CREATING A 350 KILOBASE DELETION WITHIN THE IMPRINTED REGION OF MOUSE DISTAL CHROMOSOME 7

by

Li Ma

A thesis submitted in conformity with the requirements for the degree of Masters
Graduate Department of Zoology
University of Toronto

© Copyright by Lynn Mar (1999)
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-45420-7
ABSTRACT

CREATING A 350 KILOBASE DELETION WITHIN THE IMPRINTED REGION OF MOUSE DISTAL CHROMOSOME 7

by Li Ma

A thesis submitted in conformity with the requirements for the degree of Masters in the Zoology Graduate Department at the University of Toronto, 1999

Genomic imprinting is a fascinating biological phenomenon that does not appear to obey the genetic principle of equivalence. Imprinting is essential for eutherian development and has been implicated in disease processes. This thesis describes an approach to understand genomic imprinting and its mechanisms. The imprinted region under investigation is mouse distal chromosome 7 which contains many maternally and paternally imprinted genes that are essential for development.

The experiment applies recent mouse molecular technology by creating precise deletions and duplications in embryonic stem cells. Using a combination of gene targeting and the Cre-loxP system, I have created a 350 kb deletion between Mash2 and Ins2. By studying the phenotypic effects of paternal and maternal inheritance of these rearrangements in mice, one may identify novel genes and cis-acting regulatory elements. Furthermore, mice and cell lines derived from this experiment can be used for future analysis of this imprinted region.
ACKNOWLEDGMENT

Without the encouragement and support of my family, dad, mom and Gip, the completion of this thesis would never have been possible. I want to thank you from the bottom of my heart.

I also want to acknowledge Dave and Benji for their love and affection. In particular, Dave, you have been instrumental in my growth and independence. Your subtle gestures often go unappreciated, but always needed.

I want to thank Andras Nagy, my supervisor, for taking me under his wings. You inspire much creativity and originality. Thank you too, Sue Varmuza, for taking the time out to help me.

I'm also grateful to my fellow lab mates who have been helpful and excellent teachers, Kat, Melinda, Lucile, Mika, Louis, Hilda, Karen and many others. I thank Marina, Sue and Malgosia for helping me with my experiments. A special thanks go to Elizabeth Mulroney.

Finally, I couldn't have do this without my friends, Sonia, Dave, Kat, Melinda, Celia, Virginia, Hong, Mara, Renu and Andrew. I want to also thank Brian Beckett for helping me along the way.
OUTLINE OF THESIS

Preliminary Section

Title Page i
Abstract ii
Acknowledgements iii
Outline of Thesis iv-v
List of Abbreviations vi-vii
List of Figures and a Table viii

Chapter 1 Literature Review 1-38

1.1 Genomic Imprinting 1-20
1.1.1 Introduction 1
1.1.2 Classical Experiments 1-4
1.1.3 Distal Chromosome 7 4-9
1.1.3.1 Unidentified Genes in Distal Chromosome 7 9-11
1.1.4 Mechanism of Imprinting 11-18
1.1.4.1 The Primary Imprint 11-13
1.1.4.2 Imprint Propagation 13-14
1.1.4.3 Chromatin Structure 14-15
1.1.4.4 Chromosomal Domains 15-16
1.1.4.5 Mechanisms within the Domains 16-18
1.1.4.5.1 Competition 16-17
1.1.4.5.2 Other Proposed Mechanisms 17-18
1.1.5 Function and Evolution of Genomic Imprinting 19-20
1.2 Mouse Genome Engineering 21-38
1.2.1 Introduction 20-21
1.2.2 Cre-loxP 21-25
1.2.2.1 Cre-mediated Chromosomal Rearrangements 25-37
1.3 Purpose of Thesis 37-38

Chapter 2 Materials and Methods 39-59

2.1 Materials 39-42
2.1.1 Tissue Culture 39
2.1.2 Hybridization Probes 39-41
2.1.3 Primers 41
2.1.4 Animals 42
2.2 Methods 42-
2.2.1 Pulsed Field Gel Electrophoreis 42-45
2.2.1.1 Preparation of High-Molecular-Weight Genomic DNA 43
2.2.1.2 Restriction-Enzyme Digestion 43
Pulsed-Field Gel Electrophoresis 43-44
Southern Blotting 44
Hybridization 44
DNA Size Standards 44-45
Construction of Targeting vectors 45-52
PGK-neo Cassettes 45-47
Mash2 Target Vector 48-50
Ins2 Target Vector 48-52
Embryonic Stem Cell Culture 53
Maintenance of R1 Cells 53
Transfection and Selection for Mutant ES Cells 53
ES Cell Screening 54
Southern Blot and Hybridization 54-58
Polymerase Chain Reaction 58
Generation of Mutant Chimera 58
Genotyping of Mice 58-9

Chapter 3 Results: Deletion of 350 kb Within Distal Chromosome 7 in ES Cells 60-82

The Physical Distance between Mash2 and Ins2 is 350 Kilobases 60-63
Deletion and Duplication Strategy 64-69
Construction of Targeting Vectors 69-70
Identification of ES Cells Correctly Targeted at the Mash2 Locus by Homologous Recombination 70-72
Identification of ES Cells Correctly Cre Excised at the Mash2 Locus 70-75
Identification of ES Cells Correctly Targeted at the Ins2 Locus by Homologous Recombination 73-77
Identification of ES Cells Correctly Cre Excised at the Ins2 Locus 78-80
Identification of ES Cells Correctly Deleted for 350 Kilobases 78-82
Chimeras and Mutant Mice 78-82

Chapter 4 Discussion 83-90
Physical Mapping Around Mash2 and Ins2 83
Cre/loxP Efficiency and Implications 84-88
Chromosomal Targeting and Implications 88-89
Mash2 and Ins2 Chimeras and Mutants 89
350 kb Deletion Chimeras 89
Future Prospects 89-90

References 91-103
LIST OF FIGURES

Figure 1.1  Mouse Imprinting Map.
Figure 1.2  The Physical Map of Mouse Distal Chromosome 7.
Figure 1.3  Cre-loxP Site-specific Recombination.
Figure 1.4  Strategy for Cre Mediated Translocation.
Figure 1.5  Strategy for Cre Mediated Translocation.
Figure 1.6  Strategy for Cre Mediated Deletions, Duplications, Inversions and other Rearrangements.
Figure 2.1  Construction of PGK-neo with embedded loxP cassettes.
Figure 2.2  Construction Mash2 Target Vector.
Figure 2.3  Construction Ins2 Target Vector.
Figure 2.4  Mash2.
Figure 2.5  Ins2.
Figure 3.1  Physical Map Surrounding Mash2 and Ins2.
Figure 3.2  Strategy for Cre Mediated Deletion and Duplication Within Mouse Distal Chromosome 7.
Figure 3.3  Strategy for Cre Mediated Deletion and Duplication Within Mouse Distal Chromosome 7.
Figure 3.4  Targeted Insertion at the Mash2 Locus.
Figure 3.5  Cre Mediated Deletion at the Mash2 Locus.
Figure 3.6  Targeted Insertion at the Ins2 Locus.
Figure 3.7  Cre Mediated Deletion at the Ins2 Locus.
Figure 4.1  Proposed Strategy for Cre Mediated Deletion and Duplication Within Mouse Distal Chromosome 7.

LIST OF TABLES

Table 1.1  Imprinted Genes in the Mouse Genome.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>percent</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celcius</td>
</tr>
<tr>
<td>AS</td>
<td>Angelman syndrome</td>
</tr>
<tr>
<td>bp</td>
<td>basepairs</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>BWS</td>
<td>Beckwith-Wiedeman Syndrome</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>calcium chloride</td>
</tr>
<tr>
<td>CMV</td>
<td>cytomegalovirus</td>
</tr>
<tr>
<td>Cre</td>
<td>Cre recombinase</td>
</tr>
<tr>
<td>del</td>
<td>deletion</td>
</tr>
<tr>
<td>dup</td>
<td>duplication</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl-sulfate oxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetra-acetic acid disodium salt</td>
</tr>
<tr>
<td>ES cells</td>
<td>embryonic stem cells</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>H₂O</td>
<td>water</td>
</tr>
<tr>
<td>Igf2</td>
<td>insulin-like growth factor</td>
</tr>
<tr>
<td>Igf2r</td>
<td>insulin-like growth factor receptor</td>
</tr>
<tr>
<td>Ins2</td>
<td>insulin 2</td>
</tr>
<tr>
<td>IPW</td>
<td>imprinted gene in the Prader-Willi syndrome region</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>KCl</td>
<td>potassium phosphate</td>
</tr>
<tr>
<td>LIF</td>
<td>leukemia inhibition factor</td>
</tr>
<tr>
<td>loxP</td>
<td>loxP sites</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>Mash2</td>
<td>mammalian achaete-scute homologue 2</td>
</tr>
<tr>
<td>Mb</td>
<td>megabase</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>ug</td>
<td>microgram</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>ul</td>
<td>microliter</td>
</tr>
<tr>
<td>min</td>
<td>minutes</td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>mm</td>
<td>millimeter</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>μM</td>
<td>micromolar</td>
</tr>
<tr>
<td>mmole</td>
<td>millimole</td>
</tr>
<tr>
<td>Na</td>
<td>sodium</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>sodium phosphate dibasic</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaOH</td>
<td>sodium hydroxide</td>
</tr>
</tbody>
</table>
Chapter 1  Literature Review

1.1  Genomic Imprinting

1.1.1  Introduction

Genomic imprinting is an epigenetic phenomenon that renders the maternal and paternal genomes functionally nonequivalent. This phenomenon is found in a wide variety of organisms including seed plants, some insect species, eutherians and possibly marsupials. In eutherian mammals, the basis for this nonequivalence is that some autosomal genes are expressed or repressed preferentially from either the maternal allele or the paternal allele. As a result, both parental genomes are essential for embryonic and fetal development. Excellent reviews on imprinting have been written by Barlow, 1995; John and Surani, 1996; Nakao, 1996; Bartolomei and Tilghman, 1997; Reik and Walter, 1998; Surani, 1998.

1.1.2  Classical Experiments

Eutherian imprinting was discovered by pronuclear transplantation experiments in mice. Transplantation of male or female pronuclei into enucleated zygotes demonstrated that diploid zygotes with only male or only female pronuclei cannot develop properly (Surani et al., 1986; McGrath and Solter, 1984). Reciprocal translocations showed that nine chromosomal regions (Figure 1.1), depending on their parental origin contributed differently to embryonic development and neonatal growth and behavior.
Figure 1.1

Mouse Imprinting Map.

