Western University

Scholarship@Western

Digitized Theses

Digitized Special Collections

2009

Characterization of the mouse KCNQ1OT1 noncoding RNA

Sarah Allen Lalone

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Lalone, Sarah Allen, "Characterization of the mouse KCNQ1OT1 noncoding RNA" (2009). *Digitized Theses*. 3219.

<https://ir.lib.uwo.ca/digitizedtheses/3219>

This Thesis is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

CHARACTERIZATION OF THE MOUSE *KCNQ1OT1* NONCODING RNA

(Spine title: Characterization of the mouse *Kcnq1ot1* noncoding RNA)
(Thesis format: Monograph)

by

Sarah Allen Lalone

Graduate Program in Physiology and Developmental Biology

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Sarah A. Lalone, 2009

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor

Dr. Chris Pin

Co-Supervisor

Dr. Mellissa Mann

Supervisory Committee

Dr. Andrew Watson

Dr. Thomas Drysdale

Examiners

Dr. Gerald Kidder

Dr. Dean Betts

Dr. David Rodenhiser

The thesis by

Sarah Allen Lalone

entitled:

Characterization of the mouse *Kcnqlot1* noncoding RNA

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

Date _____

Chair of the Thesis Examination Board

CHARACTERIZATION OF THE MOUSE *KCNQ1OT1* NONCODING RNA

Sarah Allen Lalone, Master of Science, 2009, Department of Physiology/Collaborative Program
of Developmental Biology, University of Western Ontario, London, Ontario, CANADA

ABSTRACT

For a small group of important developmental genes, there is preferential silencing of either the maternal or paternal allele. This process of preferential expression is called 'genomic imprinting'. Imprinted genes often cluster in imprinted domains. Interestingly, many imprinting domains are associated with a non-coding RNA (ncRNA) that may regulate imprinted gene expression across the entire domain. Disruptions in imprinting can have severe consequences for growth and development. To understand the complex regulation of genomic imprinting, studies are required to determine how early embryos set up a hierarchy of events that will establish imprinting across large chromosomal domains. The goal of this project is to characterize the mouse *Kcnq1ot1* ncRNA and its role in imprinted gene regulation. My findings indicate that the *Kcnq1ot1* ncRNA terminates at 463 kb from the transcriptional start site, and that this length is conserved in various tissues and developmental stages. This extends the boundary between the *Kcnq1ot1* and *H19* domains, downstream of *Th*, between 464 and 617 kb. shRNA and siRNA depletion of the *Kcnq1ot1* transcript at 43 kb and 460 kb, respectively, indicates that *Kcnq1ot1* is one continuous transcript as opposed to having multiple start sites along the length. In addition, the *Kcnq1ot1* transcript originates from the imprinting control region (ICR), as deletion of the paternal ICR, which contains the *Kcnq1ot1* promoter, results in loss of amplification along the entire length of the transcript. To determine whether the *Kcnq1ot1* ncRNA regulates domain-wide imprinting during early embryogenesis, RNA interference was employed to deplete *Kcnq1ot1* in embryonic and extraembryonic stem cells. Loss of >90% of the *Kcnq1ot1* transcript had no effect on imprinted expression in the domain, nor on imprinted methylation at the *Kcnq1ot1* ICR, suggesting that transcription itself may play a more important role than the transcript *per se*. Results from this study will lead to a better understanding of the role that long ncRNAs play in establishing and/or maintaining imprinted gene regulation across imprinting domains, as well as their role in human imprinted disorders.

Keywords: epigenetics, genomic imprinting, noncoding RNA, *Kcnq1ot1*

CO-AUTHORSHIP STATEMENT

All experimental work was carried out by myself except for:

Michael Golding: generation and maintenance of stem cell lines, transfection and transduction of NIH 3T3 and stem cells with siRNAs and shRNAs.

Liana Kaufmann: design of primers from 64k to 85k

Morgan McWilliam: design of primers from 85k to 300k and sequencing of pure B6 and CAST tissues for SNP analysis

Liyue Zhang and Malaika Miles: selected colony PCRs for bisulfite sequencing

ACKNOWLEDGEMENTS

First and foremost, I'd like to thank my supervisor, Dr. Mellissa Mann. It goes without saying that I could not have completed this degree without your guidance and teaching. Thank you for showing me how to think critically, how to answer questions logically, and present my findings in a clear and professional way. Not every supervisor can show a student so thoroughly how to do those things, and they are absolutely required if you are to be a scientist. Thank you for promoting a team mentality in the lab and providing us with social events to solidify our friendships and cohesiveness. Lastly, thank you for your time and commitment. I don't ever remember knocking on your door and being turned away, and I always had your help.

I'd also like to thank my co-supervisor, Dr. Chris Pin. You literally made this possible for me a couple of years ago when I was at a crossroads and truly ready to quit. Thank you for your support and for acting to turn me around immediately. Thank you also for your work and time spent on my project.

To my committee members: Drs. Andy Watson and Tom Drysdale. I very much appreciated your willingness to be part of my research, and your thought-provoking questions and suggestions that made this project that much stronger.

To my labmates, I can't thank you enough for your friendship, counsel and help during my 2 and a half years in the Mann Lab. Particularly Dr. Mike Golding, your help was absolutely instrumental in my project. Thank you for that. Liyue Zhang, thank you for your hard work and finesse with the bisulfite mutagenesis. You're a genius with protocols!

Eli and Sam... without the two of you, I wouldn't have wanted to do such a thing. It was difficult, and exhausting, and a real personal struggle to get through this project and I'm so thankful that I had the two of you at home for inspiration.

Lastly, thank you to my family. You know what "all work and no play" can do to a person! I couldn't ask for more in a support network. Thanks for making everything so clear for me!

TABLE OF CONTENTS

CERTIFICATE OF EXAMINATION	ii
ABSTRACT	iii
CO-AUTHORSHIP STATEMENT	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	x
LIST OF TABLES	xii
LIST OF ABBREVIATIONS	xiii
1 CHAPTER 1 – INTRODUCTION	1
1.1 Epigenetics	1
1.2 Genomic imprinting	3
1.2.1 Imprinting diseases in humans	9
1.3 Imprinting domains	12
1.4 Regulation of imprinting control regions	15
1.4.1 Long noncoding RNAs	16
1.5 <i>Kcnq1ot1</i> imprinting domain	19
1.5.1 Deletion and truncation studies	22
1.5.2 Temporal and spatial regulation of imprinting	24
1.5.3 Models of imprinting regulation of the <i>Kcnq1ot1</i> domain	25
1.6 <i>Kcnq1ot1</i>	27
1.6.1 <i>Kcnq1ot1</i> transcript length	27
1.7 Rationale	28
1.8 Hypotheses	29

1.9 Objectives	30
2 CHAPTER 2 – MATERIALS AND METHODS	31
2.1 Materials	31
2.1.1 Genetic crosses	31
2.1.2 Primers	31
2.1.3 siRNA	32
2.1.4 Embryo-derived stem cells	32
2.1.5 Transfection of NIH-3T3 cells with siRNA	32
2.1.6 Chemicals	33
2.2 Methods	33
2.2.1 Collection of embryonic tissues and postnatal samples	33
2.2.2 DNA and RNA isolation	34
2.2.3 Reverse transcription of RNA	34
2.2.4 Primer design	35
2.2.5 Polymerase chain reaction (PCR) amplification	38
2.2.6 Gel electrophoresis	38
2.2.7 3' Rapid amplification of cDNA ends (RACE)	39
2.2.8 Quantitative real-time PCR analysis	40
2.2.9 Statistical analysis	41
2.2.10 Sequencing analysis of PCR amplicons of <i>Kcnq1ot1</i>	41
2.2.11 Allelic discrimination using restriction endonucleases	42
2.2.12 Denistometry for allelic discrimination	42
2.2.13 Bisulfite conversion of DNA	43
3 CHAPTER 3 – RESULTS	46

3.1 Analysis of <i>Kcnq1ot1</i> transcript and expression	46
3.1.1 Determination of the transcript length of <i>Kcnq1ot1</i>	46
3.1.1.1 PCR amplification of the putative <i>Kcnq1ot1</i> ncRNA	46
3.1.2 Tissue-specific analysis of <i>Kcnq1ot1</i> length	47
3.1.3 Identification of the 3' end of <i>Kcnq1ot1</i>	50
3.2 Verification that amplified products correspond to the <i>Kcnq1ot1</i> ncRNA	55
3.2.1 Parental origin of the <i>Kcnq1ot1</i> ncRNA	55
3.2.2 RNA interference analysis	56
3.2.2.1 shRNA-mediated reduction of <i>Kcnq1ot1</i> in XEN cells	56
3.2.2.2 siRNA-mediated reduction of <i>Kcnq1ot1</i> in NIH 3T3 cells	59
3.2.3 IC2 deletion analysis	62
3.3 Imprint Status of the tyrosine hydroxylase (<i>Th</i>) gene	65
3.4 <i>Kcnq1ot1</i> -depletion effects on imprinted regulation in the <i>Kcnq1ot1</i> domain	66
3.4.1 Imprint status of genes in the <i>Kcnq1ot1</i> domain in stem cells	66
3.4.1.1 Imprinted methylation at the <i>Kcnq1ot1</i> ICR	66
3.4.1.2 Expression of imprinted genes in the <i>Kcnq1ot1</i> domain	69
3.4.2 Effects of <i>Kcnq1ot1</i> depletion in early embryonic-derived	

tissues	72
3.4.2.1 Methylation status of the <i>Kcnq1ot1</i> ICR in	
<i>Kcnq1ot1</i> -depleted cells	75
3.4.2.2 Effects of <i>Kcnq1ot1</i> -depletion on imprinted gene	
Expression	75
4 CHAPTER 4 – DISCUSSION	81
4.1 <i>Kcnq1ot1</i> length and <i>Kcnq1ot1</i> domain boundaries	81
4.2 <i>Kcnq1ot1</i> length and its function	84
4.3 Imprinted status of <i>Th</i>	85
4.4 Regulatory elements responsible for <i>Kcnq1ot1</i> domain regulation	87
4.5 Comparison of the <i>Kcnq1ot1</i> domain between mice and humans	88
4.6 Why is <i>Kcnq1ot1</i> so long?	89
4.7 Conclusion	91
REFERENCES	92
CURRICULUM VITAE	101

LIST OF FIGURES

Figure	Description	Page
1.1	Nuclear transfer studies show that both maternal and paternal compliments are necessary for normal growth and development	6
1.2	Chromosomal map of imprinted genes in the mouse showing clustered distribution	8
1.3	Schematic representation of the <i>H19</i> , <i>Kcnq1ot1</i> , and <i>Airn</i> imprinting domains	14
1.4	Organization of the <i>Kcnq1ot1</i> imprinted region	21
3.1	The putative <i>Kcnq1ot1</i> ncRNA extends 463 kb from the transcription start site	49
3.2	The length of the <i>Kcnq1ot1</i> transcript is consistent in a variety of tissues and developmental time-points	52
3.3	<i>Kcnq1ot1</i> 3'RACE Analysis	54
3.4	Paternal-specific expression of <i>Kcnq1ot1</i> at 2 kb from transcriptional start site	58
3.5	shRNA and siRNA depletion of <i>Kcnq1ot1</i> RNA leads to reduced amplification of the transcript up to 461k from the start site	61
3.6	Deletion of the paternal ICR resulting in loss of <i>Kcnq1ot1</i> transcription leads to significantly reduced amplification of the transcript up to 461 kb from the start site	64
3.7	<i>Th</i> exhibits imprinted expression in embryonic but not extraembryonic	

	tissues	67
3.8	WT as well as transduced stem cells maintain imprinted methylation	71
3.9	<i>Kcnq1ot1</i> depletion in stem cells transduced with a shRNA that targets	
	<i>Kcnq1ot1</i> at 43 kb	74
3.10	Imprinted methylation was preserved in <i>Kcnq1ot1</i> -depleted stem cells	77
3.11	Imprinted gene expression was preserved in <i>Kcnq1ot1</i> -depleted stem cells	79
4.1	Comparison of the <i>Kcnq1ot1</i> and <i>H19</i> domains in mice and humans	83

LIST OF TABLES

Table	Description	Page
1.1	Summary of known imprinting syndromes	10
1.2	Summary of ncRNAs associated with well-characterized imprinting domains	18
2.1	List of primers, reaction annealing temperature and amplicon size	36
2.2	Summary of SNP position and RE digest for imprinted genes	45

LIST OF ABBREVIATIONS

%	percent
°C	degrees Celcius
<i>Airn</i>	antisense <i>Igf2r</i> RNA gene
AS	Angelman Syndrome
<i>Ascl2/ASCL2</i>	mouse/human achaete-scute homolog 2 gene
B6	C57BL/6J
bp	base pair
BWS	Beckwith-Weidemann Syndrome
CaCl ₂	calcium chloride
<i>Cars</i>	cysteinyl-tRNA synthetase gene
CAST	<i>Mus musculus castaneus</i>
CAST7	C57BL/6J mice with two CAST chromosome 7s
<i>Cd81/CD81</i>	mouse/human CD81 antigen gene
<i>Cdkn1c/CDKN1C</i>	mouse/human cyclin-dependent kinase inhibitor 1C gene
cDNA	complimentary DNA
CH ₃	methyl group
COBRA	combined bisulfite and restriction analysis
CpG	CpG dinucleotide
C _T	cycle threshold
CTCF	CCCTC-binding factor (zinc finger protein)
d	day
Del	deletion
DEPC	diethyl pyrocarbonate
DMR	differentially methylated region

DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DNMT	DNA methyltransferase
<i>Dnmt1</i>	DNA methyltransferase 1 gene
DNMT1	DNA methyltransferase 1 protein
<i>Dnmt3a</i>	DNA methyltransferase 3a gene
<i>Dnmt3b</i>	DNA methyltransferase 3b gene
dsDNA	double-stranded DNA
DTT	dithiothreitol
Dup	duplication
E	enhancer
EDTA	ethylenediaminetetra-acetic acid disodium salt
EMB	embryo
ES cells	embryonic stem cells
F	forward
F ₁	first filial generation
FISH	fluorescence <i>in situ</i> hybridization
g	gram
gDNA	genomic DNA
<i>GNAS1</i>	guanine nucleotide binding protein, alpha stimulating gene
<i>Gtl2</i>	gene trap locus 2, also known as maternally expressed 3 gene
<i>H19/H19</i>	mouse/human hepatic cDNA clone No. 19 gene
H ₂ O	water
H3K14ac	histone 3 lysine 14 acetylation
H3K27me3	histone 3 lysine 27 trimethylation
H3K4me	histone 3 lysine 4 methylation
H3K4me3	histone 3 lysine 4 trimethylation

H3K9ac	histone 3 lysine 9 acetylation
H3K9me3	histone 3 lysine 9 trimethylation
ICR	imprinting control region
<i>Igf2/IGF2</i>	mouse/human insulin-like growth factor type 2 gene
<i>Igf2r/IGF2R</i>	mouse/human insulin-like growth factor type 2 receptor gene
<i>INS</i>	insulin
<i>Ins2</i>	insulin 2 gene
K43	short hairpin RNA targeting <i>Kcnq1ot1</i> at 43 kb from the start site
K82	ineffective short hairpin RNA targeting <i>Kcnq1ot1</i> at 82 kb from the start site
kb	kilobase
KCl	potassium chloride
<i>Kcnq1/ KCNQ1</i>	mouse/human potassium voltage-gated channel, KQT-like subfamily, member 1 gene
<i>Kcnq1ot1/KCNQ1OT1</i>	mouse/human potassium voltage-gated channel, KQT-like subfamily, member 1, overlapping transcript 1 gene
KH ₂ PO ₄	potassium phosphate
KvDMR	Kcnq1 domain differentially methylated region
L	litre
LB	Luria-Bertani medium
LiCl	lithium chloride
LINE	long interspersed nuclear element
Luc	short hairpin RNA targeting the luciferase gene
m	methylated
M	molar
Mat	maternal
Mb	megabase

mg	milligram
miRNA	micro RNA
ml	millilitre
mm	millimetre
mM	millimolar
mm ²	millimetre squared
mRNA	messenger RNA
<i>Mrpl1</i>	mitochondrial ribosomal protein L1 gene
Na	sodium
Na ₂ HPO ₄	sodium phosphate dibasic
NaCl	sodium chloride
NaOH	sodium hydroxide
<i>Nap1l4</i>	nucleosome assembly protein 1-like 4 gene
ncRNA	non-coding RNA
NDN	necdin
neg	no cDNA control
<i>Nespas</i>	neuroendocrine secretory protein antisense gene
ng	nanogram
NIH 3T3	3-day transfer, inoculum 3 x 10 ⁵ fibroblast cells
nm	nanometer
<i>Osbpl5</i>	oxysterol binding protein-like 5 gene
P	probability
Pat	paternal
PCR	polymerase chain reaction
pH	potential of hydrogen
<i>Phdla2</i>	pleckstrin homology-like domain, family A, member 2 gene
<i>Phmx</i>	tetraspanin 32 gene

piRNA	piwi interacting RNA
PLA	placenta
PWS	Prader-Willi Syndrome
qPCR	quantitative PCR
R	reverse
RACE	rapid amplification of cDNA ends
RE	restriction endonuclease
RITS	RNA-induced transcriptional silencing
RNA	ribonucleic acid
RNAi	RNA interference
RNase A	ribonuclease A
RPA	RNase protection assay
RT	reverse transcription
RT-PCR	reverse transcriptase PCR
SDS	sodium dodecyl sulphate
SEM	standard error of the mean
si460	small interfering RNA designed to target <i>Kcnqlot1</i> at 460 kb from the start site
SINE	short interspersed nuclear elements
siRNA	short interfering RNA
<i>Slc22a18</i>	solute carrier family 22 (organic cation transporter), member 18 gene
<i>Slc22a2</i>	solute carrier family 22 (organic cation transporter), member 2 gene
<i>Slc22a3</i>	solute carrier family 22 (organic cation transporter), member 3 gene
snoRNAs	small nucleolar RNAs
SNP	single nucleotide polymorphism

<i>Snrpn/SNRPN</i>	mouse/human small nuclear ribonucleoprotein-associated polypeptide N gene
SOC	super optimal broth with catabolite repression
SSIII	superscript III
TAE	Tris-acetate-EDTA
TBE	Tris-boric acid-EDTA
TE	Tris-EDTA
<i>Th/TH</i>	mouse/human tyrosine hydroxylase gene
T _m	melting temperature
<i>Tnfrsf22</i>	tumor necrosis factor receptor superfamily, member 22 gene
<i>Tnfrsf23</i>	tumor necrosis factor receptor superfamily, member 23 gene
<i>Tnfrsf26</i>	tumor necrosis factor receptor superfamily, member 26 gene
Tris	Tris(hydroxymethyl)aminomethane
<i>Trpm5</i>	transient receptor potential cation channel, subfamily M, member 5 gene
TS cells	trophectoderm stem cells
<i>Tssc4</i>	tumor-suppressing subchromosomal transferable fragment 4 gene
U	unmethylated
<i>UBE3A</i>	ubiquitin protein ligase E3A gene
<i>Ube3aas</i>	ubiquitin protein ligase E3A antisense gene
UPD	uniparental disomy
UV	ultraviolet light
V	volts
WT	wildtype
X	times
XEN cells	extraembryonic endoderm stem cells
<i>Xist</i>	X-inactive specific transcript gene

YS	yolk sac
Δ	change
Δ IC2	imprinting centre 2 deletion
μ g	microgram
μ M	micromolar

CHAPTER 1

INTRODUCTION

1.1 Epigenetics

The term “epigenetics” was first used in 1942 by C.H. Waddington, a biologist, embryologist and geneticist (Goldberg et al., 2007). He used the term to describe the “epigenetic landscape” as a metaphor for how gene *regulation* modulates development, or how genes might be interacting with their surroundings to produce a phenotype. More recently, this term has been refined to refer to the heritable, reversible alterations in gene regulation that occur with *no* change in the DNA sequence (Bestor et al., 1994; Holliday, 1994; Hendrich and Willard, 1995).

