.1 were crossed to flies mutant for *Esc* as compared to the flies crossed to w^{1118}. This indicates that the I^2 possibly works by interaction with these two chromatin regulatory/modifying proteins in a position-dependent manner.

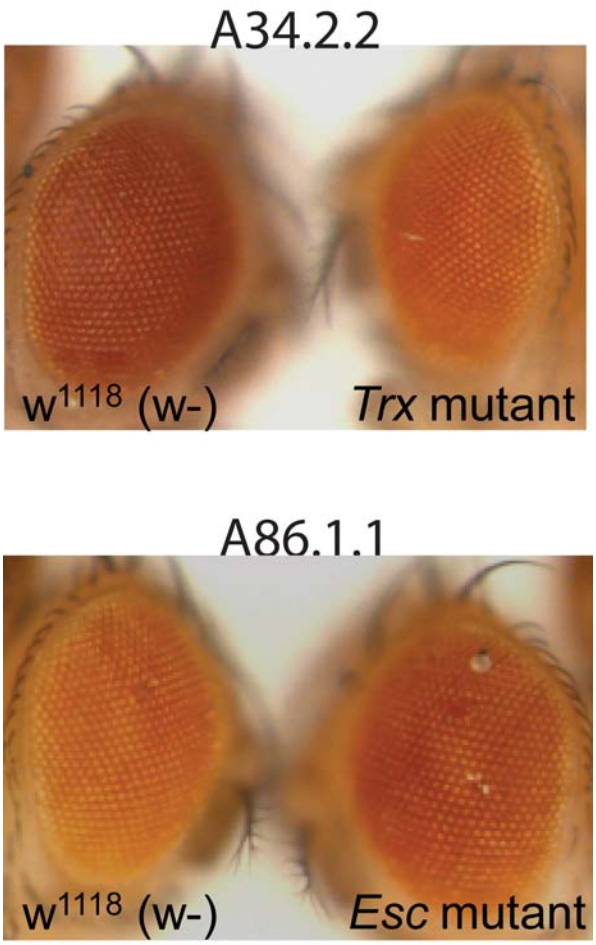


Figure 2.7: Eye color comparison for transgenic lines when crossed to w^{1118} flies vs when crossed to chromatin modulator mutants. Name of the transgenic line is indicated on top of the panel. w^{1118} – white eyed *Drosophila*. Only the crosses that showed a difference are represented. Both pictures are for 3 day-old males.

2.4 CONCLUSIONS:

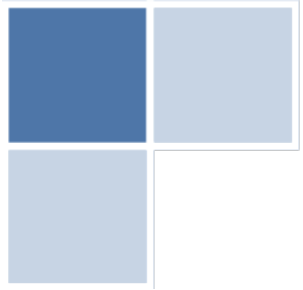
To investigate the regulatory potential of the second intron of *Neuronatin* (NNI^2) in controlling transcription, reporter gene assay in *Drosophila* using P-element transposition was performed. Various transgenic lines obtained were analysed for effect of NNI^2 on the *hsp70* promoter of *miniwhite* gene in the reporter construct. 13 unique transgenic lines were obtained, of which 7 lines showed a difference in eye color after NNI^2 was flipped-out from the construct using Cre-*LoxP* system. Importantly, all the 7 lines showed a lighter shade of red eye color when NNI^2 was absent from the construct.

This indicated that the second intron of *Neuronatin* was acting as an activator of transcription for the *miniwhite* reporter gene. The effect was same whether NNI^2 was present in (+) or (-) orientation in the construct. The remaining 6 transgenic lines did not show any difference, possibly due to the "position-effect" of the P-element.

To explore the mechanistic aspect of transcriptional activation of potential of NNI^2, the NNI^2 transgenic lines were crossed with a series of mutants of epigenetic regulators of transcription in *Drosophila*, particularly Polycomb and Trithorax group of proteins. Through various crosses, *Trx* (H3K4 methyltransferase) and *Esc* (H3K27 methyltransferase) were found to be interacting with NNI^2. The red eye color decreases or increases when the mutation is on *Trx* (activator) and *Esc* (repressor) respectively. However, this effect is not seen in case of all the transgenic lines. This indicates that the interaction might be position-dependent and holds true only when the P-element is inserted at a locus regulated by these proteins.

CHAPTER III

ANALYSIS OF MICE DELETED FOR SECOND INTRON OF NEURONATIN



3.1 INTRODUCTION:

Allele-specific transcription of imprinted genes is achieved by the interplay of several factors like epigenetic modifications, *trans*-acting proteins, non-coding RNA with several *cis*-elements within each imprinted domain like differentially methylated regions (DMRs), enhancers, and most importantly the imprinting control regions (ICRs). As discussed in Chapter I, the ICRs are regions that regulate the imprinted expression of all or several genes within a locus. All the ICRs reported till date have been shown to transcriptionally silence the imprinted genes that they control. As described in Chapter II and as shown in Sowpati et al (2008), the biochemically identified putative ICR encompassing the second intron of *Neuronatin* (NNI^2) behaves as a transcriptional activator in the *Drosophila* reporter gene assay. This could suggest that the ICR for mouse *Neuronatin* locus acts as a transcriptional activator rather than a silencer. To examine this possibility, mice deleted for this intronic region ($NN\Delta I^2$) were sought to be created.

This chapter describes the generation and characterization of mice wherein the second intron of *Neuronatin* was deleted from its endogenous locus. The aim of this experiment was to examine whether NNI^2 has a role in initiation, establishment and maintenance of imprinting at *Neuronatin* locus in its normal genomic context. To test this hypothesis, mice that carry the deletion of NNI^2 were created. These mice were used to study the effect of deletion of second intron of *Neuronatin* on its expression, and the resultant functional and phenotypic consequences. Furthermore, this strategy allowed us to look at the functional and regulatory roles of NNI^2 in its native environment.

3.2 MATERIALS AND METHODS:

3.2.1 GENERATION OF NNΔI^2 KNOCK-OUT MICE:

Annotations for the mouse *Neuronatin* locus are in reference to the GenBank sequence entry AF303656 (19,373 bp) that includes both the *Neuronatin* and *Blcap* genes. Mice carrying the deletion of second intron of *Neuronatin* were generated at the Centre for Developmental Biology, Kobe, Japan. 238 bp of the 248 bp long intron (12960nt - 13207nt) was deleted and replaced with a Neo^R cassette (1849 bp) using homologous recombination (Left arm: from 7298nt – 12969nt, Right Arm: from 13208nt – 15453nt), taking care that the splice junctions of the intron remained intact (Fig. 3.1A). The Neo^R cassette was flanked on either side by *LoxP* repeats to facilitate Cre recombinase mediated removal of the cassette when required. Sequence positions corresponding to the Neo^R cassette are referred to using # after the position.

Deletion of second intron of *Neuronatin* and integration of the Neo^R cassette at the correct genomic locus was confirmed by Southern hybridization using the probes (5' probe: from 6282nt – 6896nt, 3' probe: from 15638nt – 16506nt) described (Fig. 3.1B and 3.1C). The targeted region was PCR amplified and sequenced to confirm the presence of *LoxP* sites and splice junctions. The *Neuronatin* Knock-Out mice are hereafter mentioned as NNΔI^2 mice. The primers used for amplification of probes are:

5' PROBE: Forward – GGAGACTCCCAGATCAACAGG
Reverse – GGAGCTGGGCATAGGGAAAGG

3' PROBE: Forward – GCCAAACCAGGGAGAAAGG
Reverse - GGGAAGATTAGCTGCCAGC

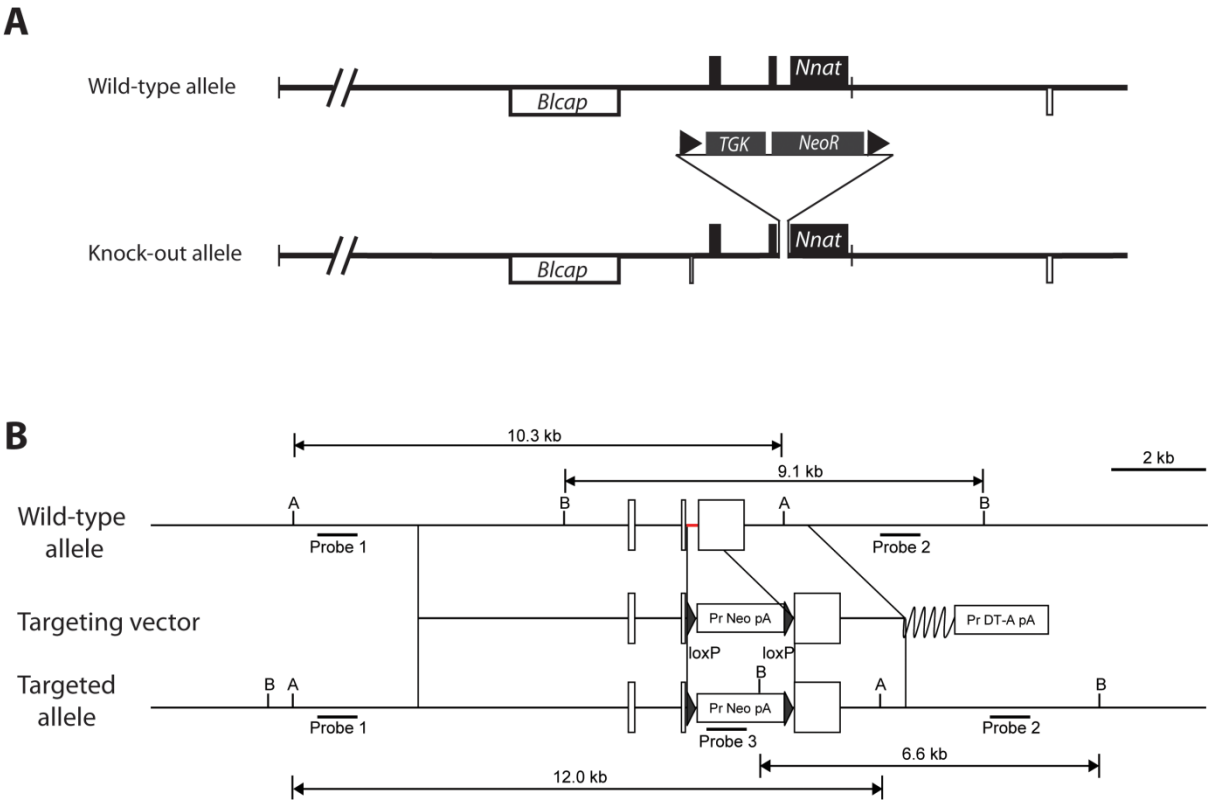


Figure 3.1: Generation of NNΔI^2 mice – Strategy. **A:** Wild type and mutant alleles of *Neuronatin*. TGK – TGK promoter that drives expression of Neo^R, Neo^R – Neomycin Resistance Cassette, filled triangles on either side represent *LoxP* sites. **B:** Constructs used for deletion of second intron of *Neuronatin*. Various probes used for Southern hybridization are indicated. Unfilled boxes indicate exons of *Neuronatin*. 'Pr Neo pA' – Neomycin Resistance Cassette flanked with *LoxP* sites (filled triangles). A – Cleavage sites for *Bst*EII, B – Cleavage sites for *Bam*HI. **C:** Southern hybridization to confirm the deletion of NNI^2 and integration of Neo^R cassette. DNA was digested either with *Bst*EII (Probes 1 and 3) or with *Bam*HI (Probe 2). +/+ indicates wild type mice, +/- indicates heterozygous NNΔI^2 mice.

3.2.2 CROSSES:

All mice were maintained as per CPCSEA guidelines, in cages with 12 hour day/night cycle (6AM – 6PM), at 22°C ± 1 and relative humidity of 60 – 70%. Food and water were replaced every 3 days. A maximum of 5 mice were kept per cage. Ideally, a single male was used to mate with two females while in some cases a 1:1 ratio for mating was maintained. Weaning was done 3 – 4 weeks after delivery. Mice heterozygous for NNΔI^2, obtained from CDB, Kobe, were crossed with the C57BL/6 strain of mice to generate maternal and paternal heterozygotes (the mice carried the deletion on maternally or paternally inherited allele respectively). NNΔI^2 heterozygous mice were self-crossed to get homozygous NNΔI^2 mice. Their wild type littermates were taken as controls.

3.2.3 GENOTYPING:

All progeny were genotyped using DNA collected from tail-tips post weaning. Mice were numbered by ear-punching, and a separate series of numbers were given for each kind of Knock-Out. Briefly, tail-tips were collected and incubated in 100 µl of Digestion buffer (50 mM KCl, 10 mM Tris pH 8.3, 0.1 mg/ml Gelatin, 0.45% NP-40, 0.45% Tween 20, 200 µg/ml Proteinase K) at 55°C for 3 hours followed by heat inactivation of Proteinase K in the buffer by incubation at 95°C for 15 min. After centrifugation at 12,000 rpm for 1 min at room temperature, supernatant was collected into a fresh tube. 1 µl of this supernatant was used in a PCR reaction. A PCR reaction contained in 20 µl 1X PCR Buffer, 200 µM dNTPs, 10 pmoles each of F1, F2, R1 primers, 1 µl of template and 1 unit of Taq Polymerase (Sigma, Cat. # D6677). Amplification was done for 30 cycles and PCR products were analyzed on an Agarose gel. The primers used for genotyping are:

```
Forward 1 – GCAGCTCTCAGCACAGTTGG
Forward 2 – CTGACCGCTTCCTCGTGCTTTACG
Reverse 1 – CCAGCTTCTGCAGGGAGTACC
```

3.2.4 DISSECTION AND TISSUE COLLECTION:

Terminal procedures on mice were done as per CPCSEA guidelines and tissues were collected immediately to minimize endogenous nuclease activity, in 2 ml microfuge tubes, and snap-frozen in liquid Nitrogen. The flask containing liquid Nitrogen and tissues was kept open at -80°C. Once the liquid Nitrogen evaporated, the tubes were transferred to a Cryo-Box™ and stored at -80°C.

3.2.5 DNA AND RNA ISOLATION:

Tissues were disrupted in liquid Nitrogen with a mortar-pestle. For cell lysis, 1 ml of Lysis buffer (100 mM Tris pH 8.0, 100 mM EDTA pH 8.0, 0.5% SDS, 20 µg/ml Proteinase K, 10 µg/ml RNase A) was used for approximately 50 mg of tissue sample. Lysis was accomplished by incubation at 55°C overnight. The proteins in the lysate were removed by extraction with Phenol:Chloroform (1:1) twice, followed by DNA precipitation using 2.5 volumes of 95% Ethanol. The spools were collected using wide-bore tips and washed in 70% Ethanol. Pellets were air-dried and resuspended in appropriate volume of 10 mM Tris pH 8.0.

RNA was isolated either using TriReagent (Sigma Cat. # T9424) or RNeasy Mini Kit (QIAGEN Cat. # 74104) following manufacturer's protocols. Briefly, 50 mg of disrupted tissue sample was used per ml of TriReagent, while 10-30 mg of tissue powder was taken per column of RNeasy Mini Kit. RNA was either resuspended or eluted in appropriate volume of RNase-free water (Sigma Cat. # 95284). The concentration and quality of DNA and RNA samples were analyzed using NanoDrop™ spectrophotometer and Agarose gel electrophoresis respectively.

3.2.6 cDNA SYNTHESIS, SEMI-QUANTITATIVE AND QUANTITATIVE REAL-TIME PCR:

500 ng of total RNA was used to synthesize cDNA. RNA was incubated with dNTPs and oligo dT at 65°C for 5 min and snap-chilled on ice. A mixture of 1X PCR Buffer (Invitrogen), 50 mM DTT, RNase-OUT (Invitrogen, Cat. # 10777-019) and SuperScript III Reverse Transcriptase (Invitrogen, Cat. # 18080-044) was added to the RNA for synthesis and incubated at 50°C for 1 hour. The enzyme was heat inactivated at 70°C for 15 min. 1 µl of cDNA mix was used for subsequent PCR reactions.

Semi-Quantitative RT PCR was done as follows: Each reaction contained 1X PCR Buffer, 0.2 mM dNTPs, 5 pmol each of Forward and Reverse primers, 1 µl of cDNA mix and 1 unit of *Taq* polymerase (Sigma, Cat. # D6677) in a total of 20 µl. Products were amplified for 30 cycles and analyzed on Agarose gels. Real-Time PCR was performed on an ABI 7500 machine and results were analyzed using Sequence Detection Software (from ABI). The primers used for expression analysis are listed in Table 3.1 (at the end of the chapter).

3.2.7 NORTHERN ANALYSIS:

10 µg of total RNA was loaded on a 1.2% MOPS-Formaldehyde Agarose Gel and electrophoresed at 120V till the Bromophenol Blue dye front reached 70% of the total gel length. The gel was denatured in 0.4N NaOH and transferred to Hybond N+ membrane (Amersham, Cat. # RPN303P) overnight, using 2X SSC (300 mM NaCl, 30 mM Sodium Citrate) as the transfer buffer. The membrane was UV-crosslinked for 2 min (2100 mJoules) and incubated in pre-hybridization buffer (6X SSC, 10% Dextran Sulfate, 5X Denhardt's solution, 10 µg/ml Salmon Sperm DNA, 0.5% SDS) for 3 hours at 60°C. The probe was prepared as follows: 50 ng of appropriate probe and random primers were denatured at 95°C and snap-chilled on ice. To the snap-chilled template, a mixture

of dNTPs (200 µM each of dATP, dGTP, dTTP, and 50 µCi of α-^{32}P dCTP), 1X reaction buffer (NEB) and 5 units of Klenow polymerase (NEB, Cat. # M0210S) was added. The reaction was incubated at 37°C for 1 hour followed by heat inactivation of Klenow at 70°C for 15 min. The probe was purified by passing the reaction mix through a G-25 sepharose (Sigma, Cat. # S5772) spin column. Purified probe was heat denatured at 95°C for 5 min, snap-chilled and added to the pre-hybridization buffer. The hybridization was carried out 60°C overnight. The membrane was washed thrice with wash buffer (1% SDS, 1X SSC) at 60°C for 20 min each, and exposed to a phosphor-imager film overnight. The results were scanned using FUJIFLM BLA-6000.