Imprinted autosomal regions were identified by reciprocal translocations. Mice inheriting duplication of an imprinted region from the indicated parental origin and a deficiency from the other parent exhibit distinct imprinting phenotypes. Imprinted regions are represented by rectangles. Other imprinted regions were identified by mapping of imprinted genes. Imprinted genes are represented by oval grey circles. Only imprinted chromosomes are depicted.

Maternal chromosomes are depicted as having white centromeres, while paternal chromosomes are depicted as having black centromeres.

Adapted from Beechey & Cattanach *Mouse Genome* (1996).
(Cattanach and Kirk, 1985; Beechey and Cattanach, 1996). The recent cloning of Neuronatin (Kagitani et al., 1997; Kikyo et al., 1997), Ins1 (Giddings et al., 1994), Htr2 (Kato et al., 1998), Impact (Hagiwara, 1997) and Grf (Gariboldi et al., 1994; Plass et al., 1996; Itier et al., 1998; Pearsall et al., 1998) revealed additional imprinted regions that were not identified by translocations. Table 1.1 lists genes currently known to be imprinted in mice. Conventionally, the imprinted allele always refers to the repressed or silent allele.

1.1.3 Mouse Distal Chromosome 7

The distal portion of mouse chromosome 7 is one of the imprinted regions identified by Searle and Beechey (1990). It contains both maternally and paternally imprinted genes that are essential for embryonic development. Embryos with maternal disomy and paternal nullisomy (MatDup.d7) are growth retarded and exhibit lethality on embryonic day 15 (E15), while embryos with paternal disomy and maternal nullisomy (PatDup.d7) exhibit overgrowth and die by E10 (Searle and Beechey, 1990). Furthermore, chimeras of PatDup.d7 cells have umbilical hernia, skeletal abnormalities and other defects (McLaughlin et al., 1997).

The Beckwith-Wiedeman syndrome (BWS) is a human syndrome characterized by prenatal overgrowth, omphalocele, macroglossia, umbilical hernia, hemi-hypertrophy, perinatal death, as well as other symptoms (reviewed by Mannens, 1995; Weksberg and Squire, 1995). These symptoms
### Table 1.1: Imprinted genes in the mouse genome.

<table>
<thead>
<tr>
<th>Imprinted gene</th>
<th>Chromosomal location</th>
<th>Imprinted allele</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nnat</td>
<td>distal 2</td>
<td>M</td>
<td>neuronatin</td>
</tr>
<tr>
<td>Gnas</td>
<td>distal 2</td>
<td>M</td>
<td>guanine nucleotide binding protein, alpha stimulating</td>
</tr>
<tr>
<td>Peg1/Mest</td>
<td>proximal 6</td>
<td>M</td>
<td>mesoderm specific transcript</td>
</tr>
<tr>
<td>Peg3/Pwo</td>
<td>proximal 7</td>
<td>M</td>
<td>paternally expressed gene 3</td>
</tr>
<tr>
<td>Snrpn</td>
<td>central 7</td>
<td>M</td>
<td>small nuclear ribonucleoprotein polypeptide N</td>
</tr>
<tr>
<td>Ndn</td>
<td>central 7</td>
<td>M</td>
<td>neddin</td>
</tr>
<tr>
<td>Dn34</td>
<td>central 7</td>
<td>M</td>
<td>mouse homolog of antisense genes DN34</td>
</tr>
<tr>
<td>Znf127(Zfp127)</td>
<td>central 7</td>
<td>M</td>
<td>mouse homolog of sense genes Znf127</td>
</tr>
<tr>
<td>Ipw</td>
<td>central 7</td>
<td>M</td>
<td>imprinted in Prader-Willi syndrome</td>
</tr>
<tr>
<td>Ube3a</td>
<td>central 7</td>
<td>M</td>
<td>E6-AP ubiquitin protein ligase 3A</td>
</tr>
<tr>
<td>H19</td>
<td>distal 7</td>
<td>P</td>
<td>H19 fetal liver mRNA</td>
</tr>
<tr>
<td>Igf2</td>
<td>distal 7</td>
<td>M</td>
<td>insulin-like growth factor 2</td>
</tr>
<tr>
<td>Igf2as</td>
<td>distal 7</td>
<td>M</td>
<td>insulin-like growth factor 2, antisense</td>
</tr>
<tr>
<td>Ins2</td>
<td>distal 7</td>
<td>M</td>
<td>insulin 2</td>
</tr>
<tr>
<td>Mash2</td>
<td>distal 7</td>
<td>P</td>
<td>mammalian achaete-scute homologue 2</td>
</tr>
<tr>
<td>Kolq1</td>
<td>distal 7</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Tapa1/Cdb1</td>
<td>distal 7</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>p57Kip2/Cdkn1c</td>
<td>distal 7</td>
<td>P</td>
<td>cyclin-dependent kinase inhibitor 1C</td>
</tr>
<tr>
<td>Impt1</td>
<td>distal 7</td>
<td>P</td>
<td>multi-membrane-spanning polypeptide transporter-like gene 1</td>
</tr>
<tr>
<td>Ipl</td>
<td>distal 7</td>
<td>P</td>
<td>imprinted in placenta and liver (Tdag51?)</td>
</tr>
<tr>
<td>Rasgrf1</td>
<td>9</td>
<td>M</td>
<td>Ras specific guanine nucleotide-releasing factor 1</td>
</tr>
<tr>
<td>Meg1/Grb10</td>
<td>proximal 11</td>
<td>P</td>
<td>growth factor receptor bound protein 10</td>
</tr>
<tr>
<td>U2af1-rs1</td>
<td>proximal 11</td>
<td>M</td>
<td>U2 small nuclear ribonucleoprotein auxiliary factor (U2AF), 35kDa, related sequence 1</td>
</tr>
<tr>
<td>Htr2a</td>
<td>distal 14</td>
<td>P</td>
<td>5 hydroxytryptamine (serotonin) receptor 2A</td>
</tr>
<tr>
<td>Igf2r</td>
<td>proximal 17</td>
<td>P</td>
<td>insulin-like growth factor 2 receptor</td>
</tr>
<tr>
<td>Impact</td>
<td>proximal 18</td>
<td>M</td>
<td>homology with yeast &amp; bacterial protein family YCR59c/yigZ</td>
</tr>
<tr>
<td>Ins1</td>
<td>19</td>
<td>M</td>
<td>insulin 1</td>
</tr>
</tbody>
</table>

adapted from [http://www.mgu.har.mrc.ac.uk/imprinting/impltables.html](http://www.mgu.har.mrc.ac.uk/imprinting/impltables.html)
are similar to characteristics of PatDup.d7 mice described above. Although most BWS cases are sporadic and have normal karyotype, less than 20% involve paternal uniparental disomy (pUPD) and trisomy with two paternal and one maternal chromosome over human chromosome 11p15, the region syntenic to mouse distal chromosome 7. In addition, this region is implicated in other embryonic tumors that are associated with imprinting, such as Wilms' tumor (WT2), hepatoblastoma, and rhabdomyosarcoma, which often occur in BWS patients (Hoovers et al., 1995; Paulsen et al., 1998).

Mouse Distal chromosome 7/human 11p15 contains many imprinted genes (Figure 1.2; Beechey and Cattanach, 1996). Currently the two closest flanking non-imprinted genes are Rpl23/L23mrp (Zubair et al., 1997) and Nap1l4/Nap2; that have been positioned 1Mb apart (Caspersy et al., 1998). Eleven imprinted genes have so far been identified within this interval: Ipl/Tssc3/BWR1c (Qian et al., 1997; Cooper et al., 1998), Impt1/Tssc5 (Dao et al., 1998), Orctl2/Orctl2S (Cooper et al., 1998), p57Kip2/Cdkn1c (Hatada et al., 1995; Zhang et al., 1997; Yan et al., 1997), Kvlqt1 (Lee et al., 1997; Jiang et al., 1998; Gould and Pfeifer, 1998), Tapal/CdSl (Andria et al. 1991), Mash2 (Guillemot et al., 1994; Guillemot et al., 1995; Rossant et al., 1998), and H19 (Bartolomei et al., 1991; Tremblay et al., 1995; Leighton et al., 1995) are paternally imprinted, while Ins2 (Giddings et al., 1994) and Igf2-AS/Igf2 (DeChiara et al., 1991; Ferguson-Smith et al., 1991; Moore et al., 1997) are maternally imprinted. In addition, at least one non-imprinted gene, TH, (Zhou et al., 1995) is also found amongst this cluster. Other genes have also
Figure 1.2

The Physical Map of Mouse Distal Chromosome 7 and Syntenic Human 11p15.

Maternally imprinted genes are represented by black squares, paternally imprinted genes by white squares, and non-imprinted genes by grey squares. Centromeres are represented by black circles, and telomeres by arrows. Transcription direction is indicated by a smaller arrow below the gene. Physical distance is indicated above the chromosomes.
MOUSE distal chromosome 7

HUMAN 11p15
been identified by positional cloning, such as BWR1a and BWR1b (Cooper et al., 1998), but whether they are imprinted is not yet established. Mouse knockouts of p57Kip2/Cdkn1c (Zhang et al., 1997; Yan et al., 1997), Mash2 (Guillemot et al., 1994; Guillemot et al., 1995), H19, Ins2 (Duvillie et al., 1997), and Igf2 (DeChiara et al., 1991; Leighton et al., 1995) demonstrated their essential role during embryonic development and fetal growth, with the exception of Tapal/Cd81 (Andria et al. 1991).

1.1.4 Unidentified Genes in Distal Chromosome 7

To deduce the possibility of unidentified genes in this region, a comparison between the defects of mice disomic for the entire distal chromosome 7 to those mutant for known genes within this region is informative. The former class is represented by PatDup.d7 mice and mice chimeric for PatDup.d7 cells. Reduced expression in Mash2 probably causes PatDup.d7 mice to die at E10 (MacLaughlin et al., 1997), while PatDup.d7 chimeras exhibit cartilage hyperplasia, organomegaly and umbilical hernia. PatDup.d7 cells are disomic for the paternal distal chromosome 7 region and nullisomic for the maternal, and are expected to overexpress Ins2 and Igf2, but to depress H19, Mash2 and p57Kip2/Cdkn1c. Mice mutant for p57kip2/Cdkn1c, H19 and Igf2 have phenotypes similar to PatDup.d7 chimeras. Significantly, PatDup.d7 chimeras have additional characteristics such as eyelid fusion, reduced limb extension, and postaxial polydactyly, (MacLaughlin et al., 1997) suggesting the existence of at least another
unknown gene over the translocated region. *Ins2* mutants have no discernable phenotype.