The most studied form of epigenetic modification is DNA methylation, which involves the addition of a methyl group (CH_3) to the cytosine of a CpG dinucleotide, converting it to 5-methylcytosine, and is generally associated with silencing of expression. Clusters of CpG dinucleotides are found throughout the genome and are referred to as CpG islands. CpG islands tend to be found around the 5' end of many regulated genes, particularly housekeeping genes (Saxanov et al., 2006). Approximately 80% of CpG islands are syntenic in human and mouse genomes (Antequera and Bird, 1993). Epigenetic changes leading to chromatin alterations have a profound impact on gene regulation.

DNA methylation is established and preserved by DNA methyltransferases (DNMTs). These molecules catalyze the transfer of a methyl group to the cytosines of CpG dinucleotides (Smith et al., 1992). There are two types of DNMTs, those that are involved in adding methyl groups *de novo* (*de novo* methyltransferases), and those that

maintain methylation patterns (maintenance methyltransferases) by adding methyl groups to hemi-methylated CpG dinucleotides (Smith et al., 1992). DNMT1, a maintenance DNMT, acts during mitotic divisions to maintain a 'cellular memory' (Kho et al., 1998). It has a high affinity for hemi-methylated DNA strands, and is ubiquitously expressed in proliferating cells. DNMT1 is critical for embryonic development, as a targeted null mutation of *Dnmt1* in mice results in genome-wide loss of methylation, and embryos do not develop past embryonic day 10.5 of embryogenesis (d10.5) (Li et al., 1992). *De novo* methylation is accomplished through other DNMTs such as *Dnmt3a* and *Dnmt3b*, which act in setting up DNA methylation patterns in the developing embryo (Okano et al., 1999, Kaneda et al., 2004; Oda et al., 2006). Loss of *Dnmt3a* leads to postnatal lethality at 4 to 8 weeks of age, as well as male sterility, and causes failure to establish methylation imprints in male and female germ lines. Loss of *Dnmt3b* results in embryonic lethality at d14.5, and is accompanied by vascular and liver defects (Okano et al., 1999). These studies emphasize that the appropriate establishment of methylation marks, as well as their subsequent maintenance, are important for mammalian development.

A second form of epigenetic regulation can occur through modifications to histone tails invoking chromatin changes that can result in active or repressive chromatin. Histones are proteins around which DNA is coiled for the purpose of DNA packaging, and also play a very important role in gene regulation (Strahl and Allis, 2000). Active chromatin (uncondensed) is characterized by histone modifications including histone 3 lysine 9 acetylation (H3K9ac), histone 2 lysine 14 acetylation, (H3K14ac), and histone 3 lysine 4 mono- or trimethylation (H3K4me, H3K4me3), which promotes more loosely packed DNA such that transcription factors have access to promoting or enhancing

sequences, thereby promoting gene expression. Repressed or “silenced” chromatin is generally characterized by H3K9me3, H3K27me3 and H4K20me3, and is tightly packed DNA such that transcription is effectively silenced in those regions (Strahl and Allis, 2000).

Epigenetic mechanisms function in biological processes such as X-inactivation, genomic imprinting, and transposon silencing, and are essential for proper mammalian development.

1.2 Genomic imprinting

In mammals, autosomal genes are diploid, i.e. there are two copies of each allele. Generally, both the maternally-inherited and paternally-inherited alleles are expressed. In a subset of genes, however, epigenetic marks restrict expression to either the maternal or paternal allele, while the other allele remains silent (McGrath and Solter 1984, Verona et al. 2003). These genes are said to be ‘imprinted’ as they were epigenetically marked or imprinted during sperm or egg development.

During gametogenesis, erasure of previous epigenetic marks restores totipotency with subsequent establishment of sex-specific epigenetic marks that will be transmitted to the offspring (Morgan et al. 2005). Following fertilization, the paternal and maternal genomes undergo global demethylation separately; the maternal genome shows passive demethylation, where methylation is lost in dividing cells over several cell cycles, while the paternal genome experiences active demethylation where methylation marks are rapidly removed from the DNA. The paternal genome undergoes further programming as protamines used for DNA packaging are exchanged for histones. After implantation,

methylation marks are restored *de novo* in the embryo. However, the imprinting marks established during gametogenesis must be maintained throughout development, and therefore, escape the global demethylation that occurs in preimplantation embryos.

Genomic imprinting was first described in mammals using nuclear transfer studies (Figure 1.1). Results showed that androgenetic embryos (reconstructed from 2 paternal pronuclei) or gynogenetic embryos (2 maternal pronuclei) failed to survive; only those made with one maternal and one paternal pronucleus were viable (McGrath and Solter, 1984a; Surani et al., 1984). It is interesting to note that not only were the resulting embryos not viable, but gynogenetic embryos were defective in extraembryonic tissue development, while androgenetic embryos were deficient in embryonic tissue. This led to the theory that maternally expressed imprinted genes were important for the development of placental tissues, whereas paternally expressed genes were important for the development of the embryo (Barton et al., 1984), although it is now accepted that the relationship is more complicated. What is clear is that both the maternal and paternal complements are necessary for mammalian growth and development, and that alterations in this genomic complement can have profound consequences.

To date, approximately 100 genes have been identified as being imprinted in the mouse (<http://www.har.mrc.ac.uk/mousebook/>) with a high level of conservation between mice and humans (Figure 1.2). The first imprinted genes to be described in the mouse were *Igf2r* (Insulin-like growth factor type 2 receptor), *Igf2* (Insulin-like growth factor type 2), and *H19* (Hepatic cDNA clone No. 19) (Barlow et al., 1991; DeChaira et al., 1990; Ferguson-Smith et al., 1991; Bartolomei et al., 1991). Loss-of-function studies

Figure 1.1. Nuclear transfer studies show that both maternal and paternal compliments are necessary for normal growth and development. Nuclear transfer produces reconstituted embryos with either one maternal complement (white circle) and one paternal complement (black circle), with two maternal complements (gynogenetic), or two paternal complements (androgenetic). Only embryos with one maternal and one paternal complement are viable.

Figure 11. Percentage of nitrogenous compounds in the total nitrogen content of various plant parts. The data are presented in a series of horizontal bar charts. The top row shows 'Total N', 'Protein N', and 'Non-protein N'. The second row shows 'Roots', 'Stems', and 'Leaves'. The third row shows 'Seeds' and 'Fruit'. The fourth row shows 'Bark' and 'Wood'. The fifth row shows 'Bark' and 'Wood'. The sixth row shows 'Bark' and 'Wood'. The seventh row shows 'Bark' and 'Wood'. The eighth row shows 'Bark' and 'Wood'. The ninth row shows 'Bark' and 'Wood'. The tenth row shows 'Bark' and 'Wood'. The eleventh row shows 'Bark' and 'Wood'. The twelfth row shows 'Bark' and 'Wood'. The thirteenth row shows 'Bark' and 'Wood'. The fourteenth row shows 'Bark' and 'Wood'. The fifteenth row shows 'Bark' and 'Wood'. The sixteenth row shows 'Bark' and 'Wood'. The seventeenth row shows 'Bark' and 'Wood'. The eighteenth row shows 'Bark' and 'Wood'. The nineteenth row shows 'Bark' and 'Wood'. The twentieth row shows 'Bark' and 'Wood'. The twenty-first row shows 'Bark' and 'Wood'. The twenty-second row shows 'Bark' and 'Wood'. The twenty-third row shows 'Bark' and 'Wood'. The twenty-fourth row shows 'Bark' and 'Wood'. The twenty-fifth row shows 'Bark' and 'Wood'. The twenty-sixth row shows 'Bark' and 'Wood'. The twenty-seventh row shows 'Bark' and 'Wood'. The twenty-eighth row shows 'Bark' and 'Wood'. The twenty-ninth row shows 'Bark' and 'Wood'. The thirtieth row shows 'Bark' and 'Wood'. The thirty-first row shows 'Bark' and 'Wood'. The thirty-second row shows 'Bark' and 'Wood'. The thirty-third row shows 'Bark' and 'Wood'. The thirty-fourth row shows 'Bark' and 'Wood'. The thirty-fifth row shows 'Bark' and 'Wood'. The thirty-sixth row shows 'Bark' and 'Wood'. The thirty-seventh row shows 'Bark' and 'Wood'. The thirty-eighth row shows 'Bark' and 'Wood'. The thirty-ninth row shows 'Bark' and 'Wood'. The fortieth row shows 'Bark' and 'Wood'. The forty-first row shows 'Bark' and 'Wood'. The forty-second row shows 'Bark' and 'Wood'. The forty-third row shows 'Bark' and 'Wood'. The forty-fourth row shows 'Bark' and 'Wood'. The forty-fifth row shows 'Bark' and 'Wood'. The forty-sixth row shows 'Bark' and 'Wood'. The forty-seventh row shows 'Bark' and 'Wood'. The forty-eighth row shows 'Bark' and 'Wood'. The forty-ninth row shows 'Bark' and 'Wood'. The fiftieth row shows 'Bark' and 'Wood'. The fifty-first row shows 'Bark' and 'Wood'. The fifty-second row shows 'Bark' and 'Wood'. The fifty-third row shows 'Bark' and 'Wood'. The fifty-fourth row shows 'Bark' and 'Wood'. The fifty-fifth row shows 'Bark' and 'Wood'. The fifty-sixth row shows 'Bark' and 'Wood'. The fifty-seventh row shows 'Bark' and 'Wood'. The fifty-eighth row shows 'Bark' and 'Wood'. The fifty-ninth row shows 'Bark' and 'Wood'. The sixtieth row shows 'Bark' and 'Wood'. The sixty-first row shows 'Bark' and 'Wood'. The sixty-second row shows 'Bark' and 'Wood'. The sixty-third row shows 'Bark' and 'Wood'. The sixty-fourth row shows 'Bark' and 'Wood'. The sixty-fifth row shows 'Bark' and 'Wood'. The sixty-sixth row shows 'Bark' and 'Wood'. The sixty-seventh row shows 'Bark' and 'Wood'. The sixty-eighth row shows 'Bark' and 'Wood'. The sixty-ninth row shows 'Bark' and 'Wood'. The seventieth row shows 'Bark' and 'Wood'. The seventy-first row shows 'Bark' and 'Wood'. The seventy-second row shows 'Bark' and 'Wood'. The seventy-third row shows 'Bark' and 'Wood'. The seventy-fourth row shows 'Bark' and 'Wood'. The seventy-fifth row shows 'Bark' and 'Wood'. The seventy-sixth row shows 'Bark' and 'Wood'. The seventy-seventh row shows 'Bark' and 'Wood'. The seventy-eighth row shows 'Bark' and 'Wood'. The seventy-ninth row shows 'Bark' and 'Wood'. The eightieth row shows 'Bark' and 'Wood'. The eighty-first row shows 'Bark' and 'Wood'. The eighty-second row shows 'Bark' and 'Wood'. The eighty-third row shows 'Bark' and 'Wood'. The eighty-fourth row shows 'Bark' and 'Wood'. The eighty-fifth row shows 'Bark' and 'Wood'. The eighty-sixth row shows 'Bark' and 'Wood'. The eighty-seventh row shows 'Bark' and 'Wood'. The eighty-eighth row shows 'Bark' and 'Wood'. The eighty-ninth row shows 'Bark' and 'Wood'. The ninetieth row shows 'Bark' and 'Wood'. The ninety-first row shows 'Bark' and 'Wood'. The ninety-second row shows 'Bark' and 'Wood'. The ninety-third row shows 'Bark' and 'Wood'. The ninety-fourth row shows 'Bark' and 'Wood'. The ninety-fifth row shows 'Bark' and 'Wood'. The ninety-sixth row shows 'Bark' and 'Wood'. The ninety-seventh row shows 'Bark' and 'Wood'. The ninety-eighth row shows 'Bark' and 'Wood'. The ninety-ninth row shows 'Bark' and 'Wood'. The hundredth row shows 'Bark' and 'Wood'.

Figure 1.2. Chromosomal map of imprinted genes in the mouse showing clustered distribution. From http://www.har.mrc.ac.uk/research/genomic_imprinting/maps.htm.

Chromosome:

Christine Williamson, Colin Beechey,
Bruce Cattanaach, Jo Patena
2009

Imprinted genes within clusters are not necessarily in correct order
imprinted genes in red are maternally expressed
imprinted genes in blue are paternally expressed
* maternally expressed small nucleolar RNAs and microRNA genes
* paternally expressed small nucleolar RNAs and microRNA genes
regions with abnormal imprinting phenotypes with maternal (Mat)
or paternal (Pat) duplication

of imprinted genes reveal that they are very important in mammalian development (Reik and Walter, 2001). For example, deletion of the expressed copy of *Igf2r* or *Igf2* results in growth defects in the embryo and placenta (DeChiara et al., 1990; Davies et al., 2002; Kono et al., 2004; Wang et al., 1994). The importance of imprinted gene regulation during development is further highlighted by the fact that imprinted expression of a gene can vary among tissues, developmental stages, and among species (Guillemot et al., 1995; Sasaki et al., 1995).

1.2.1 Imprinting diseases in humans

Disruption of imprinting can have profound consequences on growth and development of the mammalian embryo. For example, in humans, defects such as uniparental disomy (two copies of a chromosome, region of a chromosome or gene from only one parent), gene or chromosomal deletions, translocations, or rearrangements, as well as imprinting defects, result in imprinting disorders such as Prader-Willi, Angelman, and Beckwith-Wiedemann Syndromes (Table 1) (Bartolomei and Tilghman 1997, Maher and Reik 2000, Nicholls and Knepper 2001, Weksberg et al. 2003).

An interesting example of imprinting disorders occurs for the long arm of human chromosome 15q11-13 (7C in the mouse). Disruptions of this cluster of imprinted genes leads to two independent imprinting disorders, Prader-Willi Syndrome and Angelman Syndrome. Prader-Willi Syndrome (PWS) is a rare disorder characterized by hypotonia, short stature, polyphagia, small hands and feet, and mild mental retardation (Kileen, 2004). It is caused by loss of the normally paternally-expressed imprinted genes *SNRPN*

Table 1.1. Summary of known imprinting syndromes.

Disorder	Location	Gene(s)	Mutation
Prader-Willi Syndrome	15q11-q13	snoRNAs and others	mutations, deletion, translocation, UPD, imprinting defect
Angelman Syndrome	15q11-q13	<i>UBE3A</i>	mutations, deletion, translocation, UPD, imprinting defect
Beckwith-Wiedemann Syndrome	11p15.5	<i>IGF2</i> , <i>CDKN1C</i>	UPD, 11p15.5 duplication, translocation, point mutation, imprinting defects
Silver-Russell Syndrome	7p11.2	<i>IGF1R</i> , <i>GRB10</i> , and others	UPD, duplication, translocation, inversion
Silver-Russell Syndrome	11p15.5	biallelic expression of <i>H19</i> and reduced expression of <i>IGF2</i>	imprinting defect
Pseudohypoparathyroidism	20q13.2	<i>GNAS1</i>	point mutation, UPD, imprinting defect

UPD: uniparental disomy

(Small nuclear ribonucleoprotein polypeptide N) and *NDN* (Necdin), along with several small nucleolar RNAs (snoRNAs) which can result from sporadic mutations, uniparental chromosome translocations, deletions, or imprinting disruptions. Angelman Syndrome (AS) is a neurogenetic disorder characterized by defects such as developmental delay, seizures, jerky movements, and frequent laughing or smiling, and is due most commonly to a maternal deletion in the 15q11-13 region but can also result from maternal uniparental disomy, translocations, single gene mutations in *UBE3A*, and imprinting defects (Dooley, 1981). Another imprinting disorder, Beckwith-Wiedemann Syndrome (BWS), results from disruption of the 11p15.5 region (7F in the mouse) and is characterized by developmental defects including pre- and post-natal overgrowth, macroglossia, abdominal wall defects as well as a predisposition to embryonal Wilms' tumors (Maher and Reik, 2000). It can arise from paternal uniparental disomy of chromosome 11, paternal duplication of 11p15.5, maternally inherited coding mutations in cyclin-dependent kinase inhibitor 1C (*CDKN1C*), maternal chromosome rearrangements, or loss of imprinting at *KCNQ1OT1* and *H19/IGF2* (Weksberg et al., 2003). Overall, 40% of sporadic BWS patients have a hypomethylated CpG island at *KCNQ1OT1* that controls expression of the imprinted genes in 11p15.5 resulting in activation of genes from the maternal allele that are normally expressed only from the paternal allele. Patients with hypomethylation show markedly decreased expression of *CDKN1C* (Lee et al., 1999; Smilnich et al., 1999; Diaz-Meyer et al., 2003). *CDKN1C* is an important gene in both BWS and some cancers, and it is considered a putative tumor suppressor gene (Lee et al., 1995; Watanabe et al., 1998).

Thus, lack of the proper parental complement in imprinted regions has profound consequences on growth and development in humans, and a better understanding of genomic imprinting regulation is required to reduce the incidence of such diseases.

1.3 Imprinting domains

Imprinted genes are found on multiple chromosomes, and tend to be clustered in groups of 3 to 15 genes (Verona et al. 2003). These clusters of imprinted genes are referred to as imprinting domains. Several imprinting domains have been identified to date, and 6 of them have been fairly well-characterized (Figure 1.3 for selected domains). Imprinting domains are regulated by a region known as the imprinting control region (ICR). ICRs are *cis*-acting regulatory regions for the entire domain (Spahn et al., 2003; Regha et al., 2006). ICRs are responsible for gene regulation over the entire cluster, and do so over hundreds of kilobases (kb) in a bidirectional manner (Leighton et al., 1995; Wutz et al., 1997; Thorvaldson et al., 1998; Bielinska et al., 2000; Fitzpatrick et al., 2002; Lin et al., 2003). ICRs carry differential parental marks acquired during gamete development that differ between the sperm and oocyte. The parental marks have yet to be defined, but ICRs contain a region of differential methylation, usually ~2 to 5 kb in length, referred to as a differentially methylated region (DMR) (Brandeis et al., 1993; Verona et al., 2003; Reinhart and Chaillet, 2005). Generally, differential methylation is not found at individual genes within the cluster, but rather at imprinting control regions. Deletion of the ICRs associated with the *Snurf-Snrpn*, *Kcnqlot1*, and *Igf2r* genes leads to loss of imprinting for genes within the respective cluster (Verona et al., 2003). However, the exact mechanisms of how ICRs function are unknown.

Figure 1.3. Schematic representation of the *H19*, *Kcnq1ot1*, and *Airn* imprinting domains. Both the *Kcnq1ot1* and *Airn* domains are regulated by silencing through the ncRNA, whereas the *H19* domain is regulated by enhancer blocking. Maternal alleles are top strand (mat) and paternal alleles are the bottom strand (pat). Genes are indicated as either biallelically expressed (green), maternally expressed (red), maternally expressed in placenta only (pink), paternally expressed (blue), or silent (grey), with arrows above indicating the direction of transcription. Non-coding RNAs are denoted by wavy line. The ICR is indicated by a yellow box. Filled circles indicate a methylated CpG island, and empty circles indicate an unmethylated CpG island. Green pentagons indicate enhancer elements (e). Orange oval denotes CTCF protein.

1.4 Regulation of imprinting control regions

There are two main models for the regulation of imprinting domains: enhancer-blocking or “insulator” models, and those that rely on the function of an antisense-transcribed noncoding RNA (ncRNA).

Enhancer-blocker domains function by allowing or restricting access of an enhancer element to a target gene by an insulator element that is placed between them. A well-characterized example of an enhancer-blocking imprinted domain is that of *H19*. The *H19* domain spans ~80 kb and contains three paternally expressed genes, *Igf2*, *Igf2as*, and *Ins2*, a 2.5 kb-noncoding transcript, *H19*, enhancer elements located downstream of the *H19* gene, and an ICR located 2 kb upstream of *H19* (Figure 1.3). Imprinted gene expression operates through the ICR. On the paternal allele, the ICR is methylated and *H19* is silent. The enhancer has access to the imprinted genes upstream, promoting paternal expression of *Igf2* as well as *Ins2*. On the maternal allele, hypomethylation of the ICR allows for binding of CTCF-binding factor proteins to the ICR, which forms an insulator that blocks the interaction of the enhancer element and the downstream genes. The enhancer instead acts on and allows for expression of *H19* on the maternal allele. Although a ncRNA is expressed in this domain, it seems to have no inherent silencing ability on its own (Schmidt et al., 1999), and imprinted expression is regulated through the ICR acting as an insulator for enhancer elements.