3.2.8 TRANSFECTIONS:

Constructs were transfected into HeLa cells using Lipofectamine 2000 (Invitrogen, Cat. # 11668-019) and cells were harvested after 48 hours post transfection. Isolated RNA was used for Semi-Quantitative Real-Time PCR analyses. The primers used for construction of pcDNAWT and pcDNAKO are:

```
Forward – ATTGCGCCCAGCAGCGGACT
Reverse – CACTGCCCTGCAGGGTCCTA
```

3.2.9 IMMUNO-PRECIPITATION:

Frozen tissue was thawed on ice and cut in a petridish using a razor blade. The weight of tissue was determined. 100 mg of tissue was washed once with 1 ml ice cold PBS containing protease inhibitor (PI) cocktail (Roche, Cat. # 11 836 170 001) and centrifuged at 1000 rpm for 5 min at 4°C. Tissue was ground using pestle – mortar followed by passing through a 20 gauge needle fitted to a 2 ml syringe till a homogenous suspension was obtained. The homogenized tissue was centrifuged at 1000 rpm for 10 min at 4°C and the pellet was lysed in 1 ml of Fast Act Lysis buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS

and protease inhibitor cocktail). The lysate was incubated on ice for 30 min and homogenized in a 2 ml glass Dounce homogenizer on ice using 15 strokes of the B pestle to release nuclei and disperse cell clumps. Nuclei were collected by centrifugation at 4000 rpm for 10 min at 4°C and resuspended in 1 – 2 ml of PBS with PI. The nuclei were crosslinked using 1% HCHO for 10 min at room temperature (23°C) and HCHO was quenched with 0.125M Glycine 5 min at room temperature. The crosslinked nuclei were washed twice with 1X PBS containing PI. The nuclei pellet was resuspended in SDS Lysis buffer (1% SDS, 50 mM Tris pH 8.0, 1 mM EDTA) and incubated for 30 min on ice. The lysed nuclei were centrifuged at 1000 rpm for 10 min at 4°C and the supernatant was taken as chromatin. Chromatin was sheared by sonication using Bioruptor™ (Diagenode) for 15 min - 30 sec on and 30 sec off, at high power. The sheared chromatin was centrifuged at 14000 rpm for 10 min at 4°C and the concentration of chromatin was analyzed using NanoDrop™. To 10 µg chromatin, 20 µl of pre-blocked Protein A-Agarose beads (Sigma, Cat. # P-1406) were added. The mixture was diluted ten times with IP dilution buffer (167 mM NaCl, 16.7 mM Tris pH 8.0, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS). The suspension was precleared for 2 hours at 4°C with rotation. 10% of precleared chromatin was stored as input. Supernatant was taken into a fresh tube and 1 µg of antibody was added and kept at 4°C overnight with rotation. Following this, 20 µl of Protein A-Agarose beads were added and incubated for 2 hours at 4°C with rotation. The beads were collected by centrifugation at 2000 rpm for 2 min at 4°C. Supernatant was discarded and the beads were washed as follows:

i. Once with Low Salt Wash buffer at 23°C with rotation, 5 min
ii. Once with High Salt Wash buffer at 23°C with rotation, 5 min
iii. Once with LiCl Wash buffer at 23°C with rotation, 5 min
iv. Twice with 1X TE buffer at 23°C with rotation, 10 min each

The beads were eluted with 100 µl of elution buffer at 37°C for 15 min with rotation, twice. The samples and input were reverse-crosslinked by adding NaCl to a final concentration of 200 mM, 100 µg Proteinase K, and incubating overnight at 65°C. The samples were purified using Qiagen PCR purification kit and suspended in 40 µl nuclease-free water. 1 µl of purified DNA was used as template for end-point PCR analysis. The compositions of wash buffers are as follows:

Low Salt Wash buffer – 150 mM NaCl, 20 mM Tris pH 8.0, 2 mM EDTA, 1% Triton X-100, 0.1% SDS

High Salt Wash buffer – 500 mM NaCl, 20 mM Tris pH 8.0, 2 mM EDTA, 1% Triton X-100, 0.1% SDS

LiCl Wash buffer – 0.25 M LiCl, 10 mM Tris pH 8.0, 1 mM EDTA, 1% NP-40, 1% Sodium Deoxycholate

TE buffer – 10 mM Tris pH 8.0, 1 mM EDTA

3.2.10 DNA METHYLATION ANALYSIS BY SODIUM BISULFITE SEQUENCING:

Sodium Bisulfite sequencing was performed as described by Gokul et al., (2007). Briefly, 200–500 ng of genomic DNA was digested with a restriction enzyme that does not cleave within the region of interest. 21 µl of digested genomic DNA was denatured at 95°C for 5 min and snap-chilled on ice. To this, 4 µl of 2N NaOH was added and incubated on ice for further 5 min and mixed with 50 µl of 2% Low Melting Point (LMP) Agarose in water. The mix was incubated at 50°C for 15 min. DNA-Agarose beads were then prepared by pipetting 10 µl of this mix into pre-chilled heavy mineral oil (Sigma, Cat. # 33076) layered over Sodium Bisulfite solution. Care was taken that the beads are not merged together.

The Sodium Bisulfite solution was prepared as follows: 2.7 gm of Sodium Bisulfite (Sigma, Cat. # S9000) was dissolved in 3.25 ml of solution containing 2.5 ml of water and 750 µl of 2N NaOH. Hydroquinone solution was prepared by dissolving 55 mg of Hydroquinone (Sigma, Cat. # H9003) in 500 µl of water. Both the solutions were

prepared by slowly stirring in a 50°C water bath. Once dissolved, both solutions were mixed and filter-sterilized through a 0.45 µm syringe filter (Millipore, Cat. # SLHA033SS). Care was taken that the solutions do not froth, and are exposed to minimal light. Heavy mineral oil was layered over 1 ml of this solution and chilled by incubation on ice for 30 min. DNA Agarose beads were incubated in Sodium Bisulfite solution for 3.5 hours at 50°C. Following incubation, the beads were washed:

i. 3 times with 1X TE, 15 min each, on a rotatory shaker at 25°C
ii. 2 times with 0.2N NaOH, 15 min each, without rotation at 30°C
iii. 2 times with 1X TE, 10 min each, on a rotatory shaker at 25°C
iv. 2 times with water, 10 min each, on a rotatory shaker at 25°C

After the washes, the beads were either used immediately for PCR, or stored in ~50 µl of 1X TE at 4°C till further use. For PCR, the beads were melted at 70°C for 5 min and 3 µl of molten beads was used as template per 25 µl PCR reaction.

3.2.11 BIOCHEMICAL EXAMINATION OF NNΔI^2 MICE:

3.2.11a Levels of Cholesterol and Triglycerides:

All tests were performed on age-matched males. Required blood was collected by tail-tip bleeding. The mice were fasted overnight and the levels of Cholesterol and Triglycerides were measured using AccuTrend Plus (Roche).

3.2.11b Insulin and Leptin Levels:

Serum was collected from 5 age-matched males each of various NNΔI^2 mice. Approximately 200 µl of blood was collected from each mouse by orbital bleeding into a 1.5 ml microfuge tube. The tube was kept in a slant position for 1 hour at room temperature, and centrifuged at 12,000 rpm for 1 min. Supernatant was collected into a fresh tube while avoiding any aggregate formed due to hemolysis. Serum was stored at

-80°C till further use. Insulin and Leptin tests were performed using ELISA-based kits (Insulin – Millipore, Cat. # EZRMI-13K, Leptin – Millipore, Cat. # EZML-82K) following manufacturer's instructions. All samples were analysed in duplicates.

3.3 RESULTS:

3.3.1 LACK OF *NEURONATIN* EXPRESSION FROM PATERNAL ALLELE OF NNΔI² MICE:

Mice carrying the deletion of second intron of *Neuronatin* (NNΔI²) were created in Centre for Developmental Biology, Kobe, Japan (see Materials and Methods). To investigate the effect of deleting the second intron of *Neuronatin* on its expression, Semi-Quantitative RT-PCR was performed on RNA from Brain, E13.5 Embryos and Liver of wild-type, paternal, maternal heterozygous, and homozygous NNΔI² mice. The RT-PCR analysis indicated that there was no expression of *Neuronatin* in mice which inherited NNΔI² from the father and in homozygous NNΔI² mice, while the expression of *Neuronatin* remained unchanged in maternal heterozygotes (Fig. 3.2A). RT-PCR analysis on Liver tissue from all NNΔI² mice did not show any expression of *Neuronatin*. The observed expression profile was further confirmed by Northern hybridization and Quantitative Real-Time PCR analysis (Fig. 3.2B and 3.2C).

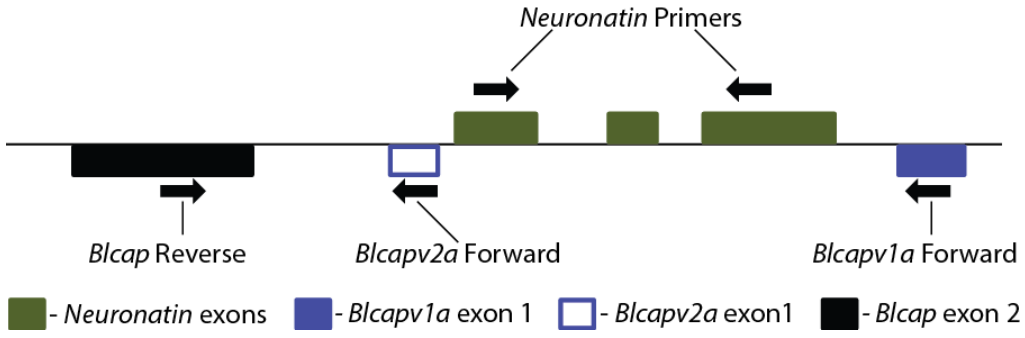


Fig. 3.2.a: Primers used for expression analysis of *Neuronatin* and *Blcap*. Primers and their orientation are represented by black arrows. Figure not drawn to scale.

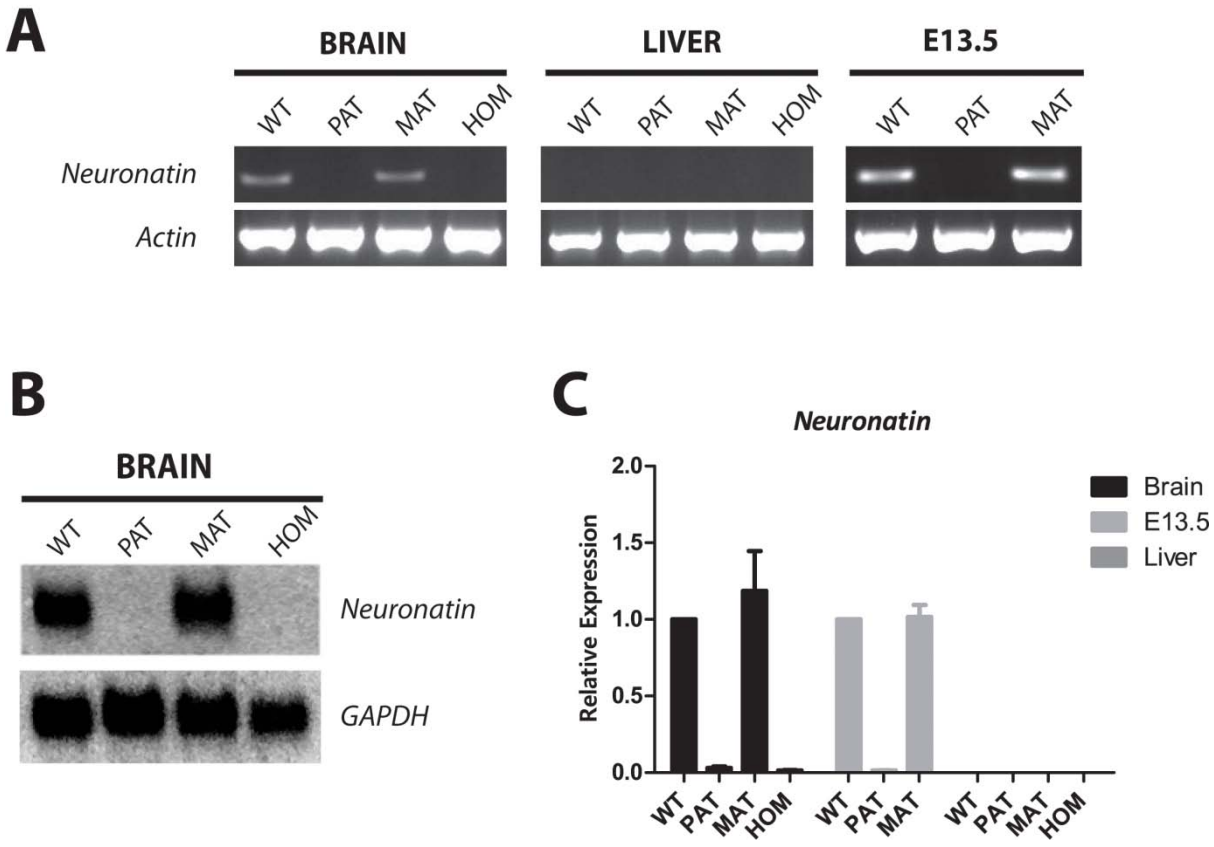


Figure 3.2: Expression analysis of *Neuronatin*. WT – wild type, PAT – paternal heterozygous, MAT – maternal heterozygous, HOM – homozygous NNΔI^2 mice. **A:** Semi-Quantitative Real Time PCR analysis of *Neuronatin* expression in Brain, Liver and E13.5 embryos of NNΔI^2 mice. β-*Actin* was used as control for cDNA concentration. Pictures shown are representative of three individual experiments. **B:** Northern analysis of *Neuronatin* expression in Brain of NNΔI^2 mice. *GAPDH* was used as control for quantity of RNA. Pictures are representative of three individual experiments. **C:** Quantitative Real-Time PCR analysis of *Neuronatin* expression in Brain, Liver and E13.5 embryos of NNΔI^2 mice. Experiments were performed thrice in duplicates. Error bars represent standard error (SEM).

3.3.2 NEO^R DOES NOT AFFECT SPLICING AND EXPRESSION OF *NEURONATIN in vitro:*

The mutant allele of *Neuronatin* carries a Neomycin Resistance Cassette in place of the intron. The observed transcriptional changes in NNΔI^2 mice could be either due the deletion of second intron or because the splicing and stability of *Neuronatin* transcript was affected by the presence of Neo^R cassette. To test whether Neo^R cassette has any effect on *Neuronatin* expression, exon 1 to end of exon 3 of *Neuronatin* gene

(excluding the A-rich tail but including the stop codon) was amplified from both wild type and homozygous NNΔI[2] mice and cloned into the *EcoR*V site of pcDNA 3.1 (+) vector (Invitrogen). The constructs were named pcDNAWT and pcDNAKO for wild type and knock-out respectively (Fig. 3.3A). The sequence and orientation of the constructs was validated by DNA sequencing. The levels of *Neuronatin* cDNA produced from the constructs were measured by Semi-Quantitative RT-PCR analysis using primers specific to mouse *Neuronatin*. The level of *Neuronatin* expression from cells transfected with pcDNAKO construct was similar to that of cells having pcDNAWT plasmid, indicating that the presence of Neo^R does not affect the splicing and stability of *Neuronatin* cDNA (Fig. 3.3B).

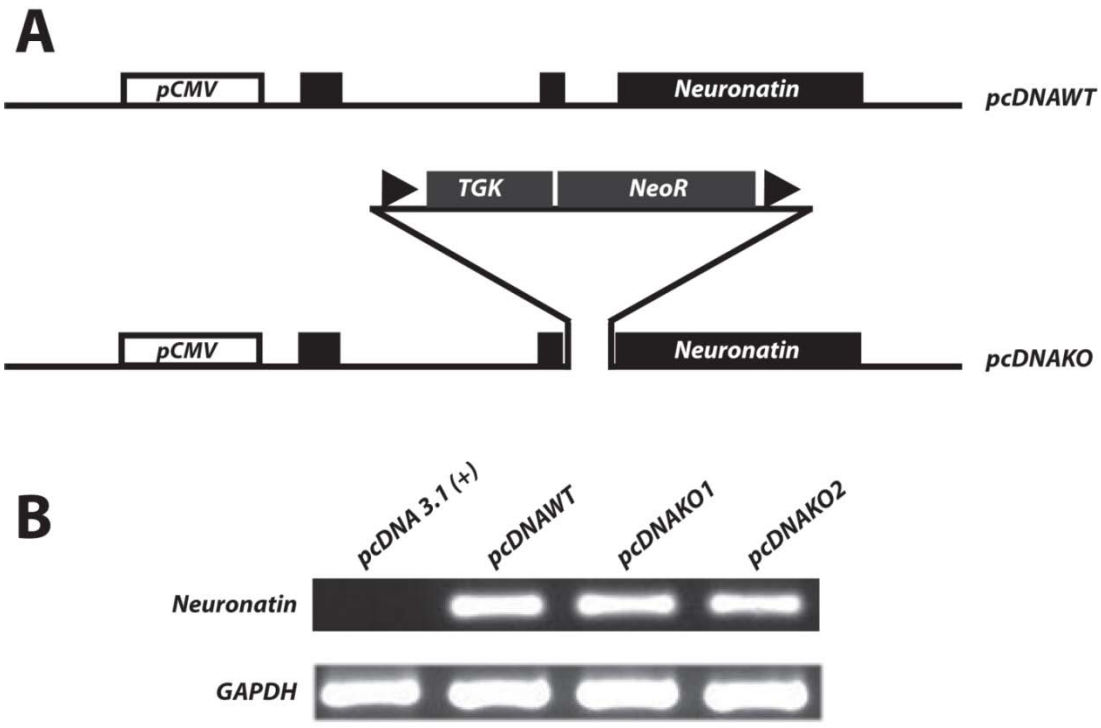


Figure 3.3: Construction of pcDNAWT and pcDNAKO A: Line diagrams representing various features of pcDNAWT and pcDNAKO constructs. pCMV – cytomegalovirus promoter, driving the expression of *Neuronatin*. TGK – TGK promoter driving expression of Neo^R, Neo^R – Neomycin Resistance Cassette. Filled triangles represent *LoxP* sites. **B:** Semi-Quantitative Real Time PCR analysis of expression of mouse *Neuronatin* in transfected HeLa cells. pcDNAKO1 and KO2 represent two individual clones of pcDNAKO. Pictures are representative of three individual experiments. *GAPDH* is used as control for cDNA concentration.