Another way to assess this possibility is to compare the characteristics of BWS patients to mice mutant for known genes within this region. Some BWS patients have translocations mapping to human chromosome 11p15 (Hoover et al., 1995; Paulsen et al., 1998) which are thought to disrupt the expression of many genes in this region (reviewed by Mannens, 1995; Weksberg and Squire, 1995). Mice chimeric for cells overexpressing *Igf2* and mice defective for *p57Kip2/Cdkn1c* display similar characteristics as BWS patients, such as overgrowth, organomegaly and umbilical hernia (Zhang et al., 1997; Yan et al., 1997; Sun et al., 1997; Eggenschwiler et al., 1997). However, many other BWS hallmarks like predisposition to childhood tumors and typical ear pits and creases (Mannens, 1995; Weksberg and Squire, 1995) are not the result of defects in these two genes. Mice mutant for *Mash2* and *Ins2* also do not have these phenotypes (Duvillie et al., 1997; Guillemot et al., 1994; Zhou et al., 1993; Leighton et al., 1995).

Imprinting may play a role in the development of embryonal tumors, such as rhabdomyosarcoma, Wilms tumor and adrenocortical carcinoma to which BWS patients are especially susceptible, since karyotyping of lesions reveals a specific loss of maternal 11p15 alleles (Hoovers et al., 1995; Rainier et al., 1993; Ogawa et al., 1993; Steenman et al., 1994). Genes associated with these biological functions have not yet been identified in this region.
Genetic analysis by Baranov in the 1930s indicates the presence of a non-imprinted gene essential for peri-implantation development in this region (Baranov, 1979). Also, other cancers such as that of lung, ovarian, and breast have also been associated with this region (Moore et al., 1998).

In conclusion, some of the recently cloned but not yet functionally characterized genes may be involved in processes detailed above, but there may be additional imprinted genes in distal chromosome 7 not yet identified.

1.1.5 Mechanism of Imprinting

1.1.5.1 The Primary Imprint

Parental-origin-specific expression of imprinted genes logically must result from imprints initially set while the genomes are separate (from gametogenesis to pronuclear fusion in the zygote). These differential allelic modifications must also be maintained during global demethylation and later, remethylation (Monk et al., 1987). In the germ cells, these imprints must be erased and set up anew for the next generation.

Modifications during oogenesis have been demonstrated to determine whether an imprinted gene is expressed or repressed. Kono and Obata (1997) demonstrated through transplantation of parthenote nuclei that such modifications take place between the ages of neonatal day 1 and 12.5. For instance, paternally expressed genes like Peg1/Mest, Peg3 and Snrpn, are modified from a potentially expressing state to a silent state. On the other hand, maternally expressing genes like Igf2r and p57Kip2/Cdkn1c are
modified from a repressed state to an expressive state. The third pattern observed in *H19* is a lack of modification during oogenesis, suggesting modifications may take place during spermatogenesis.

An alternative model proposes sex specific marks are retained through successive generations, unless the genome is passed through the germline of the opposite sex (Dittrich *et al*., 1996). The hypothesis is based on the observation of some Prader Willi syndrome/Angelman syndrome patients. These patients inherited microdeletions from a phenotypically normal parent of the opposite sex. Upon detailed examination, chromosomes carrying these microdeletions appeared to retain the same imprinting characteristics as that parent's chromosome. As a result, these patients have a normally imprinted chromosome 15 from the other parent and an abnormally imprinted region on the chromosome carrying the deletion. The deletion is thought to encompass an imprinting centre that is involved in erasing and setting of imprints (Sutcliffe *et al*., 1994; Reis *et al*., 1994; Buiting *et al*., 1995; Saitoh *et al*., 1996; Yang *et al*., 1998).

The following properties of methylation fulfill the requirements of the primary epigenetic mark that would distinguish the parental origins of imprinted genes. First, methyltransferase activity is present in gametes prior to fertilization while they are still in separate compartments (reviewed by Bestor, 1995; Mertineit *et al*., 1998). The expression pattern of methyltransferase in the developing oocyte is consistent with a role in primary imprinting (Mertineit *et al*., 1998). Second, sex specific methylated
sites can be stably maintained after many cell divisions (Bestor, 1995). Third, germline specific methylation do not occur de novo in somatic tissues. In fact, methylation associated with imprinting can only be reset in primordial germ cells (Tucker et al., 1996). Consistent with this, cultured PGC cells or embryonic germ cells have a dominant demethylating effect (Tada et al., 1997). Further support comes from DNA methyltransferase knockout mice that are not able to maintain imprinting (Li et al., 1993). Also, gamete specific differential methylation regions which fit these criteria have been identified at the Snrpn (Shemer et al., 1997), H19 (Tremblay et al., 1995) and Igf2r loci (Stoger et al., 199).

1.1.5.2 Imprint Propagation

The mechanism by which the primary imprint leads to parent-of-origin dependent expression is being intensively investigated. The following characteristics of imprinted genes and regions are providing some insights into this process. Imprinted regions replicate asynchronously (Kitsberg et al., 1993; Knoll et al., 1994; Lasalle et al., 1995; Lasalle et al., 1996) and exhibit sex-specific meiotic recombination frequencies (Paldi et al., 1995). The majority of imprinted genes cluster in distinct regions (Beechey and Cattanach, 1996), contain allele specific methylation sites (Ferguson-Smith et al., 1993; Tremblay et al., 1995; Stoger et al., 1993), and often function in growth regulation (DeChiara et al., 1993; Barlow et al., 1991). Other intriguing characteristics include, these genes have few and small introns (Neumann et
most are closely associated with regions rich in direct repeats (Shemer et al., 1997; Tremblay et al., 1995), and a significant number of these genes do not code for proteins, but some do code for antisense RNA (Bartolomei et al., 1991; Nakao et al., 1994; Wutz et al., 1997; Moore et al., 1997; Schuster-Gossler et al., 1998).

1.1.5.3 Chromatin Structure

The regulation of imprinting involves organization at the chromatin level. At non-imprinted loci, DNA replication on both homologues occur at the same time. At imprinted loci, Igf2r, Igf2 and H19, paternal DNA replication precedes that of maternal (Kitsberg et al., 1993; Knoll et al., 1994). The opposite pattern, maternal replication before paternal, can also take place, for example on human chromosomal region 15q11-q13, within which three different asynchronous patterns are present (Lasalle et al., 1995). A causative link was made in an H19 mutant that has lost its imprinting. That phenotype coincided with its loss of asynchronous replicating pattern at the H19 region (Greally et al., 1998). Moreover, the nuclear distance between homologous regions of imprinted genes are found to be shorter than for non-imprinted genes, suggesting a close association of these regions in the nucleus (Lasalle et al., 1996).

In addition, imprinted regions display higher meiotic frequency during male meiosis and lower frequency during female meiosis (Paldi et al., 1995). On average, the human female genetic map is 90% longer than the
male map which reflects the higher frequency of meiotic recombination in female germ cells. The one exception is a report documenting enhanced female meiotic recombination in mouse distal chromosome 7 (Caspary et al., 1998).

Furthermore, differential nuclease hypersensitivity, to agents such as DNAse I and restriction enzymes, is associated with imprinted regions (Ferguson-Smith et al., 1993; Szabo et al., 1998). Detailed chromatin analysis of H19 also characterized unusual chromatin structures (Szabo et al., 1998).

1.1.5.4 Chromosomal Domains

The imprinting mechanism may involve a master regulator that controls all the genes within a chromosomal domain. Imprinted genes appear to cluster as demonstrated by the cloning of over 20 imprinted genes, the majority of which reside within imprinted regions that were identified by Robertsonian and reciprocal translocations (Figure 1.1; Beechey and Cattanach, 1996). Support also comes from the conservation of these clusters, and gene synteny, transcription direction, gross physical distance within these clusters between mouse and human (Paulsen et al., 1998).

The mouse distal chromosome 7/human 11p15 imprinted regions probably operate by this mechanism. This hypothesis is based on the observation of BWS patients who exhibit biallelic IGF2 expression. In these patients, many mutations were found at the Igf2 locus, but breakpoints were also identified at the Kolqt1 locus, 500 kb away, and at other loci megabases
away (Hoover et al., 1995). Similar findings were made in another imprinted region that is associated with Prader Willi/Angelman Syndrome on human chromosome 15 (Reis et al., 1994). Mutations in Ube3A (reviewed by Dittrich and Buiting, 1995; Kerns and Giacalone, 1995; Fung et al., 1998) can affect the imprinting of Snrpn which is also located hundreds of kilobases away). These results suggest that despite the huge distance, their imprinting is affected by a common regulatory network.

An alternative idea is that distal chromosome 7 consists of multiple imprinting domains. The first observation that led to this idea was in a mutant for the H19 locus that disrupted imprinting of Igf2 and Ins2 but has no effect on the imprinting of Mash2, Kvlqt1, or p57kip2/Cdkn1c (Bartolomei and Tilghman, 1997; Caspary et al., 1998). Studies of mice mutant for other nearby genes also showed a lack of coordinate imprinting expression expected from one single domain (Caspary et al., 1998). For instance, overexpression of Igf2 in transgenic chimeras is correlated with a reduction in H19 expression, while p57Kip2/Cdkn1c expression remains constant (Sun et al., 1997). These data indicate Kvlqt1, p57Kip2/Cdkn1c and Mash2 are not in the same regulatory unit as Igf2, Ins2 and H19.

Finally, TH, which is not imprinted, is positioned between Mash2 and the H19, Igf2 and Ins2 cluster. These observations suggest that multiple regulatory domains may exist in distal chromosome 7.
1.1.5.5 Mechanisms within the Domains

1.1.5.5.1 Competition

Elegant models have been proposed to explain the mechanism by which differential expression is achieved within each domain. The competition or coordinate regulation model was proposed when Tilghman and her colleagues demonstrated that H19, Igf2 and Ins2 shared the same enhancer (Leighton et al., 1995). According to this model, the cause for allele specific expression is due to the epigenetic modification of an imprinting centre located upstream of the H19 gene. When the maternal allele is unmodified or not methylated, H19 is able to better engage the enhancer it shares with Igf2 and Ins2, and therefore prevents the expression of these genes in cis. After the imprinting centre is modified or methylated (on the paternal allele), akin to being deleted, H19 is prevented from engaging the enhancer and Igf2 and Ins2 are now able to be expressed. Consistent with this model, Igf2r and Igf2r-AS has also been shown to operate by a similar competition mechanism (Wutz et al., 1997; Barlow, 1997).