The second model involves production of a ncRNA transcribed in an antisense orientation, which acts to regulate gene expression for the entire cluster. In the case of the *Igf2r* imprinted domain, a 108 kb ncRNA, *Airn*, (antisense *Igf2r* RNA), is thought to regulate imprinting of three protein-coding genes within the domain (Sleutels

et al., 2002). On the paternal chromosome, the *Airn* ncRNA is expressed and three imprinted mRNA promoters are repressed (*Slc22a3*, *Slc22a2*, and *Igf2r*). On the maternal ICR, methylation represses *Airn* expression, thereby allowing expression of the 3 protein-coding genes. *Airn* is therefore an example of a long ncRNA that has a *cis*-restricted gene regulatory role in mammals, although how it is restricted to acting on only one allele remains unknown.

1.4.1 Long noncoding RNAs

Noncoding transcription accounts for four times more transcription of the genome than that of protein-coding transcription (Kapronov et al., 2007). Previously, the abundance of noncoding transcripts was thought to be transcriptional “noise” due to the low fidelity of RNA polymerases (Struhl, 2007). The fact that some noncoding RNAs are temporally and spatially restricted indicates that there may be functional roles for these transcripts (Amaral and Mattick, 2008). In general, ncRNAs can be grouped based on their size and function. Small regulatory ncRNAs include microRNAs (miRNA), which are between 21 and 30 base pairs (bp) in length, and regulate gene expression; short interfering RNAs (siRNA), which also effectively silence a specific gene via degradation of its mRNA through the RNA interference (RNAi) machinery; small nucleolar RNAs (snoRNA) which guide modifications of ribosomal RNAs such as methylation; and piwi-interacting RNAs (piRNA) that help to form interactions between proteins and RNA. By comparison, long ncRNAs can range in length from 2 kb to more than 1000 kb, and rather than encode proteins, act functionally as RNAs to silence genes in genomic imprinting. Noncoding RNAs are found in each of the six well-characterized

imprinting domains, and show a wide range in length (Table 1.2). Long ncRNAs associated with imprinted domains are often expressed in an antisense orientation to protein coding genes in the region (O'Neill, 2005), and the parental chromosome that carries the unmethylated ICR is the one that transcribes the associated ncRNA (Spahn and Barlow, 2003).

The location of the ncRNA transcriptional start site varies with respect to the ICR. In the case of the *H19* cluster in the mouse, which spans ~80 kb, the 2.5 kb-*H19* ncRNA transcriptional start site is located downstream of the ICR (Leighton et al., 1995; Thorvaldson et al., 1998). In two other well-known imprinting domains, potassium voltage-gated channel, KQT-like subfamily, member 1, overlapping transcript 1 (*Kcnq1ot1*) and antisense *Igf2r* RNA, the transcriptional start site of the ncRNA is located within the ICR (Stoger et al., 1993; Wutz et al., 1997; Smilnich et al. 1999).

Post-transcriptional processing of these ncRNAs differs substantially. *H19* comprises single-copy sequences and is fully spliced, whereas *Airn* is said to be unspliced and possesses a high degree of repetitive sequence (Seidl et al., 2006). *Kcnq1ot1* has yet to be thoroughly tested, but it is believed that it is unspliced and is a repetitive-rich sequence. Sub-cellular localization also varies between the ncRNAs, with *H19* being efficiently exported into the cytoplasm, and *Airn* and *Kcnq1ot1* remaining within the nucleus (Seidl et al., 2006). The degree of silencing activity is also inconsistent. It is believed that the *H19* transcript lacks any silencing function, while *Airn* and *Kcnq1ot1* possess silencing activity in their respective imprinting domains as truncation of the transcript leads to de-repression of imprinted genes within the region (Schmidt et al., 1999; Sleutels et al., 2002; Mancini-Dinardo et al., 2006).

Table 1.2. Summary of ncRNAs associated with well-characterized imprinting domains

Imprinting Domain	ncRNA	length (kb)	overlapping gene(s)	reference
<i>Airn</i>	<i>Airn</i>	108	<i>Igf2r</i>	Lyle et al., 2000
<i>Kcnqlot1</i>	<i>Kcnqlot1</i>	60+	<i>Kcnq1</i>	Mitsuya et al., 1999
<i>Snrpn</i>	<i>Ube3aas</i>	1000+	<i>Snrpn, Ube3a</i>	Landers et al., 2004
<i>Gnas</i>	<i>Nespas</i>	unknown	<i>Nesp, Gnas</i>	Wroe et al., 2000
<i>H19</i>	<i>H19</i>	2.3	none	Brannan et al., 1990
<i>Gtl2</i>	<i>Gtl2</i>	unknown	<i>none</i>	Teiriling et al., 2005
<i>Xist</i>	<i>Xist</i>	15kb	<i>Tsix</i>	Brockdorff et al., 1992

Generally, imprinting domains possessing paternally methylated ICR, that show expression of ncRNAs from the maternal allele, tend to have short ncRNAs that are spliced (*H19*, *Gtl2*) (Schmidt et al., 1999; Sleutels et al., 2002; Mancini-Dinardo et al., 2006). These domains rely on insulator function to regulate imprinting. In contrast to this, domains that have maternally methylated ICRs with expression of the ncRNA from the paternal allele tend to have long unspliced ncRNAs (*Airn*, *Kcnq1ot1*, *Ube3aas*, *Nespas*, *Xist*) that may silence the neighbouring imprinted genes in the domain (Lin et al., 1995; Fitzpatrick et al., 2002).

In general, ncRNAs are a common feature of imprinted domains, can act to regulate imprinted genes within the domain, and are responsible for chromatin inactivation in *cis*, meaning that they act locally within the domain (Lewis and Reik, 2006).

1.5 *Kcnq1ot1* imprinting domain

The *Kcnq1ot1/KCNQ1OT1* imprinting domain is located within 800 kb of the telomere on chromosome 7 in the mouse, and at 11p15.5 in the human. The *Kcnq1ot1* domain spans 1 megabase (Mb) and contains one paternally expressed ncRNA, *Kcnq1ot1*, and eight maternally expressed, protein coding genes, including *Slc22a18* (solute carrier family 22 (organic cation transporter) member 18), *Cdkn1c* (cyclin-dependent kinase inhibitor 1C), *Kcnq1* (potassium voltage-gated channel, KQT-like subfamily, member 1), and *Ascl2* (Achaete-scute homolog 2) (Onyango et al. 2000) (Figure 1.4). Based on expression studies, the boundaries of the *Kcnq1ot1* domain are thought to be located just

Figure 1A. Organization of the Karyotype imprinted region. Arrows show the direction of transcription (-) and downstream from the Karyotype non-coding RNA transcriptional start site (10 kb). Karyotype M1 is reported in the literature as ~60 kb in length (Mancini-Herrero et al., 1993). Maternally expressed genes are indicated in red or pink (genes that show imprinting in embryo and placenta, or placenta only, respectively), and paternally expressed genes are shown in blue. Paternally expressed genes are shown in green. The imprinting control region. Closed circles indicate a methylated CpG island, open circles indicate an unmethylated CpG island. Wavy lines indicate non-coding RNA. Arrows indicate distance from the transcriptional start site of Karyotype.

Figure 1.4. Organization of the *Kcnqlot1* imprinted region. Arrows show the distance upstream (-) and downstream from the *Kcnqlot1* ncRNA transcriptional start site (0 kb). *Kcnqlot1* mRNA is reported in the literature as ~60 kb in length (Mancini-Dinardo et al., 1003). Maternally expressed genes are indicated in red or pink (genes that show imprinting in embryo and placenta, or placenta only, respectively), and paternally expressed genes are shown in blue. Putative non-imprinted genes are shown in green. ICR; imprinting control region. Closed circles indicate a methylated CpG island, open circles indicate an unmethylated CpG island. Wavy lines indicate non-coding RNAs. Arrows indicate distance from the transcriptional start site of *Kcnqlot1*.

outside of the *Ospl5* and *Ascl2* genes. The tyrosine hydroxylase (*Th*) gene resides between the boundary of the *Kcnqlot1* domain and the *H19* domain.

Kcnqlot1 is the antisense ncRNA associated with this imprinted region. Its promoter is located in intron 11 of *Kcnq1*. *Kcnqlot1* is transcribed antisense to the 320 kb *Kcnq1* gene (Smilnich et al., 1999). The transcription start site of *Kcnqlot1* localizes to the *Kcnqlot1* imprinting control region, also known as KvDMR1. The *Kcnqlot1* ICR is methylated on the maternal allele, silencing *Kcnqlot1*. On the paternal, this region is unmethylated and *Kcnqlot1* is transcribed.

This domain shows bi-directional silencing, with respect to the ICR, of eight imprinted genes within the domain. Four of these genes, *Ospl5*, *Tssc4*, *Cd81*, and *Ascl2* are imprinted only in the placenta, whereas *Phlda2*, *Slc22a18*, *Cdkn1c*, and *Kcnq1* are ubiquitously imprinted, meaning they are maternally expressed in embryonic and placental tissues (Figure 1.4).

1.5.1 Deletion and truncation studies

To determine how the *Kcnqlot1* domain regulates imprinting within the domain, the entire ICR including the DMR as well as the promoter and transcriptional start site of *Kcnqlot1* on the paternal allele was deleted (2.8 kb deletion). This resulted in loss of imprinting for all imprinted genes in the domain (i.e. biallelic expression of normally imprinted genes), showing the broad control of the ICR in imprinted gene regulation of the *Kcnqlot1* domain (Fitzpatrick et al., 2002). Additionally, paternal transmission of a 3.6 kb-deletion, which encompasses the entire ICR, three hypersensitive sites in the *Kcnqlot1* promoter, and five repetitive sequences in the transcript, also led to biallelic

expression of *Ascl2*, *Kcnq1*, *Cdkn1c*, *Phlda2* and *Osbpl5*. These same effects were not seen if the deletion was maternally-transmitted (Fitzpatrick et al., 2002; Mancini-DiNardo et al., 2006). Unfortunately, these deletions removed the transcript start site and the ICR as well. Therefore, it is impossible to distinguish whether the effects are due to loss of the ICR, loss of the transcript, or both.

In an attempt to prevent transcription of the ncRNA, but leave most of the ICR intact, 244 bp of the paternal *Kcnqlot1* promoter was deleted. This deletion resulted in loss of *Kcnqlot1* expression, and biallelic expression of *Ascl2*, *Tssc4*, *Cdkn1c*, *Slc22a18*, *Phlda2*, and *Osbpl5* in d11.5 placenta (Mancini-Dinardo et al., 2006). These results show that transcript initiation and/or the transcript itself is/are required for proper imprinting of the domain, although these two factors cannot be teased apart from these experiments.

Insertion of a premature polyadenylation signal into the sequence to truncate the RNA was performed to determine whether initiation of transcription was sufficient to regulate imprinting (Mancini-Dinardo et al., 2006; Fitzpatrick et al., 2002). Activation of the normally silent paternal allele of maternally expressed genes was seen when the *Kcnqlot1* transcript itself was truncated to 1.5 kb, or 1.7 kb, indicating a functional role for the transcript in the monoallelic expression of genes within the domain. This suggests that initiation of transcription is not sufficient to confer its silencing properties but instead regulation is through the *Kcnqlot1* ncRNA (Mancini-DiNardo et al., 2006). One caveat, however, is that premature truncation may interfere with ICR function.

It is clear from these studies that both the ICR as well as the ncRNA itself are required for silencing of the imprinted genes in this imprinted domain. Additionally,

either initiation/transcription of the RNA, or the transcript itself is required for proper imprinting regulation in this domain.

1.5.2 Temporal and spatial regulation of imprinting

Additional complexity in the regulation of the *Kcnqlot1* domain lies in the fact that tissue-specific as well as temporal-specific imprinting is observed for genes in this domain. There is differential regulation of the eight imprinted genes in the *Kcnqlot1* domain: specifically *Osbp15*, *Tssc4*, *Cd81* and *Ascl2* are only imprinted in placental tissues whereas *Phlda2*, *Slc22a18*, *Cdkn1c*, and *Kcnq1* show imprinted expression in both embryonic and extraembryonic tissues (Engemann et al., 2000). Additionally, not all of the genes are imprinted throughout development and into adulthood, for example, *Cdkn1c* is maternally expressed in most embryonic and postnatal tissues, whereas *Kcnq1* is monoallelically expressed during embryonic development, but becomes biallelic in the adult (Caspari et al., 1998). *Ascl2* is imprinted and expressed in the placenta only (Guillemot et al., 1995). *Kcnqlot1* is expressed exclusively by the paternal allele as early as the 2-cell embryo, and maintains that status until the late blastocyst stage and continues to be monoallelically expressed in the placenta (Lewis et al., 2006).

Embryonic loss of function of DNA methyltransferase 1 (*Dnmt1*) results in widespread loss of methylation and imprinting, as well as embryonic lethality (Li, 1992; Li, 1993). However, loss of *Dnmt1* had no effect on placental-specific imprinted genes in the *Kcnqlot1* domain (Caspari et al., 1998, Tanaka et al., 2001, Lewis et al., 2004, Green et al., 2007), suggesting that imprinted regulation of these genes in placenta is DNA methylation-independent. Instead, imprinted genes in the *Kcnqlot1* domain are marked

by histone modifications, with active marks (H3K9ac, and H3K4me2) being associated with the expressed allele, and repressive marks (H3K9me3, and H3K27me3) associated with the silenced allele (Lewis et al., 2004; Umlauf, 2004). Paternal *Kcnqlot1* expression is required for the targeting of repressive histone modifiers to adjacent genes (Lewis et al., 2004). Allelic expression of *Kcnqlot1* has also been described in embryonic stem cells (ES cells) and trophoblast stem cells (TS cells) where it was shown that both cell lines express the *Kcnqlot1* ncRNA exclusively from the paternal allele (Lewis et al., 2006). Exactly how these variable expression patterns are maintained is unknown, but they appear to be under the control of the *Kcnqlot1* ICR and its associated ncRNA (Mancini-DiNardo et al. 2003; Fitzpatrick et al. 2002).

1.5.3 Models of imprinted regulation of the *Kcnqlot1* domain

The precise mechanism by which the hierarchy of silencing of paternal genes in the domain is initiated or maintained in the *Kcnqlot1* domain is unknown. It is clear that the ncRNA may be important, based on the truncation experiments noted above. An important question that remains is whether silencing is achieved by regulation through the imprinting control region, through the act of transcription, or the transcript itself.

First, imprinting of the domain could be achieved by the ICR acting as a silencer on the unmethylated paternal allele by binding with specific factors and/or chromatin remodeling factors that establish and/or maintain an inactive state on the paternal chromosome. This region has been seen to have silencing activity in cell culture studies (Mancini-Dinardo et al., 2003).

Second, the act of transcription of the ncRNA may be responsible for silencing within the *Kcnqlot1* domain. This model, termed expression-competition, suggests that ncRNA and protein-coding promoters compete in *cis* for one or more regulatory elements (Barlow, 1997) and that the transcript itself is dispensable. In this situation, transcription initiation and elongation would deny regulatory elements to promoters within the domain (transcription interference), leading to silenced paternal genes, whereas lack of transcription at the maternal allele allows access of transcription factors to the protein-coding genes, leading to expression. In this instance, degradation of *Kcnqlot1* post-transcriptionally, as with RNAi methods, should have no effect on the imprinting within the region, as both initiation and elongation of the transcript would occur in the presence of the RNAi.

Third, one must consider the possibility that the transcript itself is responsible for silencing within the domain. Since *Kcnqlot1* is transcribed in an antisense orientation, overlapping translated genes, it is possible that low levels of transcription from *Kcnql* could produce dsRNA intermediates and target *Kcnql* for degradation by a RNA-induced Transcriptional Silencing (RITS) complex (Rouguelle and Heard, 2002). However, there are no known nuclear RITS complexes that function in *cis* as of yet, and this model does not explain the bi-directional silencing nature of *Kcnqlot1*.

The transcript could also function through a physical interaction between the DNA and the ncRNA. The *Xist* ncRNA acts to coat, in *cis*, one X-chromosome to inactivate it, recruiting repressive histone modifications and DNA methylation (Marahrens, 1997; Thorvaldsen, 2006). There is some evidence to suggest that the *Kcnqlot1* RNA remains in close proximity to its region of transcription based on

combined RNA and DNA FISH (Murakami et al., 2007). However, proper splicing of *Xist* is required for its function, and no evidence of splicing has been described for *Kcnq1ot1* (Seidl et al., 2006; Wutz, 2002).

1.6 *Kcnq1ot1*

Based on studies examining regulation of imprinting of the *Kcnq1ot1* domain, it appears that the ncRNA transcript is important. *Kcnq1ot1* is the paternally expressed, maternally repressed long noncoding RNA associated with the *Kcnq1ot1* imprinting domain. It is believed to be a RNA polymerase II transcript that is polyadenylated, unspliced, and retained in the nucleus (Terranova et al., 2008; Pandey et al., 2008). It is paternally expressed from the *Kcnq1ot1* domain on distal chromosome 7 in the mouse (11p15.5 in the human). During development, expression is first seen at the 2-cell stage where it exhibits imprinted expression, and remains imprinted through preimplantation development, in embryo-derived stem cells, and at d9.5 and d17.5 of postimplantation development in both the embryo and placenta (Lewis et al., 2006; Umlauf et al., 2004).

1.6.1 *Kcnq1ot1* transcript length

Originally, *Kcnq1ot1* was estimated to be approximately 60 kb in length based on the high degree of sequence homology between mouse and human (Smilinich et al. 1999). Additionally, the EST database shows little alignment just after 60 kb. However, further analysis by our laboratory revealed alignment of the putative RNA sequence to the EST database at approximately 84 kb. Recently, it was reported that the *Kcnq1ot1* transcript terminates at 91.5 kb from the transcriptional start site (Pandey et al., 2008). In these

experiments, the 3' end of the gene was mapped using a strand-specific RNase protection assay (RPA) followed by 3' Rapid Amplification of cDNA Ends (RACE). However, another recent study suggests that the transcript terminates somewhere between 80 kb and 120 kb (Redrup et al., 2009). In these experiments, analysis with RACE provided four transcriptional start sites in a 20 bp range with a polyadenylation site located 121 kb downstream. Using quantitative RT-PCR on wildtype tissue compared to tissue with a 2.8 kb deletion of the *Kcnqlot1* ICR that lack expression of *Kcnqlot1*, ICR-dependent transcripts could be detected up to 80 kb but not 120 or 130 kb from the start site. However, strand-specific RT-PCR could still detect a signal as far as 120 kb from the start site.

Taken together, these data point out that the true length of *Kcnqlot1* is arguably indeterminate, and requires further examination. Because of its involvement in regulation of imprinting in the *Kcnqlot1* domain, it is important to elucidate the characteristics of the *Kcnqlot1* noncoding transcript, including its full length, to better understand its role in imprinted gene regulation as well as in human disease.

1.7 Rationale

As the length the *Kcnqlot1* transcript has not yet been determined, characterization of the *Kcnqlot1* domain is incomplete. To fully characterize the noncoding RNA and decipher its function, it is imperative that the true length of the transcript and the 3' end be identified. Since the *Kcnqlot1* transcript is longer than previously thought, and its transcription extends beyond that of *Kcnql*, it is possible that transcription through silenced genes on the paternal allele may lead to their silencing, at

least for downstream genes. Therefore, it is important to determine how long the transcript is, and what genes it may be transcribed through. Based on this, while tyrosine hydroxylase is located outside of the putative domain boundaries for the *Kcnqlot1* domain, transcription through it may lead to it being silenced and therefore classified as imprinted.