3.3.3 LACK OF TRANSCRIPTION INITIATION FROM *NEURONATIN* PROMOTER IN PATERNAL HETEROZYGOUS NNΔI^2 MICE:

RNA Polymerase II interaction with the promoters of genes is used as a mark of their transcriptional status (Gerber et al., 1995). Inactive genes lack RNA Pol II from their promoters. To assay for the interaction of RNA Polymerase II with *Neuronatin* promoter, immuno-precipitation with RNA Pol II antibody was performed on Brain nuclei from wild type, and paternal and maternal NNΔI^2 mice (see Materials and Methods). As can be seen in Figure 3.4, interaction of RNA Pol II with promoter of *Neuronatin* was detected in WT and maternal heterozygous NNΔI^2 mice, but paternal heterozygotes lacked RNA Pol II-promoter interaction. This suggested possible disruption in initiation of transcription from *Neuronatin* promoter in Brain of paternal NNΔI^2 mice.

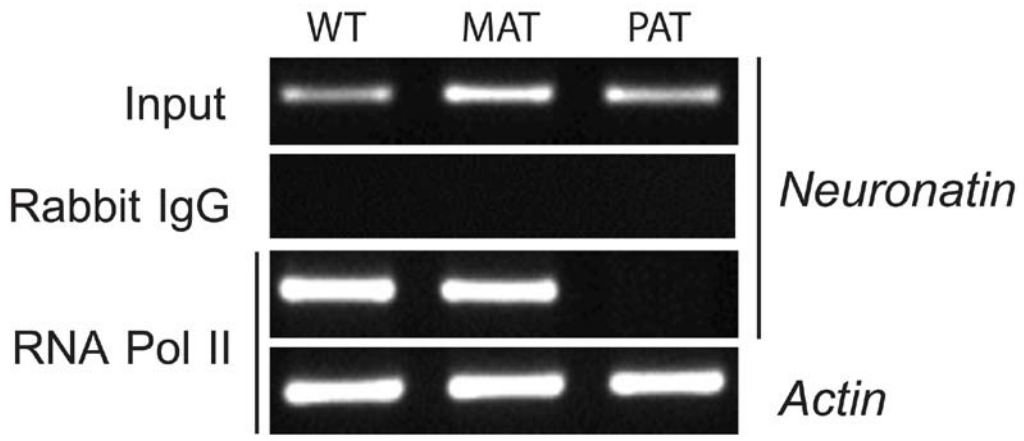


Figure 3.4: Immuno-precipitation with RNA Pol II antibody on Brain nuclei of NNΔI^2 mice. WT, MAT and PAT indicate wild type, maternal and paternal heterozygous NNΔI^2 mice respectively. β-*Actin* is used as control.

3.3.4 METHYLATION ANALYSIS OF *NEURONATIN* PROMOTER IN NNΔI^2 MICE:

DNA methylation has been shown to be involved in regulation of imprinted gene expression. Furthermore, deletion of ICRs has been shown to change the methylation profile at the respective imprinted locus. To check whether the observed change in the

expression profile of *Neuronatin* is due to a change in the methylation status of the locus, Sodium Bisulfite sequencing was done for the *Neuronatin* promoter in Brain and Liver nuclei of various NNΔI^2 mice. The DNA methylation status of *Neuronatin* promoter in maternal heterozygous NNΔI^2 mice was similar to that of wild type mice, whereas the promoter of *Neuronatin* was hypermethylated in paternal heterozygous and homozygous NNΔI^2 mice (Fig. 3.5). *Neuronatin* promoter was hypermethylated in Liver tissue of all NNΔI^2 mice, which corresponded with loss of *Neuronatin* expression. This indicated that the gain of methylation was responsible for loss of *Neuronatin* expression in paternal and homozygous NNΔI^2 mice.

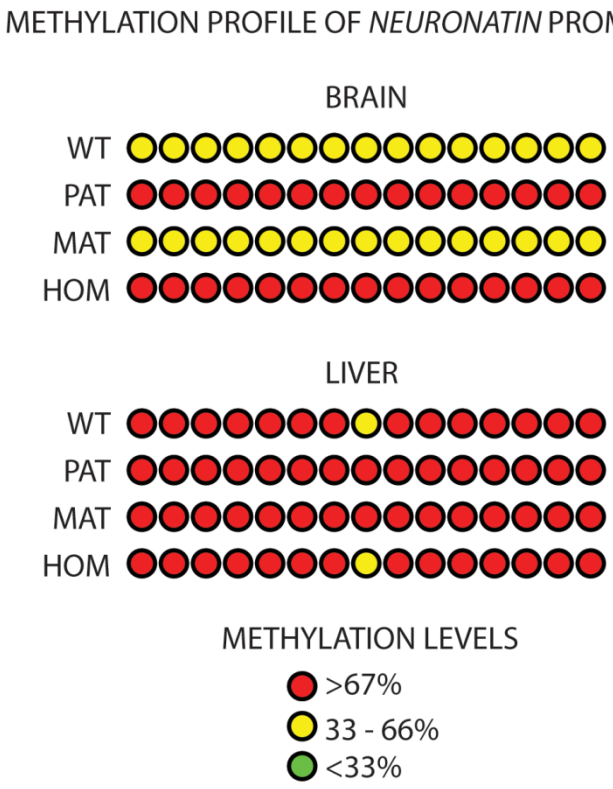


Figure 3.5: Methylation analysis of promoter region of *Neuronatin* in NNΔI^2 mice. Each circle represents a single CpG dinucleotide of the region. Percentage of methylation is calculated from a minimum of 12 clones per sample. WT – wild type, PAT – paternal heterozygous NNΔI^2, MAT – maternal heterozygous NNΔI^2, HOM – homozygous NNΔI^2 mice.

3.3.5 ABNORMAL BODY WEIGHT OF NNΔI^2 MICE:

While we were keen in identifying the phenotypic consequences arising because of the loss of *Neuronatin* expression in NNΔI^2 mice, morphologically the paternal and maternal heterozygous and homozygous mice appeared normal. However, as *Neuronatin* has been shown to be correlated with insulin secretion and obesity (Chu and Tsai, 2005; Vrang et al., 2010) and is also a known interacting partner of Leptin, which controls appetite and metabolism (Vrang et al., 2010), we decided to carefully examine the biochemical characteristics of these mice. Initially, body weights of the various NNΔI^2 mice were taken at various times during their life-time. As can be seen from Figure 3.6A, the body weight of paternal heterozygous and homozygous NNΔI^2 mice was found to be consistently higher than their wild type counterparts. Across the various age groups, the increase observed was 12%. However, it was surprising to find that maternal heterozygous NNΔI^2 mice, inheriting the deletion from the mother and having *Neuronatin* expression similar to wild type (Fig. 3.2), showed body weight that was on average 20% lower than wild type (Fig. 3.6A and 3.6B). In addition, further statistical analysis of the data suggested that with age, all NNΔI^2 mice (paternal and maternal heterozygous, and homozygous) seem to lose control on their body weight: The range of body weights for all groups was found to be similar for 2-month old mice. However, older mice (10-15 months old) showed statistically different range of body weights. While all wild type mice in an age group were very similar in body weights, paternal and maternal heterozygous and homozygous NNΔI^2 showed very divergent body weights (Fig. 3.6C).

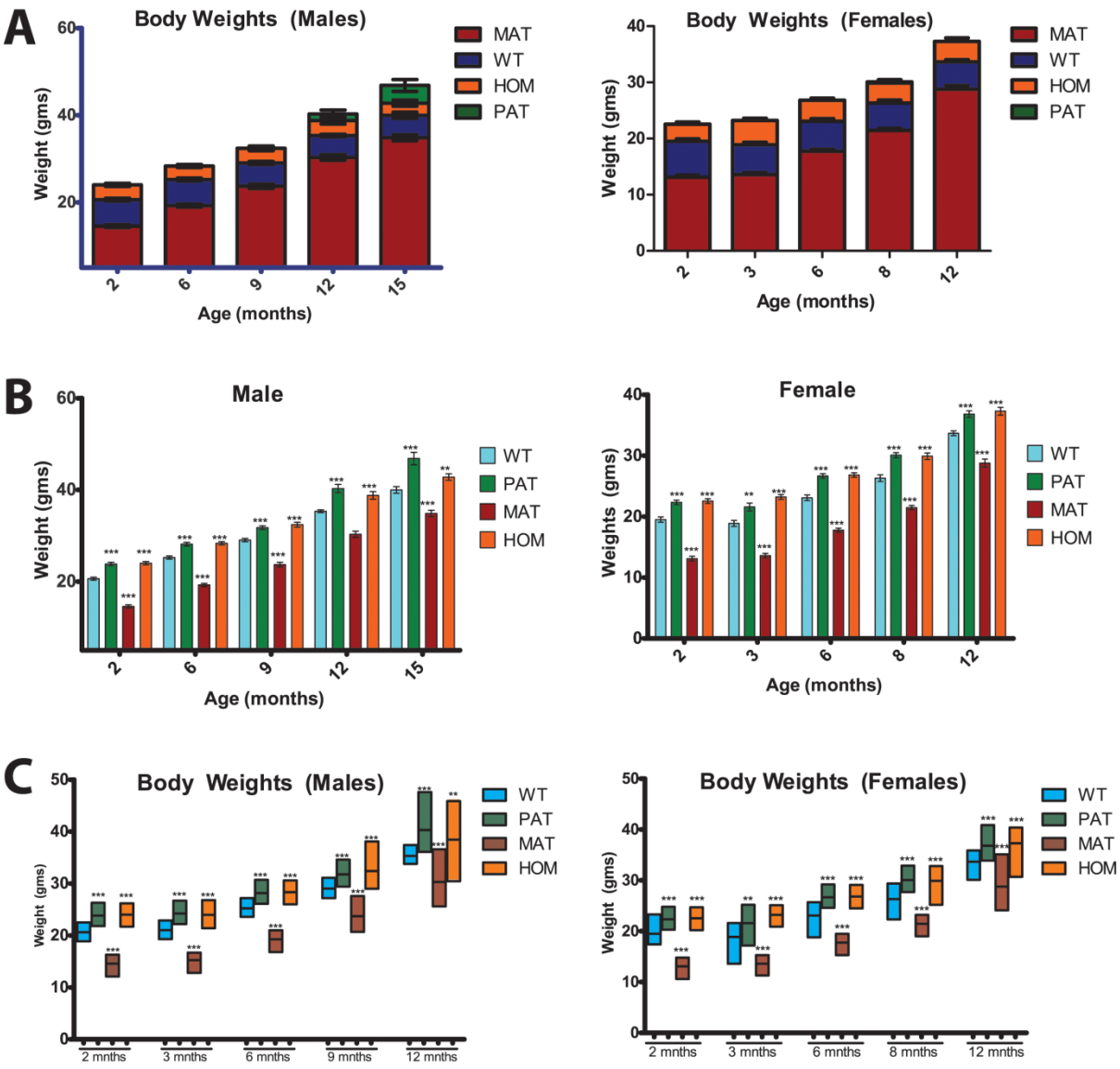


Figure 3.6: Analysis of body weights of NNΔI^2 mice. WT – wild type, PAT – paternal heterozygous NNΔI^2, MAT – maternal heterozygous NNΔI^2, HOM – homozygous NNΔI^2 mice. Weights of at least 15 mice were taken per group. Asterisks indicate significant difference (Student's t test, * - $p < 0.05$, ** - $p < 0.01$, *** - $p < 0.001$. Only representative age groups are shown. **A:** Superimposed graph depicting differences in body weights of various NNΔI^2 mice. **B:** Bar diagram showing average body weights of NNΔI^2 mice. **C:** Range bars to depict age dependent loss of control on body weights in NNΔI^2 mice.

3.3.6 BIOCHEMICAL CHARACTERIZATION OF NNΔI^2 MICE:

To understand the biochemical rationale behind the abnormal body weights of various NNΔI^2 mice, the levels of Insulin, Leptin, Cholesterol and Triglycerides were measured in age matched males of two groups: young (4 months old) and old (10

months old). No statistically significant difference (Student's t-test, $p > 0.05$) was observed in the levels of Insulin and Leptin of any NNΔI^2 mice (Fig. 3.7A). The levels of Cholesterol were very similar, both across and within the two age groups. However, a significant change in the levels of Triglycerides was observed with age in paternal NNΔI^2 mice ($p < 0.05$); young mice had relatively low levels of Triglycerides compared to wild type mice of same age, as well as paternal NNΔI^2 mice of older age (Fig. 3.7B). Though there was a decrease in the mean level of Triglycerides in older wild-type mice, it was not significant ($p > 0.05$).

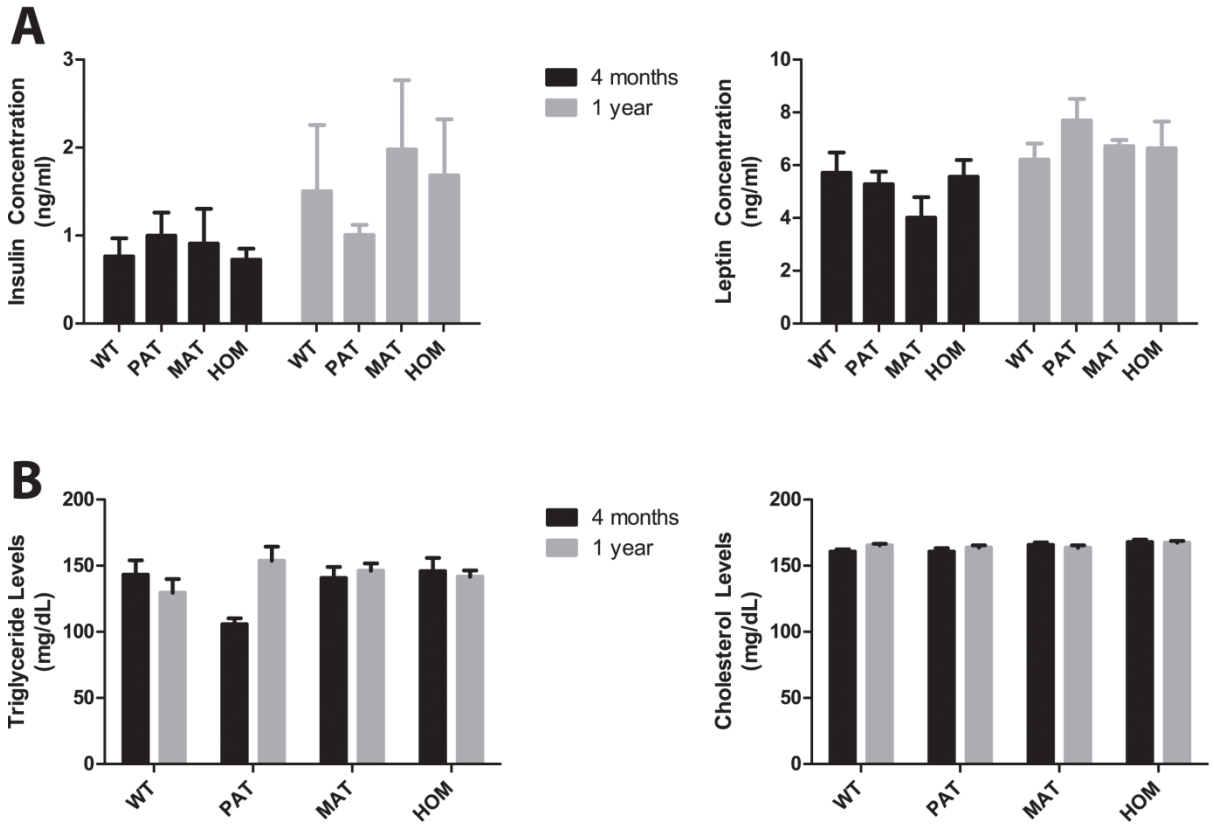


Figure 3.7: **Biochemical characterization of NNΔI^2 mice.** WT – wild type, PAT – paternal heterozygous NNΔI^2, MAT – maternal heterozygous NNΔI^2, HOM – homozygous NNΔI^2 mice. **A:** ELISA measurement of Insulin and Leptin levels of NNΔI^2 mice. Each group contained 5 age matched males. Error bars represent Standard Error (SEM). **B:** Triglyceride and Cholesterol levels of NNΔI^2 mice. A minimum of 10 mice were analysed per group. Error bars represent Standard Error (SEM).

