The Imprinted Polycomb Group Gene $Sfmbt2$ is Required for Trophoblast Maintenance and Placenta Development

by

Kamelia Miri

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Cell and Systems Biology
University of Toronto

© Copyright by Kamelia Miri (2014)
Embryos with two maternal copies of the genome and no paternal contribution (parthenogenetic embryos), die early in development due to failure to sustain a functional extraembryonic tissue of trophoblast origin. The reported phenotypes of parthenogenetic embryos have been correlated with absence of expression from an imprinted paternally expressed gene (PEG) with critical roles in maintenance of trophoblast stem (TS) cells and their derivatives. To find the genes responsible for the observed phenotypes of the murine parthenogenetic embryos, I performed a survey of known PEGs and identified the paternally expressed $Sfmbt2$, a Polycomb Group (PcG) gene, as the most fitting candidate. I subsequently showed that a loss-of-function mutation of $Sfmbt2$ results in early mid-gestational lethality due to failure to maintain the trophoblast lineage. I also showed that the knockdown of this gene in pre-implantation embryos results in a significant reduction in the derivation of TS cells in vitro. Lastly, I performed coimmunoprecipitation experiments and showed that SMFBT2 interacts with a number of transcription factors including TFDP1 whose loss-of-function mutation also results in
mid-gestational lethality due to failure in proper development of the extraembryonic tissues. Collectively, my data support the initial candidate study and suggest that SFMBT2 is the earliest acting PEG identified to date with an essential role in maintenance of the trophoblast lineage both in vivo and in vitro. By extrapolation, absence of expression of Sfmbt2 in parthenogenetic embryos is likely a major contributor to the reported phenotypes.
I dedicate this thesis to my loving parents.
“I have walked that long road to freedom. I have tried not to falter; I have made missteps along the way. But I have discovered the secret that after climbing a great hill, one only finds that there are many more hills to climb. I have taken a moment here to rest, to steal a view of the glorious vista that surrounds me, to look back on the distance I have come. But I can only rest for a moment, for with freedom comes responsibilities, and I dare not linger, for my long walk is not ended.”

— Nelson Mandela
Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Susannah Varmuza, who is a brilliant intellectual and a wonderful person, for giving me the opportunity to work alongside her, teaching me to think outside the box and for believing in me. Her encouragements helped me persevere early mornings and late nights of troubleshooting. The time in her lab taught me to be independent, patient and thorough. I will be forever grateful for everything she has done for me.

I would like to thank my doctoral committee, Dr. Andrea Jurisicova, and Dr. Vince Tropepe for their valuable insights and their open doors to help me any way they could. I would like to thank Dr. Jannet Rossant, Dr. James Ellis, Dr. Dinesh Christendat and Dr. Mellissa Mann for reagents. I would like to thank Mr. Daniel Rivero, from the Biological Sciences Facility, who is a friend and a colleague and has always gone above and beyond the call of duty to accommodate our lab. I would also like to thank the past and present members of Varmuza lab, especially: Katherine Szegal, Sangeetha Paramathas, Stella Wang, Jocelyn Wang, Jessica Gu, Lawrence Park, John Tak and Eu Ddeum Chung, whom I have had the privilege of supervising and have contributed immensely to my learning experience.

I would like to thank my friends whose love and support have made me get through it all. I especially would like to thank my dear friend, Sameer Hasan, who has been there for me from the beginning of this journey to the end, whom I knew I could count on no matter what, who helped me adopt two lovely furry companions whose presence have transformed my life. None of this would have been possible without his support.
Lastly and most importantly, I would like to thank my amazing family, my mom, dad, brother, sister and my most lovely niece and nephew whose love and support is beyond anything I could describe. I would like to especially thank my mom and dad who selflessly dedicated their lives for the wellbeing and advancement of their children. Thank you for loving me and supporting me unconditionally. Thank you for making me the person that I am today and thank you for being who you are. I could not have asked for a better family.
Table of Contents

ABSTRACT ........................................................................................................................................... II

DEDICATIONS .......................................................................................................................................... IV

ACKNOWLEDGEMENTS .......................................................................................................................... VI

TABLE OF CONTENTS ............................................................................................................................. VIII

LIST OF TABLES .......................................................................................................................................... XII

LIST OF FIGURES ........................................................................................................................................ XIII

LIST OF APPENDICES ................................................................................................................................ XVI

LIST OF ABBREVIATIONS ........................................................................................................................ XVI

CHAPTER 1: INTRODUCTION ................................................................................................................... 1

1.1 IMPRINTING: AN EPIGENETIC INHERITANCE .................................................................................. 3

  1.1.1 Functional significance of imprinted genes ............................................................................... 4
          1.1.1.1 Uniparental duplications and disomies in mice ................................................................. 4
          1.1.1.2 Knockout studies of imprinted genes .............................................................................. 5

  1.1.2 Regulation of imprinting ........................................................................................................... 6
          1.1.2.1 Germline DMR-controlled, non-coding RNA-induced imprint regulation .................... 7
          1.1.2.2 Germline DMR-controlled, chromatin conformation- facilitated imprint regulation .... 8

  1.1.3 Evolution of imprinting and its correlation with placentation in mammals ...................... 9

1.2 PLACENTATION .................................................................................................................................... 10

  1.2.1 Stages of Placental Development ............................................................................................ 10

  1.2.2 Trophoblast cell types: Specification, Maintenance and Differentiation ............................. 13
          1.2.2.1 TS cells: Specification .................................................................................................... 13
          1.2.2.2 TS cells: Maintenance .................................................................................................... 15
          1.2.2.3 TS cells: Differentiation ................................................................................................ 17

  1.2.4 Heterogeneity of TGCs and their perspective roles ............................................................... 18

1.3 EPIGENETIC REGULATORS OF THE GENOME: POLYCOMB GROUP PROTEINS ........................... 21

  1.3.1 Polycomb Repressive Complex 1 (PRC1) .............................................................................. 22

  1.3.2 Polycomb Repressive Complex 2 (PRC2) .............................................................................. 24

  1.3.3 Other Polycomb repressive complexes .................................................................................... 25

  1.3.4 PcG protein recruitment: suggestive of functional diversity ................................................ 26

  1.3.5 PcG proteins and cell fate ........................................................................................................ 28

1.4 THESIS HYPOTHESIS AND OBJECTIVES ................................................................................... 29
# Chapter 2: Materials and Methods

## 2.1 Tissue Culture

2.1.1 Pre-implantation Embryos: Harvest and culture .................................................. 32
2.1.1.1 Superovulation and pre-implantation embryo harvest ........................................ 32
2.1.1.2 Superovulation and Parthenogenetic activation of oocytes ................................... 32
2.1.2 Blastocyst outgrowth and TS cell derivation ...................................................... 33
2.1.3 TS cell maintenance and differentiation ............................................................. 34
2.1.4 HEK293T cell culture and transfection ............................................................... 34
2.1.4.1 Generation of Sfmbt2 over expression constructs .............................................. 35
2.1.5 Sfmbt2 knockdown in pre-implantation embryos ................................................. 37
2.1.5.1 shRNA containing plasmid construction ............................................................ 37
2.1.5.2 In vitro generation of lentivirus ........................................................................... 38
2.1.5.3 Lentiviral mediated knockdown in pre-implantation embryos ............................... 38

## 2.2 RNA Analysis

2.2.1 RNA extraction ..................................................................................................... 39
2.2.2 cDNA synthesis .................................................................................................. 40
2.2.3 RT-PCR and qRT-PCR ....................................................................................... 40
2.2.4 Single nucleotide polymorphism (SNP) analysis .................................................. 41

## 2.3 DNA Analysis

2.3.1 DNA extraction .................................................................................................. 41
2.3.2 PCR ................................................................................................................... 42

## 2.4 Protein Analysis

2.4.1 Protein extraction ............................................................................................... 42
2.4.2 Western blot analysis ......................................................................................... 43
2.4.3 Co-immunoprecipitation and Mass Spectrometry Analysis ................................. 43
2.4.3.1 Co-immunoprecipitation ................................................................................... 43
2.4.3.2 In-gel and In-solution tryptic digestion ............................................................. 44
2.4.3.3 Mass spectrometry ......................................................................................... 45
2.4.4 Histology ............................................................................................................ 46
2.4.4.1 Paraffin embedding and sectioning ................................................................. 46
2.4.4.2 H and E staining ............................................................................................ 46
2.4.4.3 Immunohistochemistry .................................................................................... 47
2.4.4.4 Cryoblock preparation and sectioning .............................................................. 48
2.4.4.5 Alkaline phosphatase (AP) staining ................................................................. 49
2.4.4.6 X-gal staining .................................................................................................. 49
2.4.5 Immunocytochemistry ....................................................................................... 50
2.4.5.1 Blastocyst staining .......................................................................................... 50
2.4.5.2 Adherent cell staining ...................................................................................... 50
CHAPTER 3: IDENTIFICATION AND VALIDATION OF A CANDIDATE PATERNALLY EXPRESSED GENE WITH ESSENTIAL ROLES IN TROPHOBLAST STEM CELLS

3.1 ABSTRACT ........................................................................................................... 58
3.2 INTRODUCTION ............................................................................................... 59
3.3 RESULTS ............................................................................................................ 61
  3.3.1 Parthenogenetic TS cells are indistinguishable from fertilized counterparts .......... 61
  3.3.2 PTS cells exhibit loss of imprinting (LOI) at various loci ............................... 63
  3.3.3 Candidate Study: Identification of SFMBT2 .................................................. 64
    3.3.3.1 PEGs displaying imprinted expression at the blastocyst stage (Criterion 1) .... 64
    3.3.3.2 PEGs displaying imprinted expression in TS cells and LOI in all PTS cell lines (Criterion 2 and 3) .............................................................. 66
    3.3.3.3 PEGs with knockout phenotypes resembling those of the parthenogenetic embryos (Criterion 4) ................................................................. 68
  3.3.4 Observed LOI occurs as early as the blastocyst outgrowth stage ..................... 68
  3.3.5 Spatial and Temporal Expression-pattern of Sfmbt2 Confirms a Role in Trophoblast Development ................................................................. 69
    3.3.5.1 Expression of SFMBT2 in pre-implantation stage embryos ....................... 69
    3.3.5.2 Expression of SFMBT2 at the blastocyst outgrowth stage ........................ 72
    3.3.5.3 Expression pattern of SFMBT2 in Trophoblast Cell Types .......................... 74
  3.3.6 Knockdown study is suggestive of a critical role for Sfmbt2 in establishment and/or maintenance of TS cells ................................................................. 74
3.4 DISCUSSION ....................................................................................................... 79

CHAPTER 4: SFMBT2 IS REQUIRED FOR MAINTENANCE OF TROPHOBLAST CELL-TYPES AND PLACENTA DEVELOPMENT ........................................................................ 83

4.1 ABSTRACT ........................................................................................................... 84
4.2 INTRODUCTION ............................................................................................... 85
4.3 RESULTS ............................................................................................................ 87
  4.3.1 SFMBT2 Displays Nuclear Staining in Various Cell-types of the Trophoblast Origin ........ 87
  4.3.2 Paternal Transmission of the Null Allele of Sfmbt2 is Embryonic Lethal ............... 91
  4.3.3 Paternal Inheritance of the Gene Trap Allele of Sfmbt2 Leads to Severe Defects in Extraembryonic Tissues in a Strain Dependent Manner ........................................ 95
4.3.3.1 Phenotype of the hypomorphic allele in C57BL/6 background ...................................................95
4.3.3.2 Phenotype of the hypomorphic allele in CD1 background ..........................................................98
4.3.4 SFMBT2 null mutation results in mid-gestational embryonic lethality due to failure to
maintain trophoblast cell-types ........................................................................................................... 103
4.3.4.1 Phenotype of homozygous null embryos ....................................................................................103

4.4 DISCUSSION ......................................................................................................................................110

CHAPTER 5: PCG PROTEIN SFMBT2 FORMS COMPLEXES WITH VARIOUS TRANSCRIPTION FACTORS IN TS
CELLS .......................................................................................................................................................116

5.1 ABSTRACT .........................................................................................................................................117
5.2 INTRODUCTION ...............................................................................................................................118
5.3 RESULTS ...........................................................................................................................................120
5.3.1 SFMBT2 shows a distinct nuclear localization pattern in various trophoblast cell-types and
associates with mitotic chromosomes .................................................................................................120
5.3.2 SFMBT2 coimmunoprecipitates with YY1 in both exogenous and endogenous environments.
.................................................................................................................................................................122
5.3.3 Coimmunoprecipitation experiments reveal a novel interaction between SMFBT2 and TFDP1
.................................................................................................................................................................125
5.3.4 Mass spectrometry confirms the interaction of SFMBT2 with the transcription factor CASZ1
in TS cells ...............................................................................................................................................127
5.4 DISCUSSION ......................................................................................................................................129

CHAPTER 6: DISCUSSION .......................................................................................................................136

6.1 CHROMOSOMAL LOCATION AND LOI OF SFMBT2 ........................................................................137
6.2 TISSUE SPECIFIC EXPRESSION AND LOCALIZATION PATTERN OF SFMBT2
................................................................................................................................................................141
6.3 LOSS OF FUNCTION PHENOTYPE OF SFMBT2 SUGGESTS INVOLVEMENT IN MAINTENANCE OF THE TROPHOBLAST
LINEAGE ...............................................................................................................................................142
6.4 BIOCHEMICAL NATURE AND BINDING PARTNERS OF SFMBT2 .....................................................145
6.5 CONCLUSION ....................................................................................................................................149

REFERENCES .........................................................................................................................................150

APPENDICES .......................................................................................................................................175
List of Tables

Table 1.1. Expression profile of select markers of TGCs in various trophoblast cell-types...... 19

Table 3.1. Summary of the results from the candidate study for the identification of the PEG responsible for the early phenotypes of parthenogenetic embryos. .................................................. 67

Table 3.2. Knockdown of Sfmbt2 correlates with reduction in the frequency of TS cell derivation. .................................................................................................................................................. 78

Table 4.1. Sfmbt2 Gene Trap is Paternal Lethal at Mid-Gestation in C57BL/6 background...... 93

Table 4.2. Sfmbt2 Gene Trap is Paternal Lethal in CD-1 background but with reduced penetrance. .................................................................................................................................................. 94

Table 5.1. Mass Spectrometry results from coimmunoprecipitated proteins using the anti-SFMBT2 antibody. .................................................................................................................................................. 130
<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Production of anti-SFMBT2 antibody</td>
<td>52</td>
</tr>
<tr>
<td>2.2</td>
<td>Validation of anti-SFMBT2 antibody</td>
<td>56</td>
</tr>
<tr>
<td>3.1</td>
<td>Fertilized and Parthenogenetic TS cells show similarities in morphology and gene expression.</td>
<td>62</td>
</tr>
<tr>
<td>3.2</td>
<td>Imprinting status of PEGs in TS cells.</td>
<td>65</td>
</tr>
<tr>
<td>3.3</td>
<td><em>Sfmbt2</em> exhibits LOI by blastocyst outgrowth stage.</td>
<td>70</td>
</tr>
<tr>
<td>3.4</td>
<td>SFMBT2 exhibits perinuclear staining in the TE of expanded blastocysts.</td>
<td>71</td>
</tr>
<tr>
<td>3.5</td>
<td>SFMBT2 exhibits perinuclear staining in TGCs of the blastocyst outgrowth.</td>
<td>73</td>
</tr>
<tr>
<td>3.6</td>
<td>SFMBT2 exhibits heterogeneous staining in the TS cell lineage.</td>
<td>75</td>
</tr>
<tr>
<td>3.7</td>
<td>Validation of anti-<em>Sfmbt2</em> shRNAs and lentiviral infection of early embryos.</td>
<td>76</td>
</tr>
<tr>
<td>4.1</td>
<td>SFMBT2 shows nuclear staining in various cell-types of the trophoblast origin at mid-gestation.</td>
<td>89</td>
</tr>
<tr>
<td>4.2</td>
<td>Pattern of SFMBT2 staining in the placenta at late-gestation.</td>
<td>90</td>
</tr>
<tr>
<td>4.3</td>
<td>Paternal inheritance of a null <em>Sfmbt2</em> allele is lethal due to defects in proper development of the extraembryonic tissues.</td>
<td>92</td>
</tr>
<tr>
<td>4.4</td>
<td>Paternal inheritance of a null <em>Sfmbt2</em> allele results in reduction of all trophoblast cell types in C57BL/6 background.</td>
<td>96</td>
</tr>
<tr>
<td>4.5</td>
<td>Reduction in trophoblast cell types in not due to Caspase dependent cell death.</td>
<td>99</td>
</tr>
<tr>
<td>4.6</td>
<td>Paternal inheritance of a null <em>Sfmbt2</em> allele results in anemia in the CD-1 background.</td>
<td>100</td>
</tr>
<tr>
<td>4.7</td>
<td>Paternal inheritance of a null <em>Sfmbt2</em> allele results in reduction of various trophoblast subtypes in the CD-1 background.</td>
<td>102</td>
</tr>
<tr>
<td>4.8</td>
<td>SFMBT2 null embryos display abnormalities in the extraembryonic tissues as early as e8.0.</td>
<td>104</td>
</tr>
<tr>
<td>4.9</td>
<td>SFMBT2 null embryos display reduction in various cell-types of the trophoblast origin.</td>
<td>106</td>
</tr>
<tr>
<td>4.10</td>
<td>SFMBT2 null embryos fail to expand cell-types of the trophoblast origin.</td>
<td>107</td>
</tr>
</tbody>
</table>
Figure 4.11. SFMBT2 null embryos show disappearance of tissues of the trophoblast origin by e9.75. 108

Figure 4.12. SFMBT2 null embryos fail to maintain tissues of trophoblast origin. 109

Figure 5.1. SFMBT2 shows cell-type specific localization pattern and association with mitotic chromosomes. 121

Figure 5.2. Over expressed SFMBT2 and YY1 interact in HEK293 cells. 123

Figure 5.3. SFMBT2 interacts with the transcription factors YY1 and TFDP1 in TS cells. 124

Figure 5.4. YY1 staining shows overlap with mitotic chromosomes in TS cells. 126

Figure 5.5. TFDP1 shows co-localization with mitotic chromosomes in TS cells. 128

Figure 6.1. ERR-β binding site at the transcriptional start site of Sfmbt2. 144
List of Appendices

Appendix I. Primer and oligonucleotide sequences ................................................................. 175

Appendix II. Peptide spectrum and probabilities for identification of SMFBT2 (in-gel)......... 179

Appendix III. Peptide spectrum and probabililities for identification of SMFBT2 (in-solution).
............................................................................................................................................... 181

Appendix IV. Peptide spectrum and probabililities for identification of CASZ1 (in-gel ). ....... 182

Appendix V. Peptide spectrum and probabililities for identification of CASZ1 (In-solution) .. 183
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ºC</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td>CDB</td>
<td>Contaminants Data Base</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>ChIRP</td>
<td>Chromatin Isolation by RNA Purification</td>
</tr>
<tr>
<td>ChT</td>
<td>Chorionic Trophoblast</td>
</tr>
<tr>
<td>CpG</td>
<td>Cytosine-Guanine dinucleotide</td>
</tr>
<tr>
<td>C-TGC</td>
<td>Central arterial canal-Trophoblast Giant Cell</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>ddH2O</td>
<td>double distilled water</td>
</tr>
<tr>
<td>DMR</td>
<td>Differentially Methylated Region</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>d-PTGC</td>
<td>Distal-Parietal Trophoblast Giant Cell</td>
</tr>
<tr>
<td>EEE</td>
<td>Extraembryonic Ectoderm</td>
</tr>
<tr>
<td>eGFP</td>
<td>enhance Green Fluorescent Protein</td>
</tr>
<tr>
<td>EPC</td>
<td>Ectoplacental Cone</td>
</tr>
<tr>
<td>ES cells</td>
<td>Embryonic Stem cells</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>H2AK119</td>
<td>Histone H2A Lysine 119</td>
</tr>
<tr>
<td>H3K27</td>
<td>Histone H3 Lysine 27</td>
</tr>
<tr>
<td>H3K36</td>
<td>Histone H3 Lysine 36</td>
</tr>
<tr>
<td>H3K4</td>
<td>Histone H3 Lysine 4</td>
</tr>
<tr>
<td>H3K9</td>
<td>Histone H3 Lysine 9</td>
</tr>
<tr>
<td>H4K20</td>
<td>Histone H4 Lysine 20</td>
</tr>
<tr>
<td>hCG</td>
<td>human Chorionic Gonadotropin</td>
</tr>
<tr>
<td>hr</td>
<td>hour</td>
</tr>
<tr>
<td>ICM</td>
<td>Inner Cell Mass</td>
</tr>
<tr>
<td>ICR</td>
<td>Imprinting Control Region</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosomal Entry Site</td>
</tr>
<tr>
<td>IUGR</td>
<td>Intrauterine Growth Restriction</td>
</tr>
<tr>
<td>kD</td>
<td>kilo Dalton</td>
</tr>
<tr>
<td>LOI</td>
<td>Loss of Imprinting</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen Activated Protein Kinase</td>
</tr>
<tr>
<td>MBT</td>
<td>Malignant Brain Tumour</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple Cloning Site</td>
</tr>
<tr>
<td>me1,me2,me3</td>
<td>monomethylation, dimethylation, trimethylation (in the context of histone tail modifications)</td>
</tr>
<tr>
<td>MEG</td>
<td>Maternally Expressed Gene</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
</tbody>
</table>
MMRRC  Mutant Mouse Regional Resource Center
NPC  Neural Precursor Cells
NSC  Neural Stem Cell
OG  blastocyst Outgrowth
PBS  Phosphate-Buffered Saline
PcG  Polycomb Group
PCR  Polymerase Chain Reaction
PEG  Paternally Expressed Gene
Pho-RC  Pleiohomeotic-Repressive Complex
PMSG  Pregnant Mare’s Serum
PRC  Polycomb Repressive Complex
PRC1  Polycomb Repressive Complex 1
PRC2  Polycomb Repressive Complex 2
PrE  Primitive Endoderm
PRE  Polycomb Response Element
PTGC  Parietal Trophoblast Giant cell
p-PTGC  Proximal-Parietal Trophoblast Giant Cell
PTM  Post Translational Modification
qRT-PCR  quantitative RT-PCR
RNA  Ribonucleic Acid
RT-PCR  Reverse Transcribed-PCR
SAM  Sterile Alpha Motif
SDS-PAGE  Sodium Dodecyl Sulfate-Polyacrylamide gel electrophoresis
sec  second
shRNA  Short hairpin RNA
SNP  Single Nucleotide Polymorphism
SpA-TGC  Spiral arterial TGCs -Trophoblast Giant Cell
SpT  Spongiotrophoblast
S-TGC  Sinusoidal-Trophoblast Giant Cell
SynT  Syncytiotrophoblast Trophoblast
TE  Trophectoderm
TGC  Trophoblast Giant Cell
TrxG  Trithorax Group
TS cells  Trophoblast Stem cells
uNK cell  uterine Natural Killer cell
XEN cell  Extraembryonic Endoderm cell

Note: Throughout this thesis the specified genotype reflects the maternal allele followed by the paternal allele. For example, +/gt indicates that the wild type (+) allele is inherited from the mother and the gene trap (gt) allele is inherited from the father.
Chapter 1: Introduction
The mammalian placenta serves as a site of gas, nutrient and waste exchange between the mother and the fetus with essential roles in the proper growth and survival of the developing embryo (Rossant and Cross, 2001; Watson and Cross, 2005). In addition to anchoring the embryo to the maternal tissue, the placenta provides immune protection and produces hormones that regulate maternal physiology to support the growing embryo (Hsiao and Patterson, 2012). The significance of this underappreciated tissue during the in utero stages of development makes the placenta an important topic of investigation. As such, identification and characterization of genes with critical roles in the development of the placenta are of interest.

Similar to the development of other organs, the placenta is thought to originate from proper differentiation of progenitor cells facilitated by the orchestration of spatial and temporal expression of appropriate genes (Cross, 2000). All cells of an organism however, have the same genetic blueprint; without a proper readout enabled by a process known as epigenetic regulation, multicellular organisms would not exist. Epigenetics is a heritable process mediated by modifications at the DNA level such as DNA methylation, and at the nucleosome level such as covalent modification of histone tails (Bocock & Aagaard-Tillery 2009). Epigenetic regulation of the genome has been implicated in orchestrating a number of vital processes including genomic integrity, cell-cycle regulation, cell-fate determination and organogenesis including placentation (Rice et al., 2002; Schwartz and Pirrotta, 2013; Trojer and Reinberg, 2008; Trojer et al., 2007; Trojer et al., 2007; Yohn et al., 2003). Parent-of-origin specific expression, as a result of a process known as genomic imprinting, is yet another epigenetic phenomenon. In mammals, evolution of imprinting has been correlated with that of placentation suggestive of a significant role for imprinted genes in the development of this organ.
In this chapter, concepts relevant to imprinting and its connection with placentation will be reviewed followed by an overview of murine placental development including the various stages of placentation, with a special focus on intricacies of placental progenitor cells and their derivatives. Finally, PcG proteins and their function in stem cell state and lineage commitment will be discussed.

1.1 IMPRINTING: AN EPIGENETIC INHERITANCE

The process by which heritable epigenetic marks result in parent-of-origin-dependent differential expression of genes is known as genomic imprinting (Hall 1990; Reik et al. 2003). Consequently, a subset of genes are maternally silenced displaying expression only from the paternal allele while others display expression exclusively from the maternal allele.

The concept of parental effects resulting in the elimination of the paternal genome was observed in Coccids in 1940s and was originally termed “chromosome conditioning” (White, 1977). The term “imprinting” was coined in the early 1960s by Helen Crouse, a cytogeneticist who observed selective parent-of-origin-dependent elimination of X-chromosomes in sciarid flies (Crouse, 1960). The discovery of genomic imprinting in mammals occurred in 1984, facilitated by the microsurgical transplantation of pronuclei between fertilized zygotes to create androgenetic (two paternal genomes), gynogenetic (two maternal genomes) and normal embryos, resulting in fetal death at mid-gestation with under-development (gynogenetic) or over proliferation (androgenetic) of trophoblast tissues (Barton et al., 1984; McGrath and Solter, 1984).
1.1.1 **Functional significance of imprinted genes**

Since the discovery of genomic imprinting over 100 imprinted genes have been identified (reviewed in Miri and Varmuza, 2009). Imprinted genes have been implicated in non-pathological conditions such as prenatal growth regulation of both the embryo and the extraembryonic tissues, as well as pathological conditions due to embryonic and extraembryonic developmental defects leading to death (Barton et al., 1984; Clarke et al., 1988; Kaufman et al., 1977). In addition, the discovery of genomic imprinting provided an explanation for some puzzling genetic and cytogenetic disorders in both mice and humans.

1.1.1.1 **Uniparental duplications and disomies in mice**

Uniparental disomy refers to inheritance of an entire homologous chromosome pair from a single parent and uniparental duplication (partial disomy) refers to uniparental inheritance of segments of a chromosome pair (Cattanach et al., 2004). Robertsonian (centric fusion of chromosomes) and reciprocal translocations have been used to generate uniparental disomies and duplications, respectively (Baranov, 1984; Snell, 1946). To this effect, translocation heterozygotes carrying unbalanced gametes were intercrossed to generate balanced diploid offspring with maternal or paternal duplication of partial or whole chromosomes (Searle and Beechey, 1978). Genomic imprinting is thought to play a role when offspring carrying the disomy or duplication exhibits a parent-of-origin specific phenotype, leading to the generation of an “imprinting map” (MouseBook Imprinting Catalog).

Maternal duplications such as MatDp(prox7), MatDp(dist7) (maternal duplication of the proximal region of chromosome 7 and distal region of chromosome 7,
respectively) or disomies such as mUPD12 (maternal uniparental disomy of chromosome 7) were correlated with fetal intrauterine growth restriction (IUGR) and small or abnormal placentae (Cattanach et al., 2004; Ferguson-Smith et al., 1991; Georgiades et al., 2000). Paternal duplications such as PatDp(prox11) (paternal duplication of proximal region of chromosome 11) and disomies such as pUPD11 (paternal uniparental disomy of chromosome 11) had co-occurrences with fetal overgrowth and large placentae (Cattanach et al., 2004). Collectively, these studies suggested that genes expressed exclusively from the paternally inherited chromosome are likely involved in proper development of the placenta.

1.1.1.2 Knockout studies of imprinted genes

Targeted mutations in mice have provided direct evidence of the function of imprinted genes. The consensus from knockout phenotypes of various imprinted genes is that many are involved in various aspects of placentation (Lefebvre, 2012; Miri and Varmuza, 2009; Tunster et al., 2013). Among those are the PEGs, Peg10 and Peg1/Mest, which reside on the proximal region of chromosome 6. Peg10 is one of the earliest acting genes whose mutation results in differentiation defects of various placenta progenitor cells, leading to mid-gestational lethality (Ono et al., 2006). Peg1/Mest deficient embryos however, exhibit mild abnormalities mostly limited to placental weight, and maternal behaviour as well as fetal growth retardation (Lefebvre et al., 1998). The PEGs, Igf2 and Kcnq1ot1, and the maternally expressed genes (MEGs), Ascl2/Mash2, Phlda2 and Cdkn1c residing within the highly studied distal region of chromosome 7 also show a range of placental abnormalities from placental growth restrictions to defects in various placental structures leading to mid- to late -gestational lethality (Constância et al., 2002; Fitzpatrick et al., 2002; Guillemot et al., 1995; Salas et al., 2004; Takahashi et al., 2000).
1.1.2 Regulation of imprinting

The most parsimonious explanation for the generation of imprinting is acquisition during gametogenesis (John and Lefebvre, 2011; Obata and Kono, 2002). Initial studies were technically limited to analysis of DNA methylation, but expansion of experimental tools has allowed analysis of other epigenetic marks, such as histone modification. As the number of imprinted genes expanded, mapping studies revealed that, except for a small number of single-gene imprinted domains, most imprinted genes are found in clusters. Within the same cluster, there are genes displaying preferential paternal expression (mono-allelic PEGs) as well as genes showing maternal specific expression (mono-allelic MEGs). All currently identified imprinting clusters contain protein coding genes as well non-coding RNA while only some of the imprinted single-gene domains correlate with the presence of non-coding RNA (Wan and Bartolomei, 2008). Interspersed within the cluster there may also reside genes expressed from both parental alleles (biallelic). Furthermore, as parent-of-origin-specific expression is tissue specific, some genes exhibit a monoallelic pattern of expression in one tissue and biallelic pattern in another. Imprint regulation of single-gene domains as well as multiple-gene clusters relies on a number of heritable epigenetic modifications including covalent modification of histones, and DNA, as well as non-coding RNAs (Reik et al., 2001; Strahl and Allis, 2000).

Genes within an imprinted cluster are coordinately regulated by an Imprinting Control Region (ICR), also known as imprinting centers in cis (Wan and Bartolomei, 2008). An ICR is a Cytosine-Guanine dinucleotide (CpG) rich DNA found in proximity to imprinted clusters that exhibits differential methylation on the cytosine residue of CpG dinucleotide in a parental-origin-specific manner. Differential methylation of the DNA
was implicated in imprint regulation in the early 1990s by Li and colleagues reporting loss of imprinting (LOI), a condition associated with erasure of epigenetic marks resulting in aberrant expression of imprinted genes, in DNMT1 (maintenance DNA methyltransferase) null embryos (Li et al., 1993). These findings were further consolidated when Sutcliffe and colleagues associated deletion of a differentially methylated region (DMR) with loss of Snrpn expression from the paternally inherited allele in patients with Prader-Willi syndrome (Sutcliffe, 1994). Over the last 20 years a number of reports on various imprinted clusters validate this hypothesis (Edwards and Ferguson-Smith, 2007; John and Lefebvre, 2011; Thorvaldsen et al., 1998; Williamson et al., 2006).