In summary, there is a lack of not only descriptive information about the *Kcnqlot1* ncRNA, but also functional information. It is unknown whether the ICR, the act of transcription of *Kcnqlot1*, or the presence of the transcript itself is the main element responsible for silencing within the domain. It is unknown how *Kcnqlot1* may function to establish or maintain imprinting in the domain. Analysis of the imprinted methylation and expression profile of *Kcnqlot1*, and associated imprinted genes in three cell lineages, embryonic stem (ES) cells, trophectoderm stem (TS) cells, and extraembryonic endoderm (XEN) cells may delineate differences in imprinted gene regulation. As *Kcnqlot1* regulates imprinted expression of genes in this domain, strategies that reduce *Kcnqlot1* RNA but do not affect the ICR or transcription may aid in understanding the tissue-specific regulation of this domain.

1.8 Hypotheses

My primary hypothesis is that transcription of Kcnqlot1 is required to maintain imprinted silencing within the Kcnqlot1 domain. I have three additional hypotheses based on this model:

- (1) The length of *Kcnqlot1* non-coding RNA extends beyond 60 kb to more than 460 kb.

(2) *Tyrosine hydroxylase (Th)* is imprinted in placental lineages in early development, as it is located distally to the ICR of the domain and resides in close proximity to a placentally imprinted gene, *Ascl2*.

(3) The *Kcnq1ot1* ncRNA regulates imprinting of the *Kcnq1ot1* domain in embryo-derived stem cells.

1.9 Objectives

(1) To determine the precise length of *Kcnq1ot1* ncRNA and to verify that amplified products correspond to *Kcnq1ot1* and not an unidentified transcript.

(2) To determine whether the tyrosine hydroxylase gene is imprinted.

(3) To determine whether depletion of *Kcnq1ot1* will result in activation of maternally transcribed genes from the paternal chromosome as well as loss of *Kcnq1ot1* ICR methylation.

2.1.2 Primers

All primers were customarily designed and then synthesized by Sigma Genosys (Oakville, Ontario).

CHAPTER 2

MATERIALS and METHODS

2.1 MATERIALS

2.1.1 Genetic crosses

C57BL/6J (B6) and *Mus musculus castaneus* (CAST) mice were obtained from The Jackson Laboratory, Bar Harbor, Maine. B6(CAST7) mice, that contain two CAST chromosomes 7 (CAST7) within a C57BL/6J background, were generated by Dr. Mellissa Mann at the University of Pennsylvania, and re-derived at The University of Western Ontario. Mice were housed at the Victoria Research Laboratories Vivarium.

Female B6 mice were mated to CAST male mice to create C57BL/6J X CAST F₁ embryos (referred to as B6 X CAST). Reciprocal cross embryos were obtained by mating CAST7 females with B6 males (CAST7 X B6). In all crosses, the female parent is listed first.

Embryonic day 9.5 placentae (B6 X CAST) with the KvDMR1 deletion were gifted from Drs. Michael Higgins and Louis Lefebvre. The KvDMR1 deletion is paternally inherited, and removes a 2.8 kb fragment from intron 10 of *Kcnq1* that contains the entire *Kcnq1ot1* DMR, as well as the promoter and transcriptional start site for *Kcnq1ot1* (Fitzpatrick et al., 2002).

2.1.2 Primers

All primers were manually designed and then synthesized by Sigma Genosys (Oakville, Ontario).

2.1.3 siRNA

Custom siRNAs targeting the *Kcnqlot1* region were designed using RNAi Codex (<http://codex.cshl.edu/scripts/newmain.pl>) and then synthesized by Dharmacon (ThermoScientific, Lafayette, CO).

Control siRNAs were included: siGLO Green Transfection Indicator to determine whether transfection was successful, a Cyclophilin-B targeting positive control siRNA and a scrambled-sequence non-targeting siRNA.

For dilution of siRNAs to the appropriate concentrations, 5X siRNA buffer, also obtained from Dharmacon, was used.

2.1.4 Embryo-derived stem cells

B6 X CAST extra-embryonic endoderm cells (XEN cells), embryonic stem cells (ES cells), and trophectoderm stem cells (TS cells) were produced from blastocysts and maintained in culture. Stem cells were transduced with either a shRNA targeting *Kcnqlot1* at 43 kb from its start site (K43), a non-functional shRNA homologous to *Kcnqlot1* at 82 kb (K82), or with a vector targeting the luciferase gene (Luc). All cells were collected for analysis between passages 7 to 10. NIH 3T3 cells were obtained from the American Type Culture Collection (ATCC).

2.1.5 Transfection of NIH 3T3 cells with siRNA

Lyophilized siRNAs were resuspended in 5X siRNA buffer to a final concentration of 40 μ M. Lipofectamine™ 2000 Transfection Reagent was used to

introduce the siRNAs into the NIH 3T3 cells as per the manufacturer's protocol, and cell collection for analysis was performed at passage 5 (Invitrogen, Burlington, Ontario).

All experiments were carried out by myself except for production, transfection, and transduction of stem cells and NIH 3T3 cells, which were done by Dr. Michael Golding within our lab.

2.1.6 Chemicals

All reagents used for this research were purchased from the following distributors: BioRad (Montreal, Quebec), Invitrogen (Burlington, Ontario), Fisher Scientific (Ottawa, Ontario), Gibco (Burlington, Ontario), Applied Biosciences (Streetsville, Ontario), and Sigma-Aldrich (Oakville, Ontario).

2.2 METHODS

2.2.1 Collection of embryonic tissues and postnatal samples

Two crosses were used to generate reciprocal genotypes for these experiments; B6 X CAST and CAST7 X B6. Matings were set up, and B6 X CAST F₁ embryos were retrieved at d8.5 and d9.5, and CAST7 X B6 F₁ embryos were recovered at d12.5. Embryos were dissected into three tissue types: embryo proper, yolk sac, and placenta, the latter of which was dissected from the maternal decidua. Brain tissue was isolated from d12.5 B6 X CAST F₁ pups. Recovered tissue samples were frozen on dry ice and stored at -80°C.

2.2.2 DNA and RNA isolation

DNA from day 8.5, 9.5, and 12.5 embryo, yolk sac and placenta, XEN, ES, and TS cells, and neonatal brain was extracted using the DNeasy Blood and Tissue Kit as per the manufacturer's instructions (Qiagen, Mississauga, Ontario). Eluted DNA was stored at -80°C.

Total RNA was extracted using the Roche High Pure RNA Tissue Kit as per the manufacturer's instructions (Roche Applied Science, Laval, Quebec). A second DNA degradation step was performed before synthesis of cDNA with DNase I according to the manufacturer's instructions (Invitrogen, Burlington, Ontario). RNA was stored at -80°C.

2.2.3 Reverse transcription of RNA

Complementary DNA (cDNA) was synthesized from total RNA using Superscript III (SSIII) as per the manufacturer's instructions (Invitrogen, Burlington, Ontario). Initially, reactions were primed using both random hexamer primers (Invitrogen) as well as oligo-d(T) primers (Invitrogen). To proceed with this analysis, it was necessary to ensure that successful PCR amplification was not a result of amplifying contaminant genomic DNA nor any remaining RNA in the cDNA pool. To avoid DNA contamination, after RNA extraction from tissues which includes a DNase I step, a second Invitrogen DNase I step was performed; DNases from other manufacturers (Sigma-Aldrich, Oakville, Ontario; Roche, Laval, Quebec) were unable to completely remove genomic DNA, producing amplified product in the controls containing no reverse transcriptase (no-RT). Secondly, after cDNA synthesis, samples were treated with RNase A to remove any residual RNA. When a no-RT control for a given sample had no RNA or DNA contamination, the sample was then used for PCR amplification with the designed primer

sets. cDNA was synthesized from total RNA isolated from B6 X CAST day 9.5 placentae, as it showed robust expression of *Kcnqlot1*, and CAST7 X B6 genomic DNA was used as a positive control to confirm that the primer sets worked. Once it was confirmed that equivalent cDNA quality was obtained with either method by amplification with primers to β -actin as well as primers located 2 kb from the transcriptional start site showing equivalent intensity, cDNA was produced using only random hexamer primers. cDNA was stored at -20°C.

2.2.4 Primer design

Approximately 660 kb of DNA sequence was obtained from Ensembl (<http://www.ensembl.org/index.html>) for *Mus musculus* distal chromosome 7 (150,418,839 bp to 149,758,443 bp).

Primers to amplify *Kcnqlot1* were designed to the putative RNA sequence, which is identical to that of the DNA sequence, manually, with confirmation of appropriate primer attributes by NetPrimer (Premier Biosoft International, <http://www.premierbiosoft.com/netprimer/index.html>) (Table 2.1). Failure of primer pairs to amplify cDNA but not genomic DNA controls indicated that there was no transcript present at that location. Initially primers were designed at ~40 kb increments from the start site of *Kcnqlot1* towards the 3' end as far as 617 kb from the transcriptional start site. After amplification was no longer observed at 40 kb intervals, primer pairs at ~5 kb intervals were designed to further locate the 3' end of the sequence.

Table 2.1. List of primers, reaction annealing temperature and amplicon size

Primer Name	Forward Sequence	Reverse Sequence	Tm (°C)	Amplicon Size (bp)
2k	ATTGGGAACCTGGGGTGAAGC	GGCACACGGTATGAGAAAAGATTG	54	814
61k	GGAACGGGGAGTGGGAGAG	GAAATCCCATCCCTCACTTGAT	60	268
65k	ACAGGTTTGGTAGTATGAAGG	ATCCATTGTATCCATTTTGAC	58	1328
66k	ACAGGTTTGGTAGTATGAAGG	ATCCATTGTATCCATTTTGAC	50	922
89k	GCTTGATTTTCATTGATGTGTAA	GACCTATTCCACCCACCTAA	58	378
91k	CATGTGTTGAAATCGCTG	GCAAAGTTGTCTGTGTCC	53	308
94k	AAACTAACTTTGTATTCCTGAACC	CAGTGTGAGAAGTGAGATACCC	52.8	472
99k	ACACAACCAGAAGTCACGG	AGCCATTGTCCAAGTTCAG	53	355
110k	ATGTGACTATCTAAGAGACTTCCT	ACCCCTAAAATAATGTCC	58	396
			56-	
118k	CTTCCTTGTGAGAAAAGCATT	GCTGGGGATTGGGAGTAT	58	350
127k	AGGGTCACCAGTAATATGCTT	GGCTTGCCAGAGAAAATC	53	400
136k	GGCAAAAGACCCAGAAATA	CCAGAGGTAGACAGATGTTAGAA	58	511
147k	CCCGACTTGAAGCCCA	CTTCCCAAACTCCATAGG	55	472
196k	GGTGCTGCCCTTTTAC	ATCAAGACAATCCCTAATAGACA	53	480
202k	GCCAAGAGGGTACTAAGGTC	ACGTCTAGCATCCATGAGG	56/57	407
			58-	
228k	CTCGGAGGTGGAATGGCTG	TCACGGCGGAACCTGGGA	60	293
288k	AGTGCTTATTCATGGCCAGC	TCAGGAACCTGAGCTTTGCT	60	237
307k	GGCATACCTGAATGTGAGC	TTATCTTTCTGTTCTCACAACCTCTA	57	444
380k	TGTCCTCCTCATTTCATCC	GCTGTAGCCTACCCATTAGTA	50	426
392k	TGTGATAGGGTTATTTCTACTGAG	AGTTTGTGCTGCTGTGTATTC	58	416
408k	CAAATGAGGACCTGGAAGAT	TAACAGCACAACTGAGGAAAT	60	211
410k	ATACAATAGTCAGATAATGGGG	AACAGGTTTGCTTGCTCA	60	333
430k	AACAGAGATGTGTAGCAGTAGGG	GCTCCAGTCAGCAAGCACTT	61	376
440k	CTGAGCCATCTCTCCAGCC	AACAATGAAGCCGAGGATAATG	64	434
444k F	GGCTTACAAGTTCTGAGGT		54.5	0
444k Ra		TCTACGAGTTTAGACTGGAT	54.5	863
444k Rb		CCATCTATGTCCCAGTCTC	54.5	1277; 263
447k	CCTTTTGACAATGTGGTAAC	CAGACTTGGGTGTAGTGGG	58	307
455k	CCACAAAGGGATAAAAAACAT	AACCACCAAACATTCTACCA	60	358
461k	AGGGCATTGGTGGACAGGA	ATTTGGTATTGATATTTTGCTCAGAC	60	497
463k	CCGTGAAAGGCATACTGTTGAA	GGCAATCCAACCTCTGTGTGTC	62	224
463k Rb		GACACAATAATACCAGCCTTCAAT	62	322
463k Rc		ATATTAACCAGTATTACCCTACAATCAT	62	377
463k Rd		AACACAGAACTAGACAGTACAGTGGTGA	62	456
463k Re		TACCATATTATTTTATTATTACTACCCAA	62	539
463k Rf		ATTATGCCACAATTCCAAACACTCA	62	659
463k Rg		CAACATCACCAGTACAACCTCATTCT	62	827

Table 2.1 cont.

Primer Name	Forward Sequence	Reverse Sequence	T _m ((°C))	Amplicon Size (bp)
464k	GAACCAGACAC AACATCACCAGT	TTGTATATGATAGAGTTCTGTTTTATAG	60	276
464k Rb		TGCCCACCTGCTCTCCAGTG	62	398
464k Rc		AGTCCAACCTCACCCTGAGTCCC	58	643
464k Rd		CTTCTACCACCATTATGAGATATGATTG	69	890
465k	AACAGGAATGGGCAGACAAA	GGAATATGCCAACGCTGTTT	60	341
466k	AATCAGCTTAC CATCCCAGG	GCAGCAGAAGAATGGAGGTC	60	202
466k Rb		ACTATGAGGTGCTAAATGCCAGAGACT	60	436
466k Rc		AAAAATAGAATGTGCCCACTCCAC	60	790
467k	AATGTTATCCCCCTTCCCAG	GGCTGTTGCCAGGTAAATGT	60	271
470k	TTTGGAAGGTTTTAGGTC	ATTGGAGTTTAAGACATTCTCAGAG	52	113
472k	GGTAGAGATAGTCTTGGAGAAGGA	CATTGGACACGCAAACCTTATA	60	632
473k	ATGAGAGTTTGAAGAATGCCTT	TGTCCTCCCTGTCACCTAAAA	56	215
477k	CTTCCCAGTTTTCCCTCTA	TATGCCTTTGTGGTAGTTG	58	598
479k	TGAGCAATGTTTTAGACCTA	GCAAAGTATCAATATGCCCTACCTG	60	738
484k	TTTCAGATAGCAATACAAGTT	TAAGGTTCCACTCCCAC	54	411
486k	TCTGAGAAGTGGGGTATGGTT	TGTGGGGAAGCATAGTTTGAG	63	435
488k	CGTCTATCCTGTTGATTTT	TTCTTTTATGAATGTGGGT	54	609
489k	CTCGCTATATTTGGTCATTAT	AGATAGATTACACAGTGCTTTAT	53-55	391
494k	TGGGGTATTTGTGTTATCTCTTC	AATCTTGCCCTGCTGTGG	61	630
501k	CCTCTATAAATAAAGACAGCCA	AACATCCAAGTCAGTCCTG	57	483
507k	GGTAACTGTGGGGGTGGAG	CAGGCAGAGAATGAGAAGGAG	63	332
514k	TCAGGGACTTGAAGGTTTTAT	GGATGATATGATAGTATA	50	263
521k	TCAGCAGCCCGAGAGAG	AGAGCACATAGAGTGAAGAACC	60-62	408
604k	TAAACTCCAGTGAGAACCAATAGAT	AAGCCTAAGACCTCCCTAAG	58	380
617k	CTACAATCAAAAACCATCAGCAA	GCCTCTAAAGCCTACTCATCTTC	61	646
623k	TGGGAAGGGCTGGCTGAG	GCTGGGATTTGAACTCTGGA	58	307
630k	GCACTGTGGACTCTGCCGAG	CTCGGCAGAGTCCACAGTGC	58	884
Slc22a18	ATCAACAGGACTTTTGCCCC	ACAGAATCTAGGCCAGTG	58	227
Cdkn1c	GCCAATGCGAACGACTTC	TACACCTTGGGACCAGCGTACTCC	58	364
Kcnq1	CATCGGTGCCCGTCTGAACAGG	TTGCTGGGTAGGAAGAGCTCAG	58	189
<i>Kcnq1ot1</i>	ATTGGGAAGTTGGGGTGGAAGC	GGCACACGGTATGAGAAAAGATTG	58	814
Ascl2	TGAGCATCCCAACCCCTA	CCAAACATCAGCGTCAGTATAG	58	474
Th	CTTCCGTGTGTTTCAGTGC	ACCGTGGAGAGTTTTTCAA	58	195
<i>Kcnq1ot1</i> outer	GGTTAGAAGTAGAGGTGATT	CAAAACCACCCCTACTTCTAT	58	571
<i>Kcnq1ot1</i> inner	GTGTGATTTTATTTGGAGAG	CCACTCACTACCTTAATACTAACCAC	52	207
M13	CGCCAGGGTTTTCCAGTCACGAC	TCACACAGGAAACAGCTATGAC		

k: kb from start site; R: reverse; a-g: series of reverse primers

Once a putative 3' end was narrowed to a 7 kb window, primers were designed at <1 kb intervals to locate the end of the transcript.

Primers for *Th* were manually designed to span three exons and contain a SNP for allelic discrimination. Primers for *Slc22a18*, *Cdkn1c*, *Kcnq1*, and *Ascl2* were previously designed as indicated (Table 2.1).

2.2.5 Polymerase chain reaction (PCR) amplification

Unless otherwise indicated, PCR amplification was performed using 1 μ L cDNA. illustra Ready-To-Go™ PCR Beads were used for all PCRs following the manufacturer's instructions (GE Healthcare Biosciences Corp, Piscataway, NJ). A negative cDNA control was included to ensure that there was no contamination. Samples were stored at 4°C. Unless otherwise stated, the PCR cycling parameters used were 95°C for 2 minutes, followed by 95°C for 30 seconds, T_m (variable based on primer sets used, see Table 2.1) for 30 seconds, and 72°C for 50 seconds, 35 to 45 cycles, followed by 7 minutes at 72°C, with 4°C hold.

2.2.6 Gel electrophoresis

Amplified cDNA was examined by gel electrophoresis on either agarose or polyacrylamide gels. For reactions requiring visualization on agarose gels, 1.25% agarose gels containing 1 μ g/ml ethidium bromide in 1X TAE buffer (0.04 M Tris, 1 mM ethylenediaminetetra-acetic acid disodium salt [EDTA] pH 8.0, 5.72 ml glacial acetic acid per 100 ml) were electrophoresed at 125 V for 35 minutes. For reactions requiring visualization by polyacrylamide gels, 7% polyacrylamide gels (acrylamide–bisacrylamide,

37.5:1, 40% solution in water, electrophoresis grade, 20X TBE, ammonium persulfate (APS), N,N,N',N'-tetramethyl-1,2-diaminomethane (TEMED)) in 1X TBE buffer (1.21 g/L Tris, 0.5 g/L boric acid, 0.074 g/L EDTA pH 8.0, 5.72 ml glacial acetic acid per 100 ml) were electrophoresed at 200 V for 60 minutes and then stained with 1 µg/ml ethidium bromide in 1X TBE buffer for 5 minutes. Agarose and polyacrylamide gels were examined using ultraviolet (UV) light on a BioRad Gel Doc™ XR transilluminator. Successful amplification was signified by the presence of a band of DNA at the expected size as well as blank negative controls.