3.3.7 EFFECT OF NNΔI² ON EXPRESSION OF *NEURONATIN* AND *BLCAP*:

3.3.7a Neuronatin:

Increase in body weight of paternal heterozygous and homozygous NNΔI² mice was attributable to loss of *Neuronatin* expression. However, loss of body weight in maternal heterozygous NNΔI² mice, where expression of *Neuronatin* is normal, confounded us. To investigate the reason for this change, it was decided to look at the expression of *Neuronatin* in other tissues. Analysis of *Neuronatin* expression was performed on RNA isolated from Eyes, Kidneys and Adrenal glands of various NNΔI² mice. Surprisingly, the expression profile of *Neuronatin* was found to be very different from what we observed in adult Brain and E13.5 embryos (Fig. 3.8A). For paternal heterozygous NNΔI² mice, *Neuronatin* expression was found to be similar to the wild type in Eyes and Kidneys. But the expression level was two-fold more in Adrenal glands. Maternal heterozygous NNΔI² mice on the other hand showed total loss of *Neuronatin* expression in Eyes and Adrenal gland while it was approximately 15-fold more as compared to wild type in Kidney. This indicated that the normal regulatory mechanism for the control of imprinted expression of the *Neuronatin* gene had been disrupted in various adult tissues of NNΔI² mice.

3.3.7b Blcap:

Genes located within an imprinted locus are often regulated co-ordinately. To test whether the deletion of the second intron of *Neuronatin* affects the expression of *Blcap,* the other imprinted gene of the locus, Quantitative Real-Time PCR analysis of both the transcript variants of *Blcap* was performed. As can be seen from Figure 3.8B, the expression of both the transcripts of *Blcap* had changed in an allele- and tissue-specific manner. In Brain, the expression level of *Blcapv1a*, the transcript preferentially transcribed from the maternal allele in wild type mice, was down by ~40% in maternal

heterozygous NNΔI[2] mice whereas the expression in paternal heterozygotes and homozygous NNΔI[2] mice was similar to wild type. However, *Blcapv2a* levels were found to be 50% more in paternal heterozygous mice and similar to wild type in maternal heterozygotes. Surprisingly, loss of expression of *Blcapv2a* was found in homozygous NNΔI[2] mice.

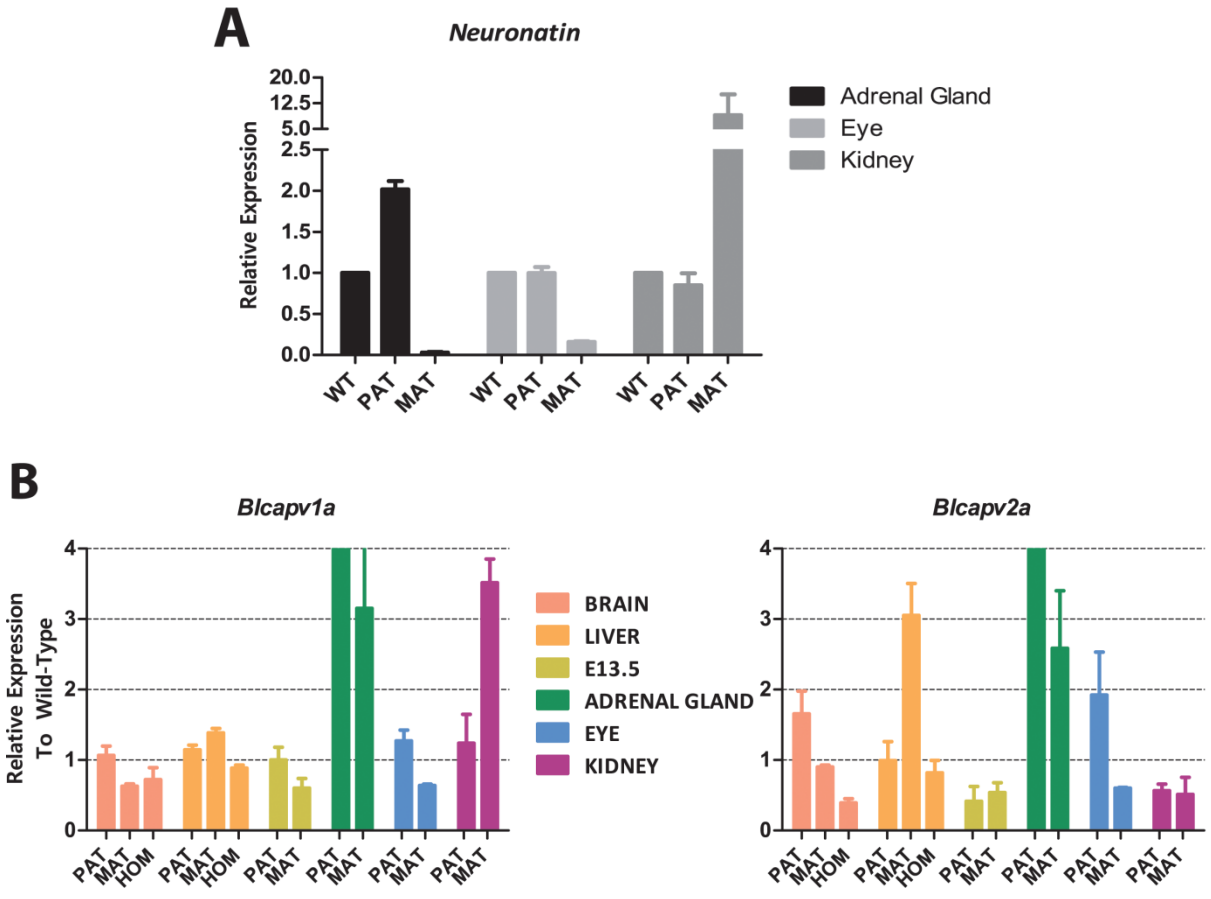


Figure 3.8: Expression analysis of *Neuronatin* and *Blcap* in NNΔI[2] mice. WT – wild type, PAT – paternal heterozygous NNΔI[2], MAT – maternal heterozygous NNΔI[2], HOM – homozygous NNΔI[2] mice. **A:** Quantitative Real Time PCR analysis of *Neuronatin* expression in Adrenal gland, Eye and Kidney of NNΔI[2] mice. Error bars represent standard error (SEM). Results shown are cumulative of at least three individual experiments. **B:** Quantitative Real-Time PCR analysis of expression of *Blcapv1a* and *Blcapv2a* in various tissues of NNΔI[2] mice. Error bars represent standard error (SEM). Results shown are cumulative of at least three individual experiments.

In E13.5 embryos, the expression of *Blcapv2a* had decreased in both paternal and maternal NNΔI^2 mice whereas that of *Blcapv1a* is only affected in mice inheriting the deletion from the mother. The expression of both transcripts has gone up several folds in both heterozygotes in Adrenal glands. In Eye, there is a reduction in expression of both transcripts in maternal NNΔI^2 mice while paternal NNΔI^2 mice show an increase only in the expression of *Blcapv2a*. In Kidney, the expression of *Blcapv1a* is three-fold higher than wild type in maternal NNΔI^2 mice while it remains unaffected in paternal NNΔI^2 mice. The expression of *Blcapv2a* in Kidney is downregulated both in maternal and paternal NNΔI^2 mice. This showed that the deletion of second intron of *Neuronatin* disrupts the expression of *Blcap* in a tissue-, transcript- and parent-of-origin specific manner.

3.3.8 EFFECT OF NNΔI^2 MICE ON EXPRESSION OF NEIGHBORING GENES:

Deletion of ICRs have been previously shown to affect the expression of genes located several hundred kilobases away. Hence, it was decided to analyse the expression of 6 genes (4 protein coding genes and 2 non-coding RNAs) located within a 250 kb region flanking *Neuronatin* (Fig. 3.9A). Semi-Quantitative Real Time PCR analysis on Brain of various NNΔI^2 mice revealed an expression profile of these genes similar to that of wild type mice, indicating that deletion of NNI^2 does not affect the expression of genes present in the vicinity of *Neuronatin* (Fig. 3.9B).

3.4 CONCLUSIONS:

This chapter primarily deals with the consequences of the deletion of a putative regulatory region in mice, the second intron of *Neuronatin*. NNI^2 acted as a transcriptional activator in *Drososphila*; hence, deletion of an activating element should result in the downregulation of its target. That was precisely what was observed when

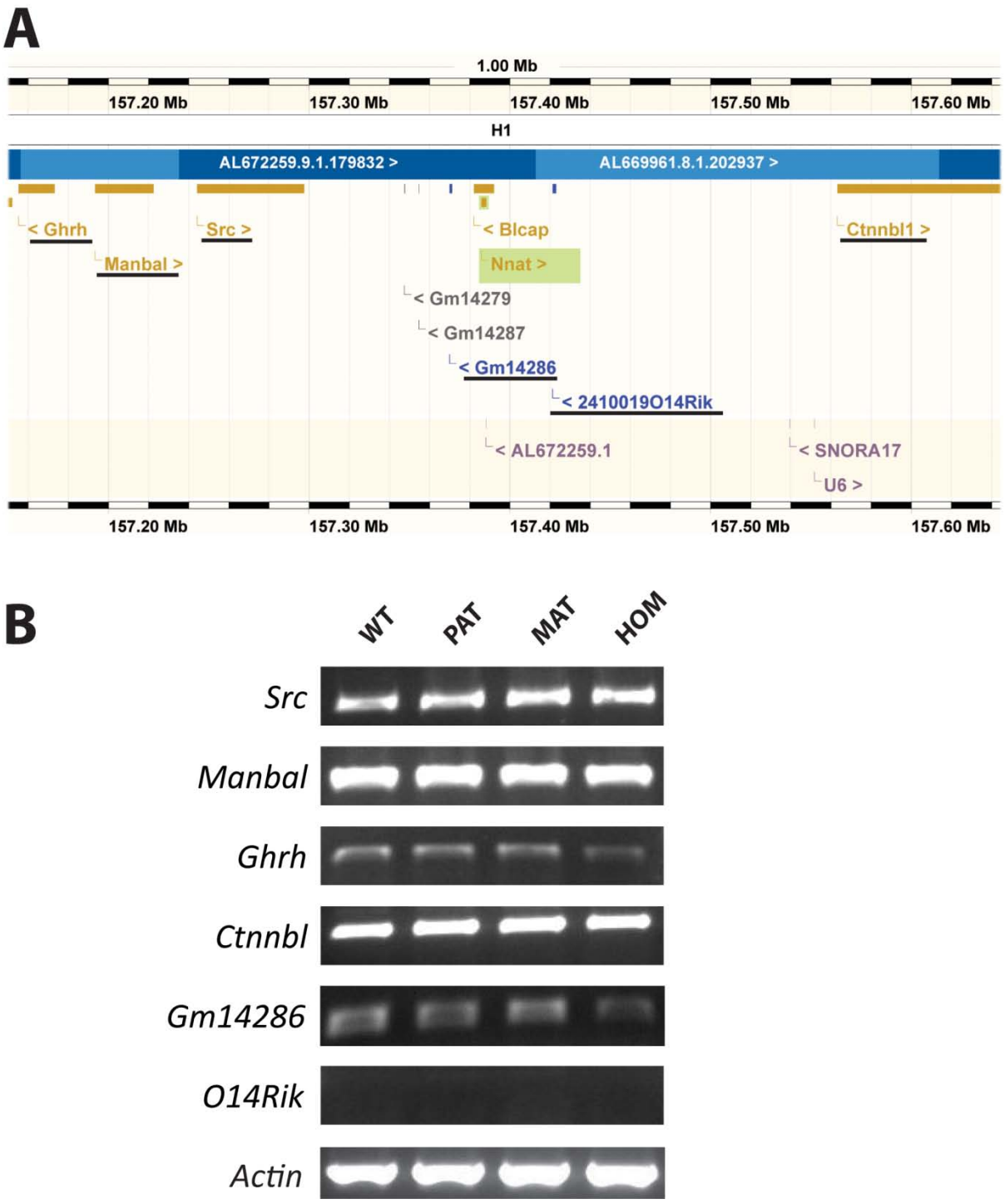


Figure 3.9: Effect of NNΔI^2 on the expression of neighbouring genes. A: Genomic region flanking *Neuronatin* in mouse depicting the location of nearby genes. **Genes** and **ncRNAs** analysed are underlined. **B:** Semi Quantitative Real Time PCR analysis of expression of surrounding genes in Brain of NNΔI^2 mice. WT – wild type, PAT – paternal heterozygous NNΔI^2, MAT – maternal heterozygous NNΔI^2, HOM – homozygous NNΔI^2 mice. β-*Actin* is used as control for cDNA concentration. Pictures are representative of two individual experiments.

NNI^2 was deleted in mice; there was no expression of *Neuronatin* in NNΔI^2 mice. Interestingly, in Brain where *Neuronatin* was first identified and attributed a functional role, this observation was only true when the deletion was present on the paternal allele, the allele from which *Neuronatin* is normally transcribed while this is not the case with other tissues as indicated by the expression profiles of various NNΔI^2 mice. However, deletion of NNI^2 results in abnormal expression of *Neuronatin*, though it is tissue- and parent-of-origin specific. Moreover, it has been proven that the observation was indubitably because of the deletion and not because the stability and splicing of *Neuronatin* is affected by the Neo^R cassette that replaces the NNI^2 in NNΔI^2 mice. In addition, it was noticed that the initiation of transcription is affected from the *Neuronatin* promoter on the paternal allele in Brain, as no occupancy of RNA Polymerase II could be detected at this region. Substantiating this, the methylation profile of *Neuronatin* promoter was concomitant with the expression pattern observed; a gain of methylation was observed at *Neuronatin* promoter in paternal and homozygous NNΔI^2 mice while the level of methylation was comparable to wild type in maternal NNΔI^2 mice. In Liver where *Neuronatin* remains unexpressed irrespective of the presence of intron, the promoter was hyper-methylated in all NNΔI^2 mice. This indicates that the control on the expression of *Neuronatin* is achieved by a change in DNA methylation.

As observed in the case of other ICRs, studies on the expression of *Blcap* showed that the deletion of NNI^2 affects *Blcap* expression in a transcript-, tissue-, and parent-of-origin-specific manner. This finding further corroborates the role of NNI^2 in regulation of gene expression at *Neuronatin* locus. A previously established role of *Neuronatin* is in energy consumption and metabolism. Therefore, a change in its expression should reflect in phenotypic effects pertaining to body metabolism. Analysis of body weights of

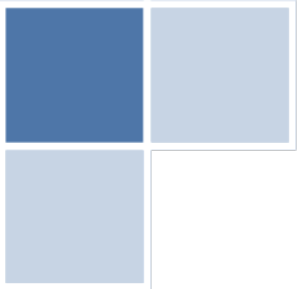
various NNΔI^2 mice substantiates this premise, as paternal and homozygous NNΔI^2 mice show an increase in their average body weight while their maternal counterparts have a lesser average body weight with respect to wild type mice. Remarkably, all NNΔI^2 mice have lost control on their body with time, a result concurrent with previous reports. Biochemical characterization of NNΔI^2 mice has thrown no further insight into the molecular mechanism through which the phenotypic differences arise, as no significant differences correlating with the phenotype are observed in the levels of Insulin, Leptin, and Cholesterol of various NNΔI^2 mice. Though a significant increase in the average levels of Triglycerides was observed in paternal heterozygous NNΔI^2 mice with age, it does not explain the change in body weights of maternal heterozygous and homozygous NNΔI^2 mice.

Gene	Primers
Neuronatin	Forward: GAAATCCTCCAGGGACACAG Reverse: CTCCTCTGACGAGGGAACAC
Blcapv1a	Forward: GGCGACGGCGACAGT Reverse: CAGCCCGGGATCACCAA
Blcapv2a	Forward: AGCCGAAGTTACAGAAGTCTCTCT Reverse: CAGCCCGGGATCACCAA
Ghrh	Forward: TGTGATCCTCATCCTCACCA Reverse: ATCCTCTCCCCTTGCTTGTT
Manbal	Forward: GAGCCATCTTCCAGCTCATC Reverse: GTCTCCTTCTTCGGCCTCTT
Src	Forward: ACACGAGGAAGGTGGATGTC Reverse: TTTCACATTTAGGCCCTTGG
Ctnnbl	Forward: AAAAGACACAACCCCAGCAC Reverse: GATGGTCCGTAAGCCAAGAA
Gm14286	Forward: GACTCAGCTGGGTTCCAGAG Reverse: TGCCTGTAATTGCAGCTTTG
O14Rik	Forward: CCTTCTTTCCTCCCAGATCC Reverse: ACAAGGCTCAGTGGCACTTC
Actin	Forward: GACATGGAGAAGATCTGGCA Reverse: GGTCTCAAACATGATCTGGGT
GAPDH (mouse)	Forward: ACAGTCAAGGCCGAGAATGG Reverse: TGGCAGTGATGGCATGGAC
GAPDH (human)	Forward: TGCTGGCGCTGAGTACGTCG Reverse: GGGTGGCAGTGATGGCATGG

Table 3.1: Primers used in this study for expression analysis.