Currently there are two known categories of DMRs, primary or germ-line DMRs and secondary or somatic DMRs. Germline DMRs adopt their differential methylation during gamete development (Obata and Kono, 2002) whereas somatic DMRs do so after fertilization (John and Lefebvre, 2011). It is believed that germline DMRs are master regulators of imprinting and somatic DMRs are dependent on germline DMRs for acquisition of their epigenetic signature (John and Lefebvre, 2011). There are a number of known mechanisms by which germline DMRs exert control over their domains.

1.1.2.1 Germline DMR-controlled, non-coding RNA-induced imprint regulation

Germline DMR-controlled non-coding transcripts are known to play a critical role in regulation of imprinted genes in cis. The non-coding RNA acts to silence surrounding genes by means of cis-acting RNA interference or recruitment of chromatin modifying enzymes to repress the genes within the cluster (Sleutels et al., 2002). An example of this regulatory mechanism is the germline DMR overlapping the promoter region of the non-
coding RNA, Air/Igf2iras, which is associated with the regulation of the Igf2r imprinted cluster. This DMR is unmethylated on the paternally inherited allele, resulting in the transcription of Air, and subsequent silencing of paternally inherited Igf2r by RNA interference in cis (Latos et al., 2012; Sleutels et al., 2002). Furthermore, non-coding RNA facilitated recruitment of histone modifying enzymes such as G9a has been associated with silencing of genes upstream of Air transcript including Slc22a2 and Slc22a3 (Nagano et al., 2008).

1.1.2.2 Germline DMR-controlled, chromatin conformation-facilitated imprint regulation

In this form of imprint regulation, absence or presence of methylation at germline DMRs dictates association or dissociation, respectively, of insulator proteins, such as the zinc-finger protein CTCF, to the chromatin which in turn results in distinct chromatin looping (Bell and Felsenfeld, 2000; Han et al., 2008; Kanduri, 2000). The resultant chromatin conformation facilitates intrachromosomal interactions between genes and chromosomal elements such as enhancers, and consequently results in differential expression of genes involved. An example of such regulation is H19-Igf2 imprinted domain. In this domain, a DMR positioned between Igf2 and H19 contains a binding site for CTCF which exhibits preferential binding to unmethylated sites. Thus, the unmethylated maternally inherited allele recruits CTCF, which in turn facilitates looping of the chromatin in such a way to position Igf2 away from an enhancer resulting in maternal silencing of this gene (Han et al., 2008).
1.1.3 Evolution of imprinting and its correlation with placentation in mammals

In mammals, many imprinted genes exhibit imprinted expression only in the extraembryonic lineage including placenta and yolk sac (Reik et al. 2003) and the most prominent defect of uniparental conceptuses is their inability to support normal placentation (Clarke et al., 1988; Varmuza et al., 1993). Furthermore, parthenogenesis results in lethality only in mammals, while intrinsic as well as induced parthenogenesis is observed in non-mammalian species (Barlow and Bartolomei, 2014; Brevini et al., 2012).

Since late 1980s, there have been several hypotheses for the evolution of imprinting. In 1988, Solter suggested that imprinting was a means to prevent parthenogenetic reproduction in mammals (Solter 1988). In 1990, imprinting was proposed as a way to control placental growth (Hall 1990). One of the most widely accepted hypotheses for the evolution of imprinting was proposed in 1991 known as the “Conflict” hypothesis (Moore & Haig 1991). The antagonistic effect of many of paternally and maternally expressed genes which result in enhancement or suppression of foetal growth respectively, led to propositioning of “Conflict” hypothesis. The acquisition of such parental “arms race” (Reik & Walter 2001) is believed to have evolved over maternal resources; the PEGs are thought to maximize resources for the offspring; meanwhile, the MEGs are believed to result in the reduction and therefore the sharing of maternal resources among all offspring.

In 1994, challenging whether the “battle of the sexes have gone molecular”, Varmuza and Mann proposed “The Ovarian Time Bomb” hypothesis (Sue Varmuza & Mann 1994). It states that imprinting is a maternal effect to ensure that trophoblast
development is prevented in spontaneously activated eggs. Furthermore, it has been highlighted that both overgrowth and growth retardation are pathological conditions, making the “Conflict” hypothesis difficult to rationalize from a fitness perspective (Miri and Varmuza, 2009). Regardless of the underlying causes of evolution of imprinting, these hypotheses share a common denominator; the evolution of imprinting appears to have implications on placental development.

1.2 PLACENTATION

1.2.1 Stages of Placental Development

The embryo is the primary source for the formation of the placenta. In mice, three days after fertilization, the cleavage stage embryo composed of totipotent blastomeres, forms the morula in a process called compaction. At the compaction stage, the blastomeres positioned on the outside are thought to give rise to the trophectoderm layer while the ones on the inside assume the inner cell mass (ICM) fate (Haffner-Krausz et al., 1999; Pedersen, 1987). At embryonic day 3.5 the first known differentiation event takes place generating two distinct lineages: the ICM, which contributes to the embryo-proper; and a simple epithelium, and the trophectoderm (TE), which primarily gives rise to the extraembryonic tissues of the trophoblast origin mainly the placenta (Cross, 2000). This differentiation event results in the formation of a blastocyst, with the trophectoderm encompassing a fluid-filled cavity called a blastocoel, and the ICM attached to one side. The second lineage specification resulting in the formation of the primitive endoderm (PrE) also occurs around day 3.5 of development (Albert & Peters 2009). PrE cells are
thought to be derivatives of the ICM and form in direct apposition to the epiblast which is the component of the ICM that gives rise to the embryo. PrE cells form extraembryonic endoderm layers of visceral and parietal yolk sacs, and are thought to be important in providing signals for the proper establishment of embryonic axes.

In mice implantation of the blastocyst follows hatching of the embryo from the zona pellucida at embryonic day 4.5 (Chai et al., 1998; Copp, 1979; Ilgren, 1983; Wilder et al., 1997a). A population of TE cells overlaying the ICM, known as the polar TE, preserve their diploid state and continue proliferating. The polar TE is thought to harbor trophoblast progenitor cells that will give rise to the extraembryonic ectoderm, and ectoplacental cone (EPC) later in gestation. The second population of TE cells known as the mural TE which are separated from the ICM by the blastocoel cavity, stop dividing and exhibit endoreduplication of their DNA, giving rise to polyploid primary trophoblast giant cells (TGCs) (Cross et al., 1994; Ouseph et al., 2012; Takahashi et al., 2000; Varmuza et al., 1988). In mice attachment to the uterine epithelium is facilitated by the mural TE.

The fibroblast growth factor (FGF) signalling pathway has been implicated in the crosstalk between the ICM and the overlaying TE. Fgf4 is expressed in the ICM at the blastocyst stage as well as the derivatives of these cells post-implantation (Chai et al., 1998). Fgf4 knockout mice survive up to the peri-implantation stage (Feldman et al., 1995) despite their ability to give rise to embryonic stem cells (Wilder et al., 1997b); the observed lethality was therefore attributed to inability to retain a proliferative TE. The FGF receptor, FGFR2, is expressed in peripheral blastomeres at the morula stage of development and maintains a high level of expression in polar TE at the blastocyst stage (Haffner-Krausz et al., 1999). Targeted mutations of Fgfr2 have been reported to result
in peri-implantation lethality (Haffner-Krausz et al., 1999) and mid-gestational lethality due to failure to maintain the trophoblast tissues (Xu et al., 1998). The role of FGF signalling in maintenance of the proliferative trophoblast was demonstrated by derivation of trophoblast progenitors, TS cells, in vitro in the presence of FGF4 and their differentiation upon withdrawal of this growth factor (Tanaka et al., 1998a).

By day 7.5 of development, extraembryonic mesoderm lines the visceral endoderm and the extraembryonic ectoderm interiorly to form the visceral yolk sac and the chorionic plate, respectively; and the allantois derived from mesoderm at the posterior end of the fetus extends into the exocoelom (Rossant and Cross, 2001). Around embryonic day 8.5, formation of components of placental vascular network is initiated by attachment of the chorionic plate and the allantois (Stecca et al., 2002). The fusion of the allantois to the chorionic plate is followed by the folding of the chorion, forming villi where embryonic blood vessels from the allantois develop. This process marks the initial stages of branching morphogenesis leading to the development of the labyrinthine layer where most of the exchange between the mother and fetus takes place (Cross et al., 2002). The labyrinth forms as the chorionic cells differentiate, giving rise to two cell-types: two layers of multinucleated syncytiotrophoblast trophoblast (SynT) cells, generated by elongation and fusion of post-mitotic chorionic trophoblast cells which juxta pose endothelium of foetal blood vessels; and mononucleated polyploid cells known as sinusoidal TGCs (S-TGC) which line the maternal blood (Stecca et al., 2002).

During its development the labyrinth receives structural support from the spongiotrophoblast (SpT) derived from the EPC (Hughes et al., 2004; Riley et al., 1998; Rossant and Cross, 2001). Compact SpT is positioned between the labyrinth and the outer layer of TGCs and contributes to the formation of the junctional zone. TGCs at the
edge of the junctional zone, overlaying the SpT on one side and juxtaposing the maternal cells on the other are known as secondary TGCs which are thought to form from the differentiation of the outer cell layer of the EPC. At mid-gestation, trophoblast cells invade both the decidua as well as the maternal spiral arteries (Adamson et al., 2002; Cross et al., 2002). It is at this stage that the endothelium lining the maternal blood spaces are replaced by trophoblast cells of the fetal origin, hence forming the hemochorial placenta. Also residing within the junctional zone are the trophoblast derived glycogen cells (Bouillot et al., 2006). These glycogen-rich, vacuolated cells are detected as early as embryonic day 12.5, reaching their maximum number by embryonic day 16.5 at which point their expansion is halted (Bouillot et al., 2006; Coan et al., 2006). By the end of gestation, following a lytic phase glycogen cells form large glycogen filled lacunae thought to serve as an energy source during the final stages of gestation and delivery (Bouillot et al., 2006).

1.2.2 Trophoblast cell types: Specification, Maintenance and Differentiation

TS cells are thought to be a population of multipotent, self-renewing progenitor cells that can give rise to various cell-types of the placenta through differentiation (Watson and Cross, 2005). TS cells have been derived in vitro from blastocysts as well as the extraembryonic ectoderm and the chorion, isolated from day 6.5 and 8.5 conceptuses, respectively (Tanaka et al., 1998a; Uy et al., 2002). These progenitor cells maintain multipotency in vitro when cultured in presence of FGF4 and can differentiate into various trophoblast cell-types upon removal of this protein.

1.2.2.1 TS cells: Specification

The earliest acting transcription factors implicated in the establishment of trophoblast
lineage are TEAD4 and CDX2. Mutation in *Tead4* results in lethality before the blastocyst stage (Nishioka et al., 2008) while *Cdx2* mutants (Strumpf et al., 2005) live up to early blastocyst stage but fail to implant, suggestive of TEAD4 acting upstream of CDX2 in the trophoblast specification signalling pathway. Further evidence for the involvement of these transcription factors comes from the commitment switch from embryonic stem (ES) cells to TS cells whereby ectopic expression of CDX2 and active TEAD4 in ES cells resulted in transition to trophectoderm (Nishioka et al., 2009; Niwa et al., 2005). Furthermore, CDX2 is believed to achieve TE specification by suppression of pluripotency factors such as OCT4 and NANOG (Strumpf et al., 2005). Specification of trophoblast fate has been linked to the position of the blastomeres as early as 8-cell stage (Marikawa and Alarcon, 2009; Nishioka et al., 2009; Suwińska et al., 2008). There are currently two signalling pathways implicated in the specification of TE lineage: the Hippo signalling pathway (Nishioka et al., 2009) and Ras-mitogen activated protein kinase (MAPK) signalling pathway (Lu et al., 2008).

Inactivation of the Hippo signalling pathway has been implicated in activation of TEAD4 and initiation of TE specification cascade in peripheral blastomeres but not inner blastomeres (Marikawa and Alarcon, 2009; Nishioka et al., 2009; Senner and Hemberger, 2010). Inner blastomeres display an active Hippo pathway whereby the Ser/Thr kinase protein LATS phosphorylates YAP (Yes kinase-associated protein 1), a co-activator of TEAD4, sequestering it in the cytoplasm and therefore inhibiting the activation of *Cdx2* and downstream genes required for establishment and maintenance of the TE lineage. In contrast, outer blastomeres, exhibiting an inactive Hippo pathway, fail to phosphorylate Yap which in turn localizes to the nucleus, binding to its coactivator TEAD4 and initiating cascades of event in TE fate-specification by activation of *Cdx2*. 

14
Involvement of the Ras-MAPK signalling pathway in TE lineage specification comes from experiments conducted by Lu and colleagues where conditional activation of this signalling pathway in ES cells resulted in the acquisition of the trophoblast fate (Lu et al., 2008). As such, Ras-MAPK-induced TS cells were able to contribute exclusively to the placenta. In addition, inhibition of this pathway was shown to result in reduced Cdx2 expression and compromised trophoblast outgrowths. Furthermore, Erk2, a downstream effector of this pathway was shown to localize to the apical surface of outer blastomeres of the 8-cell stage embryo, suggesting that Ras-MAPK pathway is involved in TE specification.

1.2.2.2 TS cells: Maintenance

TE specification is followed by expansion and maintenance of the TS cell niche which is thought to reside within the polar TE (Senner and Hemberger, 2010). Several transcription factors implicated in this process have also been associated with self-renewal of ES cells, suggesting that some of the players involved in maintenance of cell fate may act globally. Expressed in TE cells of the blastocyst, Eomes is the earliest acting transcription factor identified to date with such functions. A null mutation in this gene results in arrest at the blastocyst stage (Russ et al., 2000; Strumpf et al., 2005). Despite retaining expression of Cdx2 and Fgfr2, mutants showed undetectable levels of differentiated trophoblast markers such as Pl-I and Hand1 (Strumpf et al., 2005). In addition, homozygous null embryos of Eomes fail to form blastocyst outgrowths and TS cells.

Other transcription factors involved in maintenance of TE lineage include Tcfap2c/AP-2gamma, Ets2, Elf5 and Err-β. Both Tcfap2c (Auman et al., 2002; Werling and Schorle, 2002) and Ets2 (Yamamoto et al., 1998) mutants die around day 8.5 of gestation.
and display abnormalities in the TE derivatives. TCFAP2C appears to operate independently of CDX2 in a distinct signalling pathway, as Cdx2 deficient ES cells have been shown to acquire expression of TS cell-type markers, except for Elf5, upon induction of TCFAP2C (Kuckenberg et al., 2010). Expression of Elf5 in induced TS cells was shown to require both TCFAP2C and CDX2. Elf5 mutants also show defects in extraembryonic tissues and die by embryonic day 10.5 (Donnison et al., 2005). The transcription factor Estrogen-receptor-related receptor-β (ERR-β) belongs to the nuclear receptor superfamily (Giguère et al., 1988). This orphan nuclear receptor binds to estrogen response elements; however, its activation is not dependent on estrogen. Err-β mutants show severe defects in the development of extraembryonic tissues and die by day 10.5 of gestation (Luo et al., 1997).

Knockdown of the chromatin remodelling factor, SMARCA4 at the blastocyst stage results in derepression of Oct4 in the TE (Wang et al., 2010a). SMARCA4 was shown to coimmunoprecipitate with CDX2, hence the observed derepression was suggested to be due to failure in CDX2 recruitment. Interestingly, SMARCA4 has also been shown to occupy promoters of various pluripotency factors such as Oct4, Sox2 and Nanog, and participates in regulation of self-renewal in ES cells (Kidder et al., 2009). In concurrence, ChIP (Chromatin Immunoprecipitation)-chip data suggests binding of SMARCA4 and other transcription factors such as GATA3, ETS2, EOMES, TCFAP2C to a number of pluripotency associated genes (Kidder and Palmer, 2010). Furthermore, co-occupancy of a significant number of genes by TCFAP2C along with EOMES and SMARCA4 has been reported, suggestive of their cooperative regulation.

Mash2/Ascl2 belongs to basic helix-loop-helix (bHLH) family of transcription factors. A null mutation in this gene results in mid-gestational embryonic lethality due to
loss of both the EPC and the SpT cells where the gene is highly expressed. The authors also reported an increase in the number of TGCs, suggesting that MASH2/ASCL2 is required for maintenance of TGC precursors by preventing their premature differentiation (Guillemot et al., 1994; Scott et al., 2000; Tanaka et al., 1997). Hypoxia inducible factors (HIFs), HIF1α and HIF2α, have also been implicated in maintenance of the SpT cells. Double mutants die by embryonic day 10 and show reduced Mash/Ascl2 expression and increased TGCs (Cowden Dahl et al., 2005). HIFβ/ARNT which is thought to form a heterodimer with HIF1α or HIF2α also results in similar phenotypes when mutated (Kozak et al., 1997). TS cells in vitro, lacking HIF1α and HIF2α or HIFβ were shown to differentiate into SynT cells upon withdrawal of FGF4 while their wild type counterparts showed acquisition of the SpT fate before fully differentiating into TGCs (Cowden Dahl et al., 2005).

1.2.2.3 TS cells: Differentiation

There are currently two main classes of TGCs, primary and secondary. Originated from the TE layer of the growing blastocyst, primary TGCs are the first terminally differentiated cells that arise in development. Secondary TGCs are thought to arise from the EPC and result in various TGC subtypes discussed below. As their name implies TGCs have large nuclei that have undergone endoreduplication whereby genome replication is not subsequently followed by cell division (endomitosis) resulting in polyploidy (Gardner et al., 1973; Ouseph et al., 2012; Sherman et al., 1972; Takahashi et al., 2000; Varmuza et al., 1988; Zybina and Zybina, 1996). The significance of the observed polyploidy has been speculated to be linked with increased protein synthesis, while the post-mitotic nature of the TGCs which are intrinsically invasive is thought to serve as a preventative mechanism from tumor formation (Hu and Cross, 2010a).
1.2.2.4 Heterogeneity of TGCs and their perspective roles

A subclass of TGCs are the parietal TGCs (PTGCs) which have acquired their title due to their spatial localization; as such, PTGC outline the border between the maternal and the fetal tissues. Primary PTGCs which arise as early as the blastocyst stage of development aid in the implantation process. Although the receptive uterus is maternally primed by estrogen and progesterone, PTGCs also produce progesterone and are necessary to initiate decidualization (Bany and Cross, 2006; Zybina et al., 2000).

The mature placenta is thought to take form by day 10.5 of gestation, at which point the embryo becomes reliant on this tissue for sustenance. Before the development of the mature placenta, PTGC juxtaposing the maternal tissue are anastomosed by the maternal blood, facilitating nutrient and gas exchange through the parietal yolk sac (Simmons et al, 2007). Furthermore, both primary and secondary PTGCs express a number of paracrine and endocrine factors regulating maternal adaptation to the developing conceptus (Table 1.1). These include Placental lactogen-I (PL-I) expressed until mid-gestation; and Placental lactogen-II (PL-II) expressed from mid-gestation to term (Hu and Cross, 2010a; Jackson and Linzer, 1997; Simmons et al., 2007; Simmons et al., 2008). In the absence of the pituitary gland in pregnant female mice compensatory elevation of these two hormones from the fetus has been shown, resulting in some lactation (Hu and Cross, 2010a; Thordarson et al., 1989). In addition, to accommodate the growing fetus, rapid proliferation of decidua cells juxtaposing the PTGCs has been correlated with Plf expression from these giant cells (Nelson et al., 1995). PTGCs have also been implicated in remodeling of maternal spiral arteries. Loss of smooth muscle has been associated with dilation of spiral arteries increasing the maternal blood supply to the growing fetus (Cross et al., 2002; Müller et al., 1999; Simmons et al., 2007). PTGCs
Table 1.1. Expression profile of select markers of TGCs in various trophoblast
cell-types.

<table>
<thead>
<tr>
<th>Protein</th>
<th>In vitro</th>
<th>In vivo: Placenta</th>
<th>Period of expression</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PL-I (Prl3d1)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HAND1</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PRP (Prl7d1)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PLF (Prl2c2)</td>
<td>-</td>
<td>-</td>
<td>+/–</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>PLP-E (Prl7a1)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>PLP-A (Prl4a1)</td>
<td>+</td>
<td>-</td>
<td>+/–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PL-II (Prl3b1)</td>
<td>-</td>
<td>-</td>
<td>+/–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CTSQ</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PLP-F (Prl7a2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

+, expressed; -, not expressed; +/-, expressed in some
express Prolactin-like protein A (PLP-A) shown to regulate uterine natural killer (uNK) cells located within the metrial triangle, a region within the decidua juxtaposing the placenta, which in turn mediate loss of maternal spiral arterial smooth muscle (Adamson et al., 2002; Müller et al., 1999).

In addition to PTGCs, there are at least three subclasses of secondary TGCs with distinct expression profiles, spatial, temporal and ultimately functional characteristics including sinusoidal (S-TGC), central arterial canal (C-TGC), and spiral arterial (SpA-TGC) TGCs (Simmons et al., 2007). Secondary PTGCs are thought to arise from cells of the EPC and later the SpT cells, most of which exhibit *Tphpa* expression. However, secondary PTGCs appear to have a mixture of cells from *Tphpa* expressing precursors as well as those that do not. Simmons and colleagues have suggested that a subset of PTGCs may arise from cells of the EPC closest to the chorion where there is no detectable expression of the *Tphpa* (Simmons et al., 2007). Similar to PTGCs, C-TGCs also appear to have a mixture of *Tphpa* positive and negative precursors. SpA-TGCs appear to be almost entirely of *Tphpa* expressing precursors while S-TGCs are exclusively from the *Tphpa* negative progenitors.

Spatial localization in conjunction with double staining of TGCs with various markers and DAPI as a measure of nuclear area shows correlation between the size of the nucleus and the type of TGC (Simmons et al., 2007). In comparison to diploid trophoblast cells, S-TGC, C-TGC, SpA-TGC and PTGC have larger nuclear area as well as estimated DNA. Of the subclasses, PTGCs possess significantly larger nuclei and higher DNA content in comparison with S-TGC, C-TGC and SpA-TGC, the three of which have similar nuclear area and DNA content. Spatial organization in conjunction with *in situ* hybridization studies and immunostaining experiments of various hormones
and cytokines have also been used to shed light on the function of various TGCs (Table 1.1). SpA-TGCs line the maternal spiral arteries and facilitate their remodelling to allow flow of maternal blood into the placenta (Hu and Cross, 2010b; Simmons et al., 2007). They express PLP-A known to modulate uNK cells as well as secrete angiogenic factors such as Proliferin (PLF) and anti-angiogenic factors such as Proliferin-related protein (PRP). C-TGC refers to TGCs lining the central arterial canal; these cells have also been shown to express the angiogenic factor PLF and the Placental Lactogen protein PL-II. Finally, S-TGCs which line the maternal blood space within the labyrinth express the protease CTSQ which has been implicated in cleavage of placental lactogens such as PL-II and regulating their function before entry into fetal and maternal circulation (Ishida et al., 2004).

1.3 **EPigenetic regulators of the genome: Polycomb group proteins**

The active or repressed state of genes is believed to determine cell fate and the preservation of this state is thought to be necessary to stabilize differentiation of cell lineages through many rounds of cell division (Gould 1997). PcG and trithorax group (TrxG) proteins are thought to be major anchors of state preservation, otherwise known as cellular memory (Mahmoudi and Verrijzer, 2001). PcG proteins are thought to maintain the repressed state of genes whereas the antagonistic TrxGs are responsible for maintenance of the active state (Ringrose & Paro 2007). The regulatory role and functional significance of TrxGs are beyond the scope of this thesis and will not be further discussed.
PcG proteins received their title from early experiments linking their function to repression of the second and third thoracic male sex combs in *Drosophila* (Lewis, 1978; Struhl, 1981). It was later shown that PcG proteins are required for maintenance of silencing of *Hox* genes, involved in the organization of body plan and embryogenesis (Soshnikova & Duboule 2008). PcG repressor proteins are highly conserved in eukaryotes, accentuating their functional significance. In this section various PcG proteins and the multimeric protein complexes they participate in will be reviewed. Functional significance of these complexes will be discussed with a special focus on relevance to stem cell biology.

Two of the most extensively studied PcG complexes include the Polycomb Repressive Complex 1 (PRC1) and the Polycomb Repressive Complex 2 (PRC2) which have been shown to operate collaboratively to repress target genes (Margueron and Reinberg, 2011). Nevertheless, numerous studies suggest that not all components of what is known as the canonical PRC1 and PRC2 are required for silencing at every locus. Furthermore, a number of variants containing some of the core PRC components as well as other proteins have been identified, suggesting target specific heterogeneity of these complexes (Reviewed in Lanzuolo and Orlando, 2012; Schwartz and Pirrotta, 2013). In *Drosophila*, while PRC2 initiates catalysis of histone H3 lysine 27 trimethylation (H3K27me3) within domains targeted for repression, PRC1 is thought to maintain the silent state by association with H3K27me3 and subsequent monoubiquitination of histone H2A lysine 119 (H2AK119) (Ringrose & Paro 2007).

### 1.3.1 Polycomb Repressive Complex 1 (PRC1)

In *Drosophila*, PRC1 consists of four core PcG proteins: Polycomb (Pc), Posterior sex
combs (Psc), Polyhomeotic (Ph), and dRing/Sex combs extra (Ring/Sce) (Shao et al., 1999). Heterodimerization of dRing and Psc enhances the E3 ubiquitin ligase activity of dRing which is known to target H2A for ubiquitylation, a post-translational modification (PTM) associated with the silent chromatin (Francis et al., 2004; Wang et al., 2004). The heterodimer also acts as the nucleation site facilitating the assembly of the remaining components of the PRC1. Pc is a chromodomain containing protein thought to bind to H3K27me3 (Fischle et al., 2003). Sex comb on the midleg (Scm) containing two malignant brain tumour (MBT) domains and a sterile alpha motif (SAM), also known as the pointed domain or SPM domain, has also been shown to interact with the core PRC1 through association with Ph and plays a role in PRC recruitment to certain target genes (Wang et al., 2010b). An additional Drosophila repressive complex called dRING-associated factors (dRAF) has also been linked with H2A ubiquitylation. This complex is composed of dRing1, Psc, RYBP, and Kdm2, a histone demethylase targeting H3 lysine 36 (H3K36).

Due to whole genome duplication through vertebrate evolution, there are multiple mammalian orthologues for PcG genes; some act interchangeably while others are associated with distinct biological roles, generating at least six PRC1 variants (Gao et al., 2012). All variants have been shown to contain: one of the two mammalian orthologues of the Drosophila dRing, namely RING1/RING1A and RING2/RING1B which act interchangeably; one of the six mammalian orthologues of the Drosophila Psc, known collectively as the polycomb group RING fingers (PCGFs); and a set of unique proteins (Gao et al., 2012; Wu et al., 2013). One example is the PRC1 variant consisting minimally of RING1B, PCGF1/NSPC1 and KDM2B/FBXL10 (Wu et al., 2013, 1). The histone demethylase KDM2B, binds to unmethylated CpG islands in the promoter region.
of target genes, and demethylates histone H3K36 (Farcas et al., 2012). KDM2B has also been shown to regulate RING1B recruitment which in turn facilitates the ubiquitination of H2AK119 and subsequently silences genes. ES cells lacking KDM2B show defects in embryonic stem cell maintenance and proper differentiation suggesting that this non-canonical PRC1 variant plays a significant role in lineage commitment (He et al., 2013; Wu et al., 2013).

1.3.2 Polycomb Repressive Complex 2 (PRC2)

The *Drosophila* PRC2 contains three core polycomb proteins: Extra sex comb (Esc) or its homologue Escl which function interchangeably (Ohno et al., 2008), Enhancer of zest (E(z)), and Suppressor of zest 12 (Su(z)12). E(z) is a methyltransferase shown to di- or tri-methylate H3K27 through its SET domain (Nekrasov et al., 2007). Of these modifications tri-methylation of histone H3K27 has been reported as the stable silencing mark. E(z) also acts as the nucleation center for recruitment of the remaining PRC2 components such as Su(z)12 and Esc which have been shown to enhance the methyltransferase activity of E(z) in flies. The mammalian orthologues of E(z) include EZH1 and EZH2/KMT6, two highly homologous proteins with distinct tissue specific expression patterns (Margueron et al., 2008) and SUZ12 is the mammalian orthologue of the *Drosophila* Su(z)12. The mammalian orthologue of Esc is embryonic ectoderm development (EED) protein with four isoforms each exhibiting distinct substrate specificity (Kuzmichev et al., 2005). Additional proteins associated with the PRC2 include Nurf55/Caf1 with putative roles in chromatin remodeling and the zinc-finger protein Jing which is a non-essential member of the PRC2 with roles in complex stabilization. There are two mammalian orthologues of Caf1, including the histone
chaperone protein RBAP48/RBBP4 and RBBP7/RBAP46 and one mammalian orthologue of Jing called AEBP2, also a zinc finger protein. The *Drosophila* Polycomb-like (PCL) has also been associated with the PRC2 components. There are three tissue specific mammalian orthologues of PCL namely, PCL1/PHF1, PCL2/MTF2 and PCL3/PHF19 which have been shown to interact with EZH2, SUZ12 and stimulate PRC2 enzymatic activity and are reportedly responsible for the majority of H3K27me3 (Nekrasov et al., 2007; Sarma et al., 2008).

### 1.3.3 Other Polycomb repressive complexes

Pleiohomeotic-repressive complex (Pho-RC) and Polycomb repressive deubiquitinase (PR-DUB) are the latest additions to the repressive complexes containing PcG proteins (Otte & Kwaks 2003; Klymenko *et al.* 2006; Scheuermann *et al.*, 2010). PR-DUB is composed of the PcG protein Ubiquitin carboxy-terminal hydrolase (BAP1) encoded by the PcG gene *Calypso*, and the PcG protein ASX (Scheuermann *et al.*, 2010). PR-DUB complex has been shown to play a role in Hox gene repression through H2A de-ubiquitination.

The *Drosophila* Pho-RC is composed minimally of Pho or Pho-like (Phol) and *Drosophila* Sfmbt (dSfmbt). Pho-RC binds to DNA via the transcription factor Pho in a sequence-specific manner; notably, the only known PcG proteins with sequence-specific binding activity are Pho and Pho-like (Klymenko *et al.* 2006). Pho-RC has also been shown to associate with methylated histones (H3K9me1/2, H4K20me1/2) through the MBT domains of dSfmbt (Wu *et al.*, 2007a). The mammalian orthologues of the dSfmbt include SFMBT1 and SFMBT2; each containing four MBT domains (from l(3)malignant brain tumour in *Drosophila*) and a SAM domain.
Several studies have aimed to identify the mammalian counterpart of the Pho-RC, with mixed results. Both mouse and human SFMBT2 have been shown to interact with the transcription factor YY1, the mammalian orthologue of Pho, if over expressed (Kuzmin et al., 2008; Lee et al., 2013). However mass spectrometry analysis of affinity purified tagged SFMBT1 and SFMBT2 expressed in HEK293 cells failed to identify YY1 as one of the interacting partners (Zhang et al., 2013). SFMBT1 was shown to interact with KDM1A/LSD1, and CoREST as well as a number of PcG proteins such as L3MBTL3 and other SAM domain containing proteins; whereas, SFMBT2 was co-precipitated with a set of distinct proteins including the transcription factor CASZ1, ubiquitin-conjugating enzyme E2O (UBE2O), and PHC1B as well as a few common proteins such as L3MBTL3 and additional SAM domain containing proteins. Despite their homology, the interaction of SFMBT1 and SFMBT2 with distinct sets of proteins has been suggested to be the result of the divergence of these PcG proteins leading to distinct functional specificity.