2.2.7 3' Rapid amplification of cDNA ends (RACE)

The 3'RACE System for Rapid Amplification of cDNA Ends Kit (Invitrogen, Burlington, Ontario) was used according to manufacturer's instructions with the aim of identifying the 3' end of the *Kncqlot1* transcript. 3'RACE was carried out with the forward primer of 463k paired with the universal adapter primer provided with the 3'RACE kit which primes poly(A) sequences, on RNA from two day 9.5 placental samples. Products were electrophoresed on agarose gels, and bands of interest were excised and gel extracted using Qiaquick Gel Extraction Kit (Qiagen, Mississauga, Ontario). Once the cDNA was eluted, the amplicons were ligated into the pGEM-T EASY vector using the Promega's pGEM®-T Easy Vector System (Promega, Napean, Ontario), and transformed into competent *Escherichia coli* bacteria. Five µL of ligation reactions were added to 10 µL of bacteria and heat-shocked for 45 seconds at 42°C, followed by incubation on ice for two minutes. After recovery, 50 µL of super optimized broth with catabolite repression (SOC medium) was added to the reaction, and was

shaken at 37°C for one hour. The entire reaction was plated on agar plates containing 100 µg/mL of ampicillin. Three colonies were picked for each prospective band after a 12 hour incubation period on agar plates at 37°C. Colonies were transferred to culture tubes containing 2 mL of Luria broth (LB) containing tryptone (10 g/L), yeast extract (5 g/L), NaCl (5 g/L) and ampicillin (0.1 g/L), and were shaken at 37°C for 12 hours. Plasmids containing amplicons were then extracted from bacterial colonies using the GenElute™ Plasmid Miniprep Kit following the manufacturer's instructions (Sigma-Aldrich, Oakville, Ontario). Isolated plasmids were sent for sequence analysis to London Regional Genomics Centre (Robarts Research Institute, London, Ontario) and a BLAST analysis (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was performed to align the resulting sequences with the GenBank database.

2.2.8 Quantitative real-time PCR analysis

PCR amplification was visualized in real-time using the Opticon Real-Time PCR Detection System from Bio-Rad (Bio-Rad, Mississauga, Ontario). Commercial iQ™SYBR Green® Supermix containing all PCR constituents was used (Bio-Rad, Mississauga, Ontario). Unless otherwise indicated, 1 µl cDNA was added to each reaction, and each sample was PCR amplified in triplicate as technical replicates. All reactions were amplified with the following cycling parameters: 95°C for two minutes, followed by 35 to 40 cycles of 94°C for 20 seconds, 58°C for 30 seconds, 72°C for 45 seconds (fluorescence was detected at the end of each cycle) then 72°C for 7 minutes followed by a melting program to assess primer specificity. The cycle threshold (C_T) was determined for each reaction. One cycle difference equates to an approximate doubling

of amplicons, therefore relative fold change is calculated as $2^{\Delta C_T}$. As a normalization control, samples amplified with *Kcnqlot1* gene-specific primer sets were adjusted to the reference gene mitochondrial ribosomal protein L1 (*Mrpl1*). The change in C_T between these reactions was calculated (ΔC_T). For the *Kcnqlot1* gene-specific primers, the change in C_T for control tissue versus experimental tissues was then calculated ($\Delta\Delta C_T$) to assess relative amounts of amplicons in the reaction. For analyses, control tissue was set to a value of 1, and experimental reactions were normalized to the control. Standard error of the mean (SEM) was calculated for technical replicates to determine degree of variable.

2.2.9 Statistical analysis

One-tailed Student's t-tests were performed on Real-Time PCR $\Delta\Delta C_T$ results to determine whether the experimentally manipulated tissues were significantly different from the control cells. Differences were considered significant for $p < 0.05$.

2.2.10 Sequencing analysis of PCR amplicons of *Kcnqlot1*

For confirmation of amplicons belonging to *Kcnqlot1* and to find single nucleotide polymorphisms (SNPs) that may exist between B6 and CAST for *Kcnqlot1*, PCR was carried out using primers along the length of the *Kcnqlot1* transcript on pure B6 and pure CAST neonatal brain genomic DNA (gDNA) (Table 3). Double-stranded cDNA was then purified using the QiaQuick PCR Purification Kit according to the manufacturer's protocol (Qiagen, Mississauga, Ontario). Purified PCR product was sent to the London Regional Genomics Centre as outlined above for sequence analysis. Resulting sequences were aligned to the Ensembl sequence database by using the BLAST

function to determine whether amplicons were derived from the intended region of *Kcnqlot1*, as well to each other to identify SNPs between B6 and CAST strains of mice.

2.2.11 Allelic discrimination using restriction endonucleases

The Mouse Genome Informatics database (http://www.informatics.jax.org/strains_SNPs.shtml) was used to identify SNPs between B6 and CAST within the *Th* gene. Once polymorphisms were found, a restriction endonuclease (RE) was identified using NEBcutter v2.0 (New England Biolabs, <http://tools.neb.com/NEBcutter2/index.php>) that cut only one of the two sequences at the identified SNP (allele-specific restriction enzyme) (Table 2.2). This allows for discrimination of the B6 and CAST alleles in F₁ crosses (B6 X CAST; CAST7 X B6).

Polymorphisms between B6 and CAST, and allele-specific REs were identified previously for *Slc22a18* (Dao et al., 1998), *Cdkn1c* (Doherty et al., 2000), *Kcnq1* (Gould and Pfeifer, 1998; Jiang et al., 1998), *Kcnqlot1* (Mann, unpublished), and *Ascl2* (Mann et al., 2003).

2.2.12 Densitometry for allelic discrimination

To quantify relative amounts of amplicons from the B6 or CAST allele in F₁ embryos, placentae, and stem cells, computer-assisted densitometry was performed with QuantityOne 1-D Analysis Software (Bio-Rad, Mississauga, Ontario) to calculate the intensity/mm² for each band on a polyacrylamide gel following restriction endonuclease digest with an allele-specific RE. A background-correction was calculated by the software to yield an adjusted intensity/mm² for each band. The percent that each allele

contributed to the total intensity of the sample was then calculated to determine whether the sample was monoallelic or biallelic for the gene of interest. A mean maternal percentage of expression was calculated from replicates (n=3-5). Monoallelic expression was defined as $\geq 90\%$ from one parental allele.

2.2.13 Bisulfite conversion of DNA

Bisulfite Mutagenesis with agarose embedding was conducted on cells as described (Olek et al., 1996; Schoenherr et al., 2003), with modification (unpublished data, Mann lab). Cells were lysed with 0.1% IGEPAL (Biochemika), and 2 mg/mL Proteinase K (Sigma) in 10 μ L of lysis buffer [100 mM Tris-HCl pH 7.5 (Bioshop), 500 mM LiCl (Sigma), 10 mM EDTA pH 8.0 (Sigma), 1% LiDS (Bioshop), 5 mM DTT (Sigma)] for 1 hour at 50°C. Lysed cells were embedded in 2% low melting point agarose (Sigma) under mineral oil at 95°C. DNA/agarose beads were allowed to solidify for 10 minutes on ice. Oil was removed and denaturation of DNA was performed in 0.1 M NaOH (Sigma) at 37°C for 15 minutes with shaking. Agarose beads were placed in 2.5 M bisulfite solution [0.125 M hydroquinone (Sigma), 3.8 g sodium hydrogensulfite (Sigma), 5.5 mL water, 1 mL 3 M NaOH] at 50°C for 3.5 hours to allow bisulfite mutagenesis to occur. Following incubation, agarose beads were washed once in TE pH 7.5, and desulphonated with 0.3 M NaOH at 37°C for 15 minutes with shaking. Agarose beads were washed twice with TE pH 7.5, and twice with water. Beads were incubated under oil at 65°C and ~60 μ L of pre-warmed water was added. Agarose beads were mixed by pipetting and 20 μ L of diluted agarose was added to one Ready-to-go PCR Bead (GE) containing *Kcnqlot1* outer primers and 1 μ L of 240 ng/mL tRNA as a carrier. PCRs were split in half allowing two independent PCR reactions to be completed for each gene analyzed. Nested

PCR was performed with *Kcnqlot1* outer primers, followed by a separate PCR reaction with *Kcnqlot1* inner primers (Table 2.1) to amplify a 208 bp region of *Kcnqlot1* ICR that includes 20 CpG dinucleotides. The 2nd round PCR product was ligated into the pGEM-T EASY vector using Promega's pGEM®-T Easy Vector System (Promega, Napean, Ontario), and transformed into competent *Escherichia coli* bacteria as described above. Thirty-two colonies for each individual sample were picked from agar plates after a 12 hour incubation period on agar plates and were seeded into a PCR mix that included the M13 forward and reverse primers to amplify the region of interest (Table 2.1). PCR reactions from individual clones were then sequenced to determine the methylation status of each CpG. Sequence data were analyzed with FinchTV (<http://www.geospiza.com/Products/finchtv.shtml>). Any ambiguous base pairs (i.e. "N") were reviewed and changed manually by analysis of the chromatogram. Each sequence was examined for total number of CpGs remaining (methylated CpGs). Any non-CpG cytosines that were not converted during the bisulfite mutagenesis were calculated as a percentage of total non-CpG cytosines to determine the conversion rate. Sequences with less than 85% conversion were not considered adequately converted, and were not analyzed further. Additionally, clones that showed the exact pattern of methylation and unconverted non-CpG cytosines were considered to be identical clones and were included once in the analysis. The amplified sequence contains two SNPs that allow for allelic discrimination and were used to identify parental origin. Strands were considered hypermethylated if $\geq 50\%$ of the CpGs were methylated, and hypomethylated if $< 50\%$ were methylated. The percent of hypermethylated DNA strands was calculated for each parental allele as total hypermethylated strands/total strands.

Table 2.2. Summary of SNP position and RE digest for imprinted genes.

primer name	SNP Position	Tm (°C)	amplicon size (bp)	B6→ CAST	RE	B6 fragments	CAST fragments
Slc22a18	108, AF028739	58	227	C to T	<i>SacII</i> , 37°C	200, 27	227
Cdkn1c	1256, MMU20553	58	364	T to C	<i>TaqI</i> , 65°C	306, 58	257, 58, 49
Kcnq1	1763, MMU70068	58	189	A to C	<i>NlaIII</i> , 37°C	92, 72, 21	113, 76
Kcnq1ot1	3976, AF119385	58	814	T to G	<i>StuI</i> , 37°C	814	601, 213
Ascl2	10828, AF139595	58	474	T to C	<i>SfiI</i> , 37°C	474	267, 207
Th	1045, NM 009377.1	58	195	G to A	<i>BsrI</i> , 65°C	124, 71	195

CHAPTER 3

RESULTS

3.1 Analysis of *Kcnqlot1* transcript and expression

3.1.1 Determination of the transcript length of *Kcnqlot1*

3.1.1.1 PCR amplification of the putative *Kcnqlot1* ncRNA

The *Kcnqlot1* domain spans over 1 Mb, and contains 8 maternally expressed genes, and one paternally expressed noncoding RNA, *Kcnqlot1*. As sequence in this region is highly repetitive, harbouring 2 kb-stretches of strings of As and Gs, high AT content, and repetitive sequences, some regions were particularly difficult to amplify. As such, two sets of overlapping primers were designed to any given region of the putative RNA sequence, thereby increasing the chance of obtaining a primer set that would amplify the intended region. Failed amplification with a single primer set was insufficient to conclude that *Kcnqlot1* did not extend through this primer set. In these cases, additional primers were designed to adjacent sequences that have less repetitive stretches and are less AT-rich.

To determine whether the *Kcnqlot1* transcript extends past the putative 3' end (~60 kb), primer pairs were used to amplify cDNA from 2 to 380 kb from the start site of *Kcnqlot1* transcription (Figure 3.1). Failure of primer pairs to amplify cDNA but not genomic DNA indicated that there was no transcript present at that location. Successful amplification of cDNA was obtained with primers from 2 kb to 380 kb, indicating that *Kcnqlot1* extended at least through the *Th* allele.

Since these preliminary PCR experiments did not delineate the 3' end of the *Kcnqlot1* transcript, primer pairs were designed to putative RNA sequence in ~40 kb

increments from the end of *Th* gene towards the *H19* domain (Figure 3.1). The *H19* domain is located ~150 kb downstream of *Th*. It was considered unlikely that the *Kcnq1ot1* transcript extends through this domain, as deletion of the *Kcnq1ot1* ICR had no effect on imprinted expression of genes in *H19* domain (Fitzpatrick et al., 2002). Amplification with primer sets 365k, 408k, 447k, 486k, 521k, 565k, 617k revealed that the 3' end of *Kcnq1ot1* resided in a region between 447 to 486 kb (Figure 3.1). Primer pairs were then designed every 4-5 kb in the 39 kb region between 447 and 486 kb (primer pairs 455k, 461k, 467k, 470k, 473k, 477k, 479k, 484k). A fragment was detected at the predicted size for each of the primer pairs up to and including 461 kb. No amplification was seen in any primer pairs further downstream.

Finally, primer sets were designed at ~1 kb intervals between 461 to 467 kb (primers: 463k forward with 7 different reverse primers (a-g), 464k forward with 4 different reverse primers (a-d), 465k, and 466k). While primer sets at 461 and 463 kb successfully amplified the cDNA, primers designed from 464-467 kb failed to produce amplicons from cDNA, indicating that the end of the transcript resides within a 241 bp region between 463,930 and 464,171 bp from the transcriptional start site.

3.1.2 Tissue-specific analysis of *Kcnq1ot1* length

Of the eight maternally-expressed genes in the region, only *Phlda2*, *Slc22a18*, *Cdkn1c*, and *Kcnq1* are believed to be ubiquitously imprinted because they are maternally expressed in both embryos and placentae, *Osbp15*, *Tssc4*, *Cd81*, and *Ascl2* are imprinted

Figure 3.1. The putative *Kcnq1ot1* ncRNA extends 463 kb from the transcription start site. Selected regions were amplified from the start of the *Kcnq1ot1* transcript through the *H19* domain. Organization of the *Kcnq1ot1* and *H19* imprinting domains. The top strand corresponds to the maternal allele, while the bottom strand corresponds to the paternal allele. Maternally expressed genes are indicated in red/pink, and paternally expressed genes are shown in blue. Putative non-imprinted genes are shown in green. ICR; imprinting control region. Closed circles indicate a methylated CpG island, open circles indicate an unmethylated CpG island. Wavy lines indicate non-coding RNAs. The putative *Kcnq1ot1* ncRNA extends 463 kb (k) (vertical green line). Intronic and intergenic regions spanning from 2 to 617 kb (arrowheads) were amplified using day 9.5 placenta (P) cDNA and genomic DNA (G) as a positive control. The 3' end of *Kcnq1ot1* lies between 463 and 464 kb.

only in placentae. Additionally, some genes appear to be expressed biallelically until postimplantation whereas others are monoallelic even in preimplantation tissues, suggesting additional temporal regulation of *Kcnqlot1* (Umlauf et al., 2004; Lewis et al., 2006). Since there is variability in the expression pattern of these imprinted genes, it is possible that different transcript lengths of *Kcnqlot1* may exist within different tissues or developmental time-points. To determine if the length of *Kcnqlot1* is spatially and temporally conserved, B6 X CAST ES cells, XEN cells, TS cells and neonatal brain samples, as well as CAST7 X B6 d12.5 embryos and placentae, were examined using selected primer sets along the length of the putative *Kcnqlot1* ncRNA, and compared to amplification of d9.5 placenta from above. Amplification was observed up to and including 463 kb in all tissues, with no amplification extending beyond that point (Figure 3.2). Thus, in all three lineages (embryonic, placental, and extraembryonic endoderm), and at the selected developmental time points (peri-implantation, post-implantation, and post-natal), the transcript ends at ~463 kb.

3.1.3 Identification of the 3' end of *Kcnqlot1*

To more definitively identify the 3' end of the *Kcnqlot1* ncRNA, 3' RACE was attempted on wild-type B6 X CAST d9.5 placentae as this tissue showed robust expression of the transcript. When forward primers at 463 kb were used with the adapter primer that captures the poly(A) tail of transcripts, bands of various sizes were produced. Six potential bands at various sizes (A - ~350 bp, B - ~250 bp, C - ~500 bp, D - ~2000 bp, E - ~1500 bp, F - ~700 bp, G - ~425 bp) were gel-extracted and cloned into the pGEM-T EASY vector and then sequenced (Figure 3.3). However, results showed that

Figure 3.2. The length of the *Kcnq1ot1* transcript is consistent in a variety of tissues and developmental time-points. Selected primers along the length of the transcript were used to amplify the transcript in XEN, TS, and ES cells, d9.5 placenta (from Figure 3.1), d12.5 embryo and placenta, and neonatal brain, as well as genomic DNA (gDNA) controls. In all tissues, signal from the transcript was detected up to and including 463k; no amplification was detected at 464k through to 489k in cDNA compared to control gDNA. β -Actin and no RT reactions for each tissue show the presence of cDNA with no RNA or DNA contamination before amplification. For band sizes, see Table 2.1.

Figure 3.3. *Kcnq1ot1* 3'RACE analysis. Shown after agarose gel electrophoresis. Samples 77, 63, 56, and 74 represent different individual d9.5 placentae and the forward primer 463F was used with the reverse primer which was the universal adapter primer in the reactions. Bands were excised and gel purified at the lengths indicated by A – G (A, ~350 bp; B, ~250 bp; C, ~500 bp; D, ~2000 bp; E, ~1500 bp; F, ~700 bp; G, ~425 bp). M: marker (1 kb plus ladder). Neg; water.

The amplified products did not align with the sequence in GenBank (accession # U00000) suggesting with sequence no. 12, 19, and 18. An OR search of analysis by highly repetitive, it is clear that the forward primer is not sufficiently specific in genomic DNA and specific fragments. This suggests that more specific primers will need to be designed in the region upstream the 5' end of the transcript, or alternatively, that the forward primer will be polyadenylated, making this method of determining the 5' end ineffective.

3.2 Verification that sequences aligned to the *Kcnqlot1* ncRNA

3.2.1 Parental origin of the *Kcnqlot1* ncRNA

On the maternal allele, the *KCN* is methylated and *Kcnqlot1* is silent, while on the paternal allele is unmethylated and *Kcnqlot1* is expressed. To determine parental origin of the expressed allele, we used the *KCN* marker. With 100% and 100% of the

the amplified products did not align with the sequence on distal chromosome 7, but rather aligned with sequence on other chromosomes, such as 12, 16, and 18. As the region of analysis is highly repetitive, it is likely that the forward primer is not sufficiently specific to generate *Kcnqlot1* specific fragments. This suggests that more specific primers will need to be designed to this region to capture the 3' end of the transcript, or alternatively that the transcript may not be polyadenylated, making this method of determining the 3' end ineffective.

3.2 Verification that amplified products correspond to the *Kcnqlot1* ncRNA

While the putative *Kcnqlot1* amplicons aligned to the correct region of the genome, the possibility existed that amplicons may be from a distinct and/or previously unknown transcript. To verify that the amplified products correspond to the *Kcnqlot1* transcript, and not an unidentified transcript(s), three strategies were employed. First, parental identity of the amplified transcript was examined through allele-specific restriction endonuclease analysis. Second, RNA interference was used to determine whether amplicons were specific to the *Kcnqlot1* ncRNA. Third, placental tissue harbouring a targeted deletion of the paternal *Kcnqlot1* ICR was used to determine whether expression originated from the *Kcnqlot1* promoter.

3.2.1 Parental origin of the *Kcnqlot1* ncRNA

On the maternal allele, the ICR is methylated and *Kcnqlot1* is silent, while the paternal allele is unmethylated and *Kcnqlot1* is expressed. To determine parental origin of the expressed allele, cDNA from B6 X CAST wild-type XEN, ES, and TS cells, d8.5

and d9.5 embryo, yolk sac, and placenta, and neonatal brain was amplified using primers for a region at 2 kb from the transcriptional start site, and then digested with *StuI*, an allele-specific RE, to determine from which parental allele the transcript is expressed. Results show that the transcript is seen exclusively from the paternal allele in d8.5 and d9.5 embryos, yolk sacs, and placentae as well as ES, XEN, and TS cells (Figure 3.4). This confirms the imprinted status and paternal expression of *Kcnq1ot1* in these tissues, at least for an amplicon close to the *Kcnq1ot1* start site. Additional SNPs were identified, and future analyses will examine parental identity for additional amplicons.