CHAPTER IV

IDENTIFICATION OF PROTEIN(S) BINDING WITHIN THE SECOND INTRON OF NEURONATIN



4.1 INTRODUCTION:

The differential chromatin organization between the two alleles of *Neuronatin* within its second intron could be due to the binding of a protein(s) to one of the alleles, which could lead to non-canonical chromatin organisation. Examples of such DNA methylation-restricted protein-DNA interaction have been shown for other imprinted genes (Fitzpatrick et al., 2002; Hark et al., 2000; Kim et al., 2003). Moreover, it was previously reported in our laboratory that a 38 bp GC rich region (referred to as GC region hereafter) present within the NNI^2 was bound by protein(s) in a methylation dependent manner. CTCF and YY1 proteins have been identified in previous studies to be proteins capable of methylation restricted DNA binding (Kim, 2008). However, analysis of the sequence in and around the mouse *Neuronatin* gene using bioinformatic approaches did not identify any CTCF or YY1 binding sites. This raised the possibility of protein(s) other than CTCF and YY1 binding within the NNI^2. To identify and characterize protein(s) binding in a methylation restricted manner to the GC region, Yeast Mono Hybrid and MALDI-TOF experiments were performed, which suggested several possible proteins that bind the GC region (Doctoral thesis of Sudhish Sharma, CDFD/Manipal University, 2008). Two of those proteins, Cbx5 and hnRNPK, were selected for further investigation. The results of those experiments are presented in this chapter.

4.2 MATERIALS AND METHODS:

4.2.1 CONSTRUCTS:

GST-hnRNPK was a kind gift from David Levens' group. GST-Cbx5 was a kind gift from Eric Olson's group. Both the plasmids were constructed from pGEX backbone (Amersham Life sciences), which carried an Ampicillin resistance marker.

4.2.2 PROTEIN PURIFICATION:

GST-Cbx5 and GST-hnRNPK plasmids were transformed into *E.coli* BL21-DE3 competent cells. A single colony from each transformed plate was picked and re-streaked onto fresh plates. Glycerol stock was made from a single colony of the re-streaked plate and this stock was used for inoculation every time. Each time, the tip of glycerol stock was scratched using a loop and inoculated into 5 ml of LB containing 100 µg/ml Ampicillin. After overnight incubation at 37°C, it was used for subculture of 200 – 400 ml of LB containing 100 µg/ml Ampicillin (1% inoculum). The flask was incubated at 37°C till the $O.D._{600}$ of the culture reached 0.6, and the culture was induced with 1 mM IPTG. The cultures were left to grow for a further 3 hours, and cells were harvested by centrifugation. The efficiency of induction was analysed by running the lysates of un-induced and induced cell pellets on a 10% SDS-PAGE.

For purification of the proteins, centrifuged cells were lysed in 10 ml of lysis buffer (50 mM Tris-HCl pH 7.4, 200 mM NaCl, 1% Triton X-100, 2.0 mM EDTA, 10% Glycerol, 1 mM DTT, 0.1 mM PMSF, 100 µg lysozyme and protease inhibitor cocktail) by incubation for 30 min on ice and finally disrupted by sonication for 15 min at high power (30 sec on and 30 sec off) in Bioruptor™ (Diagenode). Cell lysate was centrifuged at 18,000g for 15 min in fixed angle rotor to remove particulate matter and cell debris. The resulting supernatant was carefully collected and 1 ml of Glutathione-Sepharose (Amersham Cat. # 17-0756-01) beads were added to the lysate. The beads were incubated at 4°C for 3 hours with rotation to facilitate binding, and then washed thrice with 10 ml of wash buffer (same as lysis buffer), 30 min each at 4°C. The protein was eluted into 2 ml of elution buffer (50 mM Tris-HCl pH 8.0, 1 mM DTT, 10 mM GSH, 1.0 mM EDTA, 10% Glycerol) containing 1X protease inhibitor cocktail. Eluted proteins were concentrated using Amicon concentrator (10 kDa cut off). In case of Cbx5, concentrated GST column

eluate was further purified using ion exchange: the eluate was applied to pre-packed Heparin Syringe column at the rate of 0.5 ml/ minute. Heparin column was washed with 10 bed volumes of wash buffer (Heparin column binding buffer (20 mM HEPES pH 7.9, 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM NaCl, 0.5 mM EDTA, 10% Glycerol, and 1 mM DTT) with 200 mM KCl). The protein was eluted with elution buffer (Heparin column binding buffer with 500 mM KCl) and the homogeneity of the protein was checked on SDS-PAGE followed by staining with Coomassie Blue. The purity of hnRNPK was also analysed using Western blotting.

4.2.3 WESTERN BLOTTING:

Western blotting was performed as described by Towbin et al., (1979). Briefly, the protein samples were electrophoresed on 10% SDS-PAGE and the gel was equilibrated in transfer buffer (25 mM Glycine, 192 mM Tris and 20% methanol). Prewet Sponges, filter papers (slightly bigger than gel), gel and membrane (Amersham Hybond™ – P, GE Lifesciences Inc. Cat. # RPN303F) were arranged in transfer apparatus [Mighty Small Transphor (GE Lifesciences Inc. Cat. # 80 6204 26)] and transfer was done according to manufacturer's instructions. Briefly, after transfer the membrane was washed with 1X PBST (3.2 mM Na_2HPO_4, 0.5 mM KH_2PO_4, 1.3 mM KCl, 135 mM NaCl, 0.05% Tween 20, pH 7.4) and incubated in blocking buffer (5% non-fat dried milk powder in 1X PBST) for 4 hours at room temperature. Diluted primary antibody was added to the membrane and further incubated for 2 hours. Membrane was washed four times (10 min each) on a rocking platform with 1X PBST. Diluted secondary antibody was added and the membrane was incubated for 45 min, following which it was washed thrice with 1X PBST. Detection of protein was done using Amersham ECL Plus Western Blotting Detection System (GE Lifesciences Cat. # RPN2132) and exposed to X-ray films (Kodak BioMax MS Cat. # 822 2648).

4.2.4 PREPARATION OF OLIGO-NUCLEOTIDE PROBES:

Methylated and unmethylated double-stranded GC regions were generated by annealing equimolar concentrations of following oligonucleotides:

GC sense – CAGGCGGGCGGGGCAGGGGGCGGGGCCGGGCGGGCAAA

GC antisense – TTTGCCCGCCCGGCCCCGCCCCCTGCCCCGCCCGCCTG

methGC sense - CAGG^{m}CGGG^{m}CGGGGCAGGGGG^{m}CGGGGC^{m}CGGG^{m}CGGGCAAA

methGC antisense - TTTGCC^{m}CGCC^{m}CGGCCC^{m}CGCCCCCTGCCC^{m}CGCC^{m}CGCCTG

^{m}C - methylated cytosine

500 pmoles of sense strand and 500 pmoles of antisense strand of the oligos were annealed in 50 µl TEN buffer (10 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 50 mM NaCl) by incubating in boiling water for 5 min followed by slow cooling to room temperature (over approximately 1 hour). The annealed oligos were radio-labelled as follows: 50 pmoles of annealed oligo was incubated in a 50 µl reaction containing 1X T4 PNK buffer, 10 µCi (3 pmoles) of γ- ^{32}P (ATP) and 10 units T4 PNK (Polynucleotide Kinase, NEB) at 37°C for 45 min. The reaction was stopped by addition of EDTA to the final concentration of 10 mM. Purification of labelled DNA from unincorporated labelled nucleotides was done using G-50 spin column (GE Lifesciences Cat. # 27-5335-01).

4.2.5 PREPARATION OF NUCLEAR PROTEIN EXTRACT:

Nuclear extract was prepared essentially as described by Lewis and Konradi (1996). All the steps were carried out on ice unless otherwise mentioned. Tissues were homogenized in Dounce homogenizer (Sigma Inc. Cat # T0691) in 3 ml tissue homogenization buffer (Buffer A) (10 mM HEPES (pH 7.9), 10 mM KCl, 0.1 mM EDTA, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10% Glycerol, 1 mM DTT) in the presence of 1X Protease Inhibitors Cocktail (Roche, Cat. # 1836170) and suspended in 15 ml buffer A. The cell

suspension was incubated for 5-10 min on ice and filtered through 1 mm strainer to remove cell debris and particulate material. Cells were collected by centrifugation at 1500g for 5 min. Cells were washed twice in buffer A, and suspended in 5 ml of buffer A with 0.2% NP-40. Cell suspension was incubated for 5 min on ice. The suspension was carefully overlaid on 1 M sucrose buffer (10 mM HEPES (pH 7.9), 10 mM KCl, 0.1 mM EDTA, 1.5 mM $MgCl_2$, 1.0 M Sucrose, 10% Glycerol, 1 mM DTT) and the centrifugation was carried out for 15 min at 4500g in swing bucket rotor. Supernatant was discarded and the pellet containing nuclei was retained. Nuclei thus obtained were washed twice with buffer A to remove the trace amount of NP-40 and suspended in 2 ml of nuclear protein extraction buffer B (20 mM HEPES (pH 7.9), 500 mM KCl, 0.4 mM EDTA, 1.5 mM $MgCl_2$, 25% Glycerol, 1 mM DTT) with 1X protease inhibitors cocktail and incubated for 30 min on ice. The suspension was centrifuged for 15 min at 21,000g at 4^0C and supernatant containing nuclear proteins was collected. The concentration of total nuclear protein was determined by spectrophotometric analysis.

4.2.6 ELECTROPHORETIC MOBILITY SHIFT ASSAY:

Stock solution of 5X EMSA binding buffer was prepared with 50 mM HEPES, 50% Glycerol and 0.5 mM EDTA and stored at -20^0C. Working solution of EMSA binding buffer (4X binding buffer, 2.5 mM Spermidine, 5 mM $MgCl_2$, 25 ng/µl Poly dI-dC , 1 mM DTT) was prepared during the experiment. Each EMSA reaction contained in 16 µl 1X EMSA binding buffer, 100 pmoles of linearized pBluescript II vector as non-specific competitor and 1 pmole of labelled GC or methGC probe. For Cbx5, 0.5, 1 or 2 pmoles of protein was used per reaction. For hnRNPK, 0.2, 0.5 or 1 pmole of protein per reaction was used. 1 pmole of GST and 100µg of Liver Nuclear Extract were used as controls. 5% native PAGE used for electrophoresing. EMSA gel was prepared by mixing of 150 µl of 20% APS and 30 µl of TEMED with 5% Acrylamide-Bis solution (37.5 : 1), 3.5%

Glycerol, 0.5X TBE. The gel was pre-run for 2 hours until the amperage had dropped to 12 mA. After electrophoresis, gels were vacuum-dried at 80^0C and phosphor-imaged using Molecular dynamics software of Typhoon-9200 system from Amersham Biosciences or exposed to X-Ray films (Kodak BioMax MS Film Cat. # 822 2648).

4.3 RESULTS:

4.3.1 CBX5 PREFERENTIALLY BINDS TO UNMETHYLATED GC PROBE:

To analyse whether Cbx5 binds to the GC region, electrophoretic mobility shift assay (EMSA) was performed on methylated and unmethylated GC probes using purified GST-Cbx5 (refer to Materials and Methods). Initial analysis revealed that Cbx5 preferentially bound to the unmethylated GC probe. To confirm this observation, reaction kinetics was performed with progressively increasing concentration of GST-Cbx5 protein (See Materials and Methods). This study revealed that the shift was indeed because of Cbx5 as the intensity of the shift increased in a Cbx5-dependent manner (Fig. 4.1A). Though a shift was also observed with methylated GC probe, it was less prominent. To analyse the significance of difference observed between binding to methylated and unmethylated GC probes, densitometric analysis was performed on two individual EMSA experiments. As can be seen in Fig.4.1B, the difference was very significant (Two-way ANOVA (GC vs meth GC), $p < 0.0001$) between both the groups. This confirms that Cbx5 binds preferentially to the unmethylated GC probe.

4.3.2 PREFERENTIAL BINDING OF HNRNPK TO METHYLATED GC PROBE:

The ability of hnRNPK to bind to the GC region was checked using electrophoretic mobility shift assay (EMSA) with methylated and unmethylated GC probes using purified GST-hnRNPK (see Materials and Methods). hnRNPK was found to

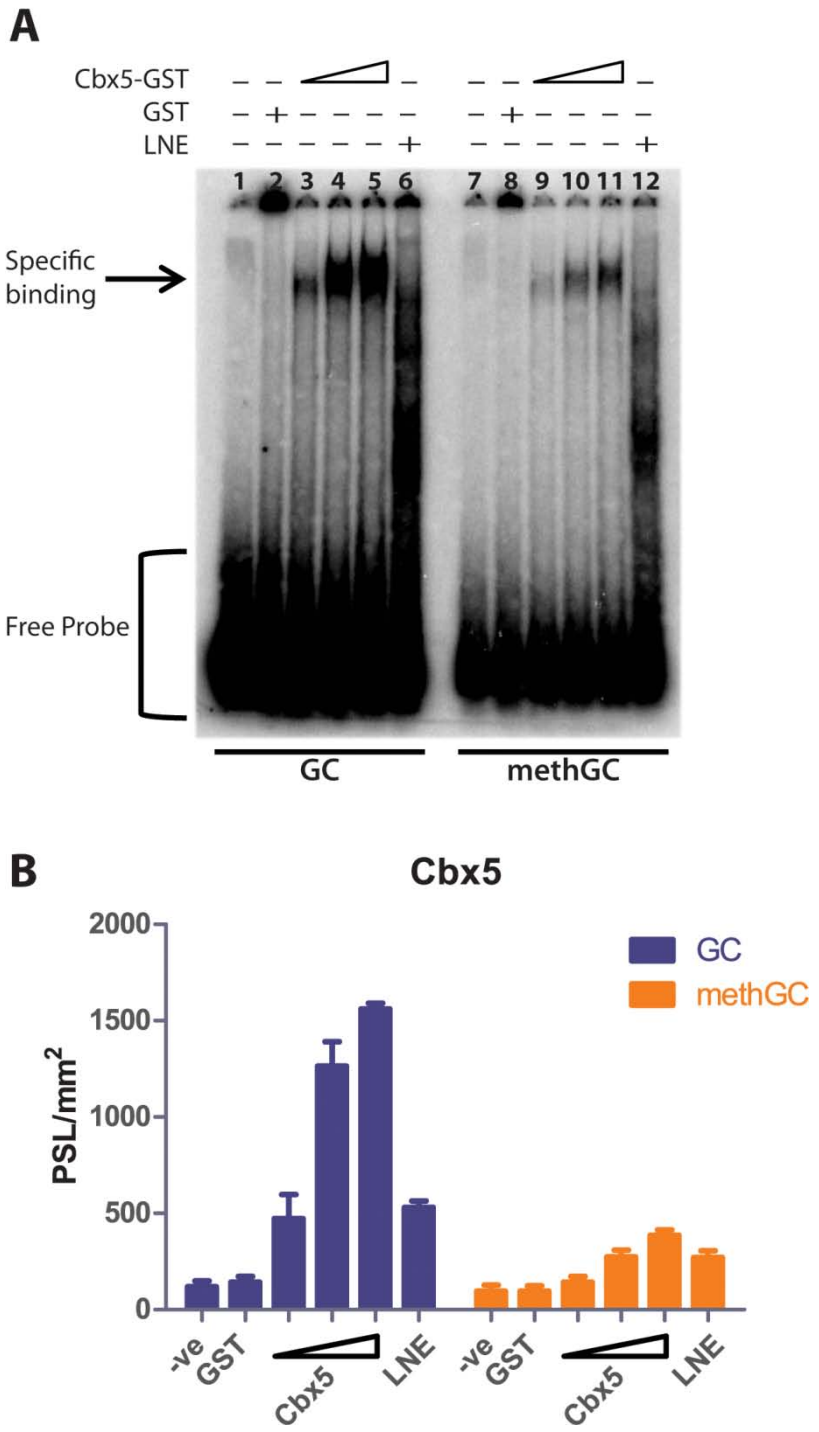


Figure 4.1: EMSA Analysis with purified GST-Cbx5. A: A representative EMSA gel showing binding of Cbx5 with GC and methGC probes. Lanes 1 and 7 represent negative controls where no protein was added. Lanes 2 and 8 contained only GST along with the probe. Lanes 3 to 5 and 9 to 11 indicate binding of GST-Cbx5 to unmethylated and methylated GC probes respectively, in a concentration dependent manner (0.5, 1 and 2 pmoles). Lanes 6 and 12 are positive controls (Liver Nuclear Extract). **B:** Densitometric analysis of EMSA experiments performed with GST-Cbx5. Graph shown is cumulative of two individual analyses. Error bars represent Standard Error of mean (SEM).