Studies of at least two imprinting centres suggest that it functions as a silencer when unmethylated (Lyko et al., 1997; Lyko et al., 1998). One proposal for the function of imprinting centres is to cause direct promoter or enhancer competition. Another idea proposed by Jones et al. (1998) is the imprinting centres may be chromatin boundary elements that in the unmodified state are able to block transactivation; modifications results in accessibility.
1.1.4.5.2 Other Proposed Mechanisms

Reik and Walter (1998) propose imprinting consists of initial establishment and later maintenance. They suggest imprinting establishment involves germ line specific differential methylated regions that interact with transactivating factors such as methylases and chromatin factors. Further maintenance may involves non-coding and antisense RNA and repetitive elements and other differential methylations.

1.1.5 Function and Evolution of Genomic Imprinting

The most fascinating aspect of genomic imprinting is the evolutionary origin of this phenomenon. Despite disadvantages of having functionally hemizygous region, many organisms adopt this mechanism.

The most popular evolutionary theory for genomic imprinting is parental conflict which was first proposed by David Haig and Mark Westoby in 1989. The basis of the theory is that males and females adopt different strategies in order to maximize the number of their offspring. For instance, a successful male passes on genes that will maximize the growth and optimize competitive/survival behaviors of his progeny, while a female is more successful when she can optimally curb the action of these genes and distribute resources evenly amongst her offspring.

This theory predicts genes involved in promoting growth or survival behaviours are expressed from the paternal allele but are silent from the maternal allele. Oppositely, genes such as growth inhibitors would be
expressed from the maternal allele but silent from the paternal allele. This theory has been borne out by experimental data obtained from transgenic mice mutant for imprinted genes such as Ctl2 (Schuster-Gossler et al., 1998), Peg1 (Kaneko-Ishino et al., 1995; Kobayashi et al., 1997; Lefebvre et al., 1997; Reule et al., 1998; Lefebvre et al., 1998), Igf2 (DeChiara et al., 1991; Ferguson-Smith et al., 1991), and Igf2r (Barlow et al., 1991; Schweifer et al., 1997).

Furthermore, Arabidopsis genetic experiments demonstrated that imprinting has similar functions on embryonic growth.

This theory also predicts that if both parents will have the same offspring and hence no conflict, for example in monogamous species, imprinting would not be positively selected during evolution. This prediction is also supported by breeding experiments between a polygamous and a monogamous species of mice (Vrana et al., 1998). They showed the monogamous specie is smaller than its polygamous counterparts again as predicted by this theory.

Contradictory data also exist to make this theory controversial. McVean and Hurt in 1997 showed cases of paternal disomyes that resulted in growth retarded babies. The fact that Mash2 is paternally imprinted but is involved in the development of the placenta and embryo (Guillemot et al., 1994; Guillemot et al., 1995) also counters Haig's hypothesis. Furthermore, it is not yet clear why many other growth factors are not imprinted, and not all imprinted genes are related to growth and behavior. In fact, MEDEA, a
paternally imprinted gene that is involved in growth inhibition was recently cloned in *Arabidopsis* (Grossniklaus *et al.*, 1998).

The following highlights some other intriguing theories (reviewed by Haig and Trivers, 1995; Hall, 1990). Since only mammals, and not other vertebrate, acquired genomic imprinting, and mammalian evolution coincided with the ontogeny of the placenta, revolves around placenta biology. Hall proposes the genomic imprinting evolved when the paternal genome within the placenta modified itself, for example by suppressing paternal antigens, so that the mother will tolerate the foreign conceptus. Varmuza and Mann (1994) noted poor trophoblast development in parthenogenotes, and suggested that imprinting prevents ovarian germ cell tumors in the mother by disabling genes involved in trophoblast development which resemble tumor progression. Other hypotheses include parthenogenesis prevention which states, since a relatively high incidence of spontaneous parthenogenesis occurs naturally, then imprinting may serve to prevent the genetically deleterious effects of homozygous recessive lethal alleles. It has also been suggested that imprinting maintains the advantages of sexual reproduction. Barlow (1993) proposes imprinting was co-opted from another pathway which functions to modify and silence foreign DNA. Finally, some think imprinting is just another level of gene regulation.
1.2 *Mouse Genome Engineering*

1.2.1 Introduction

Innovations in manipulative mouse molecular genetics during recent years allow virtually any desirable modification to the mouse genome. Consequently, novel insights into mammalian biology and medicine have been gained (Capecchi, 1989; Thomas *et al.*, 1992) and will continue to do so for many years to come.

Pivotal to the development of this powerful technology are the establishment of embryonic stem (ES) cells, efficient utilization of gene targeting and site specific recombination. ES cells are derived from preimplantation mouse embryos (blastocysts) and can be propagated and manipulated in culture (reviewed by Pility *et al.*, 1998). When properly maintained, these cells retain the pluripotency of the embryonic ectoderm, the component of the blastocyst which they are believed to represent. Thus these cells are able to contribute to nearly all cell types of a mouse, including the germ line, if allowed to develop together with a morula stage embryo or injected into blastocysts (reviewed by Nagy, 1997).

Gene targeting can be performed in ES cells and has become an efficient method to manipulate the mouse genome (Capecchi, 1989). It is based on homologous recombination and subsequent drug selection for cells that undergo these rare events. Often, targeting strategies are designed to remove the function of a gene of interest. Although such experiments are informative, incomplete insight is gained from null or loss of function...
mutants. Therefore, other more sophisticated techniques are emerging that can generate a series of other alleles. For detailed reviews, refer to Rossant and Nagy, 1995; Lobe and Nagy, 1998; Hadjantonakis et al., 1998.

1.2.2 Cre-loxP

An outstanding example of the kind of sophisticated techniques being perfected is the site specific recombination system, Cre-loxP (Sternberg, 1994; Argos et al., 1986; Sauer and Henderson, 1988). Native to the P1 phage, the Cre enzyme recognizes a 34-base pair recognition element, the loxP site (Figure 1.3), which is not present naturally in the mouse genome. If two such elements are situated some distance apart, the Cre enzyme catalyzes conservative recombination between them (Hoess and Abremski, 1985). When two loxP sites are placed in the same orientation in cis, Cre specifically excises the intervening sequence, and when two loxP sites are placed in opposite orientation, Cre specifically inverts the flanked sequence. The Cre protein is also able to catalyze intermolecular recombination. By combining the methods of gene targeting and site specific recombination, innumerable experiments can be designed to address issues never before thought possible in the mouse. The following illustrates some successful applications and suggests many possibilities.

Cre-loxP has been shown to efficiently excise sequences of up to several kilobases in ES cells (Gu et al., 1993; Nagy et al. 1998; this study). During selection of homologous recombination events, drug resistant
Figure 1.3

Cre-loxP Site-specific Recombination

A: The Cre recombinase, represented by the canoe shaped, recognizes a 34 bp consensus sequence (loxP) which is comprised of two inverted repeats and a core sequence (bold). The core sequence determines the directionality of loxP recognition site.

A-C: When the loxP sites are placed on the same DNA strand in the same orientation, Cre can catalyse a site specific recombination event leading to the deletion of the intervening DNA.
markers or GFP variants are necessarily introduced (Hadjantonakis et al., 1998). Unfortunately these sequences often interfere with the function of the endogenous gene, and obscure the true effect of a specific gene alteration (Meyers et al., 1998; Nagy et al., 1998). In other situations, subtle changes such as domain deletion or single nucleotide changes to a gene are sometimes desirable (Gu et al., 1993; Nagy et al., 1998; Pirity et al., in preparation). At other times, regulatory elements need to be modified to determine their specific function (Fiering et al., 1993). The Cre-loxP system can precisely remove short sequences at high efficiency in these scenarios so as to minimize any changes at the endogenous locus. During the design and construction of targeting vectors, loxP sites need to be properly oriented and positioned at defined locations. After identifying correctly targeted cell lines, transient expression of the Cre enzyme can excise loxP flanked sequences.

1.2.3 Cre-mediated Chromosomal Rearrangements

With Cre-loxP one is also capable of creating site-specific chromosomal and sub-chromosomal rearrangements, such as translocations, inversions, deletions, duplications, and even monosomies (Figure 1.4 - 1.6). These applications should allow better mouse models of human syndromes and cancers by recreating associated chromosomal rearrangements in the mouse (Smith et al., 1995). Investigations of large genes or large and/or distant regulatory regions may be simplified by using these methods (Li et al., 1996; Herault et al., 1998; this study). Finally, the possibility of creating large scale
The strategies generally involve two consecutive targeting steps that deliver loxP sites to the endpoints of the specific rearrangements (Figure 1.4). Then Cre is expressed transiently to catalyze recombination. Since the efficiency of this recombination reaction is thought to be extremely low, drug selection has always been used to isolate desired clones. Often mini-genes such as Hprt are divided in two and each half is strategically oriented and positioned adjacent to one loxP site during targeting events. Upon proper recombination, a functional selectable marker is reassembled. Alternatively, a functional selectable marker can be reconstructed from fragments of promoter and coding sequence of a selectable marker (this study) or a GFP variant. Sometimes, negative selection is used in a similar logistic manner (Li et al., 1996), Figure 1.5.

The first reported chromosomal rearrangement was a balanced chromosomal translocation (Smith et al. 1995). LoxP sites were targeted to two different non-homologous chromosomes in the same direction relative to the centromere, see Figure 1.4. Reassembly of Hprt mini gene was used to select for cells that correctly recombined.

Soon after, inversion, deletions and duplications of up to 3-4 cM was created (Ramirez-Solis et al..). The deletion and duplication strategy used for genes or regions of known relative orientation (Figure 1.6) involved placing
Figure 1.4

Strategy for Cre Mediated Translocation

Smith *et al.*, 1995 targeted two halves of an *Hprt* mini-gene and properly oriented and positioned loxP sites to the endpoints of a desired translocation. Once correctly targeted ES cells were identified, transient expression of Cre induces recombination between the loxP sites and creates the translocation. Cells with the recombinant chromosomes were positively selected.