1.3.4 PcG protein recruitment: suggestive of functional diversity

In Drosophila, PcG proteins are recruited to their target genes by binding to cis-acting genomic sequences called Polycomb Response Elements (PREs) (Ringrose and Paro, 2007; Schwartz et al., 2006). Consequently, deletion of PREs has been correlated with Hox gene derepression and homeotic phenotypes. Global analysis of PRC1 and PRC2 binding however, shows recruitment to additional domains throughout the genome suggestive of a sequence independent recruitment mechanism (Ku et al., 2008; Tolhuis et al., 2006). In this context, PRC2 has been shown to bind nucleosomes and regulate its own enzymatic activity dictated by pre-existing histone modifications independent of
DNA sequence (Schmitges et al., 2011).

In mammals, the cis-acting elements targeted by mammalian PcG proteins have been identified as stretches of CpG rich regions termed CpG islands found predominantly in promoters of genes associated with key developmental processes (Ku et al., 2008; Sharif et al., 2013). In Drosophila, most PRC1 and PRC2 binding sites show overlap with Pho-RC (Oktaba et al., 2008). In concurrence with this observation, Pcl has been shown to facilitate binding of PRC2 components namely E(z) to PREs of target genes by association with the transcription factor Pho and/or Phol (Savla et al., 2008). In mammals, although the mechanism anchoring the PRC1 and PRC2 variants to their target sequences is not fully elucidated, RYBP has been shown to interact with the transcription factor YY1 and RING1A/B suggesting that it may play a role in mediating the interaction between YY1 and the canonical PRC1 (Kalenik et al., 1997; Wilkinson et al., 2010). Nevertheless, a recent proteomic study has failed to confirm the interaction of YY1 with other PRC1 components (Gao et al., 2012).

In mammals, noncoding-RNAs have also been implicated in facilitating PcG recruitment to target genes in cis and in trans. In female placental mammals, one of the X-chromosomes is silenced through a process known as dosage compensation. Xist, a non-coding RNA, coats the X-chromosome in cis leading to recruitment of PcG proteins, accumulation of H3K27 methylation, heterochromatinization and consequently X-inactivation (Plath et al., 2003). EZH2, a component of PRC2 has been shown to bind the Xist transcripts to initialize the heterochromatinization. However, PRC1 has also been shown to be sufficient in maintenance of X-inactivation in embryonic stem cells in a PRC2 independent fashion (Schoeftner et al., 2006). HOTAIR is another non-coding RNA that facilitates recruitment of PcGs to target genes, albeit in trans. HOTAIR,
transcribed from the *HoxC* gene, has been shown to serve as a scaffold to tether PRC2 and CoREST repressor complex to the HOXD locus and subsequently heterochromatinize and silence the gene (Tsai et al., 2010). Chromatin isolation by RNA purification (ChIRP) followed by deep sequencing experiments performed in cancer cells show association of HOTAIR with approximately 800 genomic sites suggesting that non-coding RNA-facilitated PcG protein recruitment is a global event (Chu et al., 2011).

1.3.5 **PcG proteins and cell fate**

PcG proteins have been identified as gatekeepers of pluripotency and cellular differentiation through regulation of gene expression. Pluripotency is achieved by the plasticity associated with a number of key developmental regulators exhibiting bivalent chromatin marks consisting of the silencing H3K27me3 as well as the activating H3K4me3 modifications, co-decorating the corresponding nucleosomes (Bernstein et al., 2006; Pan et al., 2007; Zhao et al., 2007). Such bivalent epigenetic marks are thought to maintain genes in a poised state that allows subsequent lineage appropriate activation or stable repression of the key developmental regulators leading to proper differentiation and organogenesis. Upon differentiation of ES cells into neural precursor cells (NPC), a subset of genes with functions related to the NPC lineage lose their H3K27me3 mark, retaining the active H3K4me3 modification and exhibit a significant increase in their transcription. Genes with roles in unrelated lineages however were shown to retain the repressive H3K27me3 mark and lose the activating histone modifications (Mikkelsen et al., 2007). In concurrence with the above observations, knockout and knockdown studies implicate PcG proteins in maintenance of ES cell identity. To this effect, depletion of various components of PRC1 and PRC2 including RING1B (Leeb and Wutz, 2007),
BMI1 (Cui et al., 2006), SUZ12 (Pasini et al., 2007), JARID2 (Landeira et al., 2010), PCL2 (Walker et al., 2010), in ES cells has been associated with differentiation defects due to misregulation of various developmental cues.

1.4 **THESIS HYPOTHESIS AND OBJECTIVES**

In *utero* development requires the participation of both female and male genomes due to the phenomenon of genomic imprinting which results in parent-of-origin specific expression of a subset of genes. Consequently, parthenogenetic murine embryos (bimaternal genomes) die early in development and exhibit a hypoplastic trophoblast suggesting that a paternally expressed imprinted gene with essential functions in early development of this lineage is responsible for the early phenotypes.

The **first objective** of my thesis, discussed in Chapter 3, was to survey PEGs for a number of criteria to identify the most likely candidate whose loss-of-function in bimaternal embryos leads to the early defects in the trophoblast lineage. As such, the PcG protein SFMBT2, exclusively imprinted in the extraembryonic tissues where the gene is robustly expressed, was identified as a putative candidate.

The **second objective** was to evaluate the selected candidate, *Sfmbt2*, in loss-of-function studies both *in vitro* and *in vivo*. The *in vitro* aspect of this objective was achieved by knockdown of *Sfmbt2* in pre-implantation embryos followed by assessment of their ability to generate TS cells. To this effect, in the final section of Chapter 3, I show that lentivirus induced knockdown of *Sfmbt2* results in a significant reduction in the derivation of TS cells *in vitro*. The *in vivo* aspect of this objective was achieved by extensive analysis of the phenotypes associated with the null mutation of *Sfmbt2*. In
Chapter 4, I show that mutants exhibit early mid-gestational lethality due to failure to maintain the trophoblast lineage, suggesting that Sfmbt2 is required for *in vivo* development of this tissue.

The **third objective** was to further characterize the function of the candidate, SFMBT2, by investigating its interacting partners in trophoblast progenitor cells. This objective was achieved by coimmunoprecipitation and either mass spectrometry or western blot analysis, leading to the identification of three transcription factors including YY1 and CASZ1 which were previously shown to interact with SFMBT2 in HEK293 cells, and TFDP1, whose mutation has been shown to result in defects in trophoblast development.
Chapter 2: Materials and Methods
2.1 **Tissue Culture**

2.1.1 **Pre-implantation Embryos: Harvest and culture**

2.1.1.1 *Superovulation and pre-implantation embryo harvest*

CD-1 female mice of 6 to 8 weeks of age were injected intraperitoneally with 5IU of Gonadotropin from Pregnant mare’s serum (PMSG) (Sigma; G4877) between 3:30 and 4:30pm followed by 5IU of human Chorionic Gonadotropin (hCG) (Sigma; C1063) 46 hrs later and mated with CD-1 stud males. Two days after detection of a vaginal plug, females were sacrificed by cervical dislocation. The uterine horns with the oviducts still intact were dissected and transferred into M2 medium (EMD Millipore; MR-015-D). Morulae were harvested from the oviducts by flushing as described previously (Nagy, 2002). Embryos were rinsed in two changes of KSOM (EMD Millipore; MR-106-D) to remove any residual M2 medium before being transferred to a culture dish containing pre-equilibrated KSOM and cultured in a humidified incubator at 37ºC at 5% CO₂.

Morulae were either cultured for blastocyst outgrowth formation, TS cell derivation, or further manipulated by lentiviral infection, as described below.

2.1.1.2 *Superovulation and Parthenogenetic activation of oocytes*

CD-1 female mice of 6 to 8 weeks of age were injected intraperitoneally with 5IU of PMSG between 3:30 and 4:30pm followed by 5IU of hCG 46 hrs later. Mice were sacrificed by cervical dislocation 18 hours post-hCG. Oviducts were transferred to M2 medium and ruptured to release cumulus masses containing eggs. Cumulus masses were washed in pre-equilibrated KSOM and activated in 7% ethanol, also prepared in pre-equilibrated KSOM, for 5min. Cumulus masses were then thoroughly washed in five changes of KSOM to remove ethanol and transferred to KSOM supplemented with
Cytochalasin-B (10µg/ml) for 5 hrs to inhibit the second meiotic division. Following the incubation, cumulus masses were washed extensively in KSOM followed by removal of cumulus cells using Hyaluronidase (100 IU/ml). The cumulus cell-free eggs were washed in 3 to 4 changes of KSOM and cultured in a humidified incubator at 37ºC with 5% CO₂ for one hour to recover before additional manipulations. Activated oocytes were observed using phase-contract microscopy in order to select oocytes with two visible pronuclei, which were subsequently transferred to fresh pre-equilibrated KSOM and cultured in a humidified incubator at 37ºC with 5% CO₂. Embryos were cultured for blastocyst outgrowth formation as discussed below.

2.1.2 Blastoocyst outgrowth and TS cell derivation

For blastocyst outgrowth formation, day 3.5 blastocysts were cultured in the presence of TS medium (described below) until attachment and formation of blastocyst outgrowths. For TS cell derivation, individual blastocysts at embryonic day 3.5 were transferred to a single well of a 96-well plate containing a confluent layer of Mitomycin-C (10µg/ml; R&D Systems Inc.;3258/10) treated mouse embryonic fibroblast cells as described previously (Tanaka et al., 1998b). Blastocysts were cultured in TS medium until blastocyst outgrowths were formed and reached an appropriate size, 4 to 5 days after the initial plating. Blastocyst outgrowths were then disaggregated in trypsin (0.25% Trypsin with EDTA) and cultured in TS medium until TS cell colonies were visible. Culture media was changed every 48 hrs to prevent cells from differentiation.

TS medium was prepared as follows: 30% basic culture medium (RPMI1640, pH 7.2; 20% fetal bovine serum (FBS); 2mM L-glutamine; 1mM sodium pyruvate; 100µM β-mercaptoethanol; 50µg/ml penicillin and 50µg/ml streptomycin), 70% mouse
embryonic fibroblast-conditioned medium (EMFI-CM), FGF4 (25ng/ml) and heparin (1µg/ml). EMFI-CM was prepared by a 72 hr incubation of basic culture medium on a confluent plate of mouse embryonic fibroblast cells that had been treated with mitomycin C to arrest cell division.

2.1.3 TS cell maintenance and differentiation

TS cells were maintained by culture in TS medium which was replaced every 48hrs. Once cells reached confluency, media was aspirated and the monolayer was washed with PBS (Invitrogen; 14190144) and trypsinized to detach cells from the culture plate. TS cells were transferred to a fresh plate and cultured in TS medium for 2 hrs to allow heavier differentiated cells (multinucleated and endoreduplicated cell) to sediment. Cells that remained in suspension after the 2 hr incubation were transferred to a fresh well for further culture. For differentiation into TGCs, TS cells were cultured in basic culture medium (RPMI1640, pH 7.2; 20% fetal bovine serum (FBS); 2mM L-glutamine; 1mM sodium pyruvate; 100µM β-mercaptoethanol; 50µg/ml Penicillin and 50µg/ml streptomycin) without the addition of EMFI-CM or FGF4 for 3 to 6 days.

2.1.4 HEK293T cell culture and transfection

HEK293T cells were cultured in D-MEM (Invitorgen; D5796) supplemented with 10% FBS. Transfection of HEK293T cells was achieved using CaPO₄ as outlined previously (Pear et al., 1993; Qin et al., 2003). Briefly, HEK293T cells were allowed to reach 70%-90% confluency on the day of the transfection. For a 10 cm plate, plasmids were diluted in a final volume of 1095µl of ddH₂O and 155µl of 2M CaCl₂. Phosphate solution (1250µl of 2XHBS - 274mM NaCl; 10mM KCl; 1.4mM Na₂HPO₄; 15mM D-glucose
and 42mM HEPES, pH 7.05) was added dropwise with constant mixing. The mixture was incubated at room temperature for 15 to 20 min and was subsequently evenly sprinkled over the monolayer. Six to twelve hrs after the transfection, the medium was aspirated, the monolayer was gentle washed with PBS, and fresh pre-warmed culture medium was added. Cells were either harvested for western blot analysis, or fixed for immunocytochemistry 48 hrs post-transfection.

2.1.4.1 Generation of Sfmbt2 over expression constructs

The lentiviral plasmid containing the long isoform of Sfmbt2 followed by an internal ribosome entry site (IRES) and an eGFP sequence (pLV-myc-Sfmbt2 (long)) was generated using the lentiviral plasmid FUGW, a gift from Dr. Jannet Rossant, as follows. A fragment containing a multiple cloning site (MCS)-IRES-eGFP-EcoRI was constructed by PCR amplification of IRES-eGFP sequence from pMSCV-PIG plasmid using MCS-IRES-F and EGFP-EcoRI-HpaI-R primer pairs (Appendix I A). The lentiviral plasmid FUGW containing a UBC promoter and the PCR amplified sequences were digested with XbaI and EcoRI followed by gel extraction, ligation and cloning. Once the identity of the cloned sequence was verified by sequencing the modified FUGW plasmid, now containing a human Ubiquitin-C (UBC) promoter sequence followed by an MCS, followed by the IRES-EGFP sequence, the long isoform of Sfmbt2 was introduced into the plasmid using the incorporated MCS. Myc-tagged Sfmbt2 (long) was amplified from pcDNA3-myc-Sfmbt2 (Kuzmin et al., 2008) using XbaI-Myc-sf-F forward primer (contains XbaI sequence) and the reverse primer Myc-sf-L-Bmt1-R (contains Bmt1 sequence) (Appendix I A). These restriction sites were introduced in the modified FUGW plasmid as part of the MCS to accommodate this insertion. As such both the modified FUGW and the amplified flanked myc-Sfmbt2 were digested with XbaI and
Bmt1, gel extracted, ligated and cloned into STBL2 cells.

The plasmid containing the short isoform of Sfmbt2 (pLV-myc-Sfmbt2 (short)) was generated using the modified FUGW plasmid from above and a PCR-prepared sequence containing eGFP, followed by the A2 sequence, followed by the myc-tagged short isoform of Sfmbt2. First, eGFP was amplified using a forward primer, pLV-XbaI-Egfp-F (contains an XbaI site) and the reverse primer, pLV-Egfp-A2-R (contains the first 18 bases of the A2 sequence) (Appendix I A). The product was reamplified using the same forward primer but the reverse primer used, A2 Domain-R, was designed to contain the previously introduced 18 bases of the A2 sequence flanked by 86 additional bases of the A2 sequence; amplification was followed by agarose-gel purification of the PCR product. The myc-tagged short isoform of Sfmbt2 was amplified using a forward primer, pLV-A2-sf-F (contains the last 18 bases of the A2 sequence) and a reverse primer, pLV-sf-sh-R (contains a HpaI site) (Appendix I A). The product was reamplified using the same reverse primer and a forward primer, A2 domain-R, which contains the previously introduced 18 bases of the A2 sequence flanked by 82 additional bases of the A2 sequence; amplification was followed by agarose-gel purification of the PCR product. The purpose of the reamplification of the PCR products was to add the A2 sequence to the 3' end of the constructed eGFP and the 5' end of the constructed myc-Sfmbt2. These regions were then used to serve as primers in a final PCR reaction containing both of the gel purified reamplified products. The two purified PCR products were mixed in a single tube without the addition of any primers for five PCR cycles. The reaction was stopped, the forward primer, pLV-XbaI-Egfp-F, and the reverse primer, pLV-sf-sh-R, were added and the reaction was allowed to proceed for another 25 cycles. The end product containing the full length eGFP-A2-myc-Sfmbt was gel purified and sequenced before
cloning into the modified FUGW plasmid. After the validation of the insert with sequencing, both the modified FUGW plasmid and the constructed insert were digested with XbaI and HpaI, gel extracted, ligated and cloned in STBL2 cells. The final construct was verified by sequence analysis. All sequencing was performed by The Center for Applied Genomics (Toronto, Ontario, Canada).

2.1.5 *Sfmbt2* knockdown in pre-implantation embryos

2.1.5.1 shRNA containing plasmid construction

Plasmids containing anti-*Sfmbt2* short hairpin RNAs (shRNAs) and the control shRNA were constructed through a two-step cloning process following previously published protocols (Georgiades et al., 2007; Qin et al., 2003). First, oligonucleotides containing the short hairpin sequences (Invitrogen) were cloned into pBS-hU6 plasmid, the expression cassette containing the human U6-RNA promoter; the cloned shRNA sequence was subsequently subcloned into FG12, the lentiviral plasmid (see below). Sense and antisense oligonucleotides for each shRNA, listed in Appendix I B, were phosphorylated using T4 polynucleotide kinase and allowed to anneal by boiling followed by gradual cooling to room temperature. The double stranded oligos containing appropriate overhangs were then cloned into pBS-hU6 which was digested with XhoI and BpiI to create receptive overhangs for ligation of the double stranded oligos. Positive clones were sequenced to confirm appropriate incorporation of the insert before subcloning. pBS-hU6 containing the shRNA inserts were then digested with XbaI and XhoI. The digested plasmids were electrophoresed on a 1% agarose gel followed by gel extraction of the appropriate bands to be subcloned into FG12 plasmid which was also digested with the same enzymes. All positive clones were sequenced for the correct
incorporation of the insert.

2.1.5.2 *In vitro generation of lentivirus*

Third generation replication deficient lentiviruses were generated using HEK293T cells as a packaging vehicle as described previously with minor modifications (Georgiades et al., 2007). Briefly, using CaPO₄, HEK293T cells were transiently transfected with the appropriate lentiviral vectors (FUGW plasmid containing human Ubiquitin-C driven eGFP; FG12 containing either anti-Sfmbt2-shRNA1, shRNA2 or control shRNA, as well as three packaging plasmids encoding Gag-pol, Rev and an envelope protein VSV-G, gifts from Dr. James Ellis. Culture medium was replaced with 30% heat-inactivated FBS (Gibco: 10437) in D-MEM (Gibco; 11965092) 7 to 11 hrs after the initial transfection. Lentivirus-containing medium was then harvested 36 hrs later and concentrated using Lenti-X concentrator (Clontech; 631231) following the manufacturer's protocol. Virus was diluted in TS medium and stored in aliquots at -80ºC. All lentiviruses contained an eGFP marker which was used for visual confirmation of successful infections.

2.1.5.3 *Lentiviral mediated knockdown in pre-implantation embryos*

Embryonic day 2.5 morulae were harvested as described above and incubated with acid Tyrode's solution (Millipore; MR-004-D) to remove the zonae pellucidae. Embryos were cultured overnight in microdrops containing concentrated lentivirus, diluted in KSOM medium in the presence of 8μg/ml of polybrene (ABM Inc.; G062), under light mineral oil (Georgiades et al., 2007). Following incubation with the lentivirus, embryos were thoroughly washed with M2 medium and transferred individually to 96 well plates on a confluent layer of mitomycin-C treated mouse embryonic fibroblasts. Two days after the infection, blastocyst outgrowths were scored for eGFP expression. Once the eGFP
positive outgrowths reached the appropriate size, they were disaggregated and cultured until TS cell colonies emerged. Each experiment was performed in parallel with one or more of the controls to account for unspecific experimental effects.

2.2 RNA ANALYSIS

Tissue or cells were harvested by immersion in Trizol reagent (Invitrogen) followed by RNA extraction and cDNA synthesis as described below. Following cDNA synthesis, gene expression was measured by means of RT-PCR and/or quantitative RT-PCR (qRT-PCR). When required, PCR products were subsequently used to determine allelic expression by single nucleotide polymorphism (SNP) analysis.

2.2.1 RNA extraction

RNA was extracted using Trizol reagent following the manufacturer's protocol. Briefly, tissue or cells were incubated in Trizol reagent for 5 min at room temperature to allow dissociation of nucleoprotein complexes. To separate proteins from the nucleic acids, samples were mixed with 1/5 volume of chloroform and centrifuged. The aqueous phase containing the RNA was transferred to a fresh tube and precipitated using isopropanol. RNA was washed in 75% ethanol made with RNase/DNase free ddH₂O, dried briefly and re-suspended in RNase/DNase free ddH₂O. Samples were allowed to dissolve at 55°C for 10 min before storage at -80°C to minimize degradation. If low yield of RNA was expected in cases were the starting material was limited such as extraction from blastocysts or blastocyst outgrowths, 0.1 μg/μl glycogen was added prior to the isopropanol.
To remove genomic DNA contamination as well as prevent further degradation of extracted RNA, samples were treated with DNase I and RNase inhibitor (Ribolock), respectively, following manufacturer's protocol. Briefly, per final volume of 20 µl, 1µl of DNase I and 0.5 µl of RNase Inhibitor and appropriate buffer were incubated at 37°C for 30 min. DNase I was then inactivated using EDTA to a final concentration of 5 mM and incubation at 65°C for 10 min.

2.2.2 cDNA synthesis

Reverse transcription of the RNA was performed using the reverse transcriptase, SuperscriptIII (Invitrogen) using the manufacturer’s buffer, dNTPs and random primers (Fermentas). RNA was incubated with random primers at 65°C for 5 min and transferred to ice for 1 min before the addition of the buffer, dNTPs and the reverse transcriptase. The mixture was incubated at room temperature for 5 min before incubation at 42°C for 1hr. For confirmation of removal of contaminating genomic DNA, a small aliquot of the mixture was removed before the addition of the reverse transcriptase for PCR analysis. In addition, primers specific to genomic DNA amplification only were used to amplify the cDNA preparation. The resulting cDNA was stored at -20°C in aliquots to reduce degradation associated with repeated freeze-thawing.

2.2.3 RT-PCR and qRT-PCR

RT-PCR was performed with Taq DNA polymerase from Clontech (RR002A) following the manufacturer’s protocol; the list of primers used can be found in Appendix I C. qRT-PCR was performed using SYBR Advantage q-PCR premix (Clontech; 638320) following the manufacturer’s instructions using a RotorGene6000 Light Cycler (Corbett
Life Sciences). Standard curves were generated by amplification of serial dilutions of template DNA with nested primers (Appendix I D) which were also used for amplification and quantitation of the cDNA of interest. Template DNA was generated from cDNA prepared from TS cell RNA extracts and purified using a PCR clean up kit (Sigma; NA1020). The concentration of template DNA was determined using a NanoDrop spectrophotometer.

2.2.4 Single nucleotide polymorphism (SNP) analysis

The region containing the SNP of interest was amplified using PCR. In cases where the SNP overlapped with a restriction enzyme digestion site, the PCR product amplified from cDNA or genomic DNA was subsequently digested with the appropriate restriction enzyme to determine allelic expression of a particular gene or confirm the presence of an annotated SNP on genomic DNA, respectively. In cases where the SNP of interest did no overlap with a restriction enzyme digestion site, the amplified PCR product was sequenced.

2.3 DNA ANALYSIS

2.3.1 DNA extraction

Tissue or cells were incubated in Proteinase K solution (50mM Tris, pH 8.0; 100mM EDTA; 100mM NaCl; 1% SDS; and 100µg/ml Proteinase K) overnight at 60°C. The following day, samples were mixed with equal volumes of Phenol/Chloroform and centrifuged at 13,000Xg for 10 min at room temperature. DNA was precipitated using
isopropanol, washed with 75% ethanol and diluted in ddH$_2$O. Samples were stored at 4°C for short term use and -20°C for long term storage.

2.3.2 PCR

PCR was performed with Taq DNA polymerase from Clontech (RR002A) following the manufacturer’s protocol. High fidelity Taq DNA polymerase (Stratagene; 600670) was used for the construction of inserts for the purpose of cloning to minimize mutagenesis of the DNA during amplification, following the manufacturer’s instructions. PCR products were electrophoresed on 1% agarose and imaged in the presence ethidium bromide (0.2µg/ml).

2.4 PROTEIN ANALYSIS

2.4.1 Protein extraction

TS cell monolayers were washed with PBS (137mM NaCl, 2.7mM KCl, 10mM Na$_2$HPO$_4$, 1.8mM KH$_2$PO$_4$; pH 7.2) three times before protein extraction in lysis buffer (350mM NaCl; 50mM HEPES, pH 7.0; 1mM EDTA, pH 8.0; 1mM DTT; 1mM PMSF and protease inhibitor cocktail-Sigma P8340). Lysis buffer was added to TS cell monolayers and incubated on ice for 20 min. Following the incubation, cells were detached using a cell scraper and transferred to an Eppendorf tube. Four cycles of freeze-thawing were performed to facilitate extraction of nuclear proteins as follows: 5 min on a dry-ice/ethanol slurry for snap freezing followed by thawing on ice with periodic vortexing. Cell lysates were then sonicated on ice (three 15 sec bursts with 30
sec cooling period in between each set) followed by centrifugation at 13,000Xg for 20 min. Protein extracts in the supernatant were either used directly for immunoprecipitation experiments or flash frozen in aliquots for long term storage at -80°C. HEK293T cell proteins were extracted in lysis buffer by incubation of the monolayer on ice for 20 min followed by harvest and centrifugation.

2.4.2 Western blot analysis

Protein samples were electrophoresed on a conventional 8-10% SDS-polyacrylamide gel and transferred to nitrocellulose membranes. Membranes were blocked in 5% dried milk in PBS plus 0.05% Tween for 2 hrs followed by incubation with primary antibody overnight. Primary antibodies used and their respective dilutions are as follows: 1:3,000 for anti-SFMBT2; 1:4,000 for anti-YY1 (Abcam; ab10237); 1:10,000 for anti-TFDP1 (Abcam; ab124678). Membranes were washed thoroughly in wash-buffer (PBS plus 0.05% Tween) and then incubated with the horseradish peroxidase (HRP)-conjugated secondary antibody (Cell Signalling Technology; 7074S), at 1:5,000 dilution, for 1 to 2 hrs before washing in wash-buffer and developing with enhanced ECL solution (PerkinElmer, Inc.; NEL10300EA). For western blotting on immunoprecipitated samples, a conformation specific HRP-conjugated antibody (Cell Signalling Technology; L27A9) was used at a dilution of 1:2,000. Visualization was performed using the BioRad’s Molecular Imager ChemiDoc.

2.4.3 Co-immunoprecipitation and Mass Spectrometry Analysis

2.4.3.1 Co-immunoprecipitation

Co-immunoprecipitation of protein extracts was done as follows. Cell extracts were pre-
cleared by incubation with pre-washed Protein G agarose beads (Invitrogen; 15920-010) at 4°C for 1 hr with end-to-end mixing on a rotator-wheel. The pre-cleared lysate was then isolated and incubated with the primary antibody or rabbit IgG (Cell signalling; 2729) at 4°C overnight. Antibody-protein complexes were subsequently incubated with fresh pre-washed Protein G agarose beads at 4°C the following day for 4 hrs with end-to-end mixing. Unbound protein extracts were collected after centrifugation of the bead-antibody-protein complexes. Bead complexes were then washed with five changes of non-denaturing lysis buffer (20mM Tris, pH 8.0; 137mM NaCl; 1% NP-40; 2mM EDTA; 10% glycerol; 1mM PMSF and protease inhibitor cocktail) and three changes of PBS containing 1mM PMSF and the protease inhibitor cocktail. Immunoprecipitated proteins were either eluted using 0.5M Ammonium hydroxide for in-solution preparation of samples for mass spectrometry analysis, or eluted in SDS loading dye (50mM Tris, pH 6.8; 2% SDS; 0.1% bromophenol blue; 20% glycerol) for SDS-PAGE separation of the proteins followed by in-gel preparation of selected regions for mass spectrometry analysis or direct western blotting.

2.4.3.2  In-gel and In-solution tryptic digestion

Coimmunoprecipitated samples were electrophoresed on SDS-PAGE followed by silver staining following the manufacturer’s protocol (Invitrogen; LC6100). Protein bands and regions containing several bands were excised and destained in destaining solution (1 volume of 30 mM Potassium ferricyanide and 1 volume of 100mM sodium thiosulfate) for 15 min. The destaining solution was then discarded and the gel pieces were washed in ddH2O followed by incubation in 50mM ammonium bicarbonate. The gel pieces were shrunk using a solution containing 1 volume of acetonitrile and 1 volume of 25mM ammonium bicarbonate. Disulfide bond reduction and alkylation was performed with
10mM DTT for 30 min at 56°C followed by 100mM iodoacetamide for 15 min in the dark at room temperature, respectively. The gel pieces were further shrunk as described above before tryptic digestion using 13ng/µl trypsin (Sigma; T6567) in 50mM ammonium bicarbonate for 3 hrs at 37°C. Extraction of the peptides was performed by collection of supernatant and serial incubation of the gel pieces with various solutions, described below, for 10 min each and subsequent collection and combination of the supernatant collected at each step. Solutions added in sequence are as follows: 25mM ammonium bicarbonate, 5% formic acid, 100% acetonitrile, 5% formic acid, 100% acetonitrile. For whole immunoprecipitated extracts eluted directly from the beads, in-solution reduction and alkylation was followed by tryptic digestion. Solutions containing the immunoprecipitated proteins from both in-gel and in-solution tryptic digestions were subsequently dried by vacuum evaporation.

2.4.3.3 Mass spectrometry

Mass spectrometry (LC-MS/MS) analysis was performed by Center for the Analysis of Genome Evolution and Function (CAGEF, Toronto, Ontario, Canada) using thermoscientific LTQ-Orbitrap mass spectrometer with a nano-LC system. Sample analysis was performed using X! Tandem (The GPM, thegpm.org; version X! Tandem Sledgehammer (2013.09.01.1)). A fragment ion mass tolerance of 0.50 Da and a parent ion tolerance of 20 PPM were used in the X! Tandem search. Matches were searched to the contaminants database (CDB) for elimination of false positives. Validation of the peptides and protein identification was performed with Scaffold (version Scaffold_4.0.4, Proteome Software Inc., Portland, OR, USA). The Peptide Prophet algorithm was used to assign probabilities to the identified peptides (Keller et al., 2002). Protein identifications were accepted if they exhibited greater than 95% probability (Nesvizhskii et al., 2003) and contained at
least three identified peptides. Proteins with similar peptides that could not be
differentiated were grouped to satisfy principles of parsimony. Proteins that showed
significant similarity in their peptide evidence were clustered together. In addition to the
search performed by Scaffold using CDB, any protein also found in IgG-mediated
immunoprecipitation was eliminated as an interaction candidate.

2.4.4  **Histology**

2.4.4.1  **Paraffin embedding and sectioning**

Dissected tissues were fixed in neutral buffered 10% Formalin (Sigma; HT501320) for
12-24hrs, depending on size. Following fixation, tissues were washed in PBS and
dehydrated in serial dilutions of 50%, 70%, 80%, 90%, 95%, and 100% ethanol for 30 to
60 min each depending on the size of the tissue. Samples were incubated in 100%
ethanol for two additional changes before clearing in Xylenes. Incubation in Xylenes was
performed a total of three times for 30 min each and transferred to 60°C paraffin. After
three changes of paraffin samples were allowed to solidify at room temperature
overnight. Paraffin blocks were then sectioned at 5 to 7µm thickness using a Leica
RM2235 microtome.