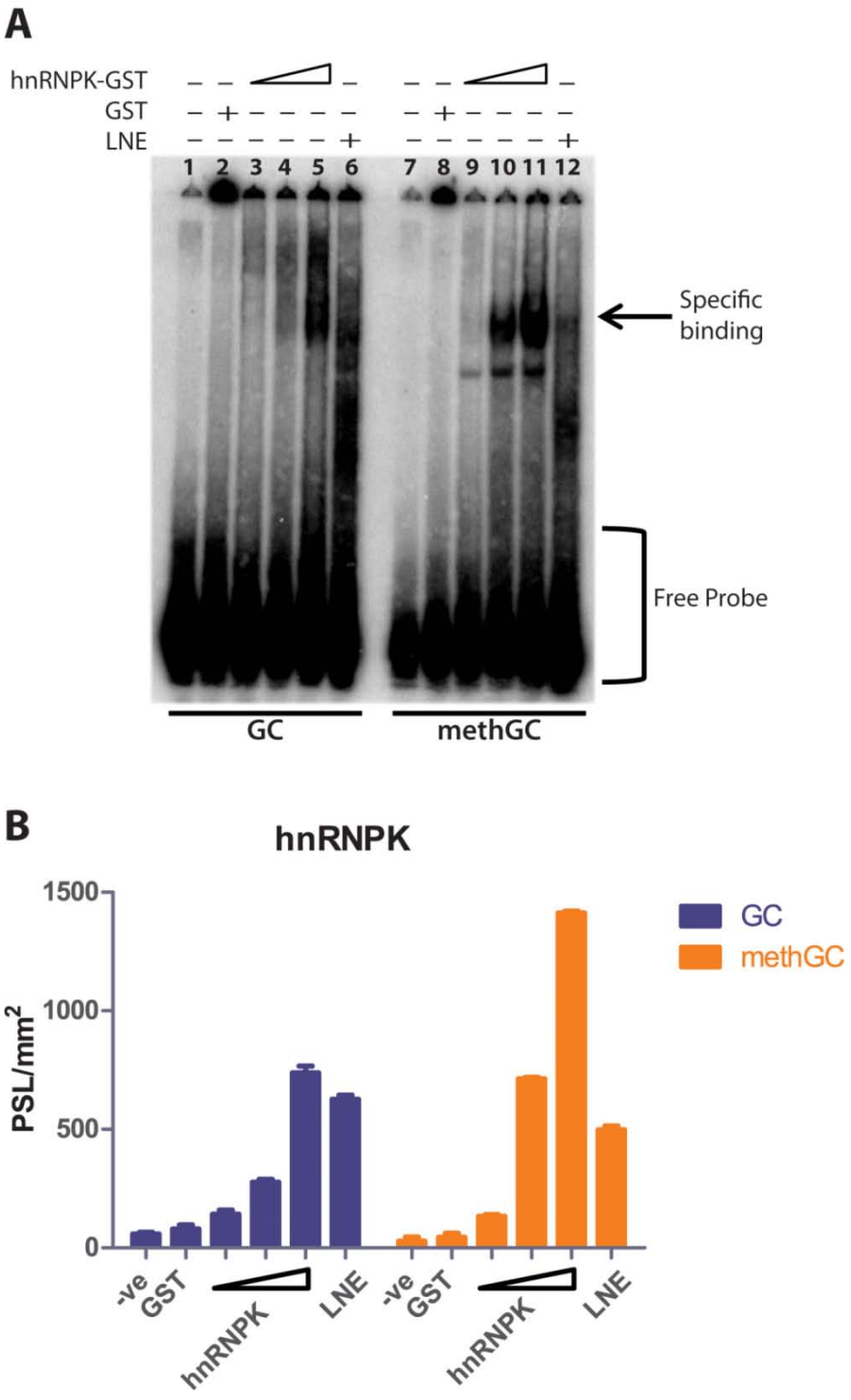


Figure 4.2: EMSA Analysis with purified GST-hnRNPK. A: A representative EMSA gel showing binding of hnRNPK with GC and methGC probes. Lanes 1 and 7 represent negative controls where no protein was added. Lanes 2 and 8 contained only GST along with the probe. Lanes 3 to 5 and 9 to 11 indicate binding of GST-hnRNPK to unmethylated and methylated GC probes respectively, in a concentration dependent manner (0.5, 1 and 2 pmoles). Lanes 6 and 12 are positive controls (Liver Nuclear Extract). **B:** Densitometric analysis of EMSA experiments performed with GST-hnRNPK. Graph shown is cumulative of two individual analyses. Error bars represent Standard Error of mean (SEM).

preferentially bind to the methylated GC probe. Reaction kinetics with increasing concentration of GST-hnRNPK indicated that the shift was hnRNPK-dependent and the intensity of the shifted complex increased with an increase in the concentration of hnRNPK (Fig. 4.2A). A lesser prominent shift was also observed with unmethylated GC probe at higher concentration of hnRNPK (1 pmole). Densitometric analysis on two individual EMSA experiments indicated that the difference observed between methylated and unmethylated probes was very significant (Fig.4.2B, Two-way ANOVA (GC vs meth GC), $p < 0.0001$), confirming the finding that hnRNPK preferentially binds to the methylated GC probe.

4.4 CONCLUSIONS:

Two proteins, Cbx5 and hnRNPK, were tested for their ability to bind the GC region present within the NNI^2 in a methylation-dependent manner. To study this, EMSA was performed using purified proteins with methylated and unmethylated GC oligos. Both the proteins have shown distinct binding profiles; Cbx5 was observed to bind preferentially to the unmethylated GC oligo, while hnRNPK preferentially bound to the methylated GC oligo probe. hnRNPK is a ribonuclear protein known to be involved in gene activation by binding to secondary DNA structures (Uribe et al., 2011). Though Cbx5 is primarily associated with heterochromatin and gene silencing, several recent reports have indicated its role in gene activation (Kwon and Workman, 2011; Piacentini et al., 2003). Whether Cbx5 and hnRNPK proteins are involved in allele-specific silencing or activation of *Neuronatin* and *Blcapv1a/v2a,* needs to be examined further. However, our results point to a complex interplay of DNA-protein interactions within NNI^2 that may be involved in allele-specific regulation of *Neuronatin* and *Blcap*.

CHAPTER V

DISCUSSION

Several mechanisms have been put forward to explain the regulation of imprinted expression at various loci. In many cases, the transcription of a non-coding RNA antisense to protein coding genes results in their perturbed expression, while in case of others, binding of insulator proteins to their target sites prevent the interaction of gene promoters with enhancers (Fitzpatrick et al., 2002; Lin et al., 2003; Thorvaldsen et al., 1998; Williamson et al., 2006; Wutz et al., 1997). However, all mechanisms studied till now suggest that the regulatory regions or ICRs of the loci bring about the silencing of actively expressed genes. The mechanisms that establish and maintain the transcription from the active allele of an imprinted gene remain ambiguous. Mechanisms that prevent silencing of the active allele are also not known. Our analysis of the imprinting control region (ICR) for the mouse *Neuronatin* locus, harbouring two imprinted genes – *Neuronatin* and *Blcap/Bc10*, points to its role as a transcriptional activator.

An ICR is a region within the imprinted locus that is important for maintaining correct allele-specific expression of imprinted genes present within a locus. Biochemically, the ICR is the only region within the imprinted locus that shows a mutual exclusiveness of DNA methylation on one allele and specialized chromatin conformation on the other allele (Delaval and Feil, 2004; Kato and Sasaki, 2005; Khosla et al., 1999). Previous work from our laboratory has shown that the second intronic region of *Neuronatin* (NNI^2) displayed such a mutual exclusiveness, with the nuclease hypersensitivity site present only on the transcribed paternal allele whereas the maternal allele is methylated (Sowpati et al., 2008). This suggested that NNI^2 could be the putative ICR for this locus. In this thesis, we tested this hypothesis by performing

assays to test the role of second intron of *Neuronatin* in i) transcriptional regulation and ii) controlling imprinting status of mouse *Neuronatin* gene.

To assess the functional role of *cis* elements present within this putative ICR, we performed a transgene reporter assay in *Drosophila.* A previous study had identified the *H19* ICR to be a silencer in a similar reporter gene assay in *Drosophila* (Lyko et al., 1997). However, in contrast to *H19* ICR, NNI^2 was found to be a transcriptional activator. Similar to *H19* ICR, other ICRs have also been shown to be silencers (Delaval and Feil, 2004; Lewis and Reik, 2006). Therefore, contrary to other ICRs, NNI^2 was found to be a transcriptional activator. This observation was further strengthened by analysis of mice deleted for NNI^2. Deletion of NNI^2 leads to loss of *Neuronatin* expression from the normally active paternal allele, whereas deletion of NNI^2 from the maternal allele had no effect on *Neuronatin* expression. Recently, *Blcap*, the other gene in this locus, was found to be transcript- and tissue-specifically imprinted (Schulz et al., 2009). As imprinting control regions are known to control the expression of more than one imprinted gene in the locus, in this thesis we had also looked at *Blcap* expression upon deletion of NNI^2. As shown in Chapter III, deletion of NNI^2 does affect allele-specific expression of *Blcapv1a* and *Blcapv2a*; In E13.5 embryos and adult brain, where there was total loss of *Neuronatin* expression from the paternal allele, *Blcapv1a* expression was reduced by approximately two-fold exclusively on the methylated maternal allele. Since *Blcapv1a* is expressed only from the maternal allele, these results again pointed to the role of NNI^2 as a transcriptional activator, albeit for *Blcapv1a* and in methylated state. This further strengthened the credentials of NNI^2 as a transcriptional activator. Interestingly, in unmethylated state, NNI^2 acted as an activator for the *Neuronatin* gene but activated the transcription of *Blcapv1a* in its methylated state. How

this duality of function is achieved by NNI^2 would throw more light on the imprinting mechanism for imprinted genes in general and the *Neuronatin*/*Blcap* locus in particular. Unlike E13.5 embryos or adult brain, the role of NNI^2 in *Neuronatin* and *Blcap* transcription in other tissues seems to be influenced by tissue-specific enhancers. Similar tissue-specific effect of NNΔI^2 was also observed on *Blcapv2a* transcription. Though *Blcapv2a*, which shares its promoter with *Neuronatin*, was also found to have allele-specific transcription, its expression in various tissues reflected a cumulative effect of tissue-specific enhancers and NNI^2 on its promoter. Further work is needed to not only identify these tissue-specific enhancers but also to understand how these enhancers influence the *Neuronatin* and *Blcap* promoters in conjugation with NNI^2.

Allele-specific expression of imprinted genes is correlated with epigenetic status of the two alleles (Li et al., 1993; Lucifero et al., 2002). DNA methylation has been associated with the silent allele of most imprinted genes (Edwards and Ferguson-Smith, 2007). Previous studies have shown the silent maternal allele of *Neuronatin* to be methylated whereas the active paternal allele lacks DNA methylation (John et al., 2001; Kagitani et al., 1997; Kikyo et al., 1997). Preliminary analysis of DNA methylation at the *Neuronatin* promoter in NNΔI^2 mice showed gain of methylation concomitant with loss of *Neuronatin* expression from the paternal allele. This indicated that i) *Neuronatin* expression is correlated with its promoter methylation and ii) NNI^2 is responsible for setting up of DNA methylation marks at *Neuronatin* promoter. Level of *Blcapv1a* expression also showed changes in NNΔI^2 mice. *Blcapv1a* promoter has previously been shown to be unmethylated on both alleles. Whether deletion of NNI^2 also changes the methylation profile of its promoter remains to be examined. In addition, investigating the histone modifications that are allele-specifically associated with *Neuronatin*/*Blcap*

locus would shed further light on the role of epigenetic modifications in the imprinted expression of *Neuronatin* and *Blcap*.

Though the role of *Neuronatin* was initially thought to be restricted to embryonic development, a few recent studies have identified its role in regulation of metabolism, especially in insulin secretion, diabetes and obesity (Joe et al., 2008; Mzhavia et al., 2008; Suh et al., 2005; Vrang et al., 2010). Vrang et al (2010) have reported a correlated decrease in levels of *Neuronatin* mRNA in hypothalamus of mice carrying deletion of *Ob*, the gene that encodes Leptin, a prominent player in hunger and appetite. NNΔI^2 mice showed change in average body weight as compared to wild type mice. This indicated that *Neuronatin* indeed was involved in some pathway regulating metabolism. Moreover, all mice have lost a control on their body weight by an age of ~1 year. However, the phenotypic differences observed between paternal heterozygous and homozygous NNΔI^2 mice (gained body weight) and maternal heterozygous NNΔI^2 mice (loss of body weight) suggested a more complicated scenario. Maternal heterozygous NNΔI^2 had *Neuronatin* expression similar to wild type mice in E13.5 embryos and adult brain. This would indicate that *Neuronatin* expression during embryogenesis (at E13.5) and adult brain is not responsible for the observed loss of body weight. It is possible that the aberrant *Neuronatin* expression in other tissues could be the responsible factor. It is also possible that change in expression of *Blcapv1a* and *Blcapv2a* were responsible for loss of body weight in maternal heterozygous NNΔI^2 mice. Additional experiments would be required to dissect out the reasons for these changes.

BIBLIOGRAPHY

1. Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F.*, et al.* (2000). The genome sequence of Drosophila melanogaster. Science *287*, 2185-2195.
2. Albrecht, U., Sutcliffe, J.S., Cattanach, B.M., Beechey, C.V., Armstrong, D., Eichele, G., and Beaudet, A.L. (1997). Imprinted expression of the murine Angelman syndrome gene, Ube3a, in hippocampal and Purkinje neurons. Nat Genet *17*, 75-78.
3. Alkema, M.J., Bronk, M., Verhoeven, E., Otte, A., van 't Veer, L.J., Berns, A., and van Lohuizen, M. (1997). Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian polycomb complex. Genes Dev *11*, 226-240.
4. Antequera, F., and Bird, A. (1993). Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci U S A *90*, 11995-11999.
5. Arnaud, P., Monk, D., Hitchins, M., Gordon, E., Dean, W., Beechey, C.V., Peters, J., Craigen, W., Preece, M., Stanier, P.*, et al.* (2003). Conserved methylation imprints in the human and mouse GRB10 genes with divergent allelic expression suggests differential reading of the same mark. Hum Mol Genet *12*, 1005-1019.
6. Barlow, D.P., Stoger, R., Herrmann, B.G., Saito, K., and Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature *349*, 84-87.
7. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.
8. Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian genomic imprinting. Cold Spring Harb Perspect Biol *3*.
9. Bartolomei, M.S., Webber, A.L., Brunkow, M.E., and Tilghman, S.M. (1993). Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. Genes Dev *7*, 1663-1673.
10. Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153-155.
11. Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482-485.
12. Benevolenskaya, E.V. (2007). Histone H3K4 demethylases are essential in development and differentiation. Biochem Cell Biol *85*, 435-443.

13. Berg, C.A., and Spradling, A.C. (1991). Studies on the rate and site-specificity of P element transposition. Genetics *127*, 515-524.
14. Bhogal, B., Arnaudo, A., Dymkowski, A., Best, A., and Davis, T.L. (2004). Methylation at mouse Cdkn1c is acquired during postimplantation development and functions to maintain imprinted expression. Genomics *84*, 961-970.
15. Bhutani, N., Burns, D.M., and Blau, H.M. (2011). DNA demethylation dynamics. Cell *146*, 866-872.
16. Bird, A.P. (1995). Gene number, noise reduction and biological complexity. Trends Genet *11*, 94-100.
17. Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E*., et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.
18. Bjorklund, S., and Kim, Y.J. (1996). Mediator of transcriptional regulation. Trends Biochem Sci *21*, 335-337.
19. Blanton, J., Gaszner, M., and Schedl, P. (2003). Protein:protein interactions and the pairing of boundary elements in vivo. Genes Dev *17*, 664-675.
20. Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Mielke, C., and Kohwi-Shigematsu, T. (1992). Biological significance of unwinding capability of nuclear matrix-associating DNAs. Science *255*, 195-197.
21. Brand, A.H., Breeden, L., Abraham, J., Sternglanz, R., and Nasmyth, K. (1985). Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell *41*, 41-48.
22. Buiting, K., Lich, C., Cottrell, S., Barnicoat, A., and Horsthemke, B. (1999). A 5-kb imprinting center deletion in a family with Angelman syndrome reduces the shortest region of deletion overlap to 880 bp. Hum Genet *105*, 665-666.
23. Bushey, A.M., Dorman, E.R., and Corces, V.G. (2008). Chromatin insulators: regulatory mechanisms and epigenetic inheritance. Mol Cell *32*, 1-9.
24. Cattanach, B.M., and Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. Nature *315*, 496-498.
25. Cavaille, J., Seitz, H., Paulsen, M., Ferguson-Smith, A.C., and Bachellerie, J.P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted

human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. Hum Mol Genet *11*, 1527-1538.

26. Chai, J.H., Locke, D.P., Ohta, T., Greally, J.M., and Nicholls, R.D. (2001). Retrotransposed genes such as Frat3 in the mouse Chromosome 7C Prader-Willi syndrome region acquire the imprinted status of their insertion site. Mamm Genome *12*, 813-821.
27. Chakravarti, D., LaMorte, V.J., Nelson, M.C., Nakajima, T., Schulman, I.G., Juguilon, H., Montminy, M., and Evans, R.M. (1996). Role of CBP/P300 in nuclear receptor signalling. Nature *383*, 99-103.
28. Charalambous, M., da Rocha, S.T., and Ferguson-Smith, A.C. (2007). Genomic imprinting, growth control and the allocation of nutritional resources: consequences for postnatal life. Curr Opin Endocrinol Diabetes Obes *14*, 3-12.
29. Choi, J.D., Underkoffler, L.A., Wood, A.J., Collins, J.N., Williams, P.T., Golden, J.A., Schuster, E.F., Jr., Loomes, K.M., and Oakey, R.J. (2005). A novel variant of Inpp5f is imprinted in brain, and its expression is correlated with differential methylation of an internal CpG island. Mol Cell Biol *25*, 5514-5522.
30. Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., and Kelsey, G. (2009). Transcription is required for establishment of germ line methylation marks at imprinted genes. Genes Dev *23*, 105-117.
31. Chu, K., and Tsai, M.J. (2005). Neuronatin, a downstream target of BETA2/NeuroD1 in the pancreas, is involved in glucose-mediated insulin secretion. Diabetes *54*, 1064-1073.
32. Clark, R.M., Marker, P.C., and Kingsley, D.M. (2000). A novel candidate gene for mouse and human preaxial polydactyly with altered expression in limbs of Hemimelic extra-toes mutant mice. Genomics *67*, 19-27.
33. Clayton, A.L., and Mahadevan, L.C. (2003). MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. FEBS Lett *546*, 51-58.
34. Comb, M., and Goodman, H.M. (1990). CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. Nucleic Acids Res *18*, 3975-3982.
35. Coombes, C., Arnaud, P., Gordon, E., Dean, W., Coar, E.A., Williamson, C.M., Feil, R., Peters, J., and Kelsey, G. (2003). Epigenetic properties and identification of an

imprint mark in the Nesp-Gnasxl domain of the mouse Gnas imprinted locus. Mol Cell Biol *23*, 5475-5488.