*Hprt* mini-gene is represented by black oval, loxP by black triangle.
Gene targeting

Cre recombinase
Figure 1.5

**Strategy for Cre Mediated Translocation**

Li *et al.*, 1995 targeted the TK gene and properly oriented and positioned loxP sites to the endpoints of desired deletion. After correctly targeted ES cells were identified, transient expression of Cre induces recombination between the loxP sites and creates the deletion. Candidate cells with the recombinant chromosomes were negatively selected.
Gene targeting

Cre recombinase
Figure 1.6

Strategy for Cre Mediated Deletions and Duplications.

Ramirez-Solis et al. generated 3-4 cM deletions and duplications by targeting two halves of an \textit{Hprt} mini-gene with loxP sites to the desired endpoints. Upon cre excision, cells positively selected are expected to contain a large deletion via intra-chromosomal recombination or deletion and duplication via inter-chromosomal recombination.

\textit{Hprt} mini-gene is represented by black oval, loxP by black triangle.

*unstable chromosomes
A

Intra-chromosomal recombination

B

Inter-chromosomal recombination

5. Cre recombinase
Strategy for Cre Mediated Deletions, Duplications, Inversions and other Rearrangements.

Ramirez-Solis et al. generated up to 3-4 cM rearrangements by targeting two opposite orientations of two halves of an Hprt mini-gene with loxP sites at desired endpoints. The four targeted patterns are diagrammed. Upon transient expression of Cre in ES cell lines containing both cis and trans chromosomal configurations (not diagrammed), eight possible rearrangements were obtained. Cells were positively selected and the hypothetical recombinant chromosomes are listed.

*unstable chromosomes
Locus I

A  hprt 5'
B  hprt 5'
C  hprt 5'
D  hprt 5'

Locus II

  hprt 3'
  hprt 3'
  hprt 3'
  hprt 3'

Cre recombinase

cis  trans
A  deletion  deletion +duplication
B  inversion  *acentromeric +dicentromeric
C  inversion  *acentromeric +dicentromeric
D  (deletion)  duplication +deletion
LoxP sites to the endpoints. If these loxP sites are targeted to the same chromosome, then a deletion would result; if loxP sites are targeted to different homologues, then a deletion and duplication would be generated. Inversions are achieved by orienting loxP in the opposite direction and positioned on the same chromosome. Using ES cell lines derived from an F1 hybrid of two different strains, one may use polymorphic markers to identify which of the two homologous chromosomes is targeted (Clerc and Avner, 1998). Alternatively, sufficiently high number of colonies derived from the second (independent chromosomal) targeting event can be screened after the Cre mediated recombination step (Ramirez-Solis et al., 1995). The more complex scenario involving genes or regions of unknown relative orientation was also successfully accomplished by Ramirez-Solis and colleagues. Figure 1.7 illustrates the strategy used by this group and a multitude of rearrangements that can be created.

Lastly, loss of an entire chromosome can be created (Ramirez-Solis et al., 1995; Lewandoski and Martin, 1997). First, Lewandoski and Martin created male mice carrying inverted loxP sites on their Y chromosome. By mating them to female mice expressing Cre ubiquitously at an early embryonic stage, they were able to derive XO females. Loss of Y chromosome probably results from unequal sister chromatid exchange during G2 phase of the mitotic cell cycle.

While many sophisticated experiments involving Cre-loxP are being perfected in ES cells, other innovative ideas are being tested and improved
in the mice themselves. One application is to direct Cre activity under tissue or temporal specific promoters, or inducible promoters (Tsien et al., 1996; Lam et al., 1997; Meyers et al., 1998). This possibility is the timely answer to the difficult question of addressing multiple functions of a gene. When a gene function is ablated, resulting mutant animals reveal the earliest defect associated with the absence of that gene's function, informing little about other possible later acting roles. Using the Cre-loxP system, removal or restoration of gene function in a specific lineage or during a specific stage is possible (Lewandoski and Martin, 1997; Meyers et al., 1998; Nagy et al., 1998).

Exciting applications are also emerging from convergence of different uses of Cre/loxP. For instance, the popularity of the Cre-loxP system has resulted in many transgenic mice each with loxP sites positioned at different regions of the genome. Simultaneously, Cre transgenic mice are also being made available (Lewandoski and Martin, 1997; Meyers et al., 1998; Nagy et al., 1998; Herault et al., 1998; Lobe et al., 1999). Using appropriate mating regimes, innumerable translocations, deletions, duplications can be created without the necessary gene targeting and Cre recombination experiments in ES cells, nor germline transmission of the desired genetic alteration. Presently, however, in vivo large scale Cre-loxP rearrangements is not a routine method because an appropriate selection system is not yet available.
1.3 Purpose of Thesis

Although the mouse imprinting map has been extremely informative, the resolution achieved using given Robertsonian translocations is not sufficient for detailed analysis of imprinting mechanisms. Sophisticated ES cell-mediated technology in genome alterations, such as combinations of gene targeting and site-specific recombination can more precisely engineer chromosomal rearrangements.

This project involves designing a defined deletion and duplication strategy, and generating these rearrangements in ES cells so that the effects of paternal and maternal inheritance of these rearrangements can be studied in mice. The interval being deleted is 350 kb between Mash2 and Ins2, a relatively large and uncharacterized region within distal chromosome 7. This experiment may reveal both the presence and the effects of unknown imprinted genes in this interval. This study will also assess the presence and effects of essential cis-acting regulatory elements that may reside in the deleted and duplicated region. This precise rearrangement will be able to address the issue of domain organization in this region. For example, which are the independent domains, and which genes are within them. By complementing this strategy with molecular analysis like gene cloning, genome sequencing of the region and transgene analysis, imprinting mechanism within distal chromosome 7 is certain to be better understood.

In addition, cell lines derived from intermediate steps of the deletion and duplication strategies can be used for creating smaller and larger
inversions, deletions and duplications within and outside this large interval. The resulting allelic series can be used to fine map interesting phenotypes, dissect the mechanism or test hypotheses of imprinting through combinations of various alleles.

Finally, this experiment will test the feasibility and efficacy of the Cre-loxP system to create rearrangements like large deletions and duplications, translocations and inversions.
Chapter 2  Materials and Methods

2.1 Materials

2.1.1 Tissue Culture

R1 embryonic stem (ES) cells derived from inner cell mass of E3.5 blastocyst stage mouse embryos were used throughout this experiment. The parental strains of this line is 129/SvJ and 129/SvCP (Nagy et al., 1993).

Complete ES cell culture media consists of Dulbecco’s modified Eagle’s medium (DMEM) with high glucose, supplemented with 50ug/mL penicillin and 50ug/mL streptomycin, 15% fetal calf serum, 1000U/mL LIF, 2mM L-glutamine, 100uM Bmercaptoethanol, 1mM sodium pyruvate, and 0.1mM nonessential amino acids. R1 cells were cultured at 37°C and at 5% CO₂ level.

Freezing media consists of the complete culture medium, plus an additional 10% fetal calf serum and 10% DMSO.

Selection drug G418 was purchased from Gibco, and gancyclovir from Syntex.

Electroporation cuvettes are available from BioRad (Gene Pulser).

Additional details have been described by Pirity et al., 1998.

2.1.2 Hybridization Probes

Mash2 probes are contained within the genomic clone 19ER1 (Guillemot et al., 1994). The following fragments were isolated from this plasmid and used in hybridization experiments: Mash2 5’ or external, EcoRI-
XhoI 1.5 kb, Mash2 3' or internals, NcoI 2.2 kb or EcoRI-NcoI 1.0 kb (Figure 2.4).

The Ins2 probe was isolated from P11 (Duville et al., 1997). The plasmid was obtained from Dr. J. Jami. It is a 1.0 kb EcoRI-BglII fragment. In addition, an Ins2 cosmid was obtained from Dr. A.-K. Hadjantonakis. Subclones and an external probe, a 1.3 kb AccI fragment, were derived from this cosmid (Figure 2.5).

2.1.3 Primers

*Mash2* genotyping (Figure 2.4)

M2B - CTT CAC ACG GCA GTT CTG TG
M2C - GTA TAC TCC AGA GTC TCC TG
M2E - TCT AGA GCT CGC TGA TCA GCC TCG ACT GTG

*Ins2* genotyping (Figure 2.5)

I2A - ACT GAA ATG TCA AGG CTT TAG GTG CTT AGG
neo.2 - GAA CCT CGC TGC AAT CCA TC
I2D - TCT CCA TGA GGA ATG ACA CAT AGC TTG TAT

*Ins2* targeting (Figure 2.5):

I2TVneo - CAT GGC GAT GCC TGC TTG CCG AAT ATC ATG
I2TV2 - TGA TAG CTG GGT TCT TAA CTC AGC CGA GTC
2.1.4 Animals

CD-1 females are available through suppliers and were routinely used.

2.2 Methods

2.2.1 Pulsed Field Gel Electrophoresis

2.2.1.1 Preparation of High-Molecular-Weight Genomic DNA

DNA samples for pulsed-field gel electrophoresis were prepared from R1 embedded in low-melting-point (LMP) agarose as described (Rommens et al., 1989). Briefly, the cells were harvested at 50-60% confluency by detaching them from culture dishes with SE buffer (75mM NaCl, 25mM EDTA, pH8). The harvested cells were collected by centrifugation at 2,000g, washed with additional SE buffer, and resuspended in fresh SE buffer. The cell suspensions were warmed to 37°C, quickly mixed with equal volumes of 1% molten LMP agarose, and poured on a glass plate to a thickness of approximately 1.5mm. The solidified cell suspensions were then cut into square blocks with a clean scalpel, then transferred to 10x vol lysis buffer (0.25M EDTA, 1% lauryl sarcosyl, pH8), and treated with 0.5mg proteinase K/mL at 50°C overnight with gentle shaking. The deproteinized blocks were then washed extensively with several changes of TE (10mM Tris-HCl, 50mM EDTA, pH8) at 37°C with gentle shaking and were stored at 4°C prior to restriction-enzyme digestion.
2.2.1.2 Restriction-Enzyme Digestion

Agarose blocks containing high-molecular-weight DNA were washed briefly in sterile double-distilled water and equilibrated overnight in 4-6 vol buffers appropriate for various restriction-enzyme digestions. Restriction digestions were then performed with 1-5 units enzyme/ug DNA in 3 vol fresh buffers under conditions recommended by the manufacturers (New England Biolabs and Boehringer/Mannheim). Reactions were terminated with the addition of EDTA (pH8) to a final concentration of 10mM, and the blocks containing digested DNA were stored in pulsed-field gel running buffer at 4°C prior to electrophoresis.