3.2.2 RNA Interference analysis

3.2.2.1 shRNA-mediated reduction of *Kcnq1ot1* in XEN cells

RNA interference is a means of reducing the amount of a given transcript *in vivo*, as the short hairpin RNA (shRNA) targets the transcript for degradation via the RNA interference pathway. As the shRNA targeting *Kcnq1ot1* should result in degradation of the entire *Kcnq1ot1* RNA, it is expected that amplicons arising from *Kcnq1ot1* will exhibit a reduction in the mRNA levels, while those from any other RNA transcript in the region will be unaffected by *Kcnq1ot1* depletion. If all amplicons from 2 kb to 463 kb represent the *Kcnq1ot1* transcript, targeting either end of the *Kcnq1ot1* ncRNA will lead to decreased amplification of all regions examined. To determine whether amplicons represented the *Kcnq1ot1* ncRNA, real-time quantitative PCR (qPCR) was performed using cDNA from XEN cells that harbour a shRNA targeting *Kcnq1ot1* at 43 kb from the start site (K43), and control XEN cells harbouring an ineffectual shRNA that does not affect the transcript (K82). XEN cells were chosen because they are functionally relevant

Figure 3.4. Paternal-specific expression of *Kcnqlot1* at 2 kb from transcriptional start site. Analysis of allelic-specific expression of *Kcnqlot1* was performed on B6 X CAST samples using PCR with primers located 2 kb from the transcriptional start site, followed by allele-specific restriction endonuclease digest with *StuI*. Red arrows show maternal B6 band sizes, blue bands show paternal CAST band sizes. neg: no cDNA control, EMB: embryo; YS: yolk sac; PLA: placenta; XEN: extraembryonic endoderm cells; TS: trophoctoderm stem cells; ES: embryonic stem cells. B6: pure B6 neonatal brain tissue, CAST: pure CAST neonatal brain tissue. M, 1 kb+ ladder.

as peri-implantation extraembryonic cells, exhibit abundant expression of *Kcnq1ot1*, and are easily transduced with the shRNA targeting vectors. Additionally, because they are primary cell cultures and not immortal cell lines, they are more functionally similar to *in vivo* tissues than to cell lines that are immortal. Real-time qPCR was performed with primers at intergenic regions, 2k, 65k, 202k, 392k, and 461k, with both K43 cells as well as control K82 XEN cells (Figure 3.5A). Amplification of *Kcnq1ot1* was significantly decreased ($p < 0.05$) in the K43 XEN cells compared to control cells (between 0% and 10% of control expression levels) for all primer pairs examined extending from 2 to 461 kb from the start of *Kcnq1ot1* transcription, indicating that all these amplicons represented the *Kcnq1ot1* ncRNA (Figure 3.5B).

3.2.2.2 siRNA-mediated reduction of *Kcnq1ot1* in NIH 3T3 cells

It is possible that transcripts of varying lengths are produced from the *Kcnq1ot1* promoter. If this were the case, then targeting *Kcnq1ot1* transcripts with a short interfering RNA (siRNA) near the 3' end of the transcript should result in only amplicons from transcripts extending to the terminal region of the ncRNA exhibiting decreased amplification. However, if a single RNA is produced, it is expected that expression will be depleted along the entire length of *Kcnq1ot1*. To test this, a short interfering RNA (siRNA) was designed to target *Kcnq1ot1* at 460 kb from the transcriptional start site (si460). Mouse embryonic fibroblast cells (NIH 3T3 cells) were transfected with either a non-targeting control siRNA, a siRNA targeting cyclophilin B as a positive control, or the siRNA targeting *Kcnq1ot1* at 460 kb. Total RNA was extracted from the transfected cells, and

Figure 3.5. shRNA and siRNA depletion of *Kcnq1ot1* RNA leads to reduced amplification of the transcript up to 461k from the start site. **A.** Schematic of the *Kcnq1ot1* and *H19* imprinted domains. Red/pink boxes or wavy lines indicate maternal expression, while blue boxes and wavy lines indicate paternal expression. Putative non-imprinted genes are shown in green. ICR; imprinting control region. Closed circles indicate a methylated CpG island, open circles indicate an unmethylated CpG island. Wavy lines indicate non-coding RNAs. Intronic and intergenic regions of analysis are indicated by arrowheads. **B.** XEN cells with a short hairpin RNA targeting the *Kcnq1ot1* transcript (K43). *Kcnq1ot1* was significantly depleted in K43 cells at 2, 65, 202, 392, and 461 kb (designated by grey arrowheads, black text in A), compared to control cells. **C.** In siRNA-transfected NIH 3T3 cells targeting the transcript at 460 kb downstream from the start site (si460), reduced amplification of the transcript was seen versus control NIH 3T3 cells at 2, 94, 118, 392, and 461 kb (designated by grey arrowheads, grey text in A). Error bars indicate the SEM.

real-time qPCR analysis was performed with selected primers along the putative length of the transcript (primers 2k, 94k, 118k, 392k, 461k). Targeting the transcript at 460k resulted in significantly reduced transcription ($p < 0.05$) in NIH 3T3 cells for all primer pairs examined (between 8% to 33% of control expression levels) (Figure 3.5C). As amplicons within the terminal region of the transcript did not exhibit greater reduction in amplification than the more 5' amplicons, it indicates that most *Kcnqlot1* transcripts extend to 460 kb from the start of *Kcnqlot1* transcription.

3.2.3 IC2 deletion analysis

As a final confirmation that the *Kcnqlot1* transcript extends 463 kb from the start site, as well as to determine whether amplicons are specific to a transcript originating from the *Kcnqlot1* promoter, a tissue that has a targeted deletion of the paternal *Kcnqlot1* ICR was analyzed (Figure 3.6A). Real-time qPCR was performed using cDNA from d9.5 B6 X CAST placenta that harbours the paternally-inherited *Kcnqlot1* Δ IC as well as wild-type d9.5 B6 x CAST placenta for primer sets 2k, 65k, 118k, 307k, 392k and 461k (Figure 3.6B). If the *Kcnqlot1* ncRNA is not contiguous, if additional transcripts are present, or if there are alternative internal promoter sites, it is expected that amplification will be detected with a subset of amplicons. However, if all primer sets demonstrate no to very low amplification, it will indicate that *Kcnqlot1* is primarily a single transcript that originated for the ICR. For all primer sets, paternal Δ IC placenta failed to produce amplification product compared to control tissues, verifying that expression originates from the *Kcnqlot1* promoter and that the length of the transcript is at least 461 kb.

Figure 3.6. Deletion of the paternal ICR resulting in loss of *Kcnqlot1* transcription leads to significantly reduced amplification of the transcript up to 461 kb from the start site. **A.** Schematic of the paternal ICR deletion in the *Kcnqlot1* domain. Red/pink boxes and wavy lines indicate maternal expression. Blue boxes and wavy lines indicate paternal expression. Non-imprinted genes are shown in green. ICR; imprinting control region. Closed circles indicate a methylated CpG island, open circles indicate an unmethylated CpG island. Δ IC2 shows the location of the 2.8 kb ICR deletion. *Kcnqlot1* transcript is shown as a dashed line to indicate that it is absent. **B.** Primer sets along the length of the transcript (2, 65, 118, 307, 392, and 461 kb; designated by arrows) were examined using real-time qPCR. All regions examined showed complete loss (>95%) of control *Kcnqlot1* expression indicating that the amplicons belonged to *Kcnqlot1* and originated from the ICR. Error bars indicate the SEM.

A

B

In summary, the *Kcnqlot1* transcript is 463 kb in length, extending the *Kcnqlot1* domain past *Th*, and terminating prior to the *H19* imprinting domain, approximately 155 kb away.

3.3 Imprint status of the tyrosine hydroxylase (*Th*) gene

Previous studies concluded that the boundary of the *Kcnqlot1* domain terminated after the *Ascl2* gene, between *Ascl2* and *Th* (Higashimoto et al., 2006). *Th* was not thought to be an imprinted gene since no parent-of-origin phenotypic effects were observed when *Th* was deleted (Zhou et al. 1995). However, a recent study using microarray technology to examine mice carrying uniparental disomies or duplications to evaluate on a larger scale candidate imprinted genes showed that *Th* is maternally expressed in d13.5 embryos and placentae (Schulz et al. 2006). The evidence that the *Kcnqlot1* transcript extends through *Th* prompted the examination of the imprint status of the *Th* gene.

As neighbouring genes exhibit imprinted gene expression in the placenta (Caspary et al. 1998), it was predicted that *Th* would exhibit imprinted expression in extraembryonic lineages, especially given its role in trophoctoderm development (Zhou et al. 1995). Therefore, RT-PCR was performed to examine allelic expression of *Th* in XEN cells, TS cells, ES cells, d8.5, d9.5 embryos, placentae, and yolk sacs derived from B6 X CAST mice, as well as in CAST7 X B6 d12.5 embryos and placentae. At least two samples of each tissue type at each stage were PCR amplified, and diagnostic restriction digestions performed to discriminate between B6 and CAST alleles for *Th* (Figure 3.7). Densitometry analysis was performed on the resulting images to determine the proportion

Figure 2.7. The relative expression of the *hsp70* gene in the liver of rainbow trout after 24 hours of exposure to 100 µg/l of 17β-estradiol. The results are expressed as the mean ± SEM of three independent experiments. The asterisk (*) indicates a significant difference (p < 0.05) between the control and the treated groups. The data are presented in Table 2.1.

Table 2.1. Relative expression of the *hsp70* gene in the liver of rainbow trout after 24 hours of exposure to 100 µg/l of 17β-estradiol. The results are expressed as the mean ± SEM of three independent experiments. The asterisk (*) indicates a significant difference (p < 0.05) between the control and the treated groups.

Group	Relative expression of <i>hsp70</i> gene
Control	1.0 ± 0.1
Treated (100 µg/l 17β-estradiol)	2.5 ± 0.2*

Figure 2.8. The relative expression of the *hsp70* gene in the liver of rainbow trout after 24 hours of exposure to 100 µg/l of 17β-estradiol. The results are expressed as the mean ± SEM of three independent experiments. The asterisk (*) indicates a significant difference (p < 0.05) between the control and the treated groups. The data are presented in Table 2.2.

Table 2.2. Relative expression of the *hsp70* gene in the liver of rainbow trout after 24 hours of exposure to 100 µg/l of 17β-estradiol. The results are expressed as the mean ± SEM of three independent experiments. The asterisk (*) indicates a significant difference (p < 0.05) between the control and the treated groups.

Group	Relative expression of <i>hsp70</i> gene
Control	1.0 ± 0.1
Treated (100 µg/l 17β-estradiol)	2.5 ± 0.2*

Figure 3.7. *Th* exhibits imprinted expression in embryonic but not extraembryonic tissues. Allele-specific expression analysis for *Th* was performed using PCR amplification followed by allele-specific RE, *Bsr*I. Digestion products were electrophoresed on acrylamide gels. Grey arrows indicate B6 band sizes, black arrow indicates CAST band size. B6 and CAST are from pure neonatal brain controls. EMB: embryo; YS: yolk sac; PLA: placenta; XEN: extraembryonic endoderm stem cells; ES: embryonic stem cells; TS: trophectoderm stem cells; neg: no cDNA control. % mat: percent of expression coming from the maternal allele as determined by densitometry analysis expressed as the mean of all replicates. d8.5 samples, d9.5 samples, XEN, ES and TS cells are genotype B6 x CAST; d12.5 samples are CAST7 x B6. M, 1 kb+ ladder.

of amplicons that were contributed by the maternal or paternal alleles. *Th* displayed monoallelic expression of the maternal allele in placenta and yolk sac samples at d8.5, d9.5, and d12.5 as well as in XEN and TS cells. However, *Th* was biallelically expressed in all embryonic tissues in d8.5, d9.5 and d12.5 samples, as well as in ES cells. These results confirm that *Th* is an imprinted gene. My results indicate that *Th* shares a similar expression pattern to its neighboring genes: *Th* is expressed from the maternal allele in extraembryonic lineages, but is biallelically expressed in embryonic tissues.

3.4 *Kcnq1ot1*-depletion effects on imprinted regulation in the *Kcnq1ot1* domain

3.4.1 Imprint status of genes in the *Kcnq1ot1* domain in stem cells

Previous studies have shown that truncation or deletion of *Kcnq1ot1* results in de-repression of the imprinted genes within the domain, however in these studies there was no delineation between transcription of *Kcnq1ot1* and the transcript itself. I hypothesized that depletion of the transcript, after it is fully transcribed, would lead to de-repression of the domain. In order to address this hypothesis, *Kcnq1ot1* was depleted in stem cells using shRNA techniques. Prior to determining whether the *Kcnq1ot1* ncRNA regulates domain-wide imprinting during early embryogenesis, imprinted expression and methylation must first be examined in wild-type ES, TS and XEN cells.

3.4.1.1 Imprinted methylation at the *Kcnq1ot1* ICR

Imprinted methylation was initially analyzed in stem cells using Combined Bisulfite Mutagenesis and Restriction Analysis (COBRA). Using this procedure, bisulfite mutagenesis was performed using a modified agarose bead method. Following

PCR amplification of converted DNA, the resulting PCR products were cut with *HinfI* to distinguish between methylated and unmethylated DNA present in the pool of cDNA (207 bp: unmethylated allele; 186 and 21 bp: methylated allele). This analysis revealed that all PCR reactions showed both alleles were present, indicating importantly that WT XEN cells, TS cells, ES cells did not spontaneously lose *Kcnq1ot1* imprinted methylation, and that there was no bias in the PCR amplification for one allele over the other (Figure 3.8). Thus, wild-type XEN cells, TS cells, ES cells exhibited the expected methylation pattern with methylated maternal DNA strands and unmethylated paternal stands.

3.4.1.2 Expression of imprinted genes in the *Kcnq1ot1* domain

Imprinted gene expression was also analyzed to determine that WT stem cell maintain *Kcnq1ot1* domain imprinting. For allelic expression analysis, four genes in the *Kcnq1ot1* domain, *Slc22a18*, *Cdkn1c*, *Kcnq1*, and *Ascl2*, were examined in B6 X CAST ES, TS, and XEN cells. SNPs between B6 and CAST were used for allelic discrimination of parental alleles. Restriction enzyme digests that take advantage of polymorphisms that exist between the B6 and CAST alleles were used to determine whether these genes were imprinted and from which allele they were being expressed (maternal or paternal). Greater than or equal to 90% expression from one parental allele was considered monoallelic expression. *Slc22a18* was found to be monoallelic, showing $\geq 90\%$ maternal

Figure 3A. WT as well as transduced stem cells maintain hypomethylated
 A schematic of the Klf4/Sox2 expression construct. The region of interest contains
 20 CpGs (exon 1) and 20 CpGs (exon 2). The CpGs were identified using the following
 criteria: CpG islands (CpG islands > 200 bp) and CpG islands (CpG islands > 200 bp)
 CpGs were present in the stem cell population. ES: embryonic stem cells; TS:
 trophectoderm stem cells; KEM: embryonic endoderm stem cells; WT: wild-type.
 ES cells transduced with a shRNA targeting the histone gene, H3, cells transduced
 with a shRNA targeting the histone gene, H3, cells transduced with a shRNA targeting
 the histone gene, H3, cells transduced with a shRNA targeting the histone gene, H3.

Figure 3.8. WT as well as transduced stem cells maintain imprinted methylation.

A. Schematic of the *Kcnq1ot1* imprinting control region. The region of interest contains 20 CpGs (expanded above as empty circles). **B.** COBRA was performed using the *Hinfl* restriction endonuclease. Both methylated (cut, 186 bp) and unmethylated (uncut, 207 bp) alleles were present in the stem cell populations. ES: embryonic stem cells; TS: trophectoderm stem cells; XEN: extraembryonic endoderm stem cells; WT: wildtype; Luc: cells transduced with a shRNA targeting the luciferase gene; K82: cells transduced with a non-functional shRNA; K43: cells transduced with a shRNA that targets *Kcnq1ot1*; M: methylated; U: unmethylated; neg: no cDNA control.

expression in both wild-type TS and XEN cells, with preferential expression toward the maternal allele observed in ES cells (83% from the maternal allele) (data shown in Figure 3.11). The same pattern of expression for wildtype ES, TS, and XEN cells was observed for *Cdkn1c*, with monoallelic expression from the B6 maternal allele in both TS and XEN cells, and ES cells showed a maternal bias in expression (78% from the maternal allele). ES cells, TS cells, and XEN cells all showed monoallelic expression from the maternal allele of *Kcnq1* ($\geq 90\%$ for all three cell types).

3.4.2 Effects of *Kcnq1ot1* depletion in early embryonic-derived tissues

It was hypothesized that depletion of *Kcnq1ot1* transcripts via RNA interference would lead to a loss of imprint maintenance at the *Kcnq1ot1* ICR, that is, a gain of methylation at the paternal *Kcnq1ot1* ICR and activation of maternally transcribed genes from the paternal chromosome. To test this hypothesis, RNA interference was used to deplete *Kcnq1ot1* RNA in B6 X CAST ES, TS, and XEN cells. Cells with a shRNA targeting *Kcnq1ot1* at 43 kb (K43), were compared to cells harbouring a shRNA that is homologous to *Kcnq1ot1* at 82 kb but that is non-functional (K82), cells transduced with a vector targeting the luciferase gene (Luc), and wildtype cells (WT) as described above. Analysis of *Kcnq1ot1* expression in stem cells demonstrated that the shRNA at 43 kb was effective at depleting *Kcnq1ot1* RNA compared to control cells (between 1 and 7% of control *Kcnq1ot1* expression) (Figure 3.9).

Figure 3. A. Knockdown of XEN cells transfected with a siRNA that targets XEN cells at 48 h. Real-time PCR analysis of XEN cells transfected with siRNA was performed using primers for XEN cells. A significant decrease in XEN cell expression was observed in XEN cells transfected with siRNA (p < 0.05, asterisk). XEN cells transfected with siRNA were compared to XEN cells transfected with a non-functional siRNA (XEN cells transfected with a siRNA that targets XEN cells). Error bars represent SEM.

Figure 3.9. *Kcnqlot1* depletion in stem cells transduced with a shRNA that targets *Kcnqlot1* at 43 kb. Real-time qPCR analysis of wildtype and shRNA transduced stem cells was performed using primers to *Kcnqlot1* at 2 kb. A significant decrease in *Kcnqlot1* amplification was observed in K43 transduced cells versus control cells ($p < 0.05$, asterisk). XEN: extraembryonic endoderm stem cells; ES: embryonic stem cells; TS: trophectoderm stem cells; WT: wildtype; Luc: cells transduced with a shRNA targeting the luciferase gene; K82: cells transduced with a non-functional shRNA; K43: cells transduced with a shRNA that targets *Kcnqlot1*. Error bars represent SEM.

3.4.2.1 Methylation status of the *Kcnq1ot1* ICR in *Kcnq1ot1*-depleted cells

To determine whether *Kcnq1ot1* depletion would result in loss of imprinting within the *Kcnq1ot1* locus, imprinting methylation and expression were assayed. The

initial COBRA analysis did not reveal any in imprinting methylation in K43 stem cells compared to WT (Figure 3.11). However, this analysis examined only one CpG within the *Kcnq1ot1* ICR. To obtain a more depth analysis, the bisulfite methylation and sequencing analysis was performed. This is a less number of CpG dinucleotides (29 CpGs) in the *Kcnq1ot1* ICR to be assayed. This region of analysis was previously

demonstrated to be representative of the ICR (Yatsuki et al., 2002). As there was no

obvious methylation difference between WT, Luc and K82 stem cells using COBRA,

only WT and K82 cells were included in this analysis. For imprinting methylation, results

showed no change in *Kcnq1ot1* ICR methylation patterns in K43 cells compared to

parental cells (WT, Luc, and K82) (Figure 3.12). A slight increase in hypermethylation was

observed in K43 cells compared to WT (Figure 3.12), indicating the depletion

of the *Kcnq1ot1* ncRNA in stem cells was sufficient to change the methylation status

of the *Kcnq1ot1* ICR. It was suggested that the ncRNA is not required in methylation maintenance

in this region of the ICR.