36. Crouse, H.V. (1960). The Controlling Element in Sex Chromosome Behavior in Sciara. Genetics *45*, 1429-1443.
37. DeChiara, T.M., Robertson, E.J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. Cell *64*, 849-859.
38. Delaval, K., and Feil, R. (2004). Epigenetic regulation of mammalian genomic imprinting. Curr Opin Genet Dev *14*, 188-195.
39. Dindot, S.V., Person, R., Strivens, M., Garcia, R., and Beaudet, A.L. (2009). Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions. Genome Res *19*, 1374-1383.
40. Dong, J.M., and Lim, L. (1996). The human neuronal alpha 1-chimaerin gene contains a position-dependent negative regulatory element in the first exon. Neurochem Res *21*, 1023-1030.
41. Drewell, R.A., Arney, K.L., Arima, T., Barton, S.C., Brenton, J.D., and Surani, M.A. (2002). Novel conserved elements upstream of the H19 gene are transcribed and act as mesodermal enhancers. Development *129*, 1205-1213.
42. Dunn, K.L., and Davie, J.R. (2003). The many roles of the transcriptional regulator CTCF. Biochem Cell Biol *81*, 161-167.
43. Edwards, C.A., and Ferguson-Smith, A.C. (2007). Mechanisms regulating imprinted genes in clusters. Curr Opin Cell Biol *19*, 281-289.
44. Engels, W.R. (1992). The origin of P elements in Drosophila melanogaster. Bioessays *14*, 681-686.
45. Engemann, S., Strodicke, M., Paulsen, M., Franck, O., Reinhardt, R., Lane, N., Reik, W., and Walter, J. (2000). Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. Hum Mol Genet *9*, 2691-2706.
46. Evans, H.K., Wylie, A.A., Murphy, S.K., and Jirtle, R.L. (2001). The neuronatin gene resides in a "micro-imprinted" domain on human chromosome 20q11.2. Genomics *77*, 99-104.

47. Farkas, G., and Udvardy, A. (1992). Sequence of scs and scs' Drosophila DNA fragments with boundary function in the control of gene expression. Nucleic Acids Res *20*, 2604.
48. Feil, R., Boyano, M.D., Allen, N.D., and Kelsey, G. (1997). Parental chromosome-specific chromatin conformation in the imprinted U2af1-rs1 gene in the mouse. J Biol Chem *272*, 20893-20900.
49. Feil, R., Walter, J., Allen, N.D., and Reik, W. (1994). Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes. Development *120*, 2933-2943.
50. Ferguson-Smith, A.C., Sasaki, H., Cattanach, B.M., and Surani, M.A. (1993). Parental-origin-specific epigenetic modification of the mouse H19 gene. Nature *362*, 751-755.
51. Fischle, W., Wang, Y., and Allis, C.D. (2003). Histone and chromatin cross-talk. Curr Opin Cell Biol *15*, 172-183.
52. Fitzpatrick, G.V., Soloway, P.D., and Higgins, M.J. (2002). Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. Nat Genet *32*, 426-431.
53. Frenkel, B., Mijnes, J., Aronow, M.A., Zambetti, G., Banerjee, C., Stein, J.L., Lian, J.B., and Stein, G.S. (1993). Position and orientation-selective silencer in protein-coding sequences of the rat osteocalcin gene. Biochemistry *32*, 13636-13643.
54. Fryer, C.J., and Archer, T.K. (1998). Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. Nature *393*, 88-91.
55. Gaszner, M., Vazquez, J., and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. Genes Dev *13*, 2098-2107.
56. Gerasimova, T.I., Byrd, K., and Corces, V.G. (2000). A chromatin insulator determines the nuclear localization of DNA. Mol Cell *6*, 1025-1035.
57. Gerber, H.P., Hagmann, M., Seipel, K., Georgiev, O., West, M.A., Litingtung, Y., Schaffner, W., and Corden, J.L. (1995). RNA polymerase II C-terminal domain required for enhancer-driven transcription. Nature *374*, 660-662.
58. Geuns, E., De Rycke, M., Van Steirteghem, A., and Liebaers, I. (2003). Methylation imprints of the imprint control region of the SNRPN-gene in human gametes and preimplantation embryos. Hum Mol Genet *12*, 2873-2879.

59. Geyer, P.K., and Corces, V.G. (1992). DNA position-specific repression of transcription by a Drosophila zinc finger protein. Genes Dev *6*, 1865-1873.
60. Giddings, S.J., King, C.D., Harman, K.W., Flood, J.F., and Carnaghi, L.R. (1994). Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. Nat Genet *6*, 310-313.
61. Gidoni, D., Dynan, W.S., and Tjian, R. (1984). Multiple specific contacts between a mammalian transcription factor and its cognate promoters. Nature *312*, 409-413.
62. Gokul, G., Gautami, B., Malathi, S., Sowjanya, A.P., Poli, U.R., Jain, M., Ramakrishna, G., and Khosla, S. (2007). DNA methylation profile at the DNMT3L promoter: a potential biomarker for cervical cancer. Epigenetics *2*, 80-85.
63. Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X., Golic, K.G., Jacobsen, S.E., and Bestor, T.H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science *311*, 395-398.
64. Green, C.M., and Almouzni, G. (2002). When repair meets chromatin. First in series on chromatin dynamics. EMBO Rep *3*, 28-33.
65. Haecker, S.A., Muramatsu, T., Sensenbaugh, K.R., and Sanders, M.M. (1995). Repression of the ovalbumin gene involves multiple negative elements including a ubiquitous transcriptional silencer. Mol Endocrinol *9*, 1113-1126.
66. Haines, T.R., Rodenhiser, D.I., and Ainsworth, P.J. (2001). Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. Dev Biol *240*, 585-598.
67. Hanel, M.L., and Wevrick, R. (2001). Establishment and maintenance of DNA methylation patterns in mouse Ndn: implications for maintenance of imprinting in target genes of the imprinting center. Mol Cell Biol *21*, 2384-2392.
68. Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486-489.
69. Hark, A.T., and Tilghman, S.M. (1998). Chromatin conformation of the H19 epigenetic mark. Hum Mol Genet *7*, 1979-1985.
70. Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol *18*, 6538-6547.
71. Hershko, A., Razin, A., and Shemer, R. (1999). Imprinted methylation and its effect on expression of the mouse Zfp127 gene. Gene *234*, 323-327.

72. Hikichi, T., Kohda, T., Kaneko-Ishino, T., and Ishino, F. (2003). Imprinting regulation of the murine Meg1/Grb10 and human GRB10 genes; roles of brain-specific promoters and mouse-specific CTCF-binding sites. Nucleic Acids Res *31*, 1398-1406.

73. Hiura, H., Obata, Y., Komiyama, J., Shirai, M., and Kono, T. (2006). Oocyte growth-dependent progression of maternal imprinting in mice. Genes Cells *11*, 353-361.

74. Hiura, H., Sugawara, A., Ogawa, H., John, R.M., Miyauchi, N., Miyanari, Y., Horiike, T., Li, Y., Yaegashi, N., Sasaki, H*., et al.* (2010). A tripartite paternally methylated region within the Gpr1-Zdbf2 imprinted domain on mouse chromosome 1 identified by meDIP-on-chip. Nucleic Acids Res *38*, 4929-4945.

75. Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300-1303.

76. Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

77. Jiang, N., Emberly, E., Cuvier, O., and Hart, C.M. (2009). Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in Drosophila melanogaster links BEAF to transcription. Mol Cell Biol *29*, 3556-3568.

78. Joe, M.K., Lee, H.J., Suh, Y.H., Han, K.L., Lim, J.H., Song, J., Seong, J.K., and Jung, M.H. (2008). Crucial roles of neuronatin in insulin secretion and high glucose-induced apoptosis in pancreatic beta-cells. Cell Signal *20*, 907-915.

79. John, R.M., Aparicio, S.A., Ainscough, J.F., Arney, K.L., Khosla, S., Hawker, K., Hilton, K.J., Barton, S.C., and Surani, M.A. (2001). Imprinted expression of neuronatin from modified BAC transgenes reveals regulation by distinct and distant enhancers. Dev Biol *236*, 387-399.

80. John, R.M., and Lefebvre, L. (2011). Developmental regulation of somatic imprints. Differentiation *81*, 270-280.

81. Jones, K.A., Yamamoto, K.R., and Tjian, R. (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell *42*, 559-572.

82. Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. Science *293*, 1068-1070.

83. Joseph, R., Dou, D., and Tsang, W. (1994). Molecular cloning of a novel mRNA (neuronatin) that is highly expressed in neonatal mammalian brain. Biochem Biophys Res Commun *201*, 1227-1234.
84. Joseph, R., Dou, D., and Tsang, W. (1995). Neuronatin mRNA: alternatively spliced forms of a novel brain-specific mammalian developmental gene. Brain Res *690*, 92-98.
85. Kagitani, F., Kuroiwa, Y., Wakana, S., Shiroishi, T., Miyoshi, N., Kobayashi, S., Nishida, M., Kohda, T., Kaneko-Ishino, T., and Ishino, F. (1997). Peg5/Neuronatin is an imprinted gene located on sub-distal chromosome 2 in the mouse. Nucleic Acids Res *25*, 3428-3432.
86. Kashiwagi, A., Meguro, M., Hoshiya, H., Haruta, M., Ishino, F., Shibahara, T., and Oshimura, M. (2003). Predominant maternal expression of the mouse Atp10c in hippocampus and olfactory bulb. J Hum Genet *48*, 194-198.
87. Kass, S.U., Landsberger, N., and Wolffe, A.P. (1997). DNA methylation directs a time-dependent repression of transcription initiation. Curr Biol *7*, 157-165.
88. Kato, Y., and Sasaki, H. (2005). Imprinting and looping: epigenetic marks control interactions between regulatory elements. Bioessays *27*, 1-4.
89. Kaufman, P.D., Doll, R.F., and Rio, D.C. (1989). Drosophila P element transposase recognizes internal P element DNA sequences. Cell *59*, 359-371.
90. Kelsey, G. (2007). Genomic imprinting--roles and regulation in development. Endocr Dev *12*, 99-112.
91. Kelsey, G., Bodle, D., Miller, H.J., Beechey, C.V., Coombes, C., Peters, J., and Williamson, C.M. (1999). Identification of imprinted loci by methylation-sensitive representational difference analysis: application to mouse distal chromosome 2. Genomics *62*, 129-138.
92. Khosla, S., Aitchison, A., Gregory, R., Allen, N.D., and Feil, R. (1999). Parental allele-specific chromatin configuration in a boundary-imprinting-control element upstream of the mouse H19 gene. Mol Cell Biol *19*, 2556-2566.
93. Kidwell, M.G., Kidwell, J.F., and Sved, J.A. (1977). Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. Genetics *86*, 813-833.
94. Kiefer, J.C. (2007). Epigenetics in development. Dev Dyn *236*, 1144-1156.

95. Kikyo, N., Williamson, C.M., John, R.M., Barton, S.C., Beechey, C.V., Ball, S.T., Cattanach, B.M., Surani, M.A., and Peters, J. (1997). Genetic and functional analysis of neuronatin in mice with maternal or paternal duplication of distal Chr 2. Dev Biol *190*, 66-77.
96. Kim, J. (2008). Multiple YY1 and CTCF binding sites in imprinting control regions. Epigenetics *3*, 115-118.
97. Kim, J., Kollhoff, A., Bergmann, A., and Stubbs, L. (2003). Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. Hum Mol Genet *12*, 233-245.
98. Kim, M.K., Lesoon-Wood, L.A., Weintraub, B.D., and Chung, J.H. (1996). A soluble transcription factor, Oct-1, is also found in the insoluble nuclear matrix and possesses silencing activity in its alanine-rich domain. Mol Cell Biol *16*, 4366-4377.
99. Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231-1245.
100. Kingston, R.E., Bunker, C.A., and Imbalzano, A.N. (1996). Repression and activation by multiprotein complexes that alter chromatin structure. Genes Dev *10*, 905-920.
101. Kobayashi, H., Yamada, K., Morita, S., Hiura, H., Fukuda, A., Kagami, M., Ogata, T., Hata, K., Sotomaru, Y., and Kono, T. (2009). Identification of the mouse paternally expressed imprinted gene Zdbf2 on chromosome 1 and its imprinted human homolog ZDBF2 on chromosome 2. Genomics *93*, 461-472.
102. Kong, S., Bohl, D., Li, C., and Tuan, D. (1997). Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. Mol Cell Biol *17*, 3955-3965.
103. Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.
104. Krauss, V., and Reuter, G. (2011). DNA methylation in Drosophila--a critical evaluation. Prog Mol Biol Transl Sci *101*, 177-191.
105. Kwon, S.H., and Workman, J.L. (2011). The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. Bioessays *33*, 280-289.

106. Laird, P.W. (1997). Oncogenic mechanisms mediated by DNA methylation. Mol Med Today *3*, 223-229.
107. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.
108. Landers, M., Bancescu, D.L., Le Meur, E., Rougeulle, C., Glatt-Deeley, H., Brannan, C., Muscatelli, F., and Lalande, M. (2004). Regulation of the large (approximately 1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. Nucleic Acids Res *32*, 3480-3492.
109. Latchman, D.S. (2008). Eukaryotic transcription factors, 5th edn (Amsterdam ; Boston, Elsevier/Academic Press).
110. Lefebvre, L., Viville, S., Barton, S.C., Ishino, F., and Surani, M.A. (1997). Genomic structure and parent-of-origin-specific methylation of Peg1. Hum Mol Genet *6*, 1907-1915.
111. Leighton, P.A., Saam, J.R., Ingram, R.S., Stewart, C.L., and Tilghman, S.M. (1995). An enhancer deletion affects both H19 and Igf2 expression. Genes Dev *9*, 2079-2089.
112. Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet *12*, 1725-1735.
113. Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151.
114. Lewis, A., Green, K., Dawson, C., Redrup, L., Huynh, K.D., Lee, J.T., Hemberger, M., and Reik, W. (2006). Epigenetic dynamics of the Kcnq1 imprinted domain in the early embryo. Development *133*, 4203-4210.
115. Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., Higgins, M., Feil, R., and Reik, W. (2004). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. Nat Genet *36*, 1291-1295.
116. Lewis, A., and Reik, W. (2006). How imprinting centres work. Cytogenet Genome Res *113*, 81-89.

117. Lewis, S.E., and Konradi, C. (1996). Analysis of DNA-Protein Interactions in the Nervous System Using the Electrophoretic Mobility Shift Assay. Methods *10*, 301-311.
118. Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.
119. Li, L.L., Szeto, I.Y., Cattanach, B.M., Ishino, F., and Surani, M.A. (2000). Organization and parent-of-origin-specific methylation of imprinted Peg3 gene on mouse proximal chromosome 7. Genomics *63*, 333-340.
120. Lin, S.P., Coan, P., da Rocha, S.T., Seitz, H., Cavaille, J., Teng, P.W., Takada, S., and Ferguson-Smith, A.C. (2007). Differential regulation of imprinting in the murine embryo and placenta by the Dlk1-Dio3 imprinting control region. Development *134*, 417-426.
121. Lin, S.P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., and Ferguson-Smith, A.C. (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. Nat Genet *35*, 97-102.
122. Ling, J., Ainol, L., Zhang, L., Yu, X., Pi, W., and Tuan, D. (2004). HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. J Biol Chem *279*, 51704-51713.
123. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M.*, et al.* (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322.
124. Liu, J., Chen, M., Deng, C., Bourc'his, D., Nealon, J.G., Erlichman, B., Bestor, T.H., and Weinstein, L.S. (2005). Identification of the control region for tissue-specific imprinting of the stimulatory G protein alpha-subunit. Proc Natl Acad Sci U S A *102*, 5513-5518.
125. Lucifero, D., Mann, M.R., Bartolomei, M.S., and Trasler, J.M. (2004). Gene-specific timing and epigenetic memory in oocyte imprinting. Hum Mol Genet *13*, 839-849.
126. Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., and Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. Genomics *79*, 530-538.

127. Lyko, F., Brenton, J.D., Surani, M.A., and Paro, R. (1997). An imprinting element from the mouse H19 locus functions as a silencer in Drosophila. Nat Genet *16*, 171-173.
128. Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S., and Tilghman, S.M. (2006). Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. Genes Dev *20*, 1268-1282.
129. McGrath, J., and Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell *37*, 179-183.
130. Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.
131. Miller, C.A., and Sweatt, J.D. (2007). Covalent modification of DNA regulates memory formation. Neuron *53*, 857-869.
132. Moore, T., Constancia, M., Zubair, M., Bailleul, B., Feil, R., Sasaki, H., and Reik, W. (1997). Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. Proc Natl Acad Sci U S A *94*, 12509-12514.
133. Muller, J. (1995). Transcriptional silencing by the Polycomb protein in Drosophila embryos. EMBO J *14*, 1209-1220.
134. Mullins, M.C., Rio, D.C., and Rubin, G.M. (1989). cis-acting DNA sequence requirements for P-element transposition. Genes Dev *3*, 729-738.
135. Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat Genet *36*, 889-893.
136. Mzhavia, N., Yu, S., Ikeda, S., Chu, T.T., Goldberg, I., and Dansky, H.M. (2008). Neuronatin: a new inflammation gene expressed on the aortic endothelium of diabetic mice. Diabetes *57*, 2774-2783.
137. Nan, X., Cross, S., and Bird, A. (1998). Gene silencing by methyl-CpG-binding proteins. Novartis Found Symp *214*, 6-16; discussion 16-21, 46-50.
138. Narlikar, G.J., Fan, H.Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475-487.