2.4.4.2  **H and E staining**

Sections were deparaffinized in three Xylenes washes, 5 min each, followed by
rehydration by immersion into serial dilutions of ethanol as follows: 100% ethanol three
times, followed by incubation in 95%, 90%, 80%, 70% and 50% ethanol, 3 min each.
Sections were then washed three times in ddH2O followed by incubation in Hematoxylin
(ScyTek laboratories Inc.; HAE-1-IFU) for 5 min. Sections were rinsed with two changes
of distilled H2O and quickly dipped in bluing reagent (ScyTek laboratories Inc.; HAE-1-
IFU) followed by two additional changes of tap water. Sections were briefly dipped into 100% ethanol before incubation in Eosin Y solution (ScyTek laboratories Inc.; HAE-1-IFU) for 20 to 30 sec. Slide were then rinsed in 100% ethanol and dehydrated followed by mounting with Cytoseal slide mounting medium.

2.4.4.3 Immunohistochemistry

Peroxidase based immunostaining was performed following the standard protocol with modifications as described below. Sections were rehydrated followed by incubation in PBS for 2 min at room temperature. Sections were incubated in antigen retrieval solution containing Tris-EDTA (10mM Tris; 1mM EDTA, 0.05% Tween; pH 9.0), boiled in a conventional vegetable steamer for 20 min, and allowed to cool down to room temperature. Following three washes in TBST (20mM Tris base, 150mM NaCl and 0.1% Tween; pH 7.6), the endogenous peroxidase activity was quenched using 3% hydrogen peroxide prepared in methanol (vol/vol) for 30 min. The slides were washed in TBST three times, 5 min each, followed by blocking in PBS containing 0.1% Triton-X, 1% BSA, and 10% normal goat serum for 1 to 2 hrs at room temperature, and were subsequently incubated with the primary antibody overnight at 4°C. The following day the slides were washed in three changes of TBST, 15 to 20 min followed by incubation with biotinylated secondary antibody (VECTASTAIN ABC Kit; PK-400); for 1 hr at room temperature. Primary and secondary antibodies were diluted in PBS containing 0.1% Triton-X, 0.5% BSA, and 5% normal goat serum.

Following the secondary antibody incubation, sections were washed three times in TBST for 5 min each followed by incubation in avidin-biotin complexes, prepared in PBS, for 1 hr. Sections were then rinsed in PBS before additional washes in TBST and incubation in 3,3’diaminobenzidine tetrahydrochloride (DAB) solution (0.05% DAB,
0.015% $\text{H}_2\text{O}_2$ in PBS; pH 7.2). To enhance the colour contrast with the nuclear counterstaining, the DAB solution used for some of staining experiments contained 0.05% $\text{CoCl}_2$ and 0.05% $\text{NiSO}_4$ resulting in a dark blue or black precipitate. DAB staining reaction was stopped in PBS followed by a rinse in ddH$_2$O and counterstained in methyl green solution (2% Methyl green prepared in 0.1M Sodium acetate; pH 4.2) for 5 min at 60°C. Followed by a quick rinse in ddH$_2$O sections were dehydrated serially in 95% ethanol and two changes of 100% ethanol (10 quick dips each). Clearing was performed in two changes of xylens twice, 1 min each, and was followed by mounting of sections with Cytoseal slide mounting medium.

2.4.4.4 Cryoblock preparation and sectioning

All steps were performed at 4°C unless otherwise stated. Placentae were fixed in neutral buffered 10% Formalin (Sigma; HT501320) for 4-6 hrs and washed in three changed of PBS, 20 min each. Sucrose and OCT (Optimum Cutting Temperature compound) infiltration was done by serial incubation of tissue with increasing concentrations of sucrose followed by introduction of OCT, as follows: 5% sucrose; 2 volumes of 5% and 1 volume of 30% sucrose; 1volume of 5% and 1 volume of 30% sucrose; 1 volume of 5% and 2 volumes of 30% sucrose, each for 30 min. Tissue was then incubated overnight in 30% sucrose. The following day, tissue was transferred to 2 volumes of 30% sucrose and 1 volume of OCT for 30 min at room temperature. Infiltrated tissue was then embedded in OCT on a dry ice/ethanol slurry. Tissue blocks were stored at -80°C before sectioning. Tissue sectioning was achieved at 10-12μm thickness using an UtracutE ultramicrotome with object temperature of -22°C. Sections were adhered to positively charged microscope slides (Fisherbrand: 1255015) and allowed to dry for 10 min before they were stored at -80°C for staining experiments including Alkaline phosphatase.
staining and X-gal staining.

2.4.4.5  Alkaline phosphatase (AP) staining

AP staining was performed to outline the maternal blood sinuses (Natale et al., 2006). Cryosections of placenta prepared sagittally were incubated in NT solution (0.15M NaCl; 0.1M Tris, pH 7.5) for 20 min at room temperature. Slides were then washed in fresh NTMT solution (0.1M NaCl; 0.1M Tris, pH 9.5; 0.05M MgCl₂; 0.1% Tween) for 10 min before incubation with an alkaline phosphatase substrate, BCIP/NBT (Sigma; B1911) for 10 to 15 min at 37°C. Following colour development, sections were counterstained with Nuclear Fast Red and mounted in 50% glycerol.

2.4.4.6  X-gal staining

For whole mount staining of tissue, pre-implantation embryos, and cells, samples were fixed in neutral buffered 10% Formalin (Sigma; HT501320) for 30 min to 1 hr depending on the size of the tissue and 10 min for cultured monolayer and pre-implantation embryos. After several changes of PBS, samples were permeabлизed using a detergent solution (2mM MgCl₂; 0.01% Sodium deoxycholate; 0.02% NP40 in PBS; pH7.2) for 10 to 30 min followed by several changes of PBS. Samples were then stained with X-gal solution (35mM Potassium ferrocyanide; 35mM Potassium ferricyanide; 1mM MgCl₂ in PBS) until desirable staining was achieved. For sagittal sectioning of X-gal stained wholemounts, samples were re-fixed in neutral buffered 10% Formalin (Sigma; HT501320) for an additional 24 hrs post-staining before embedding and processing. For cryo-prepared tissues, X-gal staining was performed after sectioning, counterstained using Nuclear Fast Red, and subsequently mounted in 50% glycerol. Imaging of wholemounts was performed on a Zeiss Lumar V12 stereomicroscope.
2.4.5  **Immunocytochemistry**

2.4.5.1  **Blastocyst staining**

Blastocysts were treated with Acid Tyrode’s solution to remove the zonae pellucidae and washed extensively in PBS before fixation using neutral buffered 10% Formalin (Sigma; HT501320) for 15 min at room temperature. The fixed blastocysts were rinsed with PBS and permeabilized in 0.1% Triton-X for 20 min at room temperature. Blastocysts were rinsed in PBS, followed by incubation in PBS containing 10% goat serum and 0.01% tween for 1 hr at room temperature in order to prevent non-specific binding. Embryos were then incubated with the primary antibody diluted in PBS containing 5% goat serum and 0.01% Tween at 4ºC overnight. The following day, embryos were washed extensively in PBS containing 0.05% Tween to remove the unbound primary antibodies and subsequently incubated with the Cy3 conjugated secondary antibody in PBS containing 5% goat serum and 0.01% Tween for 1 hr at room temperature followed by extensive washing with PBS containing 0.05% Tween. Counter-staining was performed using 4’,6-diamidino-2-phenylindole (DAPI) at a concentration of 0.5µg/ml diluted in PBS. Blastocysts were then mounted on 50% glycerol.

2.4.5.2  **Adherent cell staining**

Monolayers of adherent cells (TS cells, TGCs, HEK293 cells) and blastocyst outgrowths were washed three times with PBS before fixation with neutral buffered 10% Formalin (Sigma; HT501320) for 20 min at room temperature. Fixative was washed with three changes of PBS, followed by permeabilization with 0.1% Triton-X prepared in PBS for 15 min at room temperature. Cells were rinsed with PBS three times before incubation with PBS containing 10% goat serum and 0.01% tween for 1 hr at room temperature to block non-specific binding. Cells were then incubated with the primary antibody diluted
in PBS containing 5% goat serum and 0.01% Tween at 4°C overnight. The following day, cells were washed with PBS containing 0.05% Tween four times for 10 min at room temperature and subsequently incubated with the Cy3 labelled secondary antibody prepared in PBS containing 5% goat serum and 0.01% Tween for 1 hr at room temperature. Cells were then washed in PBS containing 0.05% Tween, 3 times 10 min. Counter-staining was performed using DAPI at a concentration of 0.5µg/ml diluted in PBS. Cells were mounted in 50% glycerol. Imaging was performed on the Olympus 1X71 inverted microscope, the Olympus BX60 upright microscope, or the Leica TCS SP5 confocal microscope.

2.5 ANTIBODY PRODUCTION

2.5.1 Plasmid construction and cloning

A poly-Histidine tagged Sfmbt2 (His-tagged Sfmbt2) cDNA amplicon encoding the short isoform of Sfmbt2, containing the first MBT domain which is also present in the long isoform (Figure 2.1 A), was cloned into a bacterial expression vector (pET-15b) and used to transform BL21 cells. Briefly, cDNA from TS cells was used to PCR amplify the region of interest using forward and reverse primers (Appendix I E) containing additional appropriate sequences to facilitate rapid cloning of His-tagged Sfmbt2 into the pET-15b plasmid. The forward primer, NcoI-His-Sfmbt2(ab)-F, was designed to contain the appropriate restriction enzyme site (NcoI; underlined); the appropriate codons for the histidine tag (bold) followed by an in-frame sequence of the short isoform of Sfmbt2. The reverse primer, Sfmbt2(ab)-XhoI-R, was designed to contain the appropriate restriction enzyme site (XhoI; underlined) at the 3’ end of the short isoform of Sfmbt2 to facilitate
Figure 2.1. Production of anti-SFMBT2 antibody.

(A) Multiple sequence alignment of the long and the short isoforms of SFMBT2 and the His-tagged short isoform of SFMBT2 used for the generation of anti-SFMBT2 antibody highlighting the common regions. Multalin was used for the multiple alignment of the protein sequences (Corpet, 1988). (B) Silver-stained SDS-PAGE showing induction of His-tagged short isoform of SFMBT2 in BL21 cells followed by its purification using nickel-NTA beads, and the dialysis of the purified protein.
the incorporation of the amplified product into pET-15b. The pET-15b plasmid and the amplicon were digested with XhoI and NcoI, electrophoresed on a 1% agarose gel and gel extracted using the QIAquick Gel Extraction Kit (Qiagen) following the manufacturer’s protocol. The extracted products were then ligated overnight using T4 DNA ligase from New England Biolabs at 16°C and used to transform BL21 cells. Clones were sequenced for appropriate incorporation of the sequence of interest.

2.5.2 Generation of His-tagged SFMBT2 protein and purification

STBL2 bacteria transformed with His-tagged-Sfmbt2-pET15b plasmid were cultured in 400ml of Luria Broth (LB) and ampicillin followed by induction with IPTG (final concentration of 1mM) for 2 hrs at 37°C. The bacterial culture was then centrifuged at 5,000Xg for 10 min at 4°C and the cell pellet was re-suspended in bacteria lysis buffer (10mM β-mercaptoethanol, 1mM PMSF, 0.2mg/ml of lysozyme, 1% Triton-X in PBS; pH 7.0) to facilitate the extraction of the proteins. Following a 20 min incubation on ice, the lysate was sonicated on ice (four 30 sec bursts with 30 sec cooling period in between each set). The lysate was then centrifuged at 10,000Xg for 5 min. The pellet was solubilized in denaturing binding buffer (8M Urea, 500mM NaCl, 20mM Sodium Phosphate; pH 7.8) and rocked on a nutator for 10 min at room temperature. The cell lysate was re-sonicated on ice (five 30 sec bursts with 30 sec cooling period in between each set). Following the sonication, the cell lysate was centrifuged at 3,000Xg for 15 min and the supernatant was collected for purification of the protein of interest.

Ni-NTA agarose beads (Invitrogen) were used for isolation of the His-tagged-SFMBT2 following the manufacturer’s instructions. Briefly, cell lysate was incubated with Ni-NTA agarose bead for 30 min at room temperature in a column kindly provided
by Dr. Dinesh Christendat. The column was then washed with denaturing wash buffers (8M Urea, 500mM NaCl, 20mM Sodium Phosphate) of decreasing pH (pH 7.0; pH 6.0; pH 5.3 and pH 5) before elution of the recombinant protein with denaturing elution buffer (8M Urea, 500mM NaCl, 20mM NaH2PO4; pH 4.0). Following the elution, gradual dialysis was performed over 48 hrs to substantially reduce the urea concentration of the recombinant protein before immunization of the animals. The dialysis solutions and times are as follows: 2M Urea in Buffer-D (50mM Tris- pH 8.0, 50mM NaCl, 1mM DTT, 1mM EDTA) for 6hrs; 1M Urea in Buffer-D 12hrs; PBS, pH 7.0 for 30 hrs with replacement with fresh PBS every 2 to 8hours. The recombinant protein was dialysed at 4°C for the duration of the dialysis. Aliquots of the purified protein were stored in -80°C to prevent denaturation of protein associated with repeated freeze-thawing.

Small aliquots were taken at every major step outlined above for diagnostics on a 12% polyacrylamide gel followed by silver-staining (Figure 2.1). A silver-staining kit (Invitrogen; LC6100) was used to perform the staining following the manufacturer’s protocol. Absence of a band in the sample labelled “flow through” and “bead after elution” indicates efficient binding and elution of the His-tagged protein to and from the nickel-NTA beads, respectively.

2.5.3 Animal immunization

Purified protein (200µg in Complete Freund's Adjuvant) was injected into rabbits followed by four additional booster injections. The first booster injection was delivered four weeks after the initial immunization with 100µg of purified protein in Incomplete Freund's Adjuvant, followed by three subsequent deliveries of 50µg in IFA every two weeks. Fourteen weeks after the initial immunization the rabbits were sacrificed by
exsanguination, and serum was collected and validated. All animal handling for antibody production was performed by staff of the Biological Services Facility of the Faculty of Arts and Sciences, and followed protocols approved by the Canadian Council on Animal Care.

2.5.4 Validation of the anti-SFMBT2 antibody

The generated anti-SFMBT2 antibody was tested in HEK293 cells transiently transfected with pLV-Myc-SFMBT2 (short and long) resulting in exogenous expression of the Myc-tagged mouse SFMBT2 (Figure 2.2). An eGFP transgene in the transfected construct (Figure 2.2 A) was used to identify successfully transfected cells (Figure 2.2 B,E). Immunocytochemistry showed anti-SFMBT2 signal specific to the eGFP-positive transfected HEK293 cells whereas eGFP-negative untransfected cells were negative for SFMBT2 signal (Figure 2.2 B-G). Western blot analysis of extracted proteins from the transfected HEK293 cells was performed to further investigate the specificity of the antibody, as well as proper transcription and processing of the transfected constructs (Figure 2.2 H). Two types of constructs containing the Myc-tagged SFMBT2 for both the long and the short isoforms were tested (Figure 2.2 A). Results indicate absence of signal in untransfected cells and presence of expected sized bands for both isoforms (Figure 2.2 H).
Figure 2.2. Validation of anti-SFMBT2 antibody.

(A) Schematic representation of the constructs containing the long and the short isoforms of SFMBT2 used for transient transfection of HEK293 cells. (B-G) Transfected HEK293 cells with pLV-myc-SFMBT2 short (B-D) and long (E-G) showing eGFP expressing cells (B and E) overlapping with SFMBT2 positive cells (C and F); DAPI was used as a nuclear stain (D and G). (B’-G’) shows magnified images from (B-G) as indicated by the dotted box. Arrows indicate cells in the mitotic stage of the cell cycle. (H) Western blot analysis of HEK293 cells transfected with constructs in (A) probed with anti-SFMBT2 antibody highlighting its specificity.
Chapter 3: Identification and Validation of a Candidate Paternally Expressed Gene with Essential Roles in Trophoblast Stem Cells

Disclaimer: Some of the data and figures presented in this chapter have been published in (Miri et al., 2013).
3.1 ABSTRACT

Uniparental mouse embryos with two maternal genomes (parthenogenetic) fail to develop functional extraembryonic tissues of trophoblast origin (tissue known to give rise to placenta), resulting in early mid-gestational lethality. This is likely due to the involvement of one or more PEGs with crucial role(s) in the development of the extraembryonic tissues and more specifically in the establishment and/or maintenance of trophoblast progenitors (TS cells). Pre-implantation embryos have been used to generate TS cells in culture; however, parthenogenetic embryos have been shown to be inefficient in generating these cells in vitro. In this chapter, I show that TS cell derivation from parthenogenetic embryos is correlated with LOI of a number of PEGs suggesting that the derivation of parthenogenetic TS cells is dependent on the reactivation of genes with essential roles in the establishment and/or maintenance of this cell type. To this effect, I surveyed PEGs using a number of criteria and identified Sfmbt2, encoding a highly conserved PcG protein, as a putative candidate. The investigation of this gene’s spatial and temporal expression in pre-implantation embryos and TS cells in addition to the knockdown of Sfmbt2 in early embryos resulting in impairment in generation of TS cell colonies in vitro, suggest that this gene is required for establishment and/or maintenance of the trophoblast progenitor cells in culture.
3.2 INTRODUCTION

In mammals biparental reproduction is crucial for proper growth and survival of the embryo, suggesting that the maternal and paternal genomes are not equivalent. Parthenogenetic/gynogenetic (bimaternial genome) and androgenetic (bipaternal genome) embryos derived from reconstituted eggs die early in development due to severe extraembryonic and embryonic abnormalities, respectively (Barton et al., 1984; Kaufman et al., 1977; McGrath and Solter, 1984; Surani et al., 1984). Furthermore, Robertsonian and reciprocal translocation studies have shown that inheritance of select chromosomal regions from a single parent can result in developmental abnormalities and even death in balanced diploid zygotes (Cattanach and Kirk, 1985; Searle and Beechey, 1978). These phenotypes have been attributed to aberrations in the expression of a group of genes that normally have distinct parent-of-origin specific expression due to a phenomenon known as genomic imprinting. Genomic imprinting is the result of epigenetic modifications established during gametogenesis, leading to differential expression of a number of genes solely from the maternally inherited allele (MEGs) or the paternally inherited allele (PEGs) (John and Lefebvre, 2011).

Since the discovery of genomic imprinting in 1984 (Barton et al., 1984; Kaufman et al., 1977; McGrath and Solter, 1984; Surani et al., 1984), numerous laboratories have been interested in identifying the genes responsible for the phenotypes observed in uniparental embryos leading to the generation of an imprinting map elucidating the genetic location of over 100 imprinted genes. Nevertheless, early embryonic death associated with parthenogenesis remains to be unravelled. The vast majority of parthenogenetic embryos die as early as day 5.5 of development with occasional survival detected at mid-gestation at the cost of severe retardation (Varmuza et al., 1993). Unlike
the androgenetic embryos that have hyperplastic extraembryonic tissues and poorly
developed embryos, parthenogenotes have severe developmental defects of the
extraembryonic tissues and relatively unhindered embryos (Barton et al., 1984; Kaufman
et al., 1977; McGrath and Solter, 1984; Surani et al., 1984). These observations suggest
that one or more PEGs have essential roles in the development of the extraembryonic
lineage, most notably those of the trophoblast origin.

Cells of the trophoblast origin are generated by the first differentiation event in as
early as morulae stage embryos, leading to the formation of a blastocyst constituted of an
ICM and an outer TE (Cross, 2000; Pedersen, 1987). The ICM is thought to give rise to
epiblast and primitive endoderm; the former develops into the embryo proper, while the
latter eventually forms the primitive endoderm layer of the extraembryonic yolk sac. The
TE assumes the trophectoderm fate giving rise to the various trophoblast cells of the
mature placenta. *In vitro* studies show that trophoblast progenitor cells known as TS
cells can be derived from blastocysts in culture (Tanaka et al., 1998a). These cells have
been reported to give rise to various cell types of the mature placenta including SynT
cells of the labyrinth, SpT cells, as well as TGCs. In light of these observations it can be
hypothesized that the trophoblast specific phenotypes observed in parthenogenetic
embryos leading to early death is due to lack of expression from an early acting PEG
with essential roles in the establishment and/or maintenance of TS cells.

Parthenogenetic blastocysts give rise to TS cells at a much lower frequency than
fertilized embryos (Miri et al, 2013). Considering previous reports on LOI associated
with embryos cultured *in vitro* (Doherty et al., 2000; Mann et al., 2004; Sasaki et al.,
1995), and X-chromosome associated LOI in parthenogenetic embryos (Endo and
Takagi, 1981; Rastan et al., 1980), I assayed for expression of PEGs in several
parthenogenetic TS (PTS) cell lines. I was able to show that establishment of parthenogenetic stem cells is correlated with LOI of a number of PEGs, including Peg1/Mest, Peg3, Mcts2, Nnat and Sfmbt2. In light of this evidence, I surveyed known PEGs for a number of criteria, in order to triage the list of approximately 50 PEGs to a manageable number. The survey allowed me to identify the PcG gene Sfmbt2 as the most likely candidate with essential roles in the development of the extraembryonic tissues, and in particular TS cells. I further showed that Sfmbt2’s temporal and spatial expression pattern is in agreement with this role. Additionally, knockdown of this gene at the morula stage using stable lentiviral induction of the cells with shRNAs directed against Sfmbt2 resulted in reduction in the frequency of derivation of TS cells suggesting that SFMBT2 is required for establishment and/or maintenance of the trophoblast lineage.

3.3 RESULTS

3.3.1 Parthenogenetic TS cells are indistinguishable from fertilized counterparts.

To confirm the identity of PTS cells, I performed morphological and cell-type specific marker analysis of five independently derived PTS cell lines in comparison with four fertilized TS (FTS) cell lines. In addition to displaying a characteristic epithelial appearance with defined edges to the colonies (Figure 3.1 A,B), PTS cells exhibit proper expression of TS cell-specific markers such as Cdx2, Fgfr2 and Mash2, and absence of the ES cell marker Oct4 (Figure 3.1 C).
Figure 3.1. Fertilized and Parthenogenetic TS cells Show Similarities in Morphology and Gene Expression.

(A,B) Phase contrast images of undifferentiated fertilized (A) and parthenogenetic (B) TS cell colonies. (C) RT-PCR of various PEGs and cell-type markers in four fertilized and five parthenogenetic TS cell lines. Exclusion of genomic contamination is shown by ØRT in which reverse transcriptase was omitted during cDNA synthesis and PCR was performed using β-actin primers. P, paternally expressed; Bi, biallelic; ?, unknown.
TS cells require supplementation of FGF4 and embryonic fibroblast conditioned media to remain undifferentiated. Withdrawal of these factors results in differentiation into TGCs (Tanaka et al., 1998b). To address whether PTS cells are competent to differentiation into TGCs, cells were grown in the absence of FGF4 and conditioned media for 6 days. Both fertilized and parthenogenetic TS cell lines were able to fully differentiate into TGCs (not shown). Together these results suggest that despite their low frequency of derivation, in vitro derived PTS cells are indistinguishable from their fertilized counterparts.

3.3.2 PTS cells exhibit loss of imprinting (LOI) at various loci

In light of previous reports on LOI in embryos cultured in vitro (Mann et al., 2004; Rivera et al., 2007), it has been hypothesized that the derivation of PTS cells is dependent on LOI of PEGs required for TS cell establishment and/or maintenance. To this effect, three PTS cell lines derived from MI oocytes (PTS1-PTS3) and two lines derived from MII oocytes (PTS4, PTS5) were examined for expression of several known imprinted genes by RT-PCR. In addition, four fertilized TS (FTS) cell lines, including F1-FTS cells derived from mating a C57BL/6 female with a Mus castaneus male, were assayed as controls to evaluate the consistency of expression (Figure 3.1 C). The analysis resulted in the categorization of these genes into three classes: those with consistent expression across all TS cell lines (Mcts2, Nnat, Peg1, Peg3, Sfmbt2 and Igf2); those with inconsistent expression in TS cells (Ndn, Rasgrf1, Mit1, Peg10, Plagl1, Sgce, Snrpn, Slc38a4 and Dlk1); and those with no detectable expression (Ins2).

Parental specific expression of various imprinted genes was previously shown to be tissue-specific (Miri and Varmuza, 2009). To address whether the observed expression
of the PEGs in PTS cell lines is due to LOI or simply their biallelism in TS cells, allelic specificity of genes with consistent expression across all TS cell lines was determined. SNP analysis in F1-FTS cells revealed paternal expression of Mcts2, Peg1/Mest, Peg3, and Sfmbt2 (Figure 3.2 A-D), confirming imprinted expression and indicating that expression in PTS cells is a function of LOI. Igf2 however, showed biallelic expression (Figure 3.2 E), indicating that the observed consistent expression in PTS cell lines is not due to LOI. I was unable to assess allelic expression of Nnat due to lack of polymorphisms among the tested strains. These results indicate that in vitro derivation of TS cells is correlated with LOI of a select number of PEGs, suggesting that the derivation of PTS cells is mediated by reactivation of such genes with essential functions in the trophoblast lineage.

3.3.3 Candidate Study: Identification of SFMBT2

In order to identify the gene responsible for the phenotype observed in parthenogenetic embryos, known PEGs were triaged for the following four criteria. 1) It was reasoned that since establishment of TS cells commences at the blastocyst stage, the candidate must exhibit imprinted expression at the blastocyst stage. 2) The candidate must display imprinted expression in TS cells. 3) All TS cell lines examined must exhibit consistent expression of the candidate; for PTS cells this represents LOI. 4) The knockout phenotype of the candidate should have severe effects on early post-implantation trophoblast development similar to those observed in parthenogenetic embryos.

3.3.3.1 PEGs displaying imprinted expression at the blastocyst stage (Criterion 1)

Currently, there are approximately fifty PEGs identified (Miri and Varmuza, 2009). In order to assess the first criterion, I examined the expression pattern of PEGs in
Figure 3.2. Imprinting status of PEGs in TS cells.

(A-C) RFLP analysis of long (A) and short (B) isoforms of Sfmbt2 as well as Peg3 (C) in F1-TS cells amplified with Sfmbt2SNP4103 F/R, Sfmbt2 short 3’end and Peg3-F1/R1 primers, respectively. Restriction enzymes used are as indicated. Note in (A) Hinf1 digests the Mus castaneus allele, and SnaBI digests the C57BL/6 allele (Kuzmin et al., 2008). (D,E) Sequence traces showing SNPs in gDNA and cDNA from F1-FTS cell line for Mcts2 (D) and Igf2 (E). BC, C57BL/6 X Mus castaneus; gDNA, genomic DNA.
parthenogenetic, androgenetic and fertilized blastocysts using previously published microarray data (GEO dataset identifier: GSE8163 and GSE1749). To ensure the integrity of the identifiers as true representatives of annotated transcripts, the sequences for each transcript, as well as all splice variants, were individually examined for matching probe sets using NetAffx probe match software available at http://www.affymetrix.com. Any gene with biallelic or lack of expression was eliminated as a candidate (Table 3.1).

In addition to the analysis of publically available microarray datasets, our collaborators, Dr. Keith Latham and Dr. Zhisheng Zhong tested the expression and imprint status of twenty-seven PEGs in androgenetic, gynogenetic and biparental blastocysts by qRT-PCR (Miri et al, 2013). Of the genes tested those showing no detectable expression and those with biallelic expression were eliminated, leaving a small number of genes including: Mcts2, Nnat, Peg1/Mest, Peg3, Peg10, Plagl1, Sfmnt2, Sgce, Snrpn and U2af1-rs1/Zrsr1. These genes showed similar levels of expression in androgenetic and biparental blastocysts compared to the marginal to no expression observed in gynogenetic blastocysts, indicating their monoallelicism at this stage of development (Table 3.1).

3.3.3.2 PEGs displaying imprinted expression in TS cells and LOI in all PTS cell lines (Criterion 2 and 3)

Criterion 2 and 3 were evaluated by assaying expression of PEGs in three FTS and five PTS cell lines (Figure 3.1 C). Several genes were found to display inconsistent expression. The remaining genes were assessed for monoallelic expression in F1-FTS cells (Figure 3.2). Allelic expression of these genes were further confirmed by reanalysis of published RNA-seq data (Calabrese et al., 2012). Summary of the results can be
Table 3.1. Summary of the results from the candidate study for the identification of the PEG responsible for the early phenotypes of parthenogenetic embryos.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Blastocyst Expression</th>
<th>TS Cell Expression</th>
<th>Lit Review</th>
<th>Final call</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Microarray</td>
<td>qRT-PCR</td>
<td>Allelic call</td>
<td>Sequencing</td>
</tr>
<tr>
<td>A19</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>nd</td>
</tr>
<tr>
<td>Airt/1g2ra</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>nd</td>
</tr>
<tr>
<td>Apeg3</td>
<td>-</td>
<td>nd</td>
<td>nd</td>
<td>-</td>
</tr>
<tr>
<td>Ddc</td>
<td>+/-</td>
<td>-</td>
<td>nd</td>
<td>-</td>
</tr>
<tr>
<td>Dio3</td>
<td>nd</td>
<td>+</td>
<td>Bi</td>
<td>Bi</td>
</tr>
<tr>
<td>Dlk1/Pref1</td>
<td>+/-</td>
<td>-</td>
<td>nd</td>
<td>+</td>
</tr>
<tr>
<td>Gab1</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
<td>Bi</td>
</tr>
<tr>
<td>Gnasxl</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Igf2</td>
<td>-</td>
<td>+</td>
<td>Bi</td>
<td>Bi</td>
</tr>
<tr>
<td>Igf2as</td>
<td>nd</td>
<td>na</td>
<td>Bi</td>
<td>nd</td>
</tr>
<tr>
<td>Impact</td>
<td>+/-</td>
<td>+</td>
<td>Bi</td>
<td>Bi</td>
</tr>
<tr>
<td>Ins1</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>-</td>
</tr>
<tr>
<td>Ins2</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>-</td>
</tr>
<tr>
<td>Ipw</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>-</td>
</tr>
<tr>
<td>Kcnq1ot1</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>nd</td>
</tr>
<tr>
<td>Mage2</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>-</td>
</tr>
<tr>
<td>Mas1</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>nd</td>
</tr>
<tr>
<td>MBII-13</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>MBII-52</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>MBII-85</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Mcts2</td>
<td>+</td>
<td>+</td>
<td>P</td>
<td>nd</td>
</tr>
<tr>
<td>Mit1/Cogp2as2</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>nd</td>
</tr>
<tr>
<td>Mknn3/Zfp127</td>
<td>-</td>
<td>+</td>
<td>Bi</td>
<td>+</td>
</tr>
<tr>
<td>Nap115</td>
<td>-</td>
<td>+</td>
<td>Bi</td>
<td>+</td>
</tr>
<tr>
<td>Ndn</td>
<td>-</td>
<td>+</td>
<td>Bi</td>
<td>+</td>
</tr>
<tr>
<td>Nespas</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>P</td>
</tr>
<tr>
<td>Nnat</td>
<td>+/-</td>
<td>+</td>
<td>P</td>
<td>+</td>
</tr>
<tr>
<td>Peg1/Mest</td>
<td>+</td>
<td>+</td>
<td>P</td>
<td>+</td>
</tr>
<tr>
<td>Peg10</td>
<td>-</td>
<td>+</td>
<td>P</td>
<td>Bi</td>
</tr>
<tr>
<td>Peg12/Frat3</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>Bi</td>
</tr>
<tr>
<td>Peg13</td>
<td>nd</td>
<td>-</td>
<td>na</td>
<td>nd</td>
</tr>
<tr>
<td>Peg3/Pw1</td>
<td>-</td>
<td>+</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Rasgrf1</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>+</td>
</tr>
<tr>
<td>Peg11/Rd11</td>
<td>nd</td>
<td>+</td>
<td>Bi</td>
<td>+</td>
</tr>
<tr>
<td>Sfmbt2</td>
<td>+</td>
<td>+</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Sgce</td>
<td>-</td>
<td>+</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Sklc38a4/Ata3</td>
<td>+/-</td>
<td>-</td>
<td>nd</td>
<td>P</td>
</tr>
<tr>
<td>Sarpm/Snurf</td>
<td>+/-</td>
<td>+</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>U2af1-rs1/Zrsr1</td>
<td>nd</td>
<td>+</td>
<td>P</td>
<td>Bi</td>
</tr>
<tr>
<td>Ube3a-as</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>-</td>
</tr>
<tr>
<td>Usp29</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>Bi (low)</td>
</tr>
<tr>
<td>Zac1/Plag11</td>
<td>+/-</td>
<td>nd</td>
<td>nd</td>
<td>P</td>
</tr>
<tr>
<td>Zifp264</td>
<td>+/-</td>
<td>+</td>
<td>Bi</td>
<td>nd</td>
</tr>
<tr>
<td>Zdbf2</td>
<td>-</td>
<td>nd</td>
<td>na</td>
<td>+</td>
</tr>
<tr>
<td>Zifp64</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
<td>Bi</td>
</tr>
</tbody>
</table>

+, expressed; -, not expressed; +/-, variable expression; Bi, biallelic; P, paternally expressed; nd, no data; na, not applicable; F, fail. * lethality at e9.5 will be discussed in chapter 4 of this thesis.
found in Table 3.1. These results narrowed the list to a small number of PEGs including Peg1/Mest, Peg3, Mcts2, Nnat and Sfmbt2 which showed consistent expression across all cell lines tested.