2.2.1.3 Pulsed-Field Gel Electrophoresis

DNA samples digested with various rare-cutting restriction enzymes were fractionated by electrophoresis using a CHEF apparatus. Generally, 1-2 ug digested genomic DNA was used per sample well of a 1% agarose gel.

Electrophoresis was carried out with a constant electric field of 6.5-8V/cm in 0.5x TBE (45mM Tris-borate-EDTA, pH8.3) running buffer, and the gel switched between two positions with a total rotation angle of 120°C at set time intervals. The duration of electrophoresis at each position (pulse) increased linearly with time; the initial and final pulses were set as indicated. The temperature of the running buffer was maintained at 14°C with the use of a cooling coil linked to a refrigeration unit.
The DNA samples in the agarose gels were visualized by staining with ethidium bromide after electrophoresis.

2.2.1.4 Southern Blotting

To facilitate Southern blot transfer of high-molecular weight DNA, the pulse field gel was depurinated with 0.25N HCl for 10 min. Subsequently the DNA fragments were denatured in 0.4N NaOH/0.6M NaCl and neutralized in 0.5M Tris-HCl(pH7.5)/1.5M NaCl for 30 min each. The DNA was transferred to Hybond-N+ membranes by capillary transfer in 20xSSC according to the method of Southern. All other gels were treated by the same method, except the depurination steps were omitted.

2.2.1.5 Hybridization

The filters were prehybridized and hybridized to radiolabelled probe in 0.5M NaPi pH6.8, 7% SDS, 1mM EDTA, at 65°C overnight. They were washed in 0.1xSSC, 0.1%SDS at 65°C for 30 min, three times. Probes were random labelled using a Boehringer Mannheim Labelling Kit. The hybridization signal was visualized with a Molecular Dynamics Phosphoimager. Membranes were reused for hybridization with other DNA probes after stripping in 0.05x SSC, 0.1% SDS, 10mM EDTA at 95°C for 5-10m, twice.
2.2.1.6 DNA Size Standards

Yeast chromosomes were used as markers for the estimation of DNA fragment sizes. Chromosomal DNA from *Saccharomyces cerevisiae* strain (gift from J. Rommens, Hospital for Sick Children) was prepared according to a method described elsewhere (Carle and Olsen, 1985).

2.2.2 Construction of Targeting vectors

2.2.2.1 PGK-neo Cassette

Toward construction of the final targeting vectors, a pair of PGK-neo selection cassettes were first prepared that included loxP sites flanking either the PGK promoter or the neomycin coding sequence and bovine polyA, Figure 2.1. Proper orientation and sequence were confirmed by sequencing. First, a 0.5 kb EcoRI - PstI (blunted) fragment from pPGKβgeo pA was cloned into the EcoRI-SmaI sites of pBS65 (Sauer and Henderson, 1988). The resultant plasmid, PGK-loxP, contained the PGK promoter followed by loxP. Then, this plasmid was digested with PstI and blunted with T4 to allow ligation of a 1.1 kb EcoRI - SpeI fragment from pneobpA, to yield PGK-loxP-neo. The insert, PGK-loxP-neo, can be excised using EcoRI and HindIII. They were used to insert both upstream and downstream of loxP sequences of pBS65, to yield the above mentioned PGK-neo selection cassettes that contained loxPs flanking the promoter or the coding sequence.
Figure 2.1

Construction of PGK-neo with embedded loxP cassettes.

For details, see Materials and Methods.
2.2.2.2 *Mash2* Target Vector

Figure 2.4 shows the genomic clone 19ER1 contains a 4 kb EcoRI-EcoRV fragment downstream of *Mash2*, which was derived from a 129 library and obtained from Dr. Francois Guillemot (Guillemot et al., 1994). A restriction map was also established. This plasmid was modified to remove the SpeI site in the multiple cloning sites. To construct the targeting vector, a 1600 bp blunt ended EcoRI-HindIII fragment containing a PGK-loxP-neo-loxP cassette was inserted into the now unique and blunted SpeI site.

2.2.2.3 *Ins2* Target Vector

PI11, containing a region upstream of *Ins2* exons, was also derived from a 129 genomic library and obtained from Dr. Jacques Jami (Duvillie et al., 1997). Figure 2.3 illustrates it contains a 5 kb fragment that was subcloned into pBluescript II KS+ (Strategene). A restriction map was established. One NcoI site 800bp from the EcoRV site was removed to ease difficulty of the next cloning step. The targeting vector was constructed by inserting the 1600 bp blunt ended EcoRI-XhoI fragment, which contained a loxP-PGK-loxP-neo cassette into the only NcoI site left. In addition, a TK cassette was inserted downstream of the 3' homology arm to improve the frequency of targeting.
Figure 2.2

Construction *Mash2* Target Vector.

For details, see Materials and Methods.
NcoI

EcoRI

EcoRV

NcoI

NcoI (blunt)

ERV-ERI 4kb
(Mash2)

NcoI

EcoRI

EcoRV

PGK promoter

PGK- loxP-neo- loxP

neo coding

HindIII

bPA

EcoRI (blunt) + HindIII (blunt)

NcoI (blunt)

ERV-ERI 4kb
(Mash2)

EcoRI

EcoRV

Mash2
Target Vector

EcoRI

3' homology

neo coding

PGK promoter

5' homology
Figure 2.3

Construction *Ins2* Target Vector.

For details, see Materials and Methods.
EcoRI - SpeI (blunt)

EcoRI - EcoRI (blunt)

EcoRI (blunt) + XhoI (blunt) SpeI (blunt)

EcoRI

TK

Ins2 Target Vector

PGK promoter

5' homology

3' homology

neo coding
2.2.3 Embryonic Stem Cell Culture

2.2.3.1 Maintenance of R1 Cells

R1 cells were maintained, passaged and split as described (Nagy et al., 1993; Pirity, 1998). Briefly, R1 cells were fed new media every day. Every other days, R1 cells at 20-40% confluency were passaged by splitting 1:5 to 1:7. Additional details are described by Pirity et al., 1998.

2.2.3.2 Transfection and Selection for Mutant ES Cells

DNA used for transfection were prepared using Qiagen Columns and targeting vectors were linearized and electroporated into R1 ES cells (Nagy et al., 1993). Transfection by electroporation routinely required $10^7$ ES cells and either 20ug linearized targeting vectors, or 40 ug supercoiled Cre recombinase expression vector as described (Pirity et al., 1998). If selection is required, two days after electroporation, cells were selected in a concentration of 150ug/ml of active G418, and four days after electroporation, in 2uM gancyclovir. Cells requiring no drug selection were plated sparsely. Individual colonies were picked after 7 days of selection onto gelatinized 96-well plates and grown to near confluency before splitting into two 96-well plates. The master plate was frozen down, and ES cell genomic DNA from the other plate was prepared and analyzed by Southern blotting for homologous recombination events. After identification of correctly targeted clones, they were thawed from the master plate and cultured as described (Pirity et al., 1998).
2.2.4 ES Cell Screening

2.2.4.1 Genomic DNA Extraction

Genomic DNA was prepared from 96 well plates according to Ramirez-Solis et al., 1992.

2.2.4.2 Hybridization Probes

Genomic DNA from potential *Mash2* targeted colonies was digested with EcoRI and probed with a 1.0 kb Xho-BamHI external probe and a 1.0 kb EcoRI-NcoI internal probe, indicated in Figure 2.4. Cre excision at the *Mash2* locus was screened using the same 1.0 kb EcoRI-NcoI internal probe on EcoRI digested genomic DNA.

Genomic DNA from potential *Ins2* targeted colonies was digested with PstI and probed with a 1.3 kb AccI external probe, or digested with EcoRI and probed with a 1.0 kb EcoRI-BglII internal probe, indicated in Figure 2.5. Cre excision of the PGK promoter at the *Ins2* locus was diagnosed by digesting genomic DNA with EcoRI and probed with the same 1.0 kb EcoRI-BglII internal probe.

Hybridizations were carried out overnight at 65°C in 0.5M NaPi/0.1%SDS/1mM EDTA.
Figure 2.4

*Mash2 Locus.*

Restriction map of the plasmid (19ER1) that contain the wildtype *Mash2* locus, and the *Mash2* target vector.

Probes and PCR primers used for identifying correctly targeted clones and genotyping are indicated. PCR primers are represented by arrows. Hybridization probes are indicated by bars.
EcoRI

PCR Primers (genotyping)

Target Vector

EcoRV

Probes

External EcoRI-XhoI 1.5kb

Internal NcoI 2.2kb

Internal NcoI 1.0 kb

M2B M2C

M2B M2neo

EcoRI

PGK neo

M2B

EcoRV

M2C
Figure 2.5

*Ins2 Locus.*

Restriction map of the cosmid and plasmid (P11) that contain the wildtype *Ins2* locus, and the *Ins2* target vector.

Probes and PCR primers used for identifying correctly targeted clones and genotyping are indicated. PCR primers are represented by arrows. Hybridization probes are indicated by bars.
2.2.4.2 Polymerase Chain Reaction

*Ins2* targeting was verified by PCR that detected the targeted allele with the I2TV2 primer which anchors outside the 3' homology arm, and I2TVneo which is specific for the neo coding region. Genomic DNA was amplified for 30 cycles (95°C for 1min, 65°C for 1min, 72°C for 3min).

2.2.5 Generation of Mutant Chimeras

Chimeras were generated by ES-morula aggregation with targeted ES lines (Pirity *et al.*, 1997). Morula are flushed from superovulated CD-1 outbred females. Chimeric males were then bred to CD1 females to establish F1 heterozygotes. Therefore, heterozygote and homozygote mice are of 129/Sv x CD1 mixed background. Embryos from intercrosses of F1 heterozygotes were genotyped either by Southern analysis or by PCR.

2.2.6 Genotyping of Mice

Around the time of weaning, 3 weeks of age, mice ears were systematically punched to provide unique identification. These skin samples were also collected and digested in proteinase K cocktails at 55°C overnight and were used as genomic template for PCR.

To distinguish wildtype and modified *Mash2* alleles, primers M2B, M2neo and M2C were used to amplify genomic DNA for 30 cycles (95°C 1m; 65°C 1min; 72°C 1min). The first two primers amplified the 400 bp modified
allele specifically, while the first and last primers amplified a 158 bp wildtype fragment (Figure 2.4).