139. Nelson-Rees, W.A. (1962). The effects of radiation damaged heterochromatic chromosomes on male fertility in the mealy bug, planococcus citri (risso). Genetics *47*, 661-683.
140. Nicholls, R.D., and Knepper, J.L. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. Annu Rev Genomics Hum Genet *2*, 153-175.
141. Noyer-Weidner, M., and Trautner, T.A. (1993). Methylation of DNA in prokaryotes. EXS *64*, 39-108.
142. O'Hare, K., and Rubin, G.M. (1983). Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. Cell *34*, 25-35.
143. Ohta, T., Gray, T.A., Rogan, P.K., Buiting, K., Gabriel, J.M., Saitoh, S., Muralidhar, B., Bilienska, B., Krajewska-Walasek, M., Driscoll, D.J.*, et al.* (1999). Imprinting-mutation mechanisms in Prader-Willi syndrome. Am J Hum Genet *64*, 397-413.
144. Okamura, K., Hagiwara-Takeuchi, Y., Li, T., Vu, T.H., Hirai, M., Hattori, M., Sakaki, Y., Hoffman, A.R., and Ito, T. (2000). Comparative genome analysis of the mouse imprinted gene impact and its nonimprinted human homolog IMPACT: toward the structural basis for species-specific imprinting. Genome Res *10*, 1878-1889.
145. Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.
146. Ono, R., Shiura, H., Aburatani, H., Kohda, T., Kaneko-Ishino, T., and Ishino, F. (2003). Identification of a large novel imprinted gene cluster on mouse proximal chromosome 6. Genome Res *13*, 1696-1705.
147. Paulsen, M., Takada, S., Youngson, N.A., Benchaib, M., Charlier, C., Segers, K., Georges, M., and Ferguson-Smith, A.C. (2001). Comparative sequence analysis of the imprinted Dlk1-Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2-H19 region. Genome Res *11*, 2085-2094.
148. Peters, J., Wroe, S.F., Wells, C.A., Miller, H.J., Bodle, D., Beechey, C.V., Williamson, C.M., and Kelsey, G. (1999). A cluster of oppositely imprinted transcripts at the Gnas locus in the distal imprinting region of mouse chromosome 2. Proc Natl Acad Sci U S A *96*, 3830-3835.

149. Piacentini, L., Fanti, L., Berloco, M., Perrini, B., and Pimpinelli, S. (2003). Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. J Cell Biol *161*, 707-714.

150. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

151. Reijnen, M.J., Hamer, K.M., den Blaauwen, J.L., Lambrechts, C., Schoneveld, I., van Driel, R., and Otte, A.P. (1995). Polycomb and bmi-1 homologs are expressed in overlapping patterns in Xenopus embryos and are able to interact with each other. Mech Dev *53*, 35-46.

152. Roberts, S.G., Ha, I., Maldonado, E., Reinberg, D., and Green, M.R. (1993). Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. Nature *363*, 741-744.

153. Rosenfeld, J.A., Wang, Z., Schones, D.E., Zhao, K., DeSalle, R., and Zhang, M.Q. (2009). Determination of enriched histone modifications in non-genic portions of the human genome. BMC Genomics *10*, 143.

154. Runte, M., Huttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001). The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. Hum Mol Genet *10*, 2687-2700.

155. Ruvkun, G., and Hobert, O. (1998). The taxonomy of developmental control in Caenorhabditis elegans. Science *282*, 2033-2041.

156. Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. Development *132*, 797-803.

157. Santoro, C., Mermod, N., Andrews, P.C., and Tjian, R. (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature *334*, 218-224.

158. Satijn, D.P., Gunster, M.J., van der Vlag, J., Hamer, K.M., Schul, W., Alkema, M.J., Saurin, A.J., Freemont, P.S., van Driel, R., and Otte, A.P. (1997). RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. Mol Cell Biol *17*, 4105-4113.

159. Schulz, R., McCole, R.B., Woodfine, K., Wood, A.J., Chahal, M., Monk, D., Moore, G.E., and Oakey, R.J. (2009). Transcript- and tissue-specific imprinting of a tumour suppressor gene. Hum Mol Genet *18*, 118-127.

160. Schuster-Gossler, K., Bilinski, P., Sado, T., Ferguson-Smith, A., and Gossler, A. (1998). The mouse Gtl2 gene is differentially expressed during embryonic development, encodes multiple alternatively spliced transcripts, and may act as an RNA. Dev Dyn *212*, 214-228.

161. Schweizer, J., Zynger, D., and Francke, U. (1999). In vivo nuclease hypersensitivity studies reveal multiple sites of parental origin-dependent differential chromatin conformation in the 150 kb SNRPN transcription unit. Hum Mol Genet *8*, 555-566.

162. Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., and Cavaille, J. (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res *14*, 1741-1748.

163. Seitz, H., Youngson, N., Lin, S.P., Dalbert, S., Paulsen, M., Bachellerie, J.P., Ferguson-Smith, A.C., and Cavaille, J. (2003). Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. Nat Genet *34*, 261-262.

164. Sharp, A.J., Migliavacca, E., Dupre, Y., Stathaki, E., Sailani, M.R., Baumer, A., Schinzel, A., Mackay, D.J., Robinson, D.O., Cobellis, G.*, et al.* (2010). Methylation profiling in individuals with uniparental disomy identifies novel differentially methylated regions on chromosome 15. Genome Res *20*, 1271-1278.

165. Shemer, R., Birger, Y., Riggs, A.D., and Razin, A. (1997). Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. Proc Natl Acad Sci U S A *94*, 10267-10272.

166. Shemer, R., Hershko, A.Y., Perk, J., Mostoslavsky, R., Tsuberi, B., Cedar, H., Buiting, K., and Razin, A. (2000). The imprinting box of the Prader-Willi/Angelman syndrome domain. Nat Genet *26*, 440-443.

167. Shibata, H., Ueda, T., Kamiya, M., Yoshiki, A., Kusakabe, M., Plass, C., Held, W.A., Sunahara, S., Katsuki, M., Muramatsu, M.*, et al.* (1997). An oocyte-specific methylation imprint center in the mouse U2afbp-rs/U2af1-rs1 gene marks the establishment of allele-specific methylation during preimplantation development. Genomics *44*, 171-178.

168. Shibata, H., Yoshino, K., Sunahara, S., Gondo, Y., Katsuki, M., Ueda, T., Kamiya, M., Muramatsu, M., Murakami, Y., Kalcheva, I.*, et al.* (1996). Inactive allele-specific methylation and chromatin structure of the imprinted gene U2af1-rs1 on mouse chromosome 11. Genomics *35*, 248-252.

169. Shimizu, M., Li, W., Shindo, H., and Mitchell, A.P. (1997). Transcriptional repression at a distance through exclusion of activator binding in vivo. Proc Natl Acad Sci U S A *94*, 790-795.

170. Simon, J. (1995). Locking in stable states of gene expression: transcriptional control during Drosophila development. Curr Opin Cell Biol *7*, 376-385.

171. Singer-Sam, J., and Riggs, A.D. (1993). X chromosome inactivation and DNA methylation. EXS *64*, 358-384.

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

173. Smilinich, N.J., Day, C.D., Fitzpatrick, G.V., Caldwell, G.M., Lossie, A.C., Cooper, P.R., Smallwood, A.C., Joyce, J.A., Schofield, P.N., Reik, W.*, et al.* (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. Proc Natl Acad Sci U S A *96*, 8064-8069.

174. Smith, E.Y., Futtner, C.R., Chamberlain, S.J., Johnstone, K.A., and Resnick, J.L. (2011). Transcription is required to establish maternal imprinting at the Prader-Willi syndrome and Angelman syndrome locus. PLoS Genet *7*, e1002422.

175. Smith, R.J., Arnaud, P., Konfortova, G., Dean, W.L., Beechey, C.V., and Kelsey, G. (2002). The mouse Zac1 locus: basis for imprinting and comparison with human ZAC. Gene *292*, 101-112.

176. Smith, R.J., Dean, W., Konfortova, G., and Kelsey, G. (2003). Identification of novel imprinted genes in a genome-wide screen for maternal methylation. Genome Res *13*, 558-569.

177. Sowpati, D.T., Thiagarajan, D., Sharma, S., Sultana, H., John, R., Surani, A., Mishra, R.K., and Khosla, S. (2008). An intronic DNA sequence within the mouse Neuronatin gene exhibits biochemical characteristics of an ICR and acts as a transcriptional activator in Drosophila. Mech Dev *125*, 963-973.

178. Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.E., Chen, J., Lazar, M.A.*, et al.* (2008). DOT1L/KMT4 recruitment and

H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Mol Cell Biol *28*, 2825-2839.

179. Stoger, R., Kubicka, P., Liu, C.G., Kafri, T., Razin, A., Cedar, H., and Barlow, D.P. (1993). Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. Cell *73*, 61-71.
180. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.
181. Stringer, K.F., Ingles, C.J., and Greenblatt, J. (1990). Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature *345*, 783-786.
182. Suh, Y.H., Kim, W.H., Moon, C., Hong, Y.H., Eun, S.Y., Lim, J.H., Choi, J.S., Song, J., and Jung, M.H. (2005). Ectopic expression of Neuronatin potentiates adipogenesis through enhanced phosphorylation of cAMP-response element-binding protein in 3T3-L1 cells. Biochem Biophys Res Commun *337*, 481-489.
183. Surani, M.A., Barton, S.C., and Norris, M.L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature *308*, 548-550.
184. Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet *9*, 465-476.
185. Szabo, P.E., Tang, S.H., Silva, F.J., Tsark, W.M., and Mann, J.R. (2004). Role of CTCF binding sites in the Igf2/H19 imprinting control region. Mol Cell Biol *24*, 4791-4800.
186. Takada, S., Paulsen, M., Tevendale, M., Tsai, C.E., Kelsey, G., Cattanach, B.M., and Ferguson-Smith, A.C. (2002). Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. Hum Mol Genet *11*, 77-86.
187. Takada, S., Tevendale, M., Baker, J., Georgiades, P., Campbell, E., Freeman, T., Johnson, M.H., Paulsen, M., and Ferguson-Smith, A.C. (2000). Delta-like and gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12. Curr Biol *10*, 1135-1138.
188. Thiagarajan, D., Dev, R.R., and Khosla, S. (2011). The DNA methyltranferase Dnmt2 participates in RNA processing during cellular stress. Epigenetics *6*, 103-113.

189. Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes Dev *12*, 3693-3702.
190. Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A *76*, 4350-4354.
191. Tremblay, K.D., Duran, K.L., and Bartolomei, M.S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. Mol Cell Biol *17*, 4322-4329.
192. Tsai, S.Y., Carlstedt-Duke, J., Weigel, N.L., Dahlman, K., Gustafsson, J.A., Tsai, M.J., and O'Malley, B.W. (1988). Molecular interactions of steroid hormone receptor with its enhancer element: evidence for receptor dimer formation. Cell *55*, 361-369.
193. Tuan, D., Kong, S., and Hu, K. (1992). Transcription of the hypersensitive site HS2 enhancer in erythroid cells. Proc Natl Acad Sci U S A *89*, 11219-11223.
194. Tucker, K.L. (2001). Methylated cytosine and the brain: a new base for neuroscience. Neuron *30*, 649-652.
195. Turner, B.M. (2000). Histone acetylation and an epigenetic code. Bioessays *22*, 836-845.
196. Ubeda, F., and Wilkins, J.F. (2008). Imprinted genes and human disease: an evolutionary perspective. Adv Exp Med Biol *626*, 101-115.
197. Ueda, T., Abe, K., Miura, A., Yuzuriha, M., Zubair, M., Noguchi, M., Niwa, K., Kawase, Y., Kono, T., Matsuda, Y.*, et al.* (2000). The paternal methylation imprint of the mouse H19 locus is acquired in the gonocyte stage during foetal testis development. Genes Cells *5*, 649-659.
198. Ueda, T., Yamazaki, K., Suzuki, R., Fujimoto, H., Sasaki, H., Sakaki, Y., and Higashinakagawa, T. (1992). Parental methylation patterns of a transgenic locus in adult somatic tissues are imprinted during gametogenesis. Development *116*, 831-839.
199. Uribe, D.J., Guo, K., Shin, Y.J., and Sun, D. (2011). Heterogeneous nuclear ribonucleoprotein K and nucleolin as transcriptional activators of the vascular endothelial growth factor promoter through interaction with secondary DNA structures. Biochemistry *50*, 3796-3806.

200. Valenzuela, L., and Kamakaka, R.T. (2006). Chromatin insulators. Annu Rev Genet *40*, 107-138.
201. Visel, A., Rubin, E.M., and Pennacchio, L.A. (2009). Genomic views of distant-acting enhancers. Nature *461*, 199-205.
202. Vrang, N., Meyre, D., Froguel, P., Jelsing, J., Tang-Christensen, M., Vatin, V., Mikkelsen, J.D., Thirstrup, K., Larsen, L.K., Cullberg, K.B.*, et al.* (2010). The imprinted gene neuronatin is regulated by metabolic status and associated with obesity. Obesity (Silver Spring) *18*, 1289-1296.
203. Wainfan, E., and Poirier, L.A. (1992). Methyl groups in carcinogenesis: effects on DNA methylation and gene expression. Cancer Res *52*, 2071s-2077s.
204. Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q.*, et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet *40*, 897-903.
205. Webster, K.E., O'Bryan, M.K., Fletcher, S., Crewther, P.E., Aapola, U., Craig, J., Harrison, D.K., Aung, H., Phutikanit, N., Lyle, R.*, et al.* (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proc Natl Acad Sci U S A *102*, 4068-4073.
206. Wijnholds, J., Chowdhury, K., Wehr, R., and Gruss, P. (1995). Segment-specific expression of the neuronatin gene during early hindbrain development. Dev Biol *171*, 73-84.
207. Williamson, C.M., Ball, S.T., Nottingham, W.T., Skinner, J.A., Plagge, A., Turner, M.D., Powles, N., Hough, T., Papworth, D., Fraser, W.D.*, et al.* (2004). A cis-acting control region is required exclusively for the tissue-specific imprinting of Gnas. Nat Genet *36*, 894-899.
208. Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G.*, et al.* (2006). Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. Nat Genet *38*, 350-355.
209. Wood, A.J., and Oakey, R.J. (2006). Genomic imprinting in mammals: emerging themes and established theories. PLoS Genet *2*, e147.
210. Wood, A.J., Roberts, R.G., Monk, D., Moore, G.E., Schulz, R., and Oakey, R.J. (2007). A screen for retrotransposed imprinted genes reveals an association between X chromosome homology and maternal germ-line methylation. PLoS Genet *3*, e20.

211. Wrange, O., Eriksson, P., and Perlmann, T. (1989). The purified activated glucocorticoid receptor is a homodimer. J Biol Chem *264*, 5253-5259.
212. Wroe, S.F., Kelsey, G., Skinner, J.A., Bodle, D., Ball, S.T., Beechey, C.V., Peters, J., and Williamson, C.M. (2000). An imprinted transcript, antisense to Nesp, adds complexity to the cluster of imprinted genes at the mouse Gnas locus. Proc Natl Acad Sci U S A *97*, 3342-3346.
213. Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature *389*, 745-749.
214. Wyatt, G.R. (1951). Recognition and estimation of 5-methylcytosine in nucleic acids. Biochem J *48*, 581-584.
215. Xie, W., Barr, C.L., Kim, A., Yue, F., Lee, A.Y., Eubanks, J., Dempster, E.L., and Ren, B. (2012). Base-Resolution Analyses of Sequence and Parent-of-Origin Dependent DNA Methylation in the Mouse Genome. Cell *148*, 816-831.
216. Yang, T., Adamson, T.E., Resnick, J.L., Leff, S., Wevrick, R., Francke, U., Jenkins, N.A., Copeland, N.G., and Brannan, C.I. (1998). A mouse model for Prader-Willi syndrome imprinting-centre mutations. Nat Genet *19*, 25-31.
217. Yatsuki, H., Joh, K., Higashimoto, K., Soejima, H., Arai, Y., Wang, Y., Hatada, I., Obata, Y., Morisaki, H., Zhang, Z.*, et al.* (2002). Domain regulation of imprinting cluster in Kip2/Lit1 subdomain on mouse chromosome 7F4/F5: large-scale DNA methylation analysis reveals that DMR-Lit1 is a putative imprinting control region. Genome Res *12*, 1860-1870.
218. Yoder, J.A., Soman, N.S., Verdine, G.L., and Bestor, T.H. (1997). DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. J Mol Biol *270*, 385-395.
219. Yoon, B.J., Herman, H., Sikora, A., Smith, L.T., Plass, C., and Soloway, P.D. (2002). Regulation of DNA methylation of Rasgrf1. Nat Genet *30*, 92-96.
220. Yu, S., Yu, D., Lee, E., Eckhaus, M., Lee, R., Corria, Z., Accili, D., Westphal, H., and Weinstein, L.S. (1998). Variable and tissue-specific hormone resistance in heterotrimeric Gs protein alpha-subunit (Gsalpha) knockout mice is due to tissue-specific imprinting of the gsalpha gene. Proc Natl Acad Sci U S A *95*, 8715-8720.

221. Zhang, Y. (2003). Transcriptional regulation by histone ubiquitination and deubiquitination. Genes Dev *17*, 2733-2740.
222. Zhao, H., and Dean, A. (2004). An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. Nucleic Acids Res *32*, 4903-4919.
223. Zilliacus, J., Wright, A.P., Carlstedt-Duke, J., and Gustafsson, J.A. (1995). Structural determinants of DNA-binding specificity by steroid receptors. Mol Endocrinol *9*, 389-400.
224. Zwart, R., Sleutels, F., Wutz, A., Schinkel, A.H., and Barlow, D.P. (2001). Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. Genes Dev *15*, 2361-2366.