3.3.3.3 PEGs with knockout phenotypes resembling those of the parthenogenetic embryos (Criterion 4)

Using criterion 4, several genes were provisionally excluded as being required for TS cell establishment and/or maintenance. Ddc, Dlk1, Gnasxl, Igf2, Airn, Ins2, Kcnqlot1, Magel2, Mas1, MB11-85/Pwcr1, Ndn, Nespos, Peg1/Mest, Peg3, Plagl1, Peg11/Rlt1, Peg3, Rasgrf1, Sgce, Snrpn, Snurf, and U2af1-rs1/Zrsr1 have all been subjected to targeted mutagenesis in which the null phenotypes have no or very mild effects on placental development (Reviewed in Miri and Varmuza, 2009). Peg10 is the only paternally expressed imprinted gene known to date whose loss-of-function results in mid-gestational lethality starting at embryonic day 10.5 (Ono et al., 2006). However, inconsistent LOI in PTS cells excluded this gene as a candidate. Finally, a null mutation of Sfmbt2, discussed in Chapter 4, results in mid-gestational lethality starting at embryonic day 8.5 due to failure to sustain trophoblast development (Miri et al, 2013), suggesting that Sfmbt2 is the most fitting candidate.

3.3.4 Observed LOI occurs as early as the blastocyst outgrowth stage

TS cells can be derived from blastocyst outgrowths in vitro (Tanaka et al., 1998a). In order for parthenogenetic embryos to give rise to TS cells, assuming that the derivation is dependent on LOI of critical PEGs, activation of the normally paternally-expressed gene(s) must take place by the blastocyst outgrowth stage. To address the timing of Sfmbt2’s LOI in parthenogenetic embryos cultured in vitro, I performed RT-PCR on
expanded blastocysts as well as blastocyst outgrowths collected after 4 days in culture. Results show the presence of Sfmbt2 transcripts at the outgrowth stage in both fertilized and parthenogenetic cell extracts but not at the expanded blastocyst stage (Figure 3.3), suggesting that Sfmbt2 undergoes LOI by the blastocyst outgrowth stage. It is important to note that extracts for these experiments were collected from pools of 6 to 10 blastocysts or blastocyst outgrowths, therefore the proportion of embryos exhibiting LOI at this stage of development is unknown.

3.3.5 Spatial and Temporal Expression-pattern of Sfmbt2 Confirms a Role in Trophoblast Development

3.3.5.1 Expression of SFMBT2 in pre-implantation stage embryos

I examined the spatial and temporal expression-pattern of Sfmbt2/SFMBT2 in early pre-implantation embryos. For this experiment I used embryos with a paternally inherited gene trap allele of Sfmbt2 in which a β-GEO cassette was inserted downstream of the first coding exon to assess RNA expression by proxy (see next chapter for details of gene trap line), and immunocytochemistry using a polyclonal rabbit antibody that I raised against SFMBT2 to assess protein distribution (experiments performed for the validation of the anti-SFMBT2 antibody are discussed in Chapter 2).

Lac-Z staining of embryos carrying a paternal copy of the gene trap allele illustrates that Sfmbt2 is expressed in most cells in morula stage and early blastocyst stage embryos (Figure 3.4 A); blastocysts show homogeneous expression including the ICM and TE cells (Figure 3.4 B). This is supported by the analysis of publically available microarray results showing steady increased expression from 8-cell to morulae and blastocyst stage embryos (Zeng et al., 2004). Wild type litter mates and
Figure 3.3. *Sfmbt2* exhibits LOI by blastocyst outgrowth stage.

RT-PCR showing expression of *Sfmbt2* and β-actin in Fertilized (F) and Parthenogenetic (P) Blastocysts (Blast.) and Blastocyst outgrowths (Blast. OG). Extracts for each sample were collected by pooling 6 to 10 embryos after four days in culture.
Figure 3.4. SFMBT2 exhibits perinuclear staining in the TE of expanded blastocysts.

(A-C) Lac-Z staining, (D,F,I) Immunostaining against SFMBT2 and (E,H) DAPI counterstaining of blastocysts at e2.5-e3.5 (A,D) and at e4.5 (B,C,E-J). (G,J) show superimposed images. E’-J’ are enlarged inlets from E-J (dashed boxes). Embryos in A and B were collected from a wild type mother (+/+) mated with a heterozygous father (gt/+). Embryos in C were collected from a heterozygous mother (gt/+), mated with a wild type father (+/+). Embryos in D to J are from wild type parents.
embryos possessing a maternally transmitted gene trap allele show no lac-Z staining indicating that $Sfmbt2$ is paternally expressed at early stages, as previously reported (Kuzmin et al, 2008). Furthermore absence of staining in wild type embryos suggests that endogenous $\beta$-galactosidase does not contribute to the blue signal (Figure 3.4 A,B).

The spatial organization of SFMBT2 at the blastocyst stage was carried out by immunocytochemistry experiments followed by conventional fluorescence- and confocal-microscopy. Immunostaining of blastocysts at e3.5 shows global distribution of protein in all cells of the blastocyst (Figure 3.4 D). Based on the positional analysis of the signal, fully expanded e4.5 blastocysts display a distinct staining pattern mostly limited to the nuclear periphery in the TE cells while the ICM staining appears to be cytoplasmic (Figure 3.4 E-J).

3.3.5.2 Expression of SFMBT2 at the blastocyst outgrowth stage

As $Sfmbt2$’s LOI occurs between expanded blastocyst stage and blastocyst outgrowth stage, I assayed for localization of SFMBT2 at the blastocyst outgrowth stage using immunocytochemistry (Figure 3.5). Similar to the observed pattern of expression at the expanded blastocyst stage, giant cells of the blastocyst outgrowth, identified based on the apparent size of their nuclei, display cytoplasmic staining with a higher intensity at the perinuclear region (Figure 3.5 C (arrow) and D), while the ICM clump shows mostly cytoplasmic staining (Figure 3.5 C (dashed area) and E). Lac-Z staining of the blastocyst outgrowth confirms the presence of $Sfmbt2$ transcripts in both the ICM clump (dashed area) as well as the TGCs (arrows) (Figure 3.5 A); albeit not all TGCs were lac-Z positive (not shown).
Figure 3.5. SFMBT2 exhibits perinuclear staining in TGCs of the blastocyst outgrowth.

(A) Bright field image of Lac-Z staining in a heterozygous (+/gt) blastocyst outgrowth obtained from a wild type mother (+/+) mated with a heterozygous father (gt/+), carrying a gene trap allele of Sfmbt2. (B,E) Immunostaining against SFMBT2 and DAPI counterstaining of wild type blastocyst outgrowths, four days after attachment. (B,C) show images collected with conventional fluorescence microscopy. (D,E) show confocal images of TGCs (D) and ICM clump (E) of a blastocyst outgrowth. Arrow indicates a TGC (A-C). Dotted line indicates the ICM clump (A-C).
3.3.5.3 Expression pattern of SFMBT2 in Trophoblast Cell Types

I assessed the spatial organization of SFMBT2 in undifferentiated TS cells as well as differentiated giant cells (Figure 3.6). TS cells show pronounced punctate nuclear staining in the majority of cells (Figure 3.6 C) with the exception of those going through mitosis (Figure 3.6 D,E). Mitotic cells, identified at metaphase (Figure 3.6 D), and anaphase (Figure 3.6 E) due to the distinct DAPI staining of their chromatin, display SFMBT2 staining associated with the mitotic chromosomes.

Differentiation of TS cells in culture into TGCs and subsequent immunostaining shows further diversity in SFMBT2 staining patterns (Figure 3.6 F) including cells that show punctate nuclear staining, filamentous cytoplasmic staining, as well as those that lack SFMBT2 staining in both the nucleus and the cytoplasm. Lac-Z staining of undifferentiated TS cells and differentiated TGCs reflects a change from uniform expression to varied levels of expression after differentiation (Figure 3.6 A,B).

3.3.6 Knockdown study is suggestive of a critical role for Sfmbt2 in establishment and/or maintenance of TS cells

I employed a lentiviral delivery system to introduce shRNAs directed against endogenous Sfmbt2 in early embryos followed by assessment of the frequency of derivation of TS cells. The efficiency of the knockdown constructs (Figure 3.7 A) were tested in HEK293T cells transfected with myc-Sfmbt2 and either anti-Sfmbt2 shRNA1, shRNA2 or scramble-shRNA. Four days after the transfection, cells were harvested and analyzed using western blotting. Results show a significant reduction of SFMBT2 with both anti-Sfmbt2 shRNAs but not scramble-shRNAs (Figure 3.7 B).
Figure 3.6. SFMBT2 exhibits heterogeneous staining in the TS cell lineage

(A,B) Lac-Z staining (C-F) Immunostaining against SFMBT2 and DAPI counterstaining of undifferentiated TS cells (A,C-E) and TGCs (B,F). Cells in A and B are from a heterozygous embryo (+/gt) collected from a wild type mother (+/+) mated with a heterozygous father (gt/+). Embryos in C-F are from wild type parents. D and E show mitotic cells at metaphase (D) and anaphase (E). F shows TGCs with filamentous (left), punctate (middle) and no (right) staining.
Figure 3.7. Validation of anti-Sfmbt2 shRNAs and lentiviral infection of early embryos.

(A) Map of shRNA-carrying lentiviral plasmid. ShRNA sequence was inserted downstream of hU6 Pol III promoter. Vectors also contain eGFP sequence driven by human UBC (hUBC) promoter. (B) Validation of anti-Sfmbt2 shRNAs shown by western blot analysis using anti-SFMBT2 antibody of whole protein lysate harvested from untransfected as well as transfected HEK293 cells with various combinations of plasmids as indicated. (C-J) Bright-field (C,E,G,I) and fluorescent (D,F,H,J) images of fertilized blastocysts 36hr post infection (C,D); blastocyst outgrowths (E,F); a TS cell colony (G,H); and XEN cells (I,J) derived from fertilized morulae infected with scrambled shRNA carrying lentivirus.
Following the validation of the shRNAs, fertilized 8-cell to morula stage embryos were infected with shRNA containing lentiviruses and were subsequently cultured for TS cell derivation. This embryonic stage was chosen to maximize the probability of a uniform transduction. A successful transduction was scored by the presence of eGFP expression 36 to 72 hrs after the initial infection (Figure 3.7 C,D). Embryos infected with control scrambled shRNA bearing lentiviruses resulted in positive outgrowths (Figure 3.7 E,F), that successfully produced transduced TS cells (Figure 3.7 G,H) as well as extraembryonic endoderm cells (XEN) cells (Figure 3.7 I,J) indicating relatively uniform infection of majority of the cells of the early embryo and minimal to no silencing of the reporter eGFP transgene.

Each knockdown experiment was performed in parallel with at least one of the following controls: uninfected morulae; infected morulae with eGFP carrying virus or infected morulae with scrambled-shRNA carrying virus. Table 3.2 outlines the results from a series of knockdown experiments using the various lentiviruses. The average number of TS colonies per positive outgrowth (eGFP expressing) was statistically lower in anti-Sfmbt2 shRNA1 (0.01) and shRNA2 (0.42) in comparison with scramble-shRNA (1.86) carrying virus. Infection with eGFP carrying virus without an shRNA resulted in a higher average number of TS colonies per positive outgrowth (3.47) however, these differences were not statistically significant. TS cell derivation from uninfected morulae (4.27) was comparable to the rates seen with eGFP lentivirus. TS cell derivation from uninfected fertilized blastocysts performed previously also appears to be comparable (Miri et al, 2013). Moreover, the average number of TS colonies per positive outgrowth in the parthenogenetic TS derivation experiment reported in Miri et al., 2013.
Table 3.2. Knockdown of Sfmbt2 correlates with reduction in the frequency of TS cell derivation.

Table shows number of uninfected and infected outgrowths (OG); number of outgrowths with TS colonies (OG with TS); total number of TS cell colonies (No. TS colonies); average number of TS cell colonies per outgrowth with TS cell colonies (TS Col./+ve OG); average number of TS cell colonies per outgrowth (TS Colonies/OG). Any blastocyst outgrowth or TS cell colony not expressing eGFP was excluded from the calculations. As such, all numbers displayed for infected embryos are from successfully transduced blastocyst outgrowths and TS cell colonies exhibiting eGFP expression. The statistical significance of differences observed in the number of TS cell colonies per outgrowth under each condition in comparison to uninfected controls, was obtained using Fisher’s exact test and Chi-square test. Similar results were obtained when either eGFP or the scramble siRNA controls were used as comparator.

<table>
<thead>
<tr>
<th></th>
<th>OG</th>
<th>OG with TS</th>
<th>No. TS Col.</th>
<th>TS Col./+ve OG</th>
<th>TS Col./OG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninfected</td>
<td>15</td>
<td>12</td>
<td>64</td>
<td>5.33</td>
<td>4.27</td>
</tr>
<tr>
<td>GFP</td>
<td>15</td>
<td>10</td>
<td>52</td>
<td>5.2</td>
<td>3.47&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Scramble</td>
<td>43</td>
<td>26</td>
<td>80</td>
<td>3.08</td>
<td>1.86&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>shRNA1</td>
<td>70</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.01&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>shRNA2</td>
<td>36</td>
<td>5</td>
<td>15</td>
<td>3</td>
<td>0.42&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> p>0.05 when compared to uninfected
<sup>b</sup> p<0.05 when compared to uninfected and infected with the scramble-shRNA
was slightly higher (0.07) than fertilized blastocysts infected with anti-\textit{smbt2} shRNA1 (0.01). This slight rescue may be due to re-activation of the maternal allele in some of the parthenogenetic embryos. These observations show that knockdown of \textit{Sfmbt2} results in reduced frequency of TS cell derivation.

3.4 \textbf{DISCUSSION}

Parthenogenetic mouse embryos die as early as embryonic day 5.5 of gestation with occasional survival up to the 25 somite stage at the cost of severe retardation (Kaufman et al., 1977; Varmuza et al., 1993). The early death associated with parthenogenesis has been linked to the absence of expression from PEGs with essential roles in proper development of extraembryonic tissues of the trophoblast origin (Clarke et al., 1988; Varmuza et al., 1993). Furthermore, parthenogenetic embryos have been shown to give rise to TS cells \textit{in vitro}, but at a much lower frequency when compared to fertilized counterparts (Miri et al 2013).

I assessed the integrity of previously established PTS cell lines and have shown that they are indistinguishable from FTS cells in both morphology as well as expression of appropriate lineage-specific markers such as \textit{Cdx2} and absence of \textit{Oct4}. These findings are inconsistent with the phenotypes observed in parthenogenetic embryos, as their lack of a functional trophoblast lineage suggests aberrations in the establishment and/or maintenance of trophoblast progenitor cells. Given this phenotype, how do parthenogenetic embryos give rise to sustained TS cells in culture?

Numerous studies have reported LOI in cells and embryos cultured \textit{in vitro}
LOI is a condition associated with erasure of epigenetic marks that normally give rise to parent-of-origin specific expression of imprinted genes resulting in aberrant expression. Blastocysts cultured for prolonged periods before transplant into the uterus exhibit LOI especially in the trophoblast tissues (Mann et al., 2004; Sasaki et al., 1995). In light of these reports I assessed the expression of a number of PEGs in five independently generated PTS cell lines and found the presence of transcripts from normally maternally silenced loci such as Peg1/Mest, Peg3, Mcts2, Sfmbt2 and others, suggesting that derivation of PTS cells is dependent on LOI of paternally expressed/maternally silenced genes with essential functions in establishment/or maintenance of TS cells. To this effect, I surveyed known PEGs for a number of criteria in order to identify the most likely candidate. It was reasoned that the candidate gene should:

1. Exhibit imprinted expression at the blastocyst stage at which point the trophoblast lineage is determined.
2. Exhibit imprinted expression in TS cells.
3. Exhibit consistent LOI in PTS cell lines.
4. Exhibit a knockout phenotype similar to that of the parthenogenetic embryo.

Most genes were eliminated as possible candidates (Table 3.1); the PcG gene Sfmbt2 was chosen as the only known PEG with a potential role in establishment/maintenance of the TS cell compartment.

I performed a knockdown assay using a lentiviral delivery system (Georgiades et al., 2007) to test the effect of SFMBT2 reduction on TS cell derivation from fertilized embryos. Knockdown of Sfmbt2 had a negative effect on the frequency of TS cell
derivation suggesting that this gene is required for establishment and/or maintenance of TS cells in culture. Attempts to rescue the frequency of PTS cell derivation from parthenogenetic embryos by introducing exogenous SFMBT2 into morula stage parthenogenetic embryos prior to TS cell derivation failed to produce any informative conclusions due to limitations of the lentiviral delivery system, in this case the large size of the insert. Technical limitations also hampered my ability to assess the effects of Sfmbt2 knockdown in established TS cells due to the hard-to-transfect nature of these cells (not shown).

Previous studies have shown that Sfmbt2 is expressed exclusively from the paternal allele in extraembryonic tissues, while showing significantly lower but biallelic expression in somatic tissues. I examined spatial and temporal expression of this protein in early embryos as well as TS cells. Consistent with microarray data, I showed that Sfmbt2 is expressed as early as the morula stage. Although the protein appears to be present in both the ICM as well as the TE lineage, the spatial organization of SFMBT2 becomes concentrated at the nuclear periphery by expanded blastocyst stage at day 4.5 of development in the TE lineage while remaining mostly cytoplasmic in the ICM. This is consistent with SFMBT2’s spatial organization in blastocyst outgrowths with the exception that in the blastocyst outgrowths some giant cells lack expression. In TS cells, SFMBT2 has a heterogeneous localization that appears to be dependent on the specific phase of the cell cycle; TS cells at the metaphase and anaphase stage of mitosis display SFMBT2 staining that overlaps with the mitotic chromosomes while the remaining cells at interphase show punctate nuclear staining.

Association of PcG proteins with the mitotic chromosomes has been previously
correlated with mitotic bookmarking, required for proper inheritance of epigenetic signatures during cell division and differentiation (reviewed in Kadauke and Blobel, 2013; Zaidi et al., 2011). In addition, MBT domain carrying proteins such as L3MBTL1 have been associated with mitotic chromosomes in facilitating cell-cycle progression (Koga et al., 1999). In the 5th chapter of this thesis, potential involvement of SFMBT2 in mitotic bookmarking will be further explored.

Many PcG proteins have been implicated in stem cell self-renewal, differentiation, and cell cycle progression. Ring1B has been associated with self-renewal in embryonic neuronal stem cells (NSCs) (Román-Trufero et al., 2009). NCSs lacking this histone H2A E3 monoubiquitin ligase exhibit upregulation of neural differentiation markers as well as down regulation of proliferation inhibitors. SUZ12, a component of PRC2, is involved in proper differentiation of ES cells (Pasini et al., 2007). When allowed to differentiate into neurons, ES cells lacking this protein were unable to repress ES cell specific markers such as Oct4, and Nanog and to activate differentiation makers such as Sox1 and Nestin. In light of these studies, and the observations from this chapter it can be hypothesized that SFMBT2 is required for TS cells either by roles in lineage commitment, self-renewal and/or cell cycle progression.
Chapter 4: *Sfmbt2* is Required for Maintenance of Trophoblast Cell-types and Placenta Development

Acknowledgement: Data collection for the genetics of gene trap allele transmission in both C57Bl/6 and CD-1 backgrounds, shown in (Table 4.1 and Table 4.2), was done by Dr. Sue Varmuza. Sangeetha Paramathas has contributed to this chapter by assisting me with some of the histology.

Disclaimer: Some of the data and figures presented in this chapter have been published in (Miri et al., 2013).
4.1 **Abstract**

Development of placenta is dependent on specification, maintenance and timely differentiation of TS cells. Aberrations in any of these processes are often catastrophic, resulting in IUGR and even embryonic death. The study of uniparental embryos with two maternal copies of the genome suggest that genes exhibiting differential expression from the paternal allele are associated with key processes of placentation. In this study, I show that paternal transmission of a null gene trap allele of the paternally expressed imprinted gene, *Sfmbt2*, results in embryonic lethality. Systematic analysis of the phenotype at various stages of mid-gestation revealed a significant role in the development of the extraembryonic tissues, mainly the placenta. I show that despite the persistence of a severely reduced population of trophoblast cells by embryonic day 8.5, *Sfmbt2* homozygous null embryos display no further expansion of the trophoblast lineage by embryonic day 9.25 and exhibit disappearance of the trophoblast cell types by day 9.75, suggesting that SFMBT2 is required for maintenance of the trophoblast lineage.
4.2 INTRODUCTION

The placenta is the most critical organ for proper development and growth of the embryo in utero. It serves as the interface between the mother and the fetus, facilitating nutrient and gas exchange, and fetal waste disposal (Rossant and Cross, 2001; Watson and Cross, 2005). The various cells of the placenta produce endocrine and paracrine hormones and cytokines to regulate the maternal physiology and accommodate the needs of the developing embryo (Cross et al., 2002; Hu and Cross, 2010b; Roberts and Fisher, 2011). Genes with essential roles in the development of the placenta result in embryonic lethality usually by embryonic day 10.5 to 11.5 at which point the embryo becomes dependent on this organ for sustenance.

The earliest differentiation events that take place during embryonic development result in the establishment of the trophectoderm, including progenitor cells of the trophoblast lineage which will later on contribute to the mature placenta (Cross, 2000). Outer totipotent blastomeres of early embryos commit to the TE fate while those positioned on the inside form the ICM giving rise to the embryo. Expansion of a cavity within the sphere of cells leads to the formation of the blastocyst with the TE cells on the periphery and the ICM attached to one side. The TE cells away from the ICM, the mural trophectoderm, endoreduplicate, generating primary TGCs which aid in the implantation processes as well provide an interphase for nutrient and gas exchange before the functional placenta has developed. Trophectoderm cells juxtaposing the ICM, called the polar trophectoderm, are thought to harbour TS cells, a population of progenitor cells that give rise to various cell types of the mature placenta. The polar trophectoderm further proliferates and forms the extraembryonic ectoderm, containing progenitors of chorionic ectoderm that later form the trophoblast cells of the labyrinthine layer, and the EPC,
which is thought to form the SpT layer as well the secondary PTGCs in contact with the maternal tissue.

Genes involved in the specification of the trophoblast lineages act as early as the morula stage, and in conjunction with blastomere positioning, establish the TE fate (Senner and Hemberger, 2010). In the outer cells of the morula, de-phosphorylation of YAP results in its translocation to the nucleus and binding its coactivator TEAD4 which subsequently activates Cdx2 (Nishioka et al., 2009). Hence, nuclear YAP sets in motion a cascade of events resulting in down regulation of pluripotency factors such as Oct4, establishing the trophoblast lineage. Ras-MAPK signalling pathway, involving the downstream effector protein ERK2 has also been implicated in the specification process (Lu et al., 2008). Null mutations of these genes show embryonic lethality as early as the blastocyst stage and often prevent implantation, suggestive of their involvement in the specification of the trophoblast lineage (Nishioka et al., 2008; Strumpf et al., 2005; Yagi et al., 2007). Essential genes involved in the maintenance of the trophoblast lineage have knockout phenotypes associated with the trophoblast tissue that often result in lethality between implantation and mid-gestation. These genes include: the earlier acting transcription factor Eomes (Russ et al., 2000), required at the blastocyst stage; Tcfap2c (Auman et al., 2002; Werling and Schorle, 2002) and Ets2 (Yamamoto et al., 1998) with mutants surviving up to embryonic day 8.5; and Elf5 (Donnison et al., 2005) and Err-β (Luo et al., 1997) with null mutations exhibiting lethality by day 10.5 of development.

Imprinted genes which exhibit parent-of-origin-specific differential expression have also been associated with proper development of the trophoblast lineage (Coan et al., 2005). Parthenogenetic embryos with two maternal genomes have hypoplastic trophoblast and die at early mid-gestation likely due to failure in maintaining the
trophoblast lineage (Clarke et al., 1988; Fundele, 1990; Varmuza et al., 1993). Nevertheless, knockout studies of PEGs have so far only identified genes with roles in trophoblast development in late-gestation (Miri and Varmuza, 2009). Hence to date, early acting PEGs responsible for the mid-gestational phenotypes observed in the parthenogenetic embryos are unaccounted for.

In this study, I show that the imprinted gene Sfmbt2 is the earliest acting PEG with essential roles in the development of the extraembryonic tissues. SFMBT2 is a PcG protein, imprinted exclusively in tissues of the extraembryonic origin, including TS cells and placenta where it shows robust expression (Kuzmin et al., 2008). I show that paternal inheritance of a gene trap null allele of Sfmbt2 results in embryonic lethality in a strain dependent manner with C57BL/6 mice displaying a more severe phenotype and mid-gestational lethality and CD-1 mice exhibiting a much more variable phenotype with some survivors to term. Both strains however show reduction in various trophoblast cell types and activation of the maternal allele in trans in presence of a paternally inherited gene trap allele of Sfmbt2. I further show that the homozygous mutants exhibit the most severe phenotype, evident as early as embryonic day 8. In addition, I performed a systematic histological analysis of null embryos at various stages of development and show that SFMBT2 is required for maintenance of the trophoblast cell-types.

4.3 RESULTS

4.3.1 SFMBT2 Displays Nuclear Staining in Various Cell-types of the Trophoblast Origin

The expression pattern of SFMBT2 was investigated at various stages of embryonic development starting from early mid-gestation to late-gestation. At approximately day
8.0 of embryonic development, before the allantoic attachment to the chorionic plate, *Sfmbt2* is exclusively expressed in the extraembryonic tissues and shows no staining in the embryo (Figure 4.1). The expression is predominantly nuclear in the trophoblast cells of the chorionic plate (Figure 4.1 D,E) and absent in the mesothelial cells of the chorion and allantois (Figure 4.1 E, arrow). The trophoblast cells proximal to the mesoderm layer of the chorionic plate show consistent staining while the distal trophoblast cells show reduced or no staining (Figure 4.1 E). PTGCs proximal to the chorionic plate show nuclear staining (Figure 4.1 D,F) with the exception of very few giant cells that are negative for the expression of *Sfmbt2* (Figure 4.1 D, arrow). Cytoplasmic staining in the decidua (maternal) and low or no staining in the cells of the EPC was used as a guide to draw a boundary between the two layers, distinguishing fetal and maternal cells (Figure 4.1 C). The distal region of EPC is devoid of SFMBT2 staining (Figure 4.1 B,C) but the proximal region closest to the chorionic plate contains the occasional cell with nuclear staining (Figure 4.1 C,F).

At embryonic day 15, SFMBT2 immunostaining in the mature placenta follows a similar pattern to that observed at early-mid-gestation (Figure 4.2). Trophoblast cells juxtaposing the negatively stained mesenchymal cells of placenta show consistent expression of SMFBT2 (Figure 4.2 B,E). S-TGCs of the labyrinthine layer, identified by their large nuclei and position in the labyrinthine layer show variable staining (Figure 4.2 C,E, black and white arrows show stained and non-stained cells respectively). The majority of SpT cells show very little staining (Figure 4.2 D) with the exception of those on the periphery which show a slightly higher nuclear staining (Figure 4.2 G). Glycogen cells of the placenta are completely devoid of SFMBT2 (Figure 4.2 F,G) and the PTGCs show punctate staining or no staining (Figure 4.2 G, black and white arrows
Figure 4.1. SFMBT2 shows nuclear staining in various cell-types of the trophoblast origin at mid-gestation.

(A-F) Immunostaining using anti-SFMBT2 antibody (black) and methyl green counterstaining (blue-green) on histological sections of e8.0 to e8.25 embryos harvested prior to allantoic attachment to the chorionic plate. A and B are two different embryos from the same litter. Arrow in D indicates a PTGC exhibiting lack of SFMBT2 staining. Arrow in E indicates the mesothelial layer of the chorionic plate with no SFMBT2 staining. De, decidua; PTGC, Parietal trophoblast giant cells; Al; allantois; Ch, chorion; EPC, ectoplacental cone; Emb, embryo. Scale bars: (A,B: 200µm); (C,D: 100µm); (E,F: 50µm).
Figure 4.2. Pattern of SFMBT2 staining in the placenta at late-gestation.

(A-G) Immunostaining using anti-SFMBT2 antibody and methyl green counterstaining on histological sections harvested from e15.5 placentae. White arrows indicate S-TGCs (C,E) and a PTGC (G) exhibiting no SFMBT2 staining. Black arrows indicates a S-TGC (C) and a PTGC (G) exhibiting SFMBT2 staining. PTGC, Parietal trophoblast giant cells; Lb, labyrinth; SpT, spongiotrophoblast; Gly, Glycogen cells; Mes, Mesenchymal cells; Al; allantois; Scale bars: (A: 200µm); (B-G: 25µm).
respectively). The dynamic changes in SFMBT2 distribution in various stages of development will be further described in a comparison of wild type and mutant embryos.

4.3.2 Paternal Transmission of the Null Allele of Sfmbt2 is Embryonic Lethal

A chimeric male with a gene trap allele of Sfmbt2 was generated by Mutant Mouse Regional Resource Center (MMRRC) using DCO755 ES cells of 129/Ola strain (Miri et al., 2013). The gene trap allele contains a β-geo cassette inserted downstream of the first coding exon of Sfmbt2, resulting in truncation of both the short and the long isoforms of this gene to the first 32 amino acids fused in-frame to β-geo coding sequence, generating a predicted null allele (Figure 4.3 A,B). The chimeric male was bred to C57BL/6 females, and produced one gene trap F1 male out of 30 ES derived offspring. Cryopreserved sperm from the original F1 male was subsequently injected into C57BL/6 oocytes, giving rise to a single male gene trap offspring (F2), out of 17 live pups. The F2 male was obtained from MMRRC and crossed with females of C57BL/6 and CD-1 strains.