3.4.2.2 Effect of *Kcnq1ot1* depletion on imprinted gene expression

To assess the effects of *Kcnq1ot1* depletion on imprinted gene expression, relative

3.4.2.1 Methylation status of the *Kcnq1ot1* ICR in *Kcnq1ot1*-depleted cells

To determine whether *Kcnq1ot1* depletion would result in loss of imprinting within the *Kcnq1ot1* domain, imprinted methylation and expression were assayed. The initial COBRA analysis did not reveal any change in imprinted methylation in K43 stem cells compared to controls (Figure 3.8). However, this analysis examined only one CpG within the *Kcnq1ot1* ICR. To conduct a more in-depth analysis, the bisulfite mutagenesis and sequencing assay was performed. This allows a larger number of CpG dinucleotides (20 CpGs) in the *Kcnq1ot1* ICR to be assayed. This region of analysis was previously demonstrated to be representative of the ICR (Yatsuki et al., 2002). As there was no obvious methylation difference between WT, Luc and K82 stem cells using COBRA, only WT and K82 cells were included in this analysis. For imprinted methylation, results showed no change in *Kcnq1ot1* ICR methylation patterns in K43 cells compared to control cells (WT, Luc, and K82); maternal DNA strands were hypermethylated while paternal DNA strands remained hypomethylated (Figure 3.10), indicating that depletion of the *Kcnq1ot1* transcript in stem cells was insufficient to change the methylation status of the ICR and suggested that the ncRNA is not required in methylation maintenance at this stage in development.

3.4.2.2 Effects of *Kcnq1ot1*-depletion on imprinted gene expression

To assess the effects of *Kcnq1ot1*-depletion on imprinted gene expression, allelic expression analysis was performed for 4 imprinted genes in *Kcnq1ot1*-deficient K43 cells as well as control (WT, K82 and Luc) ES, TS, and XEN cells (Figure 3.11). As

Figure 3.10. Imprinted methylation was determined by bisulfite sequencing and sequencing assay. Kcnp101 The methylation status of the Kcnp101 ICR in B6 X CAST K43 and control ES, TS, and Kcnp101-depleted stem cells.

Kcnp101 cells was determined by the bisulfite sequencing and sequencing assay. Kcnp101 ICR imprinted methylation was compared in K43 and control stem cells. Filled circles represent methylated CpGs and empty circles represent unmethylated CpGs. Each line denotes an individual CpG strand. Only strands with non-CpG conversion rates of $\geq 85\%$ were included. Conversion rates were indicated above each set of conversion rates were indicated once. The number of hypermethylated DNA strands / total number of DNA strands. Hypermethylated DNA strands displayed $>50\%$ methylated CpGs. WT: wildtype cells; K82: cells targeted with a vector that targets luciferase; K43: cells with a non-functional luciferase; and K43: cells with a hairpin targeting Kcnp101.

Figure 3.10. Imprinted methylation was preserved in *Kcnq1ot1*-depleted stem cells.

The methylation status of the *Kcnq1ot1* ICR in B6 X CAST K43 and control ES, TS, and XEN cells was determined by the bisulfite mutagenesis and sequencing assay. *Kcnq1ot1* ICR imprinted methylation was comparable in K43 and control stem cells. Filled circles represent methylated CpGs and empty circles represent unmethylated CpGs. Each line denotes an individual DNA strand. Only strands with non-CpG conversion rates of $\geq 85\%$ were included. Clones with identical methylation patterns and non-CpG conversion rates were included once. Percent methylation is indicated above each set of DNA strands, and calculated as the number of hypermethylated DNA strands / total number of DNA strands. Hypermethylated DNA strands displayed $>50\%$ methylated CpGs. WT: wildtype cells, Luc: cells transduced with a vector that targets luciferase, K82: cells with a non-functional hairpin, and K43: cells with a hairpin targeting *Kcnq1ot1*.

Figure 3.11. Imprinted gene expression was preserved in *Kcnq1ot1*-depleted stem cells. Imprinting expression analysis of imprinted genes within the *Kcnq1ot1* domain was performed for 4 imprinted genes in K43 and control ES, XEN and TS cells. cDNA from each cell type was PCR amplified with gene-specific primers for *Scl22a18*, *Cdkn1c*, *Kcnq1*, and *Ascl2*, then digested with restriction endonucleases that distinguish maternal and paternal alleles. Cells examined for each stem cell type were wildtype (WT), cells transduced with a hairpin that targets the Luciferase gene (Luc), cells transduced with a non-functional hairpin (K82), and *Kcnq1ot1*-depleted cells transduced with a hairpin targeting *Kcnq1ot1* (K43). Band sizes and parental origin are indicated by red arrows (maternal B6 bands) and blue arrows (paternal CAST bands). Pure B6, pure CAST and no cDNA (neg) samples were amplified as controls to confirm the amplicons were cutting as expected and that there was no contamination. % mat: percent expression from the maternal allele as determined by densitometry analysis

The *Kcnq1* locus is a complex of genes and regulatory elements. The maternal allele (mat) contains the *Kcnq1* gene and several upstream enhancers, including *Osbp15*, *Phlo2*, *Slc22a18*, and *Cdkn1c*. The paternal allele (pat) contains the *Kcnq1* gene and the *Kcnq1ot1* non-coding RNA gene, which is transcribed from the paternal ICR. The ICR (Imprinting Control Region) is located between the *Slc22a18* and *Kcnq1* genes.

bidirectional silencing of imprinted genes occurs within the *Kcnq1ot1* imprinted domain, two genes upstream of the *Kcnq1ot1* ICR (*Slc22a18* and *Cdkn1c*), and two gene with transcription start site downstream of the *Kcnq1ot1* ICR (*Kcnq1* and *Ascl2*) were chosen for analysis. Following PCR amplification, restriction endonuclease digestion was used to distinguish between the B6 and CAST alleles. Similar to the imprinted methylation results, allelic contributions from the maternal and paternal alleles of the imprinted genes analyzed were comparable between K34 stem cells and control. Densitometry analysis revealed no change in expression levels in Luc, K82 or K43 cells for any gene examined when compared to WT cells. Therefore, there was no effect on imprinted gene expression as a result of a $\geq 93\%$ reduction of the *Kcnq1ot1* ncRNA in ES, TS, or XEN cells.

In summary, RNA depletion of *Kcnq1ot1* in embryonic and extraembryonic stem cells did not result in domain-wide loss-of-imprinting.

CHAPTER 4

DISCUSSION

4.1 *Kcnq1ot1* length and *Kcnq1ot1* domain boundaries

Previously, it was thought that the *Kcnq1ot1* noncoding RNA was between 60 kb and up 120 kb in length. In this study, it was found that the *Kcnq1ot1* ncRNA extends to at least 463 kb from the transcriptional start site, exists primarily as a full length transcript, and originates from the promoter sequence within the ICR of the *Kcnq1ot1* domain. The length of *Kcnq1ot1* was conserved in embryonic and placental tissues, as well as embryo-derived stem cells and postnatal brain samples. This is in contrast to the estimated size of 60 kb based on sequence homology between mouse and human, and the published length of 80 kb or 91 kb as was previously found by other labs (Redrup et al., 2009, Pandey et al., 2008). This increased length is significant in that at 463 kb long, it may indicate that transcription through downstream genes leads to their silencing, whereas previously its putative stop site did not extend the transcript past *Kcnql*. shRNA-mediated *Kcnq1ot1* RNA depletion in stem cells had no observable effect on imprinted expression of genes within the domain, or on ICR methylation, which also adds weight to the argument that transcription is likely more important for *Kcnq1ot1* imprinted domain regulation than the transcript itself.

Figure 4.1. Comparison of the *Kcnqlot1* and *H19* domains in mice (A) and humans (B). Maternal alleles are on top (mat) and paternal alleles are on the bottom (pat). Genes are indicated as either biallelically expressed (green), maternally expressed (red), maternally expressed in placenta only (pink), paternally expressed (blue), or silent (grey), with arrows above indicating the direction of transcription. The ICR is indicated by a yellow box, and filled circles above it indicated a methylated ICR, and empty circles indicate an unmethylated ICR. An area of repetitive DNA sequence is shown by an orange square. Putative boundaries of the *Kcnqlot1* and *KCNQ1OT1* domains are denoted by a dashed green line. The *Kcnqlot1* and *H19* domains show a high degree of synteny between mice and humans.

A

B

4.2 *Kcnqlot1* length and its function

Expression patterns of imprinted genes in the *Kcnqlot1* domain are not conserved between embryonic and extraembryonic tissues as the genes closer to the ends of the domain are imprinted only in the placenta, *Osbpl5*, *Tssc4*, *Cd81* and *Ascl2* (Figure 1.4). Using RT-PCR, I observed that the length of the *Kcnqlot1* transcript remains the same in embryonic and extraembryonic lineages, and at all stages examined. In light of the fact that the length of *Kcnqlot1* ncRNA does not change between the embryo and placenta, it is unlikely that the *Kcnqlot1* mRNA length itself is producing any differential silencing ability of the transcript.

There is indication that histone modifications, and not DNA methylation act to regulate expression of imprinted genes within the placenta (Lewis et al., 2006, Umlauf et al., 2004). There may be an important interaction between the *Kcnqlot1* transcript and chromatin in the case of placental-specific imprinted genes, and imprinting of ubiquitously imprinted genes may be regulated in a different fashion through the *Kcnqlot1* RNA. Recently, it was shown that the *Kcnqlot1* RNA is significantly enriched in the chromatin precipitated from placenta as compared to fetal liver (Pandey, et al., 2008). Another group showed that *Kcnqlot1* mRNA were part of a lineage-specific nuclear domain associated with the DNA sequence of the domain. This domain is larger in placenta than in embryonic tissues (Redrup et al., 2009). Presumably, *Kcnqlot1* interacts with histone modifiers within this nuclear domain to regulate transcription based on the evidence that in the placenta, DNA methylation is not required for imprinting gene regulation (Lewis et al., 2006, Umlauf et al., 2004).

Genes that are ubiquitously imprinted within the domain tend to be closer to the ICR and transcriptional start site of *Kcnqlot1* than that of placental-specific imprinted genes. If RNA length is not contributing to these differential expression patterns, another possibility is that tissue-specific boundary elements or insulators not previously identified may be responsible for the differential imprinted expression pattern, and act to promote or restrict the action of *Kcnqlot1*. Novel silencer sequences (specifically that bind to CTCF proteins and act as insulators to inhibit access of enhancers to target sequences) were found within the ICR as well as within the *Cdkn1c* gene, located upstream of the ICR. Binding of CTCF proteins to these silencer sequences within *Cdkn1c* were confirmed *in vitro* (Du et al., 2003; Shin et al., 2008, Fitzpatrick et al, 2007). *Cdkn1c* has a gene-specific CpG island distinct from the ICR that showed differential methylation. However, this only explains *Kcnqlot1*-independent differential regulation of *Cdkn1c* with respect to the other imprinted genes, and is not a mechanism that allows for differential expression between the embryo and the placenta.

4.3 Imprinted status of *Th*

Th is an interesting gene as it is located in a region between the *Kcnqlot1* and *H19* imprinting domains, but was not previously thought to belong to either domain. A previous targeted disruption of *Th* failed to demonstrate any parental-origin effects (Zhou et al., 1995). The caveat, however, is that it was unknown which parental allele was deleted in these mice. While the authors mention that heterozygous mice showed no phenotypic effects, there was no comprehensive examination of size or morphology and the stage at which they were deemed normal is unknown.

In another study, using microarray technology to identify imprinted genes on a large scale, it was found that *Th* is likely maternally expressed (Schulz et al., 2006). My findings show that *Kcnq1ot1* is transcribed past *Th* (~400 kb from the transcriptional start site), extending the domain's boundary to include the *Th* gene. As *Kcnq1ot1* was transcribed past *Th*, it may suggest that *Th* is imprinted. Investigation into the imprint status of *Th* revealed that it is imprinted in extraembryonic tissues at peri-implantation as well as postimplantation stages. In light of extraembryonic tissue-specific imprinting and the low viability of the heterozygous mice, further studies are warranted to determine the significance of *Th* imprinting.

These results suggest that *Th* is regulated in a similar fashion to other placental-specific imprinted genes in the *Kcnq1ot1* domain. Previous experiments attempted to visualize the size of the nuclear domain organized by *Kcnq1ot1* in embryos and placentae using RNA/DNA FISH techniques and found that the domain is larger in the placenta. However, DNA probes no farther than *Ascl2* were used, which is only ~330 kb from the start site. As such, there was no examination of whether *Kcnq1ot1* is localized to the DNA as far as 400 kb from the ICR. However, examination of the imprint status of *Th* in tissue with loss of imprinting due to deletion of the ICR or truncation of the ncRNA would be necessary to determine for certain whether *Th* is under the control of the *Kcnq1ot1* domain. My experiments demonstrate that *Th* is maternally expressed in extraembryonic tissue from peri-implantation to d12.5 of development.

4.4 Regulatory elements responsible for *Kcnqlot1* domain regulation

It is unknown whether silencing in the *Kcnqlot1* domain is due to the ICR, transcription of *Kcnqlot1*, the *Kcnqlot1* transcript *per se*, or a combination of these mechanisms. Several studies have attempted to differentiate between these possibilities using mutation analysis. Deletion studies of the *Kcnqlot1* domain indicate that the ICR is required, a functional promoter is required, transcription must be initiated, and a transcript at least longer than the truncations used in experiments must be present for proper imprinting of the domain (Fitzpatrick et al., 2002; Mancini-Dinardo et al., 2006). However, the authors were unable to tease apart the putative regulatory elements. For example, deletion of the ICR also deleted the promoter and the transcript. Also, truncation of the transcript may not separate the function of the ICR from the act of transcription initiation, or transcription through important regulatory regions. For these reasons, a method that leaves the ICR and transcription start site unaltered would be more informative.

In this study, an RNAi-based approach was used to degrade the *Kcnqlot1* transcript without deletion of the promoter or ICR, and leaving transcription unaffected. This resulted in greater than 90% depletion of the *Kcnqlot1* transcript within stem cells but was unable to measurably disrupt imprinted expression of genes within the domain, or affect the imprinted methylation status of the ICR. These results suggest that the ICR may be the key element responsible for silencing of the domain, or that the act of transcription sufficiently maintains imprinting. However, there is the possibility that no matter how efficient shRNA-targeted degradation was, some *Kcnqlot1* transcripts remained and that this significantly reduced level was sufficient to maintain imprinting

within the domain in stem cells. The possibility also exists that the imprinted expression of the domain is established and locked in by the *Kcnqlot1* transcript before the stage examined in these studies (peri-implantation stem cells). This would argue for examination of the role of *Kcnqlot1* using RNAi-based methods earlier in development, optimally at the 2-cell stage when the RNA is first being transcribed. Nevertheless, it is difficult to understand why an RNA would continue to be transcribed well into postnatal development, if it is no longer required, especially considering that it would be an energetically expensive process to produce mRNA that is 463 kb long.

4.5 Comparison of the *Kcnqlot1* domain between mice and humans

There is a high degree of synteny between mouse and human for the *Kcnqlot1* domain (Figure 4.2) (Taniguchi et al., 1997; Engemann et al., 2000; Astuti et al., 2005). In both the human and the mouse, the domain is ~1 Mb in size. In humans, the ICR is methylated on the maternal allele and *KCNQ1OT1* is preferentially expressed from the paternal allele, as is the case in mice (Mitsuya et al., 1999). It was found that 50% of the patients examined with BWS showed expression of *KCNQ1OT1* from both the maternal and paternal alleles, and in an independent study, it was shown that 58% of patients exhibit loss of maternal methylation at the ICR, suggesting that aberrant antisense regulation can result in the development of this human disease (Higashimoto et al., 2003; Lee et al., 1999). Thus, similar regulatory mechanisms are thought to act in the human as in the mouse.

There are some notable differences in the domains in humans and mice. *KCNQ1OT1* is estimated to be ~60 kb in length as it had been for mouse, but this has not

been confirmed. In mouse, the region between *Th* and *Ins2* is over 200 kb, and is highly repetitive, containing a high concentration of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). In the human, the distance between *TH* and *INS* is considerably smaller (~10 kb). However, a similar area of repetitive DNA, rich in LINEs and SINEs is found between *ASCL2* and *TH* (~115 kb region); this region in the mouse is ~60 kb. It is unknown whether this particular repetitive region is functionally significant to the domain, but the end of the transcript in mice is now known to be embedded within it. This may indicate that the boundary between the *KCNQ1OT1* domain and the *H19* domain in humans may not include *TH*, but may be located within the repetitive region between *ASCL2* and *TH*. However, the boundaries of the *KCNQ1OT1* domain are unknown and therefore may not include *ASCL2* or this repetitive region. There is evidence that *ASCL2* is imprinted as hydatidiform (androgenetic) moles show no expression of *ASCL2*, suggesting that it is maternally expressed and therefore imprinted (Alders et al., 1997), but more recent evidence suggests that it escapes genomic imprinting (Miyamoto et al., 2002). This would place the imprinting boundary just after *CD81* in humans.

4.6 Why is *Kcnqlot1* so long?

Precedent exists for very long ncRNAs. For example, at the extreme *Ube3aas* extends more than 1000 kb in length. However, using the length of a noncoding RNA to infer function is not possible. For example, it is hypothesized that *Kcnqlot1* acts in a manner similar to that of *Xist*, a ncRNA involved in X chromosome inactivation (Pauler et al., 2007). While there is suggestive evidence of *Kcnqlot1* RNA localization to DNA

in the *Kcnqlot1* domain, *Xist* is only 15 kb long, and is more than sufficient to coat and effectively silence an entire chromosome (Brockdorff et al., 1992). Therefore, it is unlikely that a very long transcript is required to silence the 1 Mb *Kcnqlot1* domain in this manner.

The act of transcription through promoters of imprinted genes may act to inhibit access of polymerases to such regions and prevent their transcription. In this sense, a longer noncoding RNA would be able to silence more genes as it is transcribed through them. Silencing by *Kcnqlot1* and *Airn* is bi-directional which could not be solely accounted for by this model, and paternal expression of non-imprinted genes within the region is seen even when *Kcnqlot1* is transcribed through them. However, there could be more than one mechanism accounting for silencing in this domain, and it is likely that at least for downstream genes, transcription elongation is required for their silencing.

Lastly, it is possible that transcription initiation is the most important factor in silencing of the domain and that the transcript itself is dispensable. Truncation of *Airn* results in de-repression of all three normally silenced genes in the region (Schmidt et al., 1999). A recent study further enforces the idea that transcription of the ncRNA is essential for imprinted domain regulation as truncation of *Nespas* in the *Gnas* domain resulted in loss of methylation of maternal germ line DMRs (Chotalia et al., 2009). Although truncation experiments of *Kcnqlot1* showed that inserting premature polyadenylation signals in the sequence at 1.5 kb resulted in de-repression of imprinted genes within the region suggesting that the transcript itself was necessary for this function, truncation to 1.5 kb from 463 kb is severe and it is possible that allowing the transcript to terminate further downstream may have lead to no change in imprinted expression in the

domain. There is evidence for this in an episomal study that truncated the transcript to various lengths between 4.9 and 9.2 kb (Kanduri et al., 2006). Results showed that the longer the truncated transcript, the more effective silencing ability it had on a downstream gene. While this does not represent an *in vivo* setting, it does suggest that longer truncation studies in mouse models are required for a clearer understanding of the role that the transcript plays.

4.7 Conclusion

My study supports the theory that transcription itself is important for regulation of imprinting in the domain, as transcription was left completely intact in shRNA-mediated degradation of more than 90% of the endogenous *Kcnqlot1* transcript, and no change was seen in imprinting or methylation status within the domain. However, it is likely that a combination of *Kcnqlot1* elements leads to establishment and maintenance of imprinted regulation of the *Kcnqlot1* domain.