Similarly, Ins2 alleles can be distinguished by using primers I2A, neo.2 and I2D. The genomic DNA template was amplified for 30 cycles (95°C 1m; 60°C 1min; 72°C 1min). Again, the first two primers amplified the 1.7kb modified allele specifically, while the first and last primers amplified the 500 bp wildtype fragment (Figure 2.5).
Chapter 3  Results: Deletion of 350 kb Within Distal Chromosome 7 in ES Cells

3.1 The Physical Distance between Mash2 and Insulin2 is 350 Kilobases

Previously, the physical distance between Mash2 and Ins2 had not been accurately determined. In addition, the direction of Mash2 transcription was unclear. To address these issues, R1 DNA was singly and doubly digested with rare cutting restriction enzymes, fractionated by pulse field gel electrophoresis, Southern blotted and sequentially hybridized to Mash2 and Ins2 probes.

It was already known that an Eagl site lies between Ins2 and Igf2 (Zemel et al.; Guillemot et al., 1995). Furthermore, characterization of Mash2 revealed an MluI site in the second intron. Consistent with the previous report, Figure 3.1 shows both Ins2 and Mash2 probes hybridized to a 450 kb Eagl fragment. However, only the Ins2 probe and the Mash2 probe 3' of the MluI site hybridized to a 375 kb Eagl/MluI fragment. In contrast, the Mash2 probe 5' of the MluI recognized a fragment that is smaller than 240 kb. The MluI fragments that are greater than 1 Megabase was outside of the resolution range of the gel, did not transfer well on the Southern and hence hybridized weakly to probes. Since the Eagl site was determined to be only 10 kb from Ins2 (Zemel et al., 1992), the best estimate of the size of the Ins2 and Mash2 interval is close to 350 kb. Additionally, the direction of transcription of Mash2 and Ins2 is identical.
Figure 3.1

Physical Map Surrounding *Mash2* and *Ins2*.

A: Physical map as interpreted by hybridization results. Probes for hybridization is shown as bars above map.

B: Pulse field gel electrophoresis and hybridization analysis. Mouse genomic DNA was prepared from R1 ES cell line as described in Materials and Methods, digested with restriction enzymes as indicated. (Left to Right) 1. Stained with EtBr. 2-4. The same blot was used for all hybridizations using [32P]-labeled probes, as indicated in A, along left to right sequential order.

Marker sizes (*Saccharomyces cerevisiae* chromosomes) are given on the left. Arrow indicates *Mash2* YAC.

Y, *Mash2* YAC that contains yeast chromosomes and an artificial chromosome that contains the *Mash2* gene and surrounding region.

E, EagI; M, MluI; EM, EagI+MluI.
3.2 Deletion and Duplication Strategy

A strategy was designed to delete and duplicate a 350 kilobase interval in ES cells and it is illustrated in Figures 3.2 and 3.3. The key idea is to place the loxP sequence 3' of Mash2, and another loxP sequence 5' of Ins2, oriented in the same direction. Then, expression of Cre recombinase will either excise the intervening sequence or delete and duplicate it. To avoid mutagenizing these two anchor genes unnecessarily, the loxP sites were placed 5 kb 3' of Mash2 exon 3 and 2.5 Kb 5' of Ins2.

Stepwise details of the strategy are as follows. First, Mash2 locus is to be targeted using a construct that includes neo (neomycin) as a positive selectable marker. A pair of directly repeated loxP sites strategically flanks the neo coding sequence and neo is under the control of a PGK promoter. After identifying correctly targeted cell lines, neo-coding sequence is to be excised by transient expression of Cre recombinase. As a result, these cell lines are rendered neo sensitive again for the next set of targeting. These cell lines, with only the PGK promoter and loxP sites at the Mash2 locus, are targeted again using an Ins2 targeting construct. Neomycin is used again to enrich for Ins2 targeted cell lines. Properly targeted lines should have a modified Ins2 locus that contains a pair of properly oriented loxP flanking, in this case, the PGK promoter, and downstream of them the neo coding sequence. A second round of transient Cre expression removes the promoter from the Ins2 targeted site, leaving a loxP site and the neo coding sequence behind. Moreover, these cells become sensitive to neomycin yet again. A large
Figure 3.2

Strategy for Cre Mediated Deletion and Duplication Within Mouse Distal Chromosome 7.

This strategy consists of 5 steps. First, conventional replacement-type gene targeting is used to insert the PGK-loxP-neo-loxP cassette at the Mash2 locus. ES cells identified as correctly targeted are used for the second step. This step involves the transient expression of Cre which induces recombination between the loxP sites and excises the neo coding sequence. Thirdly, using correctly excised ES cells as substrate, another replacement-type gene targeting inserts the loxP-PGK-loxP-neo cassette at the Ins2 locus. Only the cis configuration is illustrated here. Fourthly, transient expression of Cre excises the loxP flanked promoter at the Ins2 locus. Lastly, transient expression of Cre reconstructs the PGK neo cassette by intra-chromosomal recombination. Cells with the recombinant (deleted) chromosome are positively selected in G418. Detailed description of the strategy is in the results section. Only the deletion pathway is depicted; the duplication pathway is shown in Figure 3.3.

PGK: PGK promoter
neo: neo coding sequence and bovine polyA
triangle: loxP
1. Gene targeting

2. Cre recombinase

3. Gene targeting

4. Cre recombinase

5. Cre recombinase
Figure 3.3

Strategy for Cre Mediated Deletion and Duplication Within Mouse Distal Chromosome 7.

The configuration of loxP sites that results from gene targeting of the second locus, the third step of this strategy, determines if the chromosome is deleted or deleted and duplicated.

A: LoxP sites in cis, upon Cre recombination, results in one deleted chromosome via intra-chromosomal recombination.

B: LoxP sites arranged in trans results in one deleted and one duplicated chromosome via inter-chromosomal recombination.
number of cells from these cell lines are subjected to the last round of transient Cre expression. A deletion event restores a functional PGK-loxP-neo gene and those cells are neo resistant. Figure 3.3. shows when both loxP are targeted on the same chromosome, through intrachromosomal recombination Cre creates a deletion. However, when the loxP sites are targeted on the two homologous chromosomes 7, a deletion and a duplication results from interchromosomal recombination.

3.3 Construction of Targeting Vectors

According to the strategy, targeting constructs were prepared as described in the Materials and Methods.

Briefly, PGK-neo cassettes containing loxP flanking the promoter or the coding sequence was constructed.

The Mash2 target vector was constructed by inserting the loxP-PGK-loxP-neo cassette into a 4 kb EcoRI-EcoRV fragment which contains a region downstream of Mash2. The targeting vector contained 4 kb of 5' homology and 1 kb of 3'. The PGKneo cassette was in the same transcriptional orientation compared to the endogenous Mash2 gene (Figure 2.4).

The Ins2 target vector was constructed by inserting the PGK-loxP-neo-loxP cassette into a 5 kb fragment which contains a region upstream of Ins2. In addition, a TK (thymidine kinase) cassette was inserted downstream of the 3' homology arm to improve the frequency of targeting. The resulting targeting vector contained 3 kb of 5' homology and 2.2 kb 3'. The
transcriptional orientation of PGKneo, PGKtk and the endogenous *Ins2* gene were in the same orientation (Figure 2.5).

3.4 Identification of ES Cells Correctly Targeted at the *Mash2* Locus by Homologous Recombination

Correctly targeted clones were identified at a frequency of 1/20; out of 438 informative clones, 24 were correct. The Southern screening method included an external probe that hybridizes to different EcoRI fragments. Figure 3.4 shows the wildtype EcoRI fragment is 23 kb, while the targeted is 10 kb. Positive clones were verified using other internal probes.

Careful selection against cell lines containing additional independent, random integration events in the correctly targeted lines was necessary to prevent possible genomic rearrangements that may result from subsequent Cre activity. Internal probes and vector probes were used for this purpose on genomic DNA digested with various enzymes. Clones with extra or aberrant bands were discarded (not shown).

In conclusion, six cell lines have met all these criteria and have been used to carry out subsequent genome modifications.

3.6 Identification of ES Cells Correctly Cre Excised at the *Mash2* Locus

The next step, after isolating *Mash2* targeted cells, was to remove their drug resistance against G418 so as to allow targeting of *Ins2*. A transient expression vector allowing the production of Cre recombinase from an
Figure 3.4

Targeted Insertion at the *Mash2* Locus.

*Mash2* was targeted at a frequency of 1/20.

A: Targeting and screening strategy of *Mash2*.

B: Southern blot analysis using *EcoRI*-digested genomic DNA. Hybridization probes are indicated in panel A.
ubiquitous promoter was electroporated into *Mash2* cell lines. The Cre expression construct in this experiment was under the control of a novel chicken B-actin promoter which was modified with the CMV-Intermediate-early enhancer 5' of the promoter (Niwa *et al.*, 1991). 70% (250 out of 350 informative colonies) underwent the desired Cre excision and were identified using the screening method depicted in Figure 3.5. Southern blotting of EcoRI digested DNA and hybridization using an internal probe shows wildtype bands at 23 kb, and either a band at 13 kb indicating an unexcised allele or an excised allele at 12 kb.

### 3.7 Identification of ES Cells Correctly Targeted at the *Ins2* Locus by Homologous Recombination

*Ins2* targeted clones were identified at a frequency of 1/37; out of 412 colonies, 11 were correct. Southern blot of PstI digested genomic DNA was hybridized to an external probe, Figure 3.6. Correctly targeted clones are those containing a wildtype 8 kb band, and an altered 6.5 kb band. PCR screen employed a primer outside the 3' arm of homology and another specific for the neo cassette, later verified that these clones were correct (Figure 3.7; data not shown).

Correctly targeted *Ins2* clones were also screened for independent, random integration events, carried out in a similar method as described above (not shown).
Cre Mediated Deletion at the Mash2 Locus.

The floxed neo portion of the PGK-neo cassette at the Mash2 targeted locus was excised in 70% clones.

A: Screening strategy used to identify correctly excised clones.

B: Southern blot analysis using EcoRI-digested genomic DNA. Hybridization probe is indicated in panel A.
Figure 3.6

Targeted Insertion at the \textit{Ins2} Locus.

\textit{Ins2} was targeted at a frequency of 1/37. using the depicted targeting and screening strategy.

A: Targeting and screening strategy of \textit{Ins2}.