Table 4.1 and Table 4.2 show the number of live births from reciprocal crosses of Sfmbt2 gene trap and WT mice in both C57BL/6 and CD-1 backgrounds. Paternal inheritance of the gene trap allele is mostly lethal when heterozygous (+/gt or gt/+ ) males are mated with C57BL/6 females. The only survivors were from the F2 male including a female heterozygous offspring which was subsequently used to maintain the gene trap colony. When CD-1 females are used, paternal transmission of the gene trap allele results in lethality with reduced penetrance. The maternal transmission of the null allele in both C57BL/6 and CD-1 backgrounds results in generation of gene trap offspring with expected Mendelian ratios. These data suggest that the paternal transmission of the
Figure 4.3. Paternal inheritance of a null Sfmbt2 allele is lethal due to defects in proper development of the extraembryonic tissues.

(A) Schematic representation of Sfmbt2 gene trap allele showing the insertion of a splice acceptor (SA) sequence followed by a β-geo cassette and a strong polyA signal (pA) downstream of the first coding exon. The two transcriptional start sites for Sfmbt2 are indicated by TSS1 and TSS2. (B) Schematic representation of wild type (both long and short isoform) and truncated SFMBT2. Numbers correspond with the amino acid sequence. (C,D) Gross morphological comparison of e10.5 +/- (Lac-Z negative) and +/-gt (Lac-Z positive) fetuses with (C) and without (D) yolk sac. (E-H) Histological sections of e10.5 +/- (Lac-Z negative) and (E’-H’) +/-gt (Lac-Z positive) extraembryonic tissues stained with H&E and X-gal. (F”) PTGC displaying no Lac-Z staining (arrow) magnified from (F’). De, decidua; Lb, labyrinth; SpT, spongiotrophoblast; PTGC, parietal trophoblast giant cells; PrE, parietal endoderm. (E,E’: 200µm); (F’-H’: 50µm).
Table 4.1. *Sfmbt2* Gene Trap is Paternal Lethal at Mid-Gestation in C57BL/6 background.

### A. Paternal Transmission is Lethal

<table>
<thead>
<tr>
<th></th>
<th>/+gt</th>
<th>++</th>
<th>Dead/resorbed</th>
</tr>
</thead>
<tbody>
<tr>
<td>C57BL/6♀ X F2♂</td>
<td>5</td>
<td>106</td>
<td>na</td>
</tr>
<tr>
<td>C57BL/6♀ X F3♂</td>
<td>0</td>
<td>90</td>
<td>na</td>
</tr>
<tr>
<td>F3-10♀ X C57BL/6♂</td>
<td>169</td>
<td>193</td>
<td>na</td>
</tr>
</tbody>
</table>

### B. Time of Death

<table>
<thead>
<tr>
<th></th>
<th>/+gt</th>
<th>++</th>
<th>Dead/resorbed</th>
</tr>
</thead>
<tbody>
<tr>
<td>e7.5</td>
<td>25</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>e9.5</td>
<td>17</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>e11.5</td>
<td>2</td>
<td>25</td>
<td>14*</td>
</tr>
<tr>
<td>e12.5</td>
<td>0</td>
<td>25</td>
<td>24</td>
</tr>
</tbody>
</table>

*Five arrested embryos and 7 resorptions were /+gt. The remaining 2 resorptions were too degraded for DNA extraction.*
Table 4.2. *Sfmmbt2* Gene Trap is Paternal Lethal in CD-1 background but with reduced penetrance.

<table>
<thead>
<tr>
<th></th>
<th>+/gt</th>
<th>+/+</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CD-1 ♀ X F2/F3 ♂</strong></td>
<td>76</td>
<td>140</td>
</tr>
<tr>
<td><strong>F2/F3 ♀ X CD-1 ♂</strong></td>
<td>11</td>
<td>18</td>
</tr>
</tbody>
</table>
gene trap allele results in a hypomorphic mutation, discussed below.

4.3.3 **Paternal Inheritance of the Gene Trap Allele of Sfmbt2 Leads to Severe Defects in Extraembryonic Tissues in a Strain Dependent Manner**

4.3.3.1 *Phenotype of the hypomorphic allele in C57BL/6 background*

Conceptuses generated from mating wild type C57BL/6 females with heterozygous males, recovered as early as embryonic day 10.5 and carrying a paternally inherited gene trap allele (+/gt) have morphologically small embryos and placentae when compared to wild type (+/+) counterparts (*Figure 4.3* C,D). Histological examination of +/gt extraembryonic tissues at e10.5 using Lac-Z staining (*Figure 4.3* E-H, E’-H’) and SFMBT2 immunostaining (*Figure 4.4* A-D,A’-D’) revealed that, in comparison with wild type littermates, all three trophoblast layers of the placenta are severely compromised. The labyrinthine layer in +/gt placentae appears to have undergone the early stages of chorioallantoic branching morphogenesis evident from villi projected into the chorion (*Figure 4.4* A’-C’; arrowheads); however, in comparison with the wild type littermates displaying extensive branching morphogenesis with the presence of both maternal and fetal blood (*Figure 4.4* A,C), the +/gt labyrinth is severely underdeveloped. The SpT layer is also severely reduced in +/gt placentae (*Figure 4.4* B’) when compared to its wild type counterpart (*Figure 4.4* B), although expression of SFMBT2 in these cells is faint or absent. Finally, all TGC compartments, including those associated with SpT (*Figure 4.4* B, B’), labyrinth (*Figure 4.4* C,C’) and the PTGCs juxtaposing both the placenta (*Figure 4.4* D,D’) and the parietal endoderm (*Figure 4.3* H,H’) are reduced.

Despite the reduction in the various layers of the +/gt placentae, the expression pattern of SFMBT2 remains consistent when compared to wild type littermates,
Figure 4.4. Paternal inheritance of a null Sfmbt2 allele results in reduction of all trophoblast cell types in C57BL/6 background.

(A-D and A’-D’) SFMBT2 immunostaining on e10.5 +/+ (A-D) and +/gt (A’-D’) placentae showing reduction of all trophoblast cell types including: SpT (B,B’); Lb (C,C’); and PTGCs (D,D’). Methyl green was used as a counter stain. White arrowheads indicate unstained TGCs (B,B’,D,D’). Black arrowheads indicate allantoic protrusions into the chorion shown by the dotted lines (B’,C’). (E) qRT-PCR results showing expression of various markers in +/gt and +/+ e9.5 placenta as a ratio to Tbp. Sfmbt2 (L) and (S); long and short isoforms of Sfmbt2, respectively. Bars indicate standard deviation. *p<0.05; **p<0.01; ***p<0.001. De, decidua; Ys, yolk sac; Lb, labyrinth; SpT, spongiotrophoblast; PTGC, Parietal trophoblast giant cells; Al; allantois; FB, fetal blood; MB, maternal blood; Ch, chorion; TGC, trophoblast giant cells; Mes, allantoic mesothelium. Scale bars: (A,A’: 100µm); (B-D, B’-D’: 25µm).
indicative of LOI. Strong punctate nuclear staining can be seen in TGCs within the SpT (Figure 4.4 B,C’), and in some of the PTGCs juxtaposing the decidua (Figure 4.4 D,D’), while some of the PTGCs display no staining (Figure 4.4 D,D’; white arrowheads). The SpT cells show faint, mainly cytoplasmic, staining and no nuclear staining (Figure 4.4 B; arrowheads). The trophoblast cells of the labyrinthine layer display robust signal, while endothelial cells remain unstained (Figure 4.4 C,C’).

A survey of genes expressed in various cell types of the placenta was carried out on both +/+ and +/gt placentae harvested from day 9.5 conceptuses (Figure 4.4 E). To minimize loss of tissue during dissection and maximize the integrity of the placenta and PTGCs, the entire placental unit including the decidua was used for RNA extraction and cDNA synthesis while the embryo and yolk sac was used for genotyping. All trophoblast specific markers including Cdx2 (TSC marker), GcmI (chorionic trophoblast), Tpbpa (SpT), and Pl-1 (PTGC), show statistically significant reduction in +/gt placentae in comparison with wild type counterparts. Runx1 however, expressed in the allantoic mesothelium (Mes) juxtaposing the chorionic trophoblast (Zeigler et al., 2006) did not show any reduction. In addition, qRT-PCR results show a significant reduction in both the long and the short isoforms of Sfmbt2, suggesting partial reactivation of the maternal allele in each cell or a subset of cells. These data indicate that reduction of Sfmbt2 correlates with reduction in trophoblast specific cell-type markers suggesting that SFMBT2 has a critical role in all cell types of trophoblast origin found at this stage of development. Nevertheless normalization of these results with a pan-trophoblast marker such as Cytokeratin would be informative to elucidate whether the observed reduction in the markers of various trophoblast cell-types is a function of reduced levels per cell or a general reduction of various trophoblast cell-types.
Histological sections of e10.5 wild type and heterozygous conceptuses with a paternally inherited gene trap allele were immunostained with Ki67 and Caspase-3 to determine whether the severe reduction in trophoblast is due to impairment in proliferation or Caspase-3 dependent cell death, respectively (Figure 4.5). Ki67, a nuclear protein associated with cellular proliferation, appears to be present in both +/+ and +/-gt placentae (Figure 4.5 E,E’,F,F’). The placentae of both wild type and heterozygous conceptuses appear to be devoid of Caspase-3 positive cells (Figure 4.5 C,C’,D,D’) suggesting that the reduction in trophoblast layers is not due to Caspase-3 dependent cell death. Nevertheless, extensive staining in +/-gt fetuses (Figure 4.5 C”,D”) is indicative of embryonic death at this stage of development, likely due to secondary effects of placental failure.

4.3.3.2 Phenotype of the hypomorphic allele in CD1 background

Gross morphological analysis of conceptuses generated from mating CD-1 females with heterozygous males revealed a more moderate phenotype associated with the hypomorph mutation. Heterozygotes with a paternally transmitted gene trap allele (+/gt) could be readily distinguished at embryonic day 12.5 of gestation (Figure 4.6 A). Embryos show growth retardation and appear anemic in comparison to their wild type litter-mates and placentae from heterozygous embryos appear smaller with noticeably reduced SpT layer.

Rodents have hemochorial placentae whereby maternal vasculature in the placenta is remodelled by trophoblast cells of fetal origin as opposed to endothelial cells (Rossant and Cross, 2001). Normal development and growth of the embryo has been associated with proper placental vascularization. The trophoblast layer lining the maternal blood space has been shown to exhibit alkaline phosphatase activity.
Figure 4.5. Reduction in trophoblast cell types in not due to Caspase dependent cell death.

Histological sections of +/+ (A,A’,C,C’,E,E’) and +/gt (B,B’,D,D’,F,F’) placentae; and +/+ (C’’) and +/gt (D’’) embryos stained with H&E (A,A’,B,B’), Capsase-3 (C,C’,C’’,D,D’,D’’), and Ki67 (E,E’’,F,F’’). Methyl green was used as a counterstain. De, decidua; Pl, placenta; Emb, embryo; Lb, labyrinth; SpT, spongiotrophoblast; PTGC, parietal trophoblast giant cells. Scale bars: (A-F 200μm); (A’-F’: 50μm).
Figure 4.6. Paternal inheritance of a null Sfmbt2 allele results in anemia in the CD-1 background.

(A) Gross morphological comparison of e12.5 +/- and +/-gt fetuses and placentae with (top) and without (bottom) the yolk sac. Bottom images indicate the visibly reduced SpT in the +/-gt placenta and anemia in the +/-gt embryo. (B-D) Histological sections of e12.5 +/- (B) and +/-gt (C,D) placentae stained with alkaline phosphatase (B,C) and X-gal (D). Nuclear fast red was used as a counterstain. Lb, labyrinth; SpT, spongiotrophoblast; Pl, placenta; Ys, yolk sac. Scale bars: (B-D: 200µm); (B’-D’ and B”-D”: 50µm).
Heterozygous placentae at embryonic day 12.5 stained for alkaline phosphatase appear to be similar to their wild type littermates (Figure 4.6 B,C). Lac-Z staining of the heterozygous placentae illustrates a staining pattern similar to that of the alkaline phosphatase (Figure 4.6 C,D), highlighting that the expression of Sfmbt2 is limited to the trophoblast lineage. Furthermore, comparison of the cross sections of wild type and heterozygous conceptuses at embryonic day 12.5 shows a noticeable reduction in the SpT layer in +/gt placentae (Figure 4.6 B’,C’,D’).

Gross morphological analysis of embryonic day 14.5 wild type and heterozygous conceptuses of the CD-1 background shows variability in the development of the heterozygotes (Figure 4.7 A) with some embryos showing significant retardation (top embryo) and others showing mild phenotypes (middle embryo). In some cases genotyping showed that embryos exhibiting wild type morphology were in fact heterozygous for the gene trap allele (paternally inherited)-image not included. Histological comparison of e14.5 and e18.5 heterozygous placentae showed reduction in the SpT layer. Glycogen cells of the junctional zone however do not seem to be negatively affected (Figure 4.7 B-E). The distribution of glycogen cells in wild type placentae appears to be limited to small pockets within the SpT layer at e14.5 (Figure 4.7 B’) and as expected (Coan et al., 2006) shows reduction by e18.5 (Figure 4.7 D’); in heterozygotes however, the glycogen cells are clustered together (Figure 4.7 C’) and do not exhibit an apparent reduction at embryonic day 18.5 (Figure 4.7 E’).

The expression profile of markers of various placental cell types of e14.5 wild type and heterozygous placentae were measured using qRT-PCR (Figure 4.7 F). Expression of Ctsq (S-TGCs) and Runx1 (endothelial cells of the mesoderm origin) did not show a significant change in +/gt placentae when compared to wild type litter mates.
Figure 4.7. Paternal inheritance of a null Sfmbt2 allele results in reduction of various trophoblast subtypes in the CD-1 background.

(A) Gross morphological comparison of e14.5 +/- and +/-gt conceptuses with the placenta and the yolk sac attached showing variation in the penetrance of the hypomorphic mutation. (B-E and B’-E’) Histological sections of e14.5 +/- (B,B’) and +/-gt (C,C’) and e18.5 +/- (D,D’) and +/-gt (E,E’) placentae stained with H&E. Whole placentae were stained with X-gal prior to sectioning, however due to low penetration of this compound only the cells of the periphery show X-gal staining (C,E’; arrow). Dashed area in B to E indicates the boundary between the labyrinth layer and the junctional zone. Dashed area in B’ to E’ indicates the location of the Glycogen cells. (F) qRT-PCR results showing expression of various markers in +/-gt and +/- e14.5 placenta as a ratio to Tbp. Sfmbt2 (L) and (S); long and short isoforms of Sfmbt2, respectively. Bars indicate standard deviation. *p≤0.05; **p≤0.01. Lb, labyrinth; SpT, spongiotrophoblast; Gly, Glycogen cells; Pl, placenta; Ys, yolk sac; TGC, trophoblast giant cell; Mes, allantoic mesothelium; STGC, Sinusoidal TGC. Scale bars: B-E: 200µm; B’-E’: 50µm.
There was a significant reduction in the SpT maker Tbpba as well as placental prolactin family of hormones such as PL-II, PRP and PLF (Simmons et al., 2008) expressed from TGCs. These results are consistent with the observed morphological differences between wild type and heterozygous placentae, in that the most severely affected tissues and cells in the hypomorphs include the SpT layer as well as the TGCs.

4.3.4 SFMBT2 null mutation results in mid-gestational embryonic lethality due to failure to maintain trophoblast cell-types

4.3.4.1 Phenotype of homozygous null embryos

Reactivation of the maternal allele associated with the paternal transmission of the Sfmbt2 gene trap allele may partially rescue the phenotype associated with nullizygosity. I performed a systematic analysis of the phenotype associated with null SFMBT2 conceptuses at various stages of mid-gestation to further elucidate the role of this gene in the development of the trophoblast cell-types.

Homozygous null conceptuses were obtained by intercrosses of heterozygous maternal carriers of the gene trap allele. Histological analysis of +/+ and gt/gt conceptuses was performed at embryonic day 8.0 of gestation prior to the chorioallantoic attachment. The genotype of the gt/gt embryos were readily identified by the complete lack of SFMBT2 signal following immunohistochemistry. Figure 4.8 shows a null gene trap embryo (gt/gt) (Figure 4.8 C-F) that appears to be slightly delayed in development in comparison with the wild type littermates (Figure 4.8 A,B). Despite the similarity in the protrusion of the allantois into the exocoelom and towards the chorionic plate in both the +/+ and gt/gt embryos, the development of placenta in the gt/gt embryo is slightly delayed (Figure 4.8 A’-D’; arrows). Histological analysis of gt/gt embryos at e8.5, after
Figure 4.8. SFMBT2 null embryos display abnormalities in the extraembryonic tissues as early as e8.0.

Histology of e8.0 +/+ (A,A’,B,B’) and gt/gt (C-F, C’-F’) conceptuses showing H&E staining (A,A’,C,C’); SFMBT2 staining (B,B’,D,D’); pan-cytokeratin staining (E,E’); and CDX2 staining (F,F’). Methyl green was used as a counter stain in immunostained sections. Arrows indicate the chorionic plate (A’-B’) and the EEE (C’-D’). De, decidua; PTGC, Parietal trophoblast giant cells; Al; allantois; Ch, chorion; EPC, ectoplacental cone; EEE, extraembryonic ectoderm; Emb, embryo. Scale bars: (A-E: 200µm); (A’-E’: 50µm).
chorioallantoic attachment, further revealed a reduction in various lineages of the extraembryonic tissues, namely the EPC (Figure 4.9 E,E’), the chorionic plate (Figure 4.9 I,I’) and the PTGC layer (Figure 4.9 J,J’). This observation was confirmed by pan-cytokeratin staining outlining the trophoblast cell types (Figure 4.8 E,E’ and Figure 4.9 G,G’). Despite this reduction, trophoblast cells of the gt/gt embryos display robust Cdx2 expression, a marker of trophoblast (Figure 4.8 F,F’ and Figure 4.9 H,H’).

Histological analysis of gt/gt embryos at e9.25 and e9.75 was performed to refine the observed phenotype (Figure 4.10, Figure 4.11, respectively). Results indicate that despite the persistence of the chorionic plate by e9.25 of development (Figure 4.10), there is no evident advancement in the development of the placenta from e8.5 embryos (Figure 4.9). By e9.75, when the chorioallantoic branching morphogenesis is well underway in wild type embryos (Figure 4.11 A,C), gt/gt embryos show disappearance of trophoblast cells in the chorionic plate (Figure 4.10 A’,C’). In addition, at e9.75, unlike their wild type littermates (Figure 4.11 D,E), gt/gt embryos retain only rare PTGCs (Figure 4.11 D’). A composite image of +/+ and gt/gt conceptuses at various stages of development from e8.0 to e9.75 suggest that SFMBT2 plays a critical role in the maintenance and expansion of various trophoblast cell types (Figure 4.12).

Despite lack of expansion of the trophoblast lineage, immunostaining experiments using trophoblast markers such as Cdx2 and Cytokeratin indicate persistence of their expression by day 9.25 of development (Figure 4.10 E-H and E’-H’). Absence of Caspase-3 staining in the chorionic plate and EPC cells of the e9.25 gt/gt conceptuses suggests that failure to maintain and expand these tissues does not correlate with Caspase-3 dependent cell death (Figure 4.10 J’). Positive Ki67 staining in the chorionic cells of the gt/gt embryos suggests that they are not quiescent (Figure 4.10 L’).
Figure 4.9. SFMBT2 null embryos display reduction in various cell-types of the trophoblast origin.

Histology of e8.5 +/+ (A-L) and gt/gt (A'-L') embryos showing H&E staining (A,E,I,A',E',I'); SFMBT2 immunostaining (B,F,J,B',F',J'); pan-cytokeratin immunostaining (C,G,K,C',G',K'); CDX2 immunostaining (D,H,L,D',H',L'). Methyl green was used as a counter stain for immunostained sections. Dotted lines (F-H, E’-H’) show the border between the fetal and the maternal tissues. De, decidua; PTGC, Parietal trophoblast giant cells; Al, allantois; Ch, chorion; EPC, ectoplacental cone; Emb, embryo. Scale bars: (A-D,A'-D': 200µm); (E-H,E'-H': 100µm); (I-L,I'-L': 25µm).
Figure 4.10. SFMBT2 null embryos fail to expand cell-types of the trophoblast origin.

Histology of e9.25 +/+ (A,C,E,G,I,K) and gt/gt (B,D,F,H,J,L) embryos showing H&E staining (A,B); SFMBT2 (C,D); pan-cytokeratin (E,F); CDX2 (G,H); Caspase-3 (I,J); and Ki67 (K,L) immunostaining. Methyl green was used as a counter stain for immunostained sections. De, decidua; PTGC, Parietal trophoblast giant cells; Al; allantois; Ch, chorion; EPC, ectoplacental cone; Emb, embryo. Scale bars: (A-L: 200µm); (A’-L’: 100µm).
Figure 4.11. SFMBT2 null embryos show disappearance of tissues of the trophoblast origin by e9.75.

Histology of e9.75 +/- (A-E) and gt/gt (A'-D') extraembryonic tissues in decidua. H&E staining (A,A'); SFMBT2 immunohistochemistry (B-E,B'-D'); Arrowheads indicate PTGCs (C',D). Arrow indicates chorionic trophoblast cells (C'). Methyl green was used as a counter stain. De, decidua; Lb, labyrinth; SpT, spongiotrophoblast; PTGC, Parietal trophoblast giant cells; Al, allantois; Ch, chorion, FB, fetal blood; Emb, embryo. Scale bar: (A,A': 200µm); (B,D,B': 100µm); (C,C',D': 50µm); (E: 25µm).
Figure 4.12. SFMBT2 null embryos fail to maintain tissues of trophoblast origin.

Composite histological images of e8.0 to e9.75 +/+ and gt/gt conceptuses as labeled, stained with H&E and immunostained with SFMBT2 to highlight the genotype. Arrow shows the chorionic plate. Scale bar: (A-H and A’-F’: 200µm); (G’,H’: 50µm).
4.4 **DISCUSSION**

In this study I have characterized the function of the paternally expressed PcG protein SFMBT2 during gestation and show that SFMBT2 is the earliest acting PEG with essential roles in maintenance of the trophoblast lineage and proper development of the placenta. In mice, the mature placenta is formed by day 10.5 of gestation (Cross et al., 2002) by timely differentiation of trophoblast progenitor cells, namely TS cells, into the trophoblast subtypes that contribute to the three main layers of the placenta. Self-renewal and maintenance of TS cells *in vivo* is therefore critical for proper development of this organ. *In vitro* studies show that TS cells can be derived from blastocysts, extra-embryonic ectoderm of e6.5 embryos as well as the chorionic plate of e8.5 conceptuses, indicative of their presence at these stages of development (Tanaka et al., 1998a; Uy et al., 2002). Furthermore, differentiation of *in vitro* derived TS cells is correlated with expression of markers of various cell-types of placenta, suggestive of their potential to mirror the *in vivo* events and therefore their multipotent integrity (Tanaka et al., 1998).

At approximately day 8.5 of gestation, SFMBT2 displays a robust nuclear signal in the trophoblast cells of the chorionic plate, especially in those proximal to the mesodermal layer and the embryo (Figure 4.1). SFMBT2 staining in the distal chorion appears to be reduced while the EPC staining is limited to few TGCs. Self-renewal and maintenance of TS cells have been associated with the TGF-β related protein NODAL and FGF4 signalling known to originate from the ICM at the blastocyst stage and the ICM derivatives such as the epiblast at early mid-gestation *in vivo* (Dardik et al., 1992; Guzman-Ayala, 2004). Derivation and maintenance of diploid TS cells *in vitro* is also dependent on NODAL/TGF-β factors as well as FGF4 which are supplemented in the medium TS cells are derived and maintained in (Erlebacher et al., 2004; Tanaka et al.,
Consistent expression of *Sfmbt2* in the trophoblast cells of the chorionic plate proximal to the ICM derivatives is suggestive of a putative role in the TS cell pool of this tissue.

Although the PTGCs bordering the EPC, hereafter referred to as the distal-PTGCs (d-PTGCs), show low or no SFMBT2 staining (*Figure 4.1 D, Figure 4.9 J*), PTGCs immediately next to the chorionic plate, hereafter referred to as proximal-PTGCs (p-PTGCs) display robust nuclear staining at e8.0 to e8.75 (*Figure 4.1 F, Figure 4.9 J*). This variation in the pattern of staining of SFMBT2 in the PTGCs may be explained by two potential scenarios. The first possibility entails reduction of SFMBT2 protein with progressive differentiation to the cells of the EPC and subsequently d-PTGCs and reactivation of *Sfmbt2* in a subset of PTGCs positioned in close proximity to the chorion, p-PTGCs. This assumes that all PTGCs come from EPC progenitors and upon terminal differentiation into giant cell fate they assume a distinct gene expression profile dictated by their spatial organization. The second and more probable scenario also entails reduction of SFMBT2 with progressive differentiation giving rise to d-PTGCs but assumes spontaneous differentiation of TS cells resulting in p-PTGCs and hence retention of the SFMBT2 signal. The latter scenario is consistent with studies by Simmons and colleagues using a *Tpbpa* derived Cre-recombinase line crossed with a lac-Z/alkaline phosphatase dual reporter line showing that not all secondary TGCs of the placenta come from progenitor cells expressing *Tpbpa*, a marker of the EPC (Simmons et al., 2007). Furthermore, Retinoic acid treatment of TS cells cultured in the presence of proliferating factors (FGF4 and feeder conditioned media) has been shown to result in terminal differentiation into giant cells without assuming an intermediate EPC fate, assessed by lack of *Tpbpa* expression (Yan et al., 2001). Additionally, Retinoic acid
administration during pregnancy resulted in a substantial increase in \textit{Pl-I} expressing p-PTGCs and a reduction of \textit{Tpbpa} positive EPC cells. In further support of the latter scenario, spontaneous differentiation into TGCs in the presence of FGF4 has been reported in culture.

Consistent with the staining pattern of SFMBT2, homozygous null mutants of this gene exhibit severe abnormalities in the development of the extraembryonic tissues of trophoblast origin. Histological analysis of e8.75 conceptuses, processed subsequent to the attachment of the allantois to the chorionic plate, show that \textit{Sfmbt2} null embryos exhibit reduced trophoblast cells in the chorionic plate, EPC and PTGCs (Figure 4.9). At e9.25 when branching morphogenesis is underway in wild type embryos, null mutants display thinning of all trophoblast layers, including the chorionic plate, EPC and PTGCs and no branching morphogenesis (Figure 4.10). \textit{Cdx2} and Cytokeratin staining suggests that the trophoblast subtypes are present at e8.75 and e9.25, however trophoblast layers containing cells stained with either protein are substantially reduced in null mutants (Figure 4.9, Figure 4.10). Nonetheless, absence of Caspase-3 staining in mutant placentae at e9.25 indicates that the reduction in the trophoblast layers is not due to caspase dependent cell death (Figure 4.10). By day 9.75 of embryonic development however, disappearance of the trophoblast cell types of \textit{Sfmbt2} null conceptuses is observed with the exception of a few PTGCs, albeit with visually smaller nuclei (Figure 4.11). These observations suggest that \textit{Sfmbt2} is required for proper maintenance of the trophoblast lineage (Figure 4.12).

\textit{Ki67} immunostaining, undertaken to elucidate the proliferative capacity of \textit{Sfmbt2} null extraembryonic tissues showed a similar pattern in mutant embryos and their wild type littermates with cells of the chorionic plate showing positive staining and cells
of the EPC showing no staining (Figure 4.10). This observation suggests that in both wild type and mutant placentae, the EPC cells are post-mitotic while cells of the chorionic plate exhibit a proliferative signature. Nevertheless, Ki67 protein is present during all phases of the cell cycle except for G0 (Scholzen and Gerdes, 2000) hence it is unclear whether the mutant cells can progress through the cell cycle similar to their wild type counterparts. Based on the observation that the EPC cells are post-mitotic, it can be hypothesized that the maintenance of the EPC and its derivatives is facilitated by a pool of progenitors whose differentiation leads to the expansion of the EPC and later the SpT layer. Considering the small size of the null mutant chorionic plate, SFMBT2 deficient TS cells are likely unable to self-renew their population. Furthermore, contrary to mutant phenotypes of trophoblast-maintenance-associated genes such as Mash2 exhibiting premature differentiation into giant cells (Guillemot et al., 1994; Scott et al., 2000; Tanaka et al., 1997), Sfmbt2 null embryos exhibit a severely reduced TGC layer followed by disappearance of these cells by e9.25 gt/gt embryos suggesting that SFMBT2 has roles in proper differentiation of progenitor cells.

Experiments performed by Simmons et al (2007) suggest that PTGCs have different developmental origins with approximately half differentiating from EPC progenitors expressing Tphpa and the other half from a district origin (Simmons et al., 2007). Furthermore, various subtypes of TGCs have been shown to exhibit distinct expression patterns of various prolactin hormones including placental lactogens. In situ hybridization experiments performed by Simmons and colleagues (2008) show that PLP-E, a placental hormone thought to promote blood cell production (Bhattacharyya et al., 2002; Zhou et al., 2005), is exclusively expressed in the p-PTGCs and absent in the EPC lineage (Simmons et al., 2008). Unlike the d-PTGCs and the EPC cells of the placenta,
the p-PTGCs show robust SFMBT2 nuclear staining. In addition, SFMBT2 deficient conceptuses show severe reduction of the p-PTGCs. Furthermore, qRT-PCR analysis on various paracrine and endocrine hormones expressed in TGCs of the placenta, including PL-I, PL-II, PLF, and PRP, indicates a significant reduction in placentae harvested from heterozygous mutants carrying a paternally inherited null Sfmbt2 allele (discussed below) when compared to wild type litter mates (Figure 4.4 E, Figure 4.7 F). Taken together, these observations suggest that SFMBT2 may play a role in regulating gene expression in a subset of TGCs.

Sfmbt2 is exclusively expressed from the paternal allele in the extraembryonic tissues of wild type embryos, suggesting that the paternal transmission of a null allele of Sfmbt2 would result in phenotypes resembling those of the homozygous null mutants. However, morphological and histological analyses of embryos with a paternally transmitted null allele of Sfmbt2 exhibit strain specific hypomorphic phenotypes. Heterozygous C57BL/6 mice carrying a paternal null allele of Sfmbt2 exhibited a more severe phenotype than those of the CD-1 background. In both C57BL/6 and CD-1 backgrounds, Sfmbt2 expression was detected from the maternal allele suggestive of LOI. Nevertheless, qRT-PCR showed only partial reactivation of the maternal allele in both strains.

In addition to placental Prolactin hormones discussed above, expression profiles of other trophoblast specific genes including, Cdx2, GcmI, and Tpbpa also showed a significant reduction in the hypomorphs, whereas Runx1, expressed in the endothelial cells of the mesoderm, was not affected. Ctsq was the only tested trophoblast specific marker that showed no variation in heterozygous mutants when compared to wild type littermates, suggesting that Sfmbt2 may not be required for this subclass of giant cells.
Consistent with these observations, histological analysis of placentae of e14.5 and e18.5 of the CD-1 background heterozygotes (+/gt) also identified the SpT as well as the PTGCs as the most severely affected trophoblast subtype. In contrast, the labyrinth layer where the Ctsq expressing S-TGCs reside appears to be morphologically similar in both wild type and heterozygous mutants. Furthermore, immunostaining experiments at e15.5 show that SFMBT2 does not exhibit consistent nuclear staining in S-TGCs of the labyrinth. Glycogen cells of the placenta also appear to be independent of SFMBT2 evident by the lack of SFMBT2 staining in wild type placentae and abundance of these cells in the junctional zone of the heterozygous mutants.