REFERENCES

- Alders M, Hodges M, Hadjantonakis AK, Postmus J, van Wijk I, Blik J, de Meulemeester M, Westerveld A, Guillemot F, Oudejans C, Little P, Mannens M: The human Achaete-Scute homologue 2 (ASCL2, HASH2) maps to chromosome 11p15.5, close to IGF2 and is expressed in extravillous trophoblasts. *Hum. Molec. Genet.* 6:859-867. 1997
- Amaral PP, Mattick JS: Noncoding RNA in development. *Mamm. Genome.* 19(7-8):454-92. 2008
- Antequera F, Bird A: Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci USA* 90(24): 11995-11999. 1993
- Astuti D, Latif F, Wagner K, et al.: Epigenetic alteration at the DLK1-GTL2 imprinted domain in human neoplasia: analysis of neuroblastoma, pheochromocytoma and Wilms' tumour. *Br J Cancer.* 92(8):1574-80. 2005
- Barlow DP, Stöger R, Herrmann BG, Saito K, Schweifer N: The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *i.* 349:84-87. 1991
- Barlow DP: Competition--a common motif for the imprinting mechanism? *EMBO J.* 16(23):6899-905. 1997
- Bartolomei MS, Tilghman SM: Genomic imprinting in mammals. *Annu Rev Genet.* 31:493-525, 1997
- Bartolomei MS, Zemel S, Tilghman SM: Parental imprinting of the mouse H19 gene. *Nature* 351(6322):153-5. 1991
- Barton SC, Surani MA, Norris ML: Role of paternal and maternal genomes in mouse development. *Nature* 405:482-485. 1984
- Bestor TH, Chandler VL, Feinberg AP: Epigenetic effects in eukaryotic gene expression. *Dev. Genet.* 15:458-462. 1994
- Bielinska B, Blaydes SM, Buiting K, Yang T, Krajewska-Walasek M, Horsthemke B, Brannan CI: De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. *Nat Genet.* 25:74-78, 2000
- Brandeis M, Kafri T, Ariel M, Chaillet JR, McCarrey J, Razin A, Cedar H: The ontology of allele-specific methylation associated with imprinted genes in the mouse. *EMBO J* 12:3669-3677. 1993

- Brannan CI, Dees EC, Ingram RS, Tilghman SM: The product of the H19 gene may function as an RNA. *Mol. Cell Biol.* 10:28-36. 1990
- Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S: The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell.* 71(3):515-26. 1992
- Casparly T, Cleary MA, Baker CC, Guan XJ, Tilghman SM: Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster. *Mol Cell Biol.* 18:3466-3474. 1998
- Davies K, Bowden L, Smith P, Dean W, Hill D, Furuumi H, Sasaki H, Cattanach B, Reik W: Disruption of mesodermal enhancers for Igf2 in the minute mutant. *Development.* 129(7):1657-68. 2002
- DeChiara TM, Efstratiadis A, Robertson EJ: A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature.* 345(6270):78-80. 1990
- Diaz-Meyer N, Day CD, Khatod K, Maher ER, Cooper W, Reik W, Junien C, Graham G, Algar E, Der Kaloustian VM, Higgins MJ: Silencing of CDKN1C (p57KIP2) is associated with hypomethylation at KvDMR1 in Beckwith-Wiedemann syndrome. *J Med. Genet.* 40(11):797-801. 2003
- Dooley JM, Berg JM, Pakula Z, MacGregor DL: The puppet-like syndrome of Angelman. *Am J Dis Child* 135(7): 621-4. 1981
- Du M, Beatty LG, Zhou W, Lew J, Schoenherr C, Weksberg R, Sadowski PD: Insulator and silencer sequences in the imprinted region of human chromosome 11p15.5. *Hum. Mol. Genet.* 12(15):1927-39. 2003
- Engemann S, Strodicke M, Paulsen M, Franck O, Reinhardt R, Lane N, Reik W, Walter J: Sequence and functional comparison in the beckwith-wiedemann region: Implications for a novel imprinting centre and extended imprinting. *Hum Mol Genet.* 9:2691-2706, 2000
- Ferguson-Smith AC, Cattanach BM, Barton SC, Beechey CV, Surani MA: Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* 20;351(6328):667-70. 1991
- Fitzpatrick GV, Pugacheva EM, Shin JY, Abdullaev Z, Yang Y, Khatod K, Lobanenkov VV, Higgins MJ: Allele-specific binding of CTCF to the multipartite imprinting control region KvDMR1. *Mol. Cell. Biol.* 27(7):2636-47. 2007
- Fitzpatrick GV, Soloway PD, Higgins MJ: Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat Genet.* 32:426-431, 2002

Geuns E, Hilven P, Van Steirteghem A, Liebaers I, De Rycke M: Methylation analysis of KvDMR1 in human oocytes. *J Med Genet.* 44:144-147, 2007

Green K, Lewis A, Dawson C, Dean W, Reinhart B, Chaillet JR, Reik W: A developmental window of opportunity for imprinted gene silencing mediated by DNA methylation and the *Kcnqlot1* noncoding RNA. *Mamm. Genome.* 18(1):32-42. 2007

Goldberg A.D, Allis CD, Bernstein, E: Epigenetics: A landscape takes shape. *Cell.* 128: 635-638. 2007

Guillemot F, Caspary T, Tilghman SM, Copeland NG, Gilbert DJ, Jenkins NA, Anderson DJ, Joyner AL, Rossant J, Nagy A: Genomic imprinting of *Mash2*, a mouse gene required for trophoblast development. *Nat Genet.* 9:235-242. 1995

Hendrich BD, Willard HF: Epigenetic regulation of gene expression: the effect of altered chromatin structure from yeast to mammals. *Hum. Mol. Genet.* 4:1765-1777. 1995

Higashimoto K, Urano T, Sugiura K, Yatsuki H, Joh K, Zhao W, Iwakawa M, Ohashi H, Oshimura M, Niikawa N, Mukai T, Soejima H: Loss of CpG methylation is strongly correlated with loss of histone H3 lysine 9 methylation at DMR-LIT1 in patients with Beckwith-Wiedemann syndrome. *Am. J Hum. Genet.* 73(4):948-56. 2003

Higashimoto K, Soejima H, Saito T, Okumura K, Mukai T. Imprinting disruption of the *CDKN1C/KCNQ1OT1* domain: the molecular mechanisms causing Beckwith-Wiedemann syndrome and cancer. *Genome Res.* 113: 1345-1352. 2006

Holliday R: Epigenetics: An overview. *Dev. Genet.* 15:453-457. 1994

Kanduri C, Thakur N, Pandey RR: The length of the transcript encoded from the *Kcnqlot1* antisense promoter determines the degree of silencing. *EMBO J.* 25:2096-2106, 2006

Kaneda M, Sado T, Hata K, Okano M, Tsujimoto N, Li E, Sasaki H: Role of de novo DNA methyltransferases in initiation of genomic imprinting and X-chromosome inactivation. *Cold Spring Harb Symp Quant Biol.* 69:125-9. 2004

Kapranov P, Willingham AT, Gingeras TR: Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* 8(6):413-23. 2007

Kho MR, Baker DJ, Layoon A, Smith SS: Stalling of Human DNA (Cytosine-5) Methyltransferase at Single Strand Conformers form a Site of Dynamic Mutation. *Journal of Molecular Biology* 275: 67-79. 1998

Killeen AA: Genetic Inheritance. *Principles of Molecular Pathology.* Humana Press. pp. 41. 2004

- Kono T, Obata Y, Wu Q, Niwa K, Ono Y, Yamamoto Y, Park ES, Seo JS, Ogawa H: Birth of parthenogenetic mice that can develop to adulthood. *Nature*. 428(6985):809-11. 2004
- Landers M, Bansescu DL, Le Meur E, Rougeulle C, Glatt-Deeley H, Brannan C, Muscatelli F, Lalande M: Regulation of the large (~1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. *Nucleic Acids Res*. 32:3480-3492. 2004
- Lee M P, DeBaun, M R, Mitsuya K, Galonek H L, Brandenburg S, Oshimura M, Feinberg A P: Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc. Natl. Acad. Sci. USA* 96(9):5203-8. 1999
- Lee MH, Reynisdóttir I, Massagué J: Cloning of p57KIP2, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution. *Genes. Dev*. 9(6):639-49. 1995
- Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM: Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature*. 375:34-39, 1995
- Lewis A, and Reik W: How imprinting centres work. *Cytogenet. Genome Res*. 113:81-89
- Lewis A, Green K, Dawson C, Redrup L, Huynh KD, Lee JT, Hemberger M, Reik W: Epigenetic dynamics of the Kcnq1 imprinted domain in the early embryo. *Development*. 133:4203-4210, 2006
- Lewis A, Mitsuya K, Umlauf D, Smilth P, Dean W, Walter J, Higgins M, Feil R, Reik W. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nature Genetics*. 36:1291-1295, 2004
- Li, E., T.H. Bestor, and R. Jaenisch: Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*. 69:915-926. 1992
- Lin SP, Youngson N, Takada S, Seitz H, Reik W, Paulsen M, Cavaille J, Ferguson-Smith AC: 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, 2003
- Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc. Nat. Acad. Sci*. 96: 5203-5208. 1999
- Lyle R., Watanabe D, te Vrugte D, Lerchner W, Smrzka OW, Wutz A, Schageman J, Hahner L, Davies C, Barlow DP: The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat. Genet*. 25:19-21. 2000

- Maher ER, Reik W: Beckwith-wiedemann syndrome: Imprinting in clusters revisited. *J Clin Invest.* 105:247-252, 2000
- Mancini-DiNardo D, Steele SJ, Ingram RS, Tilghman SM: A differentially methylated region within the gene *Kcnq1* functions as an imprinted promoter and silencer. *Hum Mol Genet.* 12:283-294, 2003
- Mancini-Dinardo D, Steele SJ, Levorse JM, Ingram RS, Tilghman SM: Elongation of the *Kcnqlot1* transcript is required for genomic imprinting of neighboring genes. *Genes Dev.* 20:1268-1282, 2006
- Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R: Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* 11(2):156-66. 1997
- McGrath J, Solter D: Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell.* 37:179-183, 1984
- Mitsuya K, Meguro M, Lee MP, Katho M, Schultz TC, Kugoh H, Yoshida MA, Niikawa N, Feinberg AP, Oshimura M: LIT1, an imprinted antisens RNA in the human Kv1QT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. *Hum. Mol. Genet.* 8:1209-1217. 1999
- Miyamoto T, Hasuike S, Jinno Y, Soejima H, Yun K, Miura K, Ishikawa M, Niikawa N: The human ASCL2 gene escaping genomic imprinting and its expression pattern. *J Assist Reprod Genet.* 19(5):240-4. 2002
- Miyamoto T, Hasuike S, Jinno Y, Soejima H, Yun K, Miura K, Ishikawa M, Niikawa N: The human ASCL2 gene escaping genomic imprinting and its expression pattern. *J Assist Reprod Genet.* 19:240-244, 2002
- Morgan HD, Santos F, Green K, Dean W, Reik W: Epigenetic reprogramming in mammals. *Hum Mol Genet.* 14 Spec No 1:R47-58. 2005
- Murakami K, Oshimura M, Kugoh H: Suggestive evidence for chromosomal localization of non-coding RNA from imprinted LIT1. *J Hum Genet.* 52:926-933, 2007
- Nicholls RD, Knepper JL: Genome organization, function, and imprinting in prader-willi and angelman syndromes. *Annu Rev Genomics Hum Genet.* 2:153-175, 2001
- Oda M, Okano M: DNA methylation during mouse early embryogenesis. *Tanpakushitsu Kakusan Koso.* 51(14 Suppl):2027-34. 2006
- Okano M, Takebayashi S, Okumura K, Li E: Assignment of cytosine-5 DNA methyltransferases Dnmt3a and Dnmt3b to mouse chromosome bands 12A2-A3 and 2H1 by in situ hybridization. *Cytogenet. Cell Genet.* 86(3-4): 333-4. 1999

- Olek, A., Oswald, J., and Walter, J. (1996) A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res.*, **24**, 5064-5066.
- O'Neill MJ: The influence of non-coding RNAs on allele-specific gene expression in mammals. *Hum Mol Genet.* 14 Spec No 1:R113-20, 2005
- Onyango P, Miller W, Lehoczky J, Leung CT, Birren B, Wheelan S, Dewar K, Feinberg AP: Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. *Genome Res.* 10:1697-1710, 2000
- Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C: *Kcnqlot1* Antisense Noncoding RNA Mediates Lineage-Specific Transcriptional Silencing through Chromatin-Level Regulation. *Cell.* 32: 232-246, 2008
- Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C: *Kcnqlot1* antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Dev. Cell.* 15(5):637-8. 2008
- Pauler FM, Koerner MV, Barlow DP: Silencing by imprinting noncoding RNAs: is transcription the answer? *Trends in Genet.* 23(6): 284-292. 2007
- Redrup L, Branco MR, Perdeaux ER, Krueger C, Lewis A, Santos F, Nagano T, Cobb BS, Fraser P, Reik W: The long noncoding RNA *Kcnqlot1* organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development.* 136(4):525-30. 2009
- Regha K, Latos PA, Spahn L: The imprinted mouse *Igf2r/Air* cluster--a model maternal imprinting system. *Cytogenet. Genome. Res.* 113(1-4):165-77. 2006
- Reik W, Walter J: Genomic imprinting: parental influence on the genome. *Nat Rev Genet.* 2(1):21-32. 2001
- Reinhart B, Chaillet JR: Genomic imprinting: cis-acting sequences and regional control. *Int. Rev. Cytol.* 243:173-213. 2005
- Rouguelle C, Heard E: Antisense RNA in imprinting: spreading silence through Air. *Trends Genet.* 18:434-437. 2002
- Sasaki H, Ferguson-Smith AC, Shum AS, Barton SC, Surani MA: Temporal and spatial regulation of H19 imprinting in normal and uniparental mouse embryos. *Development.* 121(12): 4195-4202. 1995

Saxonov S, Berg P, Brutlag DL: A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci USA* 103 (5): 1412–1417. 2006

Schoenherr, C.J., Levorse, J.M., and Tilghman, S.M. (2003) CTCF maintains differential methylation at the Igf2/H19 locus. *Nature Genet.*, **33**, 66-69.

Schmidt JV, Levorse JM, Tilghman SM: Enhancer competition between H19 and Igf2 does not mediate their imprinting. *Proc. Natl. Acad. Sci. USA*. 96(17):9733-8. 1999

Schultz R, Menhenniott TR, Woodfine K, Wood AJ, Choi JD, Oakey RJ. Chromosome-wide identification of novel imprinted genes using microarrays and uniparental disomies. *Nucleic Acids Research*. 34:e88, 2006

Seidl CI, Stricker SH, Barlow DP: The imprinted air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J*. 25:3565-3575, 2006

Shin JY, Fitzpatrick GV, Higgins MJ: Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *EMBO J*. 27(1):168-78. 2008

Sleutels F, Zwart R, Barlow DP: The non-coding air RNA is required for silencing autosomal imprinted genes. *Nature*. 415:810-813, 2002

Smilnich NJ, Day CD, Fitzpatrick GV, Caldwell GM, Lossie AC, Cooper PR, Smallwood AC, Joyce JA, Schofield PN, Reik W, Nicholls RD, Weksberg R, Driscoll DJ, Maher ER, Shows TB, Higgins MJ: 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, 1999

Smith SS, Kaplan BE, Sowers LC, Newman EM: Mechanism of Human Methyl-directed DNA methyltransferase and the Fidelity of Cytosine Methylation. *Proceedings of the National Academy of Science, U.S.A.* 89: 4748-4744. 1992

Spahn L, Barlow DP: An ICE pattern crystallizes. *Nat Genet*. 35:11-12, 2003

Stoger R, Kubicka P, Liu CG, Kafri T, Razin A, Cedar H, Barlow DP: Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. *Cell*. 73:61-71, 1993

Strahl BD, Allis CD: The language of covalent histone modifications. *Nature* 403(6765): 41–5. 2000

Struhl K: Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* 14(2):103-5. 2007

Surani MA, Barton SC, Norris ML: Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* 308(5959): 548-50. 1984

Tanaka K, Shiota G, Meguro M, Mitsuya K, Oshimura M, Kawasaki H: Loss of imprinting of long QT intronic transcript 1 in colorectal cancer. *Oncology*. 60(3):268-73. 2001

Taniguchi T, Okamoto K, Reeve AE: Human p57(KIP2) defines a new imprinted domain on chromosome 11p but is not a tumour suppressor gene in Wilms tumour. *Oncogene* 14(10):1201-6. 1997

Terranova R, Yokobayashi S, Stadler M, Otte A, van Lohuizen M, Orkin S, Peters AHFM: Polycomb Group Proteins Ezh2 and Rnf2 Direct Genomic Contraction and Imprinted Repression in Early Mouse Embryos. *Developmental Cell*. 15:1-12, 2008

Thorvaldsen JL, Duran KL, Bartolomei MS: Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes Dev*. 12:3693-3702, 1998

Thorvaldsen JL, Verona RI, Bartolomei MS: X-tra! X-tra! News from the mouse X chromosome. *Dev. Biol.* 298(2):344-53. 2006

Tierling S, Dalbert S, Schoppenhorst S, Tsai CE, Oliger S, Ferguson-Smith AC, Paulsen M, Walter J: High-resolution map and imprinting analysis of the Gtl2-Dnchc1 domain on the mouse chromosome 12. *Genomics* 87:225-235. 2005

Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, Feil R: Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of polycomb group complexes. *Nat Genet*. 36:1296-1300, 2004

Verona RI, Mann MR, Bartolomei MS: Genomic imprinting: Intricacies of epigenetic regulation in clusters. *Annu Rev Cell Dev Biol*. 19:237-259, 2003

Wang ZQ, Fung MR, Barlow DP, Wagner EF: Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature*. 372(6505):464-7. 1994

Watanabe H, Pan ZQ, Schreiber-Agus N, DePinho RA, Hurwitz J, Xiong Y: Suppression of cell transformation by the cyclin-dependent kinase inhibitor p57KIP2 requires binding to proliferating cell nuclear antigen. *Proc. Natl. Acad. Sci. USA*. 95(4):1392-7. 1998

Weksberg R, Smith AC, Squire J, Sadowski P: Beckwith-wiedemann syndrome demonstrates a role for epigenetic control of normal development. *Hum Mol Genet*. 12 Spec No 1:R61-8, 2003

Wroe SF, Kelsey G, Skinner JA, Bodle D, Ball ST, Beechey CV, Peters J, Williamson CM: An imprinted transcript, antisense to Nesp, adds complexity to the cluster of

imprinted genes at the mouse *Gnas* locus. *Proc. Natl. Acad. Sci. USA* 97:3342-3346. 2000

Wutz A, Rasmussen TP, Jaenisch R: Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* 30(2):167-74. 2002

Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP: Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature*. 389:745-749, 1997

Yatsuki H, Joh K, Higashimoto K, Soegima H, Arai Y, Wang Y, Hatada I, Obata Y, Morisaki T. Domain regulation of imprinting cluster in Kip2/Lit1 subdomain of mouse chromosome 7F4/F5: large-scale DNA methylation analysis reveals that DMR-Lit1 is a putative imprinting control region. *Genome Res.* 12:1860-1870. 2002

Zhou QY, Quaife CJ, Palmiter RD: Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development. *Nature*. 374(6523):640-3. 1995

AWARDS

SUMMARY OF SCIENTIFIC CAREER

- Fellow, National Academy of Sciences (2002)
- Member of the National Institute of Health (NIH) (2000-2002)
- Member of the National Institute of Mental Health (NIMH) (2000-2002)
- Member of the National Institute of Child Health and Human Development (NICHD) (2000-2002)
- Member of the National Institute of Environmental Health Sciences (NIEHS) (2000-2002)
- Member of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (2000-2002)
- Member of the National Institute of General Medical Sciences (NIGMS) (2000-2002)
- Member of the National Institute of Neurological Disorders and Stroke (NINDS) (2000-2002)
- Member of the National Institute of Allergy and Infectious Diseases (NIAID) (2000-2002)
- Member of the National Institute of Cancer Research (NICR) (2000-2002)