B: Southern blot analysis using \textit{PstI}-digested genomic DNA. Hybridization probe is indicated in panel A.
A

External probe

\[ \text{PGK} \quad \text{neo} \]

Wildtype allele

Targeted allele

B
3.8 Identification of ES Cells Correctly Cre Excised at the Ins2 Locus

To remove G418 resistance so as to allow recombination between the two loxP sites at Mash2 and Ins2. Cre recombinase was transiently expressed as described above. Southern of PstI digested genomic DNA was hybridized to the same external probe. Figure 3.7 shows wildtype allele is detected at 8 kb, targeted allele at 6.5 kb, and cre excised allele at 6.0 kb. 40% of clones have deleted the PGK promoter; out of 365 colonies, 134 colonies contained the deletion. 10% of clones appeared identical to the parental clones; 38 out of 365 informative colonies.

3.9 Identification of ES Cells Correctly Deleted the 350 Kilobases

Surprisingly, 50% of clones candidate for being correctly cre excised at the Ins2 locus contained the deletion of the interval between Mash2 and Ins2. Out of 365 colonies, 169 contained the large deletion. Using EcoRV, Mash2 probes normally detect wildtype alleles at 5 kb or loxP allele at 5.5 kb. On the other hand, Ins2 probes only detect wildtype alleles at 16 kb or loxP allele at 17 kb. Deletion clones shown in Figure 3.8, however, contained wildtype bands plus, a novel 10 kb band that is commonly recognized by both Mash2 and Ins2 probes.

3.10 Transgenic mice

Chimeras have been generated by aggregating morulae (Nagy et al., 1993) with Mash2 and Ins2 independently targeted cell lines. These chimeras
Cre Mediated Deletion at the Ins2 Locus.

The floxed PGK promoter at the Ins2 locus was excised in 40% of clones.

A: Screening strategy used to identify correctly excised clones.

B: Southern blot analysis using EcoRI-digested genomic DNA. Hybridization probe is indicated in panel A.
A

External probe

PstI


PstI

PstI

Ins2

PstI

PstI

PGK

neo

Cre recombinase

PstI

PstI

neo

8 kb

6.5 kb

6.0 kb

Wildtype allele

Targeted allele

Cre excised allele

B

8kb

6.5kb

6.0kb
Figure 3.8

**Cre Mediated Deletion Between *Mash2* and *Ins2*.**

The 350kb interval was excised in 50% of clones.

A: Screening strategy used to identify correctly excised clones.

B: Southern blot analysis using EcoRV-digested genomic DNA. Hybridization probes are indicated in panel A.
A

Wildtype allele
LoxP targeted allele
Deleted allele

B

Mash2
Ins2

10 kb
5.5 kb
5 kb
16 kb
17 kb
10 kb
17 kb
have transmitted these mutations to their progeny. Both heterozygous and homozygous mice for *Mash2* and *Ins2* show no obvious phenotype. *Mash2* heterozygote intercrosses resulted in mendelian ratio of homozygote wildtype, heterozygote and homozygote mutant mice, respectively 31, 53 and 21. Mice heterozygous for *Ins2* also showed similar ratio, 11 wildtype, 34 heterozygotes and 14 homozygous mutant. All analyses were carried out on a mixed CD1/129Sv background.

Chimeras are still being generated using ES cell containing the 350 kb deletion. However, few chimeras have resulted and those chimeras have low contribution by the ES cells.
Chapter 4  Discussion

4.1  Physical Mapping Around \textit{Mash2} and \textit{Ins2}

Pulse field gel electrophoresis determined that the physical distance between \textit{Mash2} and \textit{Ins2} is approximately 350 kb in mouse (Figure 3.1). This measurement also estimates the size of the deletion obtained from this experiment.

Previously, it was reported that the distance between \textit{Mash2} and \textit{TH} in the mouse is 25 kb (Caspary \textit{et al.}, 1998). Therefore, the distance between \textit{TH} and \textit{Ins2} is over 300 kb. It is interesting to note that in human the physical distance between \textit{TH} and \textit{Ins2} is only 12 kb (Paulsen \textit{et al.}, 1998). Imprinting differences between mouse and man can be found in distant genes, although the imprinting, gene order and transcription direction of adjacent genes appear conserved. For instance, \textit{Kvlqt1} is imprinted in all human fetal tissues, but in the mouse is only imprinted in extraembryonic tissues and early embryos. This apparent evolutionary divergence between mouse and human may be significant and require further characterization.

4.2  Cre/\textit{loxP} Efficiency and Implications

Reliably high Cre-mediated recombination frequency in ES cells was obtained in this experiment. Typically 70\% of clones correctly excised small fragments of up to 1 kb. In contrast, others reported only approximately 2-4\% excision efficiency for fragments up to 6 kb (Sauer and Henderson, 1989; Gu
et al., 1993). Possible explanations for these differences are the following. The Cre expression construct in this experiment was highly proficient as a result of the novel chicken B-actin promoter (Niwa et al., 1991). Alternatively, the transfection method used is superior. Finally, the use of R1 cells was critical for the high transfection efficiency.

Unexpectedly, the large deletion was obtained at a high frequency compared to similar large scale rearrangements. Deletion of 350 kb was observed at a stunning 50% efficiency, in the absence of selection. Smith et al. reported a translocation between the c-myc and immunoglobulin heavy chain genes on chromosomes 15 and 12 occurring at a frequency of $5 \times 10^{-8}$. Van Deursen et al. created a translocation between the Dek gene on chromosome 13 and the Can gene on chromosome 2 that occurred at a frequency of $4.2 - 8.3 \times 10^{-4}$. Ramirez-Solis et al. also reported low efficiency in deletions, duplications, inversions and other rearrangements they created. These groups required the use of drug selection to obtain clones with the desirable rearrangement.

This result suggests cre-mediated large scale deletion experiments in ES cells are relatively easy to create, and it may not require complicated strategies nor selection. Using the same experiment described in this thesis as an example, the strategy will only require two steps of targeting and a third step of Cre mediated site specific recombination (Figure 4.1). The targeting vector should use two different positive selectable markers, such as neomycin and puromycin. LoxP sites should be introduced appropriately so
Figure 4.1

Proposed Strategy for Cre Mediated Deletion Within Mouse Distal Chromosome 7.

This proposed strategy consists of only 3 steps. First, two conventional replacement-type gene targeting are used to insert the PGK-loxP-neo-loxP and loxP-PGK-loxP-neo cassette at the Mash2 and Ins2 locus. ES cells identified as correctly targeted are then used for transient expression of Cre. Recombination catalyzed by Cre results in a deleted chromosome.

PGK: PGK promoter  
neo: neo coding sequence and bovine polyA  
puro: puromycin gene  
triangle: loxP
1. Gene targeting

2. Gene targeting

3. Cre recombinase
that upon final deletion only one loxP sequence is left at the junction of the deletion. Less manipulations to ES cells may improve the final outcome (discussed later).

4.3 Chromosomal Targeting and Implications

In theory, there is an equal chance that Mash2 and Ins2 are targeted onto the same or different homologous chromosomes, i.e., in cis or in trans (Figure 3.2 and 3.3). Upon cre excision, a deleted and a normal chromosome results from targeting onto the same chromosome; a deleted and a duplicated chromosome would result if targeting was on different homologous chromosomes. From three independent targeting experiments performed, the same chromosome was targeted. Although three is statistically insignificant and more experiments need to be performed, here are some speculations. One possibility is genomic DNA used for homologous recombination was cloned from a strain that has more resemblance to one parental strain of the ES cell (Simpson et al., 1997) and thus led to biased chromosomal targeting. Alternatively, the chromatin structure of this imprinted region in ES cells may have biased the targeting of one parental chromosome over the other.

If targeting of genes is chromosomally biased depending on DNA sequence differences between strains, strategies can be designed to obtaining specific rearrangements. For instance, to obtain duplication events in ES cells, genomic DNA may need to be cloned from a different strain in order to
bias targeting onto specific chromosomes. Alternatively, ES cells can be derived from F1 hybrids of different strains. The homologue that has been targeted can be determined by nearby and informative polymorphic markers (Clerc and Avner, 1998).

If targeting was biased as a result of chromosome structural differences, alternative strategies are needed. For instance, to create the duplication by site-specific recombination in vivo. First, generate mice with properly positioned loxP sites. Then, mate them so that the loxP sites are in trans. By crossing Cre transgenic mice with their progeny, the resulting litters should contain deletion and duplication rearrangements. Three existing Cre transgenic mice may work well. The first is the deletor mouse created in Gail Martin's lab (Lewandoski and Martin, 1997; Meyers et al., 1998). Another transgenic mouse that expresses specifically in the germ cell is available in Andras Nagy's lab (Lobe et al., 1999). A third mouse expresses Cre under the Sycep promoter and demonstrated deletion and duplication capacity in vivo (Herault et al., 1998).

4.4  

Mash2 and Ins2 Chimeras and Mutants

Chimeras and mutant mice were generated from cell lines that were targeted at the Mash2 and Ins2 loci, ie the first step of the strategy. This experiment provides controls for the final deletion and duplication phenotypic analysis. For example, to ensure that Mash2 expression has not
been disturbed unnecessarily. Modifications to both Mash2 and Ins2 did not appear to have any phenotypic effect on mice development or viability.

Another reason for generating these mice was to determine the contribution ability of targeted cell lines that were derived from multiple steps of the strategy (Nagy et al., 1993), so that only good cell lines are used to carry out the subsequent modifications of the genome.

4.4 350 kb Deletion Chimeras

Strong chimeras so far has not been obtained from cell lines that were deleted for the Mash2 to Ins2 interval. Two possible explanations follows. First, ES cell lines have been passaged too many times. It is no longer pluripotent and cannot contribute at a high percentage to a mouse.

Alternatively, loss of at least one gene in the deleted interval is preventing ES cells from contributing at a high percentage to a mouse. To distinguish between these two possibilities, the parental and sister cell lines used to derive the large deletion are also being aggregated into chimeras.

4.5 Future Prospects

This thesis shows that deletions can be created with relative ease, and that deletion and duplication strategies is feasible in the future. Cell lines derived from intermediate steps can serve as anchors for smaller and larger rearrangements within and outside this large interval and future molecular manipulations within this imprinted interval. The resultant allelic series
can be used to carefully dissect mechanisms through combinations of various generated alleles. Mice that were generated can also serve a useful purpose to obtain rearrangements *invivo* using Cre transgenic mice as described above.
REFERENCES


Baranov, V.S. (1979) New Marker Reciprocal Translocations T(7;14)2Iem and T(7;14;15)3Iem in Laboratory Mice. Genetika 15;1651-60.


97


98