In conclusion, the function of Sfmbt2 appears to be most critical in trophoblast progenitor cells, known to contribute to various cell types of the placenta and is likely linked to their self-renewal and expansion.
Chapter 5: PcG Protein SFMBT2 Forms Complexes with Various Transcription Factors in TS cells

Acknowledgement: Katherine Szelag and Sangeetha Paramathas have contributed to this chapter by assisting me with the coimmunoprecipitation experiments.
5.1 **Abstract**

All cells of an organism have the same genetic makeup; without a proper readout facilitated by a process known as epigenetics, multicellular organisms would not exist. Operating in multi-subunit complexes, PcG proteins have been shown to play a significant role in the epigenetic regulation of the genome and have essential functions in cell fate and lineage commitment. SFMBT2 is a poorly characterized member of this group of proteins whose interacting partners in the extraembryonic lineage where the gene is predominantly expressed, are not known. By coimmunoprecipitation experiments, I showed that SFMBT2 forms a complex with a number of transcription factors, including YY1, CASZ1 and TFDP1. Immuno-labelling of the trophoblast progenitor cells with anti-SFMBT2, YY1 and TFDP1 antibodies suggest mitotic association of these proteins and a putative involvement in mitotic bookmarking, a process involved in preservation of cell fate through cell division, lineage commitment and differentiation.
5.2 INTRODUCTION

Higher-order structure of DNA is achieved through association with histones and their respective post-translational modifications (PTMs), leading to regulation of a number of biological processes including transcription, DNA repair, cell cycle and cellular identity (Rice et al., 2002; Schwartz and Pirrotta, 2013; Trojer and Reinberg, 2008; Trojer et al., 2007; Trojer et al., 2007; Yohn et al., 2003). Initially discovered in *Drosophila melanogaster*, PcG proteins are a highly conserved group of repressors that have been shown to associate with distinct PTMs of histone tails and regulate acquisition and maintenance of higher-order DNA structures (Beuchle et al., 2001; Schwartz and Pirrotta, 2013; Sharif et al., 2013). In *Drosophila*, spatiotemporal expression of homeotic genes dictates segmental expression of genes required for proper differentiation and anteroposterior body patterning (Lanzuolo and Orlando, 2012; Schwartz and Pirrotta, 2013). Loss-of-function mutations in PcG genes have been associated with ectopic expression of homeotic genes suggesting that their segmental appropriate repression is facilitated by PcG proteins.

PcGs function in multimeric protein complexes including the well studies Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2); dRing Associated Factors (dRAF) complex; and two additional complexes currently verified only in *Drosophila*: Polycomb Repressive deubiquitinase (PR-DUB) complex; and the lesser studied Pleiohomeotic-Repressive Complex (PHO-RC). The mode of action of these proteins has been linked to their association with the chromatin through post-translational modifications of histones, rendering the chromatin silent and ultimately resulting in transcriptional repression.

In *Drosophila*, Pho-RC consists minimally of Pho or Pho-like and dSfmgbt
Mutants of the Pho-RC components display HOX gene derepression and die at the larval stage. Pho, containing a zinc finger domain, is the only known PcG protein with DNA tethering ability. This transcription factor has been shown to direct dSfmbt to cis-regulatory sequences known as Polycomb response elements (PRE) for downstream repression of homeotic genes. dSMFBT consists of a zinc finger domain, four MBT domains and a SAM domain. The MBT domains of dSFMBT have been shown to interact with mono- and di-methylated Histone H4-Lysine 20 (H4K20me1 and H4K20me2) and monomethylated Histone H3 Lysine 9 (H3K9me1) in vitro (Grimm et al., 2009; Klymenko et al., 2006; Wu et al., 2007b, 1; Zhang et al., 2013). Some of the mammalian homologues of dSfmbt include SFMBT1, SFMBT2 and L3MBTL (Bonasio et al., 2010).

Recent efforts by several groups have paved the road to finding the mammalian counterpart of the Drosophila Pho-RC (Kuzmin et al., 2008; Lee et al., 2013; Zhang et al., 2013). However much of this work is carried out with over-expression constructs in exogenous settings. The first study aiming to identify the mammalian Pho-RC showed coimmunoprecipitation of mouse SFMBT2 and YY1 in HEK293 cells when over-expressed (Kuzmin et al., 2008). The interaction of over-expressed human SFMBT2 and YY1 has also been confirmed in HEK293 cells (Lee et al., 2013). Recently, tandem affinity purification of Flag-HA tagged SFMBT2 followed by mass spectrometry identified a number of interacting partners including the transcription factor CASZ1, and PcG proteins such as L3MBTL3 and RING2/RING1B/RNF2 but failed to coprecipitate YY1 (Zhang et al., 2013). Furthermore, affinity tagged human SFMBT2 was shown to coimmunoprecipitate with a number of proteins including the transcription factors CASZ1 and TFDPI as well as the PcG protein L3MBTL3 but not YY1 (Greenblatt,
SFMBT2 is robustly expressed in the extraembryonic tissues including the placenta (Kuzmin et al., 2008; Miri et al., 2013). The study of the protein networks involving SFMBT2 would therefore be most informative in the extraembryonic tissues. In this chapter, I show that SFMBT2 coimmunoprecipitates with a number of transcription factors including YY1, TFDP1 and CASZ1 in placenta progenitors, TS cells, highlighting the importance of context with respect to protein-protein interactions. Furthermore, I show that SFMBT2’s cellular localization is cell-type specific and shows overlap with mitotic chromosomes. The localization of SFMBT2 and its interacting partners YY1 and TFDP1 with the condensed mitotic chromosomes and the previously shown binding of SFMBT2 with mitotically linked histone PTMs, suggest that SFMBT2 may be involved in mitotic bookmarking.

5.3 RESULTS

5.3.1 SFMBT2 shows a distinct nuclear localization pattern in various trophoblast cell-types and associates with mitotic chromosomes.

Analysis of SFMBT2 localization in TS cells at various stages of differentiation was performed to shed light on the potential function of this protein in the trophoblast lineage (Figure 5.1). Immunocytochemistry experiments indicate that the majority of the undifferentiated TS cells display strong punctate staining in the nucleus (Figure 5.1 A). The punctate pattern of staining of SFMBT2 appears to co-localize with the heterochromatic regions showing strong DAPI staining (Figure 5.1 A, SFMBT2/DAPI).
Figure 5.1. SFMBT2 shows cell-type specific localization pattern and association with mitotic chromosomes.

Immunocytochemistry of (A) undifferentiated TS cells (B) differentiated TS cells on the third day of FGF4 withdrawal and (C) TGCs with anti-SFMBT2 antibody and DAPI counterstaining as indicated.
Undifferentiated TS cells in the mitotic stage of the cell cycle, evident from DAPI staining, show SFMBT2 co-localization with the mitotic chromosomes. By the third day of differentiation in the absence of FGF4 and feeder conditioned medium, mitotic TS cells are still detected and display co-localization of SFMBT2 with mitotic chromosomes (Figure 5.1 B); however, in contrast with undifferentiated TS cells, the majority of interphase cells at the third day of differentiation show diffuse nuclear staining with minor co-localization with the heterochromatin. Some TGCs display little or no SFMBT2 staining in the nucleus while showing filamentous cytoplasmic staining and others have a punctate staining pattern similar to undifferentiated TS cells (Figure 5.1 C).

5.3.2 SFMBT2 coimmunoprecipitates with YY1 in both exogenous and endogenous environments.

SFMBT2 has been previously shown to interact with YY1 in HEK293 cells transiently transfected with pcDNA-Myc-Sfmbt2 (short and long) and pcDNA-YY1 constructs (Kuzmin et al., 2008). These experiments were performed with anti-MYC and anti-human-YY1 antibodies. Using anti-SFMBT2 and anti-mouse-YY1 antibodies, interaction of SFMBT2 and YY1 in transiently transfected HEK293 cells was retested confirming that both isoforms of SFMBT2 interact with YY1 (Figure 5.2). However the short isoform of SFMBT2 does not appear to interact with YY1 as strongly as the long isoform. Interaction between these two proteins was further investigated in the endogenous setting using TS cells. Coimmunoprecipitation experiments in TS cells were performed reciprocally, confirming the interaction of the long isoform of SFMBT2 with YY1 (Figure 5.3 A,B). However, following immunoprecipitation a significant amount
**Figure 5.2. Over expressed SFMBT2 and YY1 interact in HEK293 cells.**

Image shows western blot analysis of whole extracts harvested from HEK293 cells transfected with the long or the short isoform of SFMBT2 and YY1 indicated as “input” and of coimmunoprecipitated (IP) samples with anti-SFMBT2 or anti-YY1 antibodies as indicated. The blot was probed with the anti-SFMBT2 antibody.
Figure 5.3. SFMBT2 interacts with the transcription factors YY1 and TFDP1 in TS cells.

(A-C) Western blot analysis of whole lysate (input) and coimmunoprecipitated samples with IgG, anti-SFMBT2, anti-YY1 and anti-TFDP1 antibodies as indicated. Blots were probed with appropriate antibodies as shown. “Ub” indicates the unbound fraction of the cell extracts and “B” indicates the bound fraction after coimmunoprecipitation.
of protein was found in the unbound fraction suggestive of inefficacies in the immunoprecipitation procedure and/or that both SFMBT2 and YY1 have other interaction partners independently of one another. The short isoform of SFMBT2 was not detected or immunoprecipitated in TS cells. This is supported by qRT-PCR experiments performed on the two isoforms as well as previously published publically available microarray data reporting significantly lower transcript levels of the short isoform when compared to the long (Miri et al., 2013).

Immunocytochemistry using the anti-mouse YY1 antibody indicates a diffuse pattern of staining in TS cells of various differentiation stages including undifferentiated, three-day differentiated and TGCs (Figure 5.4). Localization of YY1 in interphase nuclei does not appear to associate with the heterochromatic regions however YY1 colocalizes with the mitotic chromosomes in undifferentiated as well as three-day differentiated TS cells (Figure 5.4 A,B). Together with the coimmunoprecipitation data, these experiments support the existence of a mammalian Pho-RC.

5.3.3 Coimmunoprecipitation experiments reveal a novel interaction between SFMBT2 and TFDP1

Coimmunoprecipitation experiments performed in HEK293 cells transfected with tagged human SFMBT2 followed by mass spectrometry analysis have previously shown that SFMBT2 associates with a list of proteins including L3MBTL3 and TFDP1 (Greenblatt, submitted). I assessed this list for proteins with extraembryonic specific phenotypes and identified TFDP1 as the most likely candidate to interact with SFMBT2 in vivo. TFDP1 is a transcription factor exhibiting mid-gestation lethality due to aberrant extraembryonic tissue development (Kohn et al., 2003) similar to those observed in the SFMBT2 null
Figure 5.4. YY1 staining shows overlap with mitotic chromosomes in TS cells.

Immunocytochemistry of (A) undifferentiated TS cells, (B) differentiated TS cells on the third day of FGF4 withdrawal, and (C) a TGC (arrow) immuno-labelled using anti-YY1 antibody. DAPI was used as a nuclear counterstain.
mutants. Coimmunoprecipitation experiments in TS cells revealed that SFMBT2 immunoprecipitates with TFDP1 and vice versa, indicating the two proteins likely form a complex in vivo (Figure 5.3 A,C). A substantial amount of the non-immunoprecipitated protein is found in the unbound fraction indicating the inefficiency of coimmunoprecipitation and/or presence of other binding partners for both proteins. Immunocytochemistry experiments in TS cells using anti-TFDP1 antibody revealed little or no staining during the interphase stage of the cell cycle and strong staining during the mitotic stage (Figure 5.5). The staining during mitosis appears to be co-localized with the condensed chromosomes, as was observed for SFMBT2 and YY1. TGCs displayed no detectable TFDP1 staining by confocal microscopy (Figure 5.5 C, arrow).

5.3.4 Mass spectrometry confirms the interaction of SFMBT2 with the transcription factor CASZ1 in TS cells

Polycomb group proteins are generally found in complexes interacting with a number of other proteins. To elucidate the role of SFMBT2 with respect to other known complexes in the trophoblast derivatives, the anti-SFMBT2 antibody was used to immunologically isolate putative SFMBT2 interacting partners for mass spectrometry analysis using undifferentiated TS cells. SDS-PAGE was performed to separate the immunoprecipitated proteins by size followed by silver-staining to identify prominent bands. Individual bands and sections containing multiple bands were subsequently excised and processed for mass spectrometry analysis (In-gel). In addition, whole coimmunoprecipitated samples were eluted from Protein G beads and directly processed for mass spectrometry (In-solution). HEK293 cells transiently transfected with pLV-Myc-Sfmbt2 (long) were also subjected to immunoprecipitation followed by in-gel processing and mass spectrometry.
Figure 5.5. TFDP1 shows co-localization with mitotic chromosomes in TS cells.

Immunocytochemistry of (A) undifferentiated TS cells, (B) differentiated TS cells on the third day of FGF4 withdrawal, and (C) a TGC (arrow) immunostained with anti-TFDP1 antibody. DAPI was used as a nuclear counterstain.
For each coimmunoprecipitation experiment, extracts were immunoprecipitated with IgG to identify false positives due to non-specific interactions. Table 5.1 shows the results of the mass spectrometry after peptide validation and protein identification by Scaffold. The identified proteins listed in Table 5.1 were accepted as “true” predictions if they met the algorithm’s threshold (three or more unique peptides each exhibiting greater than 95% probability per identified protein also with greater than 95% probability); did not appear in the contaminants database (CDB); and did not coimmunoprecipitate with IgG. Only SFMBT2 and the zinc finger protein Castor homologue 1 (CASZ1) withstood these criteria. See Appendix II, Appendix III, Appendix IV, and Appendix V for peptide sequences and spectrum details.

5.4 DISCUSSION

The observed variation in the staining pattern and hence localization of SFMBT2 in various trophoblast cell types is suggestive of distinct biological roles at various stages of differentiation and lineage commitment. Identification and confirmation of interaction networks involving SFMBT2 in the trophoblast lineage, and more specifically in TS cells, would shed light on the function of this protein. Consequently, cell extracts harvested from TS cells were used to confirm interactions between SFMBT2 and previously identified transcription factors. SFMBT2 coimmunoprecipitates with the transcription factor YY1 under endogenous conditions suggestive of the existence of a mammalian Pho-RC. Furthermore, coimmunoprecipitation experiments also confirm that in TS cells SFMBT2 interacts with two other transcription factors, CASZ1 and TFDP1.
Table 5.1. Mass Spectrometry results from coimmunoprecipitated proteins using the anti-SFMBT2 antibody.

List of proteins generated by Scaffold after mass spectrometry of coimmunoprecipitated proteins using anti-SFMBT2 antibody and rabbit IgG in wild type TS cells and 293T-SF (HEK293T cells transiently transfected with pLV-myc-Sfmbt2 (long)). Proteins found in the contaminants database (CDB) and immunoprecipitated with IgG were categorized as contaminants. Percentages indicate the probability that the identified protein is a true prediction. Numbers in parenthesis show the number of peptides identified for each predicted protein. MW, Molecular Weight; IP, immunoprecipitates; In-solution, whole immunoprecipitated proteins eluded from Protein G agarose beads and processed directly; In-gel, SDS-PAGE of immunoprecipitated proteins followed by excision of band or sections; >65, 35-65, 30, <30, and 100-120 represent the regions of the polyacrylamide gel that was excised in kilo Daltons and prepared individually for mass spectrometry analysis; *, indicates peptides predicted with greater than 80% probability.

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>MW (kDa)</th>
<th>TS cells-IP</th>
<th>293T-SF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>In-solution</td>
<td>293T-SF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In-gel</td>
<td>SFMBT2-IP</td>
</tr>
<tr>
<td>Trypsin</td>
<td>24</td>
<td>100% 100% 100% 100% 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>KRT10</td>
<td>60</td>
<td>100% 100% 100% 100% 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>KRT14</td>
<td>52</td>
<td>6% 99% 100% 100% 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>KRT19</td>
<td>45</td>
<td>99% 0 0 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>KRT1</td>
<td>66</td>
<td>100% 100% 100% 100% 0 0 0 0 0 0</td>
<td>CDB, IgG-IP</td>
</tr>
<tr>
<td>KRT2</td>
<td>66</td>
<td>100% 100% 100% 100% 0 0 0 0 0 0</td>
<td>CDB, IgG-IP</td>
</tr>
<tr>
<td>KRT5</td>
<td>62</td>
<td>99% 0 100% 100% 0 0 0 0 0 0</td>
<td>CDB, IgG-IP</td>
</tr>
<tr>
<td>KRT7</td>
<td>63</td>
<td>99% 97% 99% 86% 0 0 0 0 0 0</td>
<td>CDB, IgG-IP</td>
</tr>
<tr>
<td>KRT9</td>
<td>62</td>
<td>100% 100% 100% 100% 0 0 0 0 0 0</td>
<td>CDB, IgG-IP</td>
</tr>
<tr>
<td>KRT9</td>
<td>53</td>
<td>89% 100% 0 0 0 0 0 0 0 0</td>
<td>IgG-IP</td>
</tr>
<tr>
<td>TAP15</td>
<td>59</td>
<td>100% 5% 0 0 0 0 0 0 0 0</td>
<td>IgG-IP</td>
</tr>
<tr>
<td>BSA</td>
<td>69</td>
<td>0 100% 0 0 0 0 0 0 0 0</td>
<td>CDB, IgG-IP</td>
</tr>
<tr>
<td>Kappa-casein</td>
<td>19</td>
<td>93% 100% 97% 100% 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>SFMBT2</td>
<td>110</td>
<td>98% (2)* 100% (4) 0 0 0 0 0 0 0 0 99% (2)*</td>
<td>NA 0</td>
</tr>
<tr>
<td>HNRNPA</td>
<td>17</td>
<td>100% 0 100% 0 0 0 0 0 0 0 0</td>
<td>IgG-IP</td>
</tr>
<tr>
<td>ACT1</td>
<td>41</td>
<td>92% 100% 92% 100% 0 0 0 0 0 0 0 0</td>
<td>CD8</td>
</tr>
<tr>
<td>ACT1B1L2</td>
<td>41</td>
<td>0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>CD8</td>
</tr>
<tr>
<td>HSTH44A</td>
<td>11</td>
<td>100% 0 0 0 0 0 0 0 0 0 0 0</td>
<td>IgG-IP</td>
</tr>
<tr>
<td>CASZ1</td>
<td>126</td>
<td>98% (1) 100% (3) 0 0 0 0 0 0 0 0 0 0</td>
<td>NA</td>
</tr>
</tbody>
</table>
TFDP1 forms a heterodimer with the E2F family of transcription factors which exhibit numerous roles in DNA replication and cell cycle progression, as well as lineage specification and differentiation (Müller et al., 2001). Loss of TFDP1 in mice results in embryonic lethality by day 12.5 of gestation due to failure in proper development and maintenance of the extraembryonic tissues (Kohn et al., 2003). Reduction of all trophoblast derivatives including the EPC and the TGCs was reported as early as day 7.5 of development. Trophoblast specific aberrations were linked with reduction of precursor cells due to either proliferative defects and/or premature and inappropriate differentiation of the trophoblast progenitors. Furthermore, reduced BrdU incorporation in TGCs was suggested to be due to endoreduplication defects resulting in reduced DNA copy-number assessed by the size of the nuclei. It was therefore suggested that TFDP1 in conjunction with its interacting partner E2F may play a significant role in regulation of genes required for S phase entry. These observations are consistent with the SFMBT2 knockout phenotype discussed in Chapter 4, strongly supporting the hypothesis that SFMBT2 and TFDP form a functional complex \textit{in vivo}.

Interaction of SFMBT2 with various transcription factors may indicate the heterogeneity of the repressive complexes in which SFMBT2 is the recurring member. This type of heterogeneity has been shown in other repressive complexes whereby binding partners of PcG proteins are context dependent and may vary based on the particular gene being repressed as well as the chromatin state, cell type and cell cycle (Schwartz and Pirrotta, 2013). The PcG protein RING2/RING1B/RNF2 has been shown to appear in a number of PRCs including the canonical PRC1 containing RING2, BMI1, CBX4 and PHC1 as well as other PRC1 variants such as RING2-RYBP core complex, RING2-KM2B complex, RING2-L3MBTL2 complex and RING2/FBRS complex.
In both mammals and flies, there are a number of MBT-domain containing proteins, each with a distinct functional significance. MBT domains have been shown to selectively associate with mono- and di-methylated Histone H4 lysine 20 (H4K20me1/me2). These PTMs have been previously linked with cell-cycle progression, genomic integrity and DNA repair, DNA replication, chromosome compaction and transcriptional repression (Jorgensen et al., 2013; Rice et al., 2002; Trojer and Reinberg, 2008; Trojer et al., 2007). PR-SET7/SET8/KMT5A is the only known histone methyltransferase involved in catalysis of monomethylation of histone H4K20 (Fang et al., 2002; Nishioka et al., 2002). Cell cycle dependent accumulation of this enzyme mirrors that of monomethylation of H4K20. PR-SET is proteolytically degraded during G1 and S phase (Abbas et al., 2010) resulting in low levels of H4K20me1 followed by phosphorylation mediated stabilization of the enzyme during the G2 phase of the cell cycle (Wu et al., 2010), subsequently resulting in global accumulation of H4K20me1 reaching its peak during mitosis (Pesavento et al., 2008). The mitotic relevance of H4K20me1 is of significance as SFMBT2 and its interacting partners also appear to associate with mitotic chromosomes. Furthermore, through in vitro studies mammalian SFMBT2 and its Drosophila orthologue dSFMBT have been shown to interact preferentially with H4K20me1 and H4K20me2 (Klymenko et al., 2006; Zhang et al., 2013).

In addition to binding with H4K20me1 and H4K20me2 in vitro, Zhang and colleagues (Zhang et al., 2013) reported the association of SFMBT2 with H3K9me1, also a repressive histone mark (Metzger et al., 2005). The heterogeneity of PTMs associated with SFMBT2 makes characterization of this protein more challenging. It has been
hypothesized however that the specificity of the MBT domain containing proteins may be facilitated by their interacting partners with consensus sequence binding motifs such as zinc finger domains found in many transcription factors (Trojer and Reinberg, 2008).

In preparation for cell division, dramatic changes incurred by the cell and the chromosomes facilitate entry and progression into mitosis. These changes include disintegration of the nuclear envelope, dissociation of the majority of transcriptional regulators and global transcriptional arrest, acquisition of mitotically associated histone PTMs, and condensation of the chromosomes (Gottesfeld and Forbes, 1997). Upon completion of mitosis and re-entry into the growth phase of the cell cycle (G1), appropriate mounting of the transcriptional regulatory machinery has been linked with retention of lineage commitment as well as providing a window of opportunity for lineage specification. The orchestration of this phenomenon has been linked with mitotic bookmarking. Through this process a number of transcription factors such as RUNX2 and chromatin proteins such as BRD4 remain bound to the mitotic chromosomes to allow rapid transcription of lineage appropriate genes upon re-entry into the growth phase (Kadauke et al., 2012; Yang et al., 2008; Young et al., 2007).

Maintenance of lineage commitment is thought to be facilitated not only by transcription of lineage specific genes but also by maintenance of repression of lineage inappropriate genes. The transcription factor GATA1 has been shown to associate with mitotic chromosomes and occupy key hematopoietic regulatory genes (Kadauke et al., 2012). Mitotic specific depletion of GATA1 results in delayed transcriptional reactivation of a number of these genes upon completion of mitosis, suggesting that GATA1 is a mitotic bookmarker. Furthermore, mitotic-specific destruction of GATA1 was shown to correlate with de-repression of a number of GATA1-bound repressed genes, such as
This suggests that in addition to a role in maintenance of the active state of genes, mitotic bookmarking is likely involved in the maintenance of the silent state.

The staining pattern of SFMBT2 and its interacting partners as well as the interaction of SFMBT2 with transcription factors such as TFDP1 with roles in cell cycle progression suggests that SFMBT2 may function as a mitotic bookmark to sustain repression of lineage inappropriate and/or cell cycle inappropriate genes. Although the characterization of SFMBT2 complexes as repressors without any further experimentation may be premature, evidence in DU145 prostate cancer cells is supportive of this hypothesis whereby siRNA knockdown of Sfmbt2 resulted in de-repression of Hoxb13 (Lee et al., 2013). Similarly, reporter assays carried out with human SFMBT1, a homologue of SFMBT2, illustrates that this PcG protein functions as a repressor and furthermore requires all four MBT domains for its function (Wu et al., 2007b).

Taken together, I hypothesize that SFMBT2 through association with various transcription factors and histone H4K20me1/me2 and Histone H3K9me1 retains silencing of lineage inappropriate genes as cells traverse through mitosis. Consequently, the association of SFMBT2 with mitotic chromosomes is lineage specific and dependent minimally on the presence of both the interacting transcription factors as well as the appropriate PTMs of the histones associated with the repressed genes whose silencing is maintained through mitosis and re-entry into the G1 phase of the cell cycle. By this model absence of any of these cues must result in lack of association of SFMBT2 with the mitotic chromosomes. In concurrence with this hypothesis, exogenous overexpression of SFMBT2 in HEK293 cells does not result in association with condensed mitotic chromosomes (Figure 2.2 C’,F’). Nevertheless, it cannot be ruled out at this
point that the tagged SFMBT2 fusion protein used for these experiments may inhibit the function of this protein and prevent chromatin binding. Furthermore, putative PTMs of SFMBT2 may also play a role in its function and localization, making further study of this protein in its endogenous environment necessary.
Chapter 6: Discussion
Severe abnormalities in the development of the extraembryonic tissues resulting in early mid-gestational lethality in parthenogenetic embryos is thought to be the consequence of lack of expression from early acting PEGs with essential roles in the development of the trophoblast lineage. In order to identify such genes, I surveyed known PEGs for a number of criteria and identified Sfmbt2 as the best candidate. The survey was followed by characterization of this gene in the extraembryonic tissues in vivo and in vitro suggesting that SFMBT2 is the earliest acting PEG with essential functions in the development of the trophoblast lineage and likely a major contributor to the early phenotypes of the parthenogenetic embryos. Furthermore, I showed that SFMBT2 forms a complex with a number of transcription factors including TFDP1 which has also been shown to have a critical role in the development of extraembryonic tissues. In this final chapter, I will review my overall findings in light of relevant literature and close by discussing some unanswered questions and prospective directions for this research.

6.1 CHROMOSOMAL LOCATION AND LOI OF SFMBT2

Reciprocal translocation studies show that mice with maternal duplication of proximal chromosome 2 and paternal duplication of proximal chromosome 11 (MatDp(prox2)/PatDp(prox11)) exhibit embryonic lethality (Cattanach et al., 2004). Embryos with paternal disomy of chromosome 11 show large placentae suggesting that the observed phenotype of the MatDp(prox2)/PatDp(prox11) fetuses is likely associated with the presence of an early acting PEG in the proximal region of chromosome 2 with essential functions in the trophoblast lineage. Nevertheless other translocations containing this region, including the T(2;8)2Wa translocation consisting of the maternal duplication of proximal chromosome 2, can produce viable pups, although the placentae
are reduced in size. It was therefore hypothesized that the observed phenotype in MatDp(prox2)/PatDp(prox11) embryos is due to interactions between imprinted genes on proximal 2 and proximal 11 as well as potential strain specific effects. However, recent allelic analysis of Sfmbt2, the only known PEG within this chromosomal segment (Wang et al., 2011), in placenta harvested from embryos with two maternal copies of proximal chromosome 2 from the T(2;8)2Wa translocation showed that this gene exhibits LOI (Miri et al., 2013). This suggests that the mild phenotypes associated with this translocation may be due, at least in part, to LOI of Sfmbt2. Expression profile of Sfmbt2 in other translocations containing the maternal duplication of proximal chromosome 2 would be informative.

In 1996, Kono and colleagues reported the generation of reconstructed parthenogenetic embryos using haploid pronuclei from non-growing (ng) oocytes obtained from 1-day old females and fully-grown (fg) oocytes that survived up to embryonic day 13 (Kono et al., 1996). In 1998, the same group showed that the reconstructed parthenogenetic embryos exhibited LOI of a number of PEGs including: Peg1, Peg3 and Snrpn but not Igf2, suggesting that the previously reported survival to embryonic day 13 may be due to LOI (Obata et al., 1998). Subsequently, restoration of expression from Dlk1 as well as Igf2 in these embryos were shown to result in survival to term (Kawahara et al., 2007), suggesting that of the three imprinted domains regulated by paternal methylation of the germ-line DMRs, only H19-Igf2 and Dlk1-Gtl2 imprinted domains have gestation roles. However, these roles appear to be limited to late gestation and therefore PEGs from these loci such as Dlk1 and Igf2 are likely not responsible for the early parthenogenetic phenotypes. These observation are in agreement with Robertsonian and reciprocal translocations (Georgiades et al., 2000). In addition, the
apparent LOI of the reported PEGs suggest that the genome harvested from ng oocytes is not yet programmed; thus the events taking effect during the maturation of the oocyte set the stage for differential expression of a number of imprinted genes with essential early-mid-gestation functions. Consequently, the PEGs responsible for the phenotypes associated with the conventional parthenogenetic embryos must also acquire their imprinting signature during the oocyte growth phase. Hence, re-investigation of these reconstructed embryos for LOI of PEGs such as Peg10 and Sfmbt2 is of interest.

Potential correlation of LOI of Sfmbt2 and early rescue of the parthenogenetic phenotype is further strengthened by TS cell derivation experiments from parthenogenetic blastocysts (Miri et al., 2013). The phenotype of the parthenogenetic embryos is indicative of a failure to establish or maintain the trophoblast progenitor cells suggesting that derivation of PTS cells from parthenogenetic blastocyst outgrowths is unlikely. Nevertheless, data show that PTS cells can be derived but at a much lower frequency when compared with fertilized counterparts. The derivation of the PTS cells is correlated with LOI of a number PEGs, including Sfmbt2, suggesting that establishment and/or maintenance of this lineage in parthenogenetic embryos may be dependent on LOI of an early acting PEG with an essential role in this tissue.

LOI has been associated with culture conditions in a number of reports (Doherty et al., 2000; Mann et al., 2004; Sasaki et al., 1995). Blastocysts cultured for prolonged periods before transplantation into the uterus exhibit LOI especially in the trophoblast tissues; in this case the locus tested was H19-Igf2, where H19, a maternally expressed gene, displayed biallelic expression (Sasaki et al., 1995). Furthermore, in situ hybridization on androgenetic embryos showed expression of H19 only within the
trophoblast lineage, indicating LOI of this gene in a tissue specific manner suggesting that the trophoblast lineage is more labile than embryonic tissues which maintained the imprints. Mann and colleagues reported LOI of a number imprinted genes including, Peg3, H19, Snrpn, Ascl2/Mash2 as well the X-inactivation transcript Xist selectively in placentae from embryos cultured in Whitten’s medium until the blastocyst stage (Mann et al., 2004). The observed LOI of a group of imprinted genes in culture conditions is suggestive of a shared mechanism of imprint maintenance. Investigation of the epigenetic modifications associated with these genes after LOI would shed light on this mechanism. Furthermore, such studies would be beneficial for advancements in assisted reproductive technologies such as in vitro fertilization.

Paternal transmission of the gene trap allele of Sfmbt2 results in post-fertilization LOI of the normally silent maternal allele in trans. The gene trap construct contains a strong polyA signal, and is predicted to cause transcript termination before the polymerase reaches a block of miRNAs (in intron 10) previously correlated with evolutionary acquisition of imprinting of Sfmbt2 (Wang et al., 2011). The presence of miRNAs within imprinted domains is not limited to Sfmbt2; the Dlk1 locus, the Snrpn locus also contain large blocks of miRNAs or snoRNAs, or both. What distinguishes Sfmbt2 is the absence of a differentially methylated region. LOI in embryos with a paternal gene trap allele of Sfmbt2 suggests that the miRNAs, or related snoRNAs processed from the primary transcript, may be involved in maintaining the silent state of the maternal allele. Testing this hypothesis by providing mutant embryos with miRNAs from the cluster in trans, or by preventing their genesis in wild type embryos, may shed some light on the role of RNA interference in transcriptional silencing at this gene.
6.2 **TISSUE SPECIFIC EXPRESSION AND LOCALIZATION**

**PATTERN OF SFMBT2**

*SFmbt2* has been previously shown to exhibit robust expression from the paternal allele in early embryos and extraembryonic tissues and minimal biallelic expression in the somatic lineage such as the brain (Kuzmin et al., 2008). In concurrence with these findings, immunostaining experiments in TS cells *in vitro* and conceptuses *in vivo* show strong nuclear localization of SFMBT2 in the various cell types of the extraembryonic tissues of the trophoblast origin. Furthermore, the observed staining pattern of SFMBT2 in TS cells is recapitulated in immunostained histological sections collected at various stages of placental development. The strong nuclear staining observed in trophoblast cells of the chorionic plate adjacent to the mesodermal layer of allantoic origin at embryonic day 8.5 is reminiscent of the strong punctate staining observed in undifferentiated TS cells. This region of the chorionic plate is thought to harbor trophoblast progenitor cells, maintained through NODAL/TGF-β signalling originating from the mesodermal layer. The chorionic plates of e8.5 placentae have been previously used to derive TS cells *in vitro*. The reduced diffuse nuclear staining observed in TS cells 3 days after FGF4 withdrawal mirrors the distal chorionic cells positioned further away from the mesodermal layer and the vast majority of the cells of the EPC at e8.5. The heterogeneity in the staining pattern of TGCs *in vitro* also resembles the variability observed in TGCs at various stages of placental development *in vivo*. Coimmunolabeling of histological sections at different stages of embryonic development and TS cells at various stages of differentiation with trophoblast cell-type markers along with SFMBT2 would shed light on further characterization of this protein as well as provide further insights into the biology of placental development.
Subcellular localization of proteins has previously been shown to exert an additional layer of regulation and is often mediated by post-translational modifications (PTM), such as phosphorylation, ubiquitination, and sumoylation (Niessen et al., 2009). The establishment of the trophectoderm fate at the blastomere stage involves dephosphorylation of YAP resulting in its translocation to the nucleus and subsequent activation of TEAD4 which in turn induces the transcriptional activation of Cdx2 (Marikawa and Alarcon, 2009; Nishioka et al., 2009; Senner and Hemberger, 2010). A number of PcG proteins have also been shown to exhibit context specific PTMs (Niessen et al., 2009). The observed variability in the intracellular localization of SFMBT2 may be, at least in part, due to PTMs and/or presence or absence of cofactors. Treatment of nuclear extracts from TS cells with alkaline phosphatase following by western blot analysis for a mobility shift of SFMBT2 would be suggestive of phosphorylation of this protein similar to that of M33, a mammalian homologue of the Drosophila Ph which exhibits nuclear translocation upon phosphorylation (Noguchi et al., 2002). Sumoylation is yet another candidate for PTM of SFMBT2; western blot analysis of SFMBT2 immunoprecipitate from TS cells with anti-SUMO antibody would provide answers on whether SFMBT2 is sumoylated in vivo.

6.3 **LOSS OF FUNCTION PHENOTYPE OF SMFBT2 SUGGESTS INVOLVEMENT IN MAINTENANCE OF THE TROPHOBLAST LINEAGE**

Homozygous null mutants of Sfmbt2 fail to maintain the trophoblast lineage and die at early mid-gestation. Histological analysis of conceptuses show that despite the presence of trophoblast cell types, indicated by positive cytokeratin staining at e8.5, disappearance
of the trophoblast tissue is observed by e9.75. The presence of CDX2, a marker of trophoblast progenitor cells, in the trophoblast tissue of the null conceptuses suggests that SFMBT2 is not required for the establishment of the trophoblast fate. At this juncture however, it is unclear whether the observed trophoblast tissue at day 8.75 of gestation, positively stained for Cytokeratin and CDX2, also exhibits the appropriate spatial and temporal expression of other trophoblast specific markers. Immunostaining and/or in situ hybridization experiments for various cell type specific markers would shed light on these unanswered questions. Furthermore, transcriptome analysis of e7.5 extraembryonic ectoderm and EPC of SFMBT2 null conceptuses would help identify genes with aberrant expression and consequently suggest putative targets of SFMBT2 in early trophoblast development. Confirmation of direct regulation of the candidate genes by SFMBT2 could be further elucidated by ChIP experiments at the single gene level or at the global level.

The nuclear hormone receptor Estrogen-receptor-related receptor-β (ESRR-β) has been shown to be expressed in the progenitor cells of the chorionic plate in early mid-gestation (Pettersson et al., 1996). Homozygous mutation of this gene results in embryonic lethality by day 10.5 of development due to severe abnormalities of early placental development, a phenotype that was rescued by tetraploid aggregation experiments (Luo et al., 1997). The reported phenotype associated with loss of ERR-β exhibits striking similarity with that of Sfmbt2 null mutants. Intriguingly, ChIP-seq analysis of ERR-β in embryonic stem cells was shown to bind the transcriptional start site of Sfmbt2 (Chen et al., 2008) (Figure 6.1). Diethylstilbestrol (DES), a synthetic estrogen, has been shown to inhibit the transcriptional activity of ERR-β in TS cells resulting in differentiation into TGCs, and expression of the TGC marker Pl-1. It remains
Figure 6.1. ERR-β binding site at the transcriptional start site of Sfmbt2.

ChIP experiments (Chen et al., 2008) show that the binding of the transcription factor ERR-β overlaps with Sfmbt2’s transcriptional start site (arrow). Image source: www.ensembl.org.
to be tested whether treatment of TS cells with DES would diminish the expression of *Sfmbt2*. Furthermore, a survey of the TGC makers in the differentiated cells after the treatment in comparison with untreated TGCs would shed light on the biological nature of these giant cells. Considering the observed phenotype of null mutants of *Sfmbt2* exhibiting PTGC, one would hypothesize that the differentiation of TS cells in presence of DES is likely to result in formation of the primary and not the secondary TGCs.

Due to the early lethality of null mutants of SFMBT2 the function of this gene in embryogenesis is not yet understood. To elucidate the role of SFMBT2 in embryonic tissues, tetraploid rescue experiments should be undertaken. Nevertheless, absence of SFMBT2 staining in embryonic tissues at mid-gestation, in conjunction with previously reported RT-PCR results on the somatic tissues showing much lower expression when compared to trophoblast tissues (Kuzmin et al., 2008), suggest that SFMBT2 is likely not required for embryonic development. In addition, experiments aiming to rescue the parthenogenetic phenotype with the introduction of exogenous SFMBT2 would be informative. I have attempted to rescue the parthenogenetic phenotype by lentiviral infection of the early blastocysts with inserts carrying both the long and short isoforms of *Sfmbt2*. However, due to the length of these inserts, lentiviral infection was not successful.

### 6.4 BIOCHEMICAL NATURE AND BINDING PARTNERS OF SFMBT2

SFMBT2 contains four MBT domains and a SAM domain. MBT domains are found in a number of *Drosophila* PcG proteins such as the dScm, dSfmbt, dL(3)mbtl and their respective mammalian orthologs. The SAM domain is found in a number of fly proteins
including PcG proteins such as dScm, dL3mbtl, dSfmbt and majority of their mammalian homologues and a number of ETS family of transcription factors (Bonasio et al., 2010; Kim and Bowie, 2003). SAM domains have been reported to induce homo- and hetero-
dimerization of proteins (Kim and Bowie, 2003; Kim et al., 2005; Peterson et al., 1997) whereas MBT domains have been shown to bind preferentially to mono- and di-
methylated lysine residues of histone tails (Li et al., 2007). Although there are currently only three MBT domain containing proteins in Drosophila, numerous mammalian orthologs with divergent roles have been identified, making the interspecies functional characterization of these proteins without direct biochemical studies difficult. In a review article by Bonasio and colleagues, CLASTALW analysis of various MBT domain containing proteins shows that despite the presence of four MBT domains, similar to that of the dSfmbt, the mammalian SFMBT1 and SFMBT2 are closer on the evolutionary tree to dL(3)mbtl which has three MBT domains (Bonasio et al., 2010). As discussed below, results from this thesis and relevant literature places SFMBT2 somewhere in between the two Drosophila homologues.

In chapter 5, I showed that SFMBT2 forms a complex with a number of transcription factors, including YY1, TFDP1 and CASZ1 in TS cells where Sfmbt2 is highly expressed. Similarly, dSFMBT has been shown to interact with Pho, the Drosophila ortholog of mammalian YY1, forming a complex called Pho-RC (Klymenko et al., 2006). Drosophila dSfmbt has also been shown to interact with dScm (Grimm et al., 2009) and the sumoylated transcription factor Sp3 (Stielow et al., 2008), suggesting that the heterogeneity of the interacting partners of SFMBT2 is not limited to mammals. In concurrence with these data, genome-wide analysis of dSfmbt and Pho binding sites show overlap in only half of the identified loci (Oktaba et al., 2008). Furthermore, the
remaining dSfmbt binding sites lack the Pho binding consensus sequence, suggestive of the involvement of other transcription factors and/or a transcription factor-independent mechanism of recruitment. In addition, 45% of the Pho-RC binding sites were shown to overlap with PRC1 and PRC2. SFMBT2 has also been shown to interact with components of PRC2; however the same study did not identify YY1 as an interacting partner but showed strong interaction with CASZ1 (Zhang et al., 2013). Tissue specificity of the interaction may explain the observed variation in the different studies.

In vivo, SFMBT2 is predominantly expressed in extraembryonic tissues, while the experiments performed by Zhang and colleagues (2013) were performed in HEK293 cells.

Both dSfmbt and dl(3)mbtl have been shown to selectively bind to mono- and di-methylated histone H4 lysine 20 (H4K20me1/2). Similarly, SFMBT2 has been shown to bind predominately to H3K9me1, H4K20me1/2, and to some extent H4K20me3 (Zhang et al., 2013). In HeLa cells, H4K20me1 and H3K9me1 show overlapping staining patterns with both diffuse and punctate staining, are excluded from the nucleoli and show association with the inactive X-chromosome (Sims et al., 2006). H4K20me2 and H3K9me2 also show overlap in Hela cells; however, they were shown to have mostly a diffuse pattern of staining suggestive of their dispersed localization. Furthermore, the distribution of H4K20me2 has been shown to be mutually exclusive with that of H4K20me1 suggestive of distinct functional significance; however, both modifications were shown to be excluded from regions of active transcription. H4K20me3 and H3K9me3 were shown to also co-localize and exhibit a noticeably punctate pattern of staining, limited to the pericentric regions showing dense DAPI staining suggestive of the heterochromatic nature of the co-localization site.
The pattern of localization association with each methylation state of H4K20 and H3K9 is reminiscent of SFMBT2 staining in various stages of TS cell differentiation with undifferentiated TS cells showing predominately punctate staining and often exhibiting overlap with the dense DAPI stained puncta. It remains to be elucidated whether the observed similarities are due to co-localization of these moieties. Co-labelling experiments with various antibodies against the aforementioned modifications of histone tails and SFMBT2 followed by high resolution confocal imaging would shed light on these unanswered questions. In addition, co-labelling of TS cells at various stages of differentiation for SFMBT2 and PolIII would address whether SFMBT2 co-localizes with predominately transcriptionally inactive regions of the genome. Notably, despite the minor variations observed amongst the MBT domain containing proteins, many have been shown to facilitate transcriptional repression (Boccuni et al., 2003; Klymenko et al., 2006; Roseman et al., 2001; Wu et al., 2007b).

Similarities between SFMBT2 and L3MBTL extend to their pattern of localization during mitosis as both proteins appear to co-localize with the condensed mitotic chromosomes. Most factors associated with the chromosomes including various transcription factors and chromatin binding proteins such as PcG proteins have been previously shown to dissociate from their targets during the chromatin condensation phase of the dividing cells. Mitotic bookmarks are amongst those proteins that remain behind to mediate retention of cellular memory and facilitate a smooth transition into interphase. Whether SFMBT2 participates in mitotic bookmarking is an intriguing question that warrants further investigation. Mitotic cell sorting (Kadauke et al., 2012) followed by coimmunoprecipitation accompanied by mass spectrometry analysis would shed light on SFMBT2’s interacting partners during mitosis. Furthermore, the identity of
the target genes during mitosis unveiled by ChIP-seq experiments may further solidify SFMBT2’s role as a mitotic bookmarker.

6.5 CONCLUSION

The highly conserved PcG protein SFMBT2 plays a critical role in the development of the murine placenta. Colocalization of SFMBT2 and its associated proteins with the mitotic chromosomes in trophoblast progenitor cells, in conjunction with failure to maintain the trophoblast tissue observed in Sfmbt2 null mutants in vivo, and the observed defects in establishment and/or maintenance of TS cells in vitro after the knockdown of this gene in early embryos, suggest that SFMBT2 is required for proliferation and maintenance of the trophoblast lineage including the trophoblast progenitors. Elucidation of the epigenetic control exerted by this protein on the genome in trophoblast, and possibly other lineages revealed by tetraploid rescue, would open up a new line of investigation into epigenome regulation of development.
References


Kuckenberg, P., Buhl, S., Woynecki, T., Fürden, B. van, Tolkunova, E., Seiffe, F.,


**MouseBook Imprinting Catalog** MouseBook Imprinting Catalog.


**Strumpf, D., Mao, C.-A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F. and Rossant, J.** (2005). Cdx2 is required for correct cell fate specification and


differentially methylated domain results in loss of imprinted expression of H19

**Tolhuis, B., de Wit, E., Muijters, I., Teunissen, H., Talhout, W., van Steensel, B. and
van Lohuizen, M.** (2006). Genome-wide profiling of PRC1 and PRC2 Polycomb

malignant brain tumor (MBT) protein L3MBTL1 impacts chromatin structure.

**Trojer, P., Li, G., Sims, R. J., Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D.,

**Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y.,


**Uy, G. D., Downs, K. M. and Gardner, R. L.** (2002). Inhibition of trophoblast stem cell
potential in chorionic ectoderm coincides with occlusion of the ectoplacental


**Walker, E., Chang, W. Y., Hunkapiller, J., Cagney, G., Garcha, K., Torchia, J.,


Wilder, P. J., Kelly, D., Brigman, K., Peterson, C. L., Nowling, T., Gao, Q. S.,
FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their

Wilkinson, F., Pratt, H. and Atchison, M. L. (2010). PcG recruitment by the YY1

Williamson, C. M., Turner, M. D., Ball, S. T., Nottingham, W. T., Glenister, P., Fray,
(2006). Identification of an imprinting control region affecting the expression of

Wu, S., Trieval, R. C. and Rice, J. C. (2007a). Human SFMBT is a transcriptional
repressor protein that selectively binds the N-terminal tail of histone H3. FEBS

Wu, S., Trieval, R. C. and Rice, J. C. (2007b). Human SFMBT is a transcriptional
repressor protein that selectively binds the N-terminal tail of histone H3. FEBS

Wu, S., Wang, W., Kong, X., Congdon, L. M., Yokomori, K., Kirschner, M. W. and
Rice, J. C. (2010). Dynamic regulation of the PR-Set7 histone methyltransferase
is required for normal cell cycle progression. Genes Dev. 24, 2531–2542.

repressive complex 1 to CpG islands and regulates H2A ubiquitylation. Mol. Cell
49, 1134–1146.

Xu, X., Weinstein, M., Li, C., Naski, M., Cohen, R. I., Ornitz, D. M., Leder, P. and
reciprocal regulation loop between FGF8 and FGF10 is essential for limb

Yagi, R., Kohn, M. J., Karavanova, I., Kaneko, K. J., Vullhorst, D., DePamphilis, M.
L. and Buonanno, A. (2007). Transcription factor TEAD4 specifies the


Appendices

Appendix I. Primer and oligonucleotide sequences.

A. Primer sequences for construction of Sfmbt2 overexpression constructs.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-iRES-F</td>
<td>TTCGAATCTAGAACGCGTGATCCGCTAGCGTTAACCGACCCCATCATCCAAATTCGCC</td>
</tr>
<tr>
<td>EGFP-EcoRI-Hpal-R</td>
<td>ATCGAATTCTTTAACTTTACTTTGACATCTCGTCATGCGA</td>
</tr>
<tr>
<td>Xbal-Myc-sf-F</td>
<td>GACTCTAGAGACCCAAAGCTATGGAGCAAAAGCT</td>
</tr>
<tr>
<td>Myc-sf-L-Bmt1-R</td>
<td>TCGGCTAGCTGCAATGCAGATTAGCTGCCC</td>
</tr>
<tr>
<td>pLV-Xbal-Egfp-F</td>
<td>TTCGAATCTAGAATGTTGAGCAAGGCGAGGAGC</td>
</tr>
<tr>
<td>pLV-Egfp-A2-R</td>
<td>GTTTGTTGTCTGGCTTCGAATTCTGTGTCAGCTGTCATGCC</td>
</tr>
<tr>
<td>A2 Domain-R</td>
<td>CATGGGCCCAGGGTTTGGAGCTCGACGTCTCCGCAAACTTGA GAAGGTCAAAAATTCAAAGTCTGTTTCACCGGTGCCCACAATTTTCTGTGTGTCTGGCT</td>
</tr>
<tr>
<td>pLV-A2-sf-F</td>
<td>CGTCGAGTCAAACCTGGCCCATGGGAGCAAAAGCTCATCTCAG</td>
</tr>
<tr>
<td>pLV-sf-sh-R</td>
<td>ACGCCTGTGTAACCTATAATAGATGAATAAAGCA</td>
</tr>
<tr>
<td>A2 domain-R</td>
<td>CATGGGCCCAGGGTTTGGAGCTCGACGTCTCCGCAAACTTGA GAAGGTCAAAAATTCAAAGTCTGTTTCACCGGTGCCCACAATTTTCTGTGTGTCTGGCT</td>
</tr>
<tr>
<td>pLV-Xbal-Egfp-F</td>
<td>TTCGAATCTAGAATGTTGAGCAAGGCGAGGAGC</td>
</tr>
<tr>
<td>pLV-sf-sh-R</td>
<td>ACGCCTGTGTAACCTATAATAGATGAATAAAGCA</td>
</tr>
</tbody>
</table>

B. Oligonucleotide sequences used for the shRNA cloning.

<table>
<thead>
<tr>
<th>Oligonucleotide Name</th>
<th>Sense strand</th>
<th>Antisense strand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-Sfmbt2-shRNA1</td>
<td>acctGGAGATCGACATTCCGGAGTAacagagaTTACTCCGATCTGCTCATTtttc</td>
<td>tcgaagaaaaatGGAGATCGACATTCCGGAGTAatctctgtTACTCGGAAATGCTGACTCC</td>
</tr>
<tr>
<td>Anti-Sfmbt2-shRNA2</td>
<td>acctGGCAGTCTGAGTACTTGCTTcaagagaAAGGTAAGTACTGCAAGGAtttttc</td>
<td>tcgagaaaaatTCCAGTCTGAGTACTTGCTTtctcttgAAAGAAGTACTGCAAGG</td>
</tr>
<tr>
<td>Control-shRNA</td>
<td>acctGGCAGTCTGAGTACTTGCTTcaagagaAAGGTAAGTACTGCAAGGAtttttc</td>
<td>tcgagaaaaatTCCAGTCTGAGTACTTGCTTtctcttgAAAGAAGTACTGCAAGG</td>
</tr>
</tbody>
</table>
C. Primers used for RT-PCR on paternally expressed genes and cell-type makers.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cdx2</td>
<td>TCAAGAAGAGCAAGCAGCAGG</td>
<td>GCAAGGAGGTCACAGGACTC</td>
</tr>
<tr>
<td>Dlk</td>
<td>TCAGGCAACTTCTGTGAGATGCG</td>
<td>CGCAGTTGGTAGAAGACGAGG</td>
</tr>
<tr>
<td>Fgfr2</td>
<td>CAAATCTCCCAACCAGAAGG</td>
<td>GTCCCCCATAAGCACTGTCG</td>
</tr>
<tr>
<td>Igf2-F2/R2</td>
<td>TCCTCCAGCAGGAGGACTG</td>
<td>TTGGGTGTCTTAGGACCAATC</td>
</tr>
<tr>
<td>Ins2</td>
<td>AACATGGCCTTGAGATG</td>
<td>TTGGGTGTGCTTAGGACCAATC</td>
</tr>
<tr>
<td>Mash2</td>
<td>TGAATGCAAGCTTTGAGAGG</td>
<td>CCAAGTTTACCAGTCTTACG</td>
</tr>
<tr>
<td>Mcts2</td>
<td>TGCCACTCCTCTGTGAG</td>
<td>CCACAGCCCCTCATTAGATA</td>
</tr>
<tr>
<td>Megal2</td>
<td>TCAACCCCTCCACACTAGA</td>
<td>AGGCTCAAAGCACCATTCAAG</td>
</tr>
<tr>
<td>Mit1</td>
<td>TCTCCCTCCTTGTGAGGAGG</td>
<td>TGAACCCTTGGAAACTCCAG</td>
</tr>
<tr>
<td>Ndn-F2/R2</td>
<td>AGGCGACGTGCTGATGAAAG</td>
<td>ATCTCTGCTGGAGCAGAAC</td>
</tr>
<tr>
<td>Nnat</td>
<td>AGATCGAACAGGCCCACATCAG</td>
<td>GCAAGGAGGGAATAATGGT</td>
</tr>
<tr>
<td>Oct4</td>
<td>ACCAGGCTCAGAGGTATTGG</td>
<td>CCCTCGAGTGTTGCTTACC</td>
</tr>
<tr>
<td>Peg1-F3/R3</td>
<td>AAAGGCACTTGGGTTTGGGCTG</td>
<td>TCAGACCCATTTGCTCTTAG</td>
</tr>
<tr>
<td>Peg10</td>
<td>GTGGAGTTGGCAGGGCTAGG</td>
<td>CTTCAAGGGCAGTCATCAG</td>
</tr>
<tr>
<td>Peg13</td>
<td>TCTGTCCTTGTGATTCATTCC</td>
<td>TGGTCTGGGAGGCTGATTC</td>
</tr>
<tr>
<td>Peg3-F2/R1</td>
<td>CCGGACTTCACTCAAGACTG</td>
<td>CCAGATTCAAAGACAGAC</td>
</tr>
<tr>
<td>Plagl1</td>
<td>TTTGGGAAAGTCACCTCCTGAC</td>
<td>GTTATCTATGGGACAGTG</td>
</tr>
<tr>
<td>Rasgrf1</td>
<td>GTGCAGTGCGATCTCCTCAGG</td>
<td>GTCGTAATCTTGGGTTGTC</td>
</tr>
<tr>
<td>Sfmbt2SNP4103</td>
<td>CACATTGGAAAGCAAGCACA</td>
<td>CTGATTAGTCTCACAATGAC</td>
</tr>
<tr>
<td>Sfmbt2 short 3'end</td>
<td>TTCTGTTATTAACTGAATTCCATA</td>
<td>TGCTGTTCAACTTGGGCATT</td>
</tr>
<tr>
<td>Sgce</td>
<td>ACAAAGGACTGGTCACAGCTATC</td>
<td>CGGCAGCACATGATAAGC</td>
</tr>
<tr>
<td>Slc38a4</td>
<td>CTCCACCCCCAGATTAACAG</td>
<td>GATCCTCACAGGGACAGC</td>
</tr>
<tr>
<td>Snrpn</td>
<td>ATGGCTCGTGGCTCCTTGTG</td>
<td>GGCCTCAAACCTCCTGAGC</td>
</tr>
</tbody>
</table>
## D. Primers used for qRT-PCR experiments.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>q-Cdx2Fo/Ro (template)</td>
<td>TAGCCCACGGTCTCAATGGT</td>
<td>CTGCGGTTCTGAAACCAAAT</td>
</tr>
<tr>
<td>q-Cdx2F/R (qRT-PCR)</td>
<td>CAGTCCCTAGGAAGCCAAA GTGAAA</td>
<td>AAGTGAAACTCCTTTTCTCCAGCTCC</td>
</tr>
<tr>
<td>q-Tpbpa-Fo/Ro (template)</td>
<td>ATAGTCCCTGAAGCGCAGTT</td>
<td>TTGCCCAAGTTCGATCTTCT</td>
</tr>
<tr>
<td>q-Tpbpa-Fi/Ri (qRT-PCR)</td>
<td>GAAGGCTCACAATAGAAAGTTAAG</td>
<td>CTTCAAAACATTGGGTGTAACGAG</td>
</tr>
<tr>
<td>q-Gcml-Fo/Ro (template)</td>
<td>CATCTACAGCTCGGACGACA</td>
<td>AACTTCCCTGACTCGGATT</td>
</tr>
<tr>
<td>q-Gcml-Fi/Ri (qRT-PCR)</td>
<td>CTTCATCCTTTCCAGCTCAAAAGTTAAG</td>
<td>CACTTTCTCAGTCGTCCTCTTTG</td>
</tr>
<tr>
<td>q-PLI-Fo/Ro (template)</td>
<td>CGAGAGGAAAGTCACCAGAAC</td>
<td>AACTGAGGAGGGGAACAGAAT</td>
</tr>
<tr>
<td>q-PLI-Fi/Ri (qRT-PCR)</td>
<td>GTCTTTTCAAGAACAAGTTTGAGTG</td>
<td>AAGCATTACAAACGTTCTGGTCAACA</td>
</tr>
<tr>
<td>q-Sfmbt2-S-Fo/Ro (template)</td>
<td>GCAATCTCCATAAAAACAGG</td>
<td>CGTCTTTTCAACCCAGTGTCAA</td>
</tr>
<tr>
<td>q-Sfmbt2-S-Fi/Ri (qRT-PCR)</td>
<td>AACTCTGTCTCTAAGGGTACTGCTG</td>
<td>ATCTGTATCTTATCCGGACAACAACA</td>
</tr>
<tr>
<td>q-Sfmbt2-L-Fo/Ro (template)</td>
<td>TCAGGACCTCTGAGTTTCTTCCGTT</td>
<td>GTCTCTCTGGACCTCGTACAAGT</td>
</tr>
<tr>
<td>q-Sfmbt2-L-Fi/Ri (qRT-PCR)</td>
<td>AGAATACCTGTCCATCTTGGTTCTCAGTC</td>
<td>ACTCTAAATCAATGGTCTGTCACAGT</td>
</tr>
<tr>
<td>q-Runx1-Fo/Ro (template)</td>
<td>CCTCTCTGCAGAACCTTCTTCTTCTT</td>
<td>GCTTTGTTGAGGGGCTAGCTTCTGCTAT</td>
</tr>
<tr>
<td>q-Runx1-Fi/Ri (qRT-PCR)</td>
<td>CTCTCGCTACCACACTCACCTAGT</td>
<td>AGGCCGCGTATAGATAGGTTTAG</td>
</tr>
<tr>
<td>q-TBP-Fo/Ro (template)</td>
<td>ATCTACGCTGAATCTTGGGCTGTTG</td>
<td>GGAGAAAATCTCCTGGGTTGAA</td>
</tr>
<tr>
<td>q-TBP-Fi/Ri (qRT-PCR)</td>
<td>GAATACAGAGAGCCACAGCAA</td>
<td>TTCTTGCTGCTAGTCTGGATTG</td>
</tr>
<tr>
<td>q-PLII- Fo/Ro (template)</td>
<td>CCAAACGTGGATTGTTGCGTT</td>
<td>GCACCGCCATAAGGTTCTAAA</td>
</tr>
<tr>
<td>q-PLII- Fi/Ri (qRT-PCR)</td>
<td>CACACTGCTGCAATCTCTACAACCCT</td>
<td>GCTTCTCCAGGCTTGGTACAA</td>
</tr>
<tr>
<td>q-PLF- Fo/Ro (template)</td>
<td>AGGAGCCATGATTTTGTGAGT</td>
<td>TGAGTCAAACAAACATGGAAGA</td>
</tr>
<tr>
<td>q-PLF- Fi/Ri (qRT-PCR)</td>
<td>TATAACATCTCAAGGCCCAACAGA</td>
<td>CAGGGTTCTCTTCTTCTCTTCACTTC</td>
</tr>
<tr>
<td>q-PRP- Fo/Ro (template)</td>
<td>TTCTCTGGGAGAGGATTGCTCTG</td>
<td>TCGAAGCGTTCGGTGATTGCT</td>
</tr>
<tr>
<td>Primer Name</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>NcoI-His-Sfmbt2(ab)-F</td>
<td>GATATCCATGGATCATCATCATCATCATCACATGGAGAGATATTTGCAGTTTC</td>
<td></td>
</tr>
<tr>
<td>Sfmbt2(ab)-Xhol-R</td>
<td>AGCCGGATCCCTCGAGCTATAATAGATGAATAATAGCA</td>
<td></td>
</tr>
</tbody>
</table>

E. Primers used for construction of His-tagged Sfmbt2 overexpression construct, to be used for generation of anti-SFMBT2 antibody.
Appendix II. Peptide spectrum and probabilities for identification of SMFBT2 (in-gel).

A. SFMBT2 protein sequence showing the sequence and location of the predicted peptides. Results are from coimmunoprecipitation with anti-SFMBT2 antibody and in-gel (>65kD) preparation for mass spectrometry. B. Percent probability of the predicted peptides. C. Individual spectrum for each predicted peptide. Information and graphs shown in this figure were generated by Scaffold (see materials and methods).
Appendix II. C continued:
Appendix III. Peptide spectrum and probabilities for identification of SMFBT2 (insolution).

A. SFMBT2 protein sequence showing the sequence and location of the predicted peptides. Results are from coimmunoprecipitation with anti-SFMBT2 antibody and insolution preparation for mass spectrometry. B. Percent probability of the predicted peptides. C. Individual spectrum for each predicted peptide. Information and graphs shown in this figure were generated by Scaffold (see materials and methods).
Appendix IV. Peptide spectrum and probabilities for identification of CASZ1 (in-gel).

A. CASZ1 protein sequence showing the sequence and location of the predicted peptides. Results are from coimmunoprecipitation with anti-SFMBT2 antibody and in-gel (>65kD) preparation for mass spectrometry. B. Percent probability of the predicted peptides. C. Individual spectrum for each predicted peptide. Information and graphs shown in this figure were generated by Scaffold (see materials and methods).
Appendix V. Peptide spectrum and probabilities for identification of CASZ1 (In-solution).

A. CASZ1 protein sequence showing the sequence and location of the predicted peptides. Results are from coimmunoprecipitation with anti-SFMBT2 antibody and in-solution preparation for mass spectrometry. B. Percent probability of the predicted peptides. C. Individual spectrum for each predicted peptide. Information and graphs shown in this figure were generated by Scaffold (see materials and methods).