Telomere Position Effect in Embryonic Stem Cells: heterogeneity, imprinting, and modifier screen

by

Rochelle Justina Szeto

B.Sc. The University of British Columbia, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE in

The Faculty of Graduate and Postdoctoral Studies (Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA (Vancouver)

December 2017

© Rochelle Justina Szeto, 2017
Abstract

Telomeres are repetitive sequences found at the end of linear chromosomes. Their main function is to protect chromosome ends from degradation and to ensure proper DNA replication. Telomere position effect (TPE) refers to the epigenetic phenomenon of a gene being stochastically active or silenced as a consequence of its proximity to telomeric heterochromatin. TPE is a subset of position effect variegation (PEV), which also involves variable transcriptional silencing due to the proximity to centromeric heterochromatin or transposable elements. We have observed a TPE-like effect on GFP expression from the DelTel7 allele in embryonic stem cell (ESC) lines. The DelTel7 allele is an engineered chromosome truncation carrying a GFP reporter next to an array of telomere repeats. The truncation breakpoint is in the middle of a large cluster of imprinted genes found on distal mouse chromosome 7. Our results suggest that the GFP reporter is regulated by TPE in undifferentiated DelTel7/+ ESCs. The studies described in this thesis first addressed the relationship between GFP heterogeneity seen in DelTel7/+ ESCs and the previously described phenotypic heterogeneity existing in undifferentiated ESCs grown in serum. TPE was found to be active in all ESCs grown in serum, irrespective of their state. Growth in KSR+2i serum-free medium, which forces ESCs into a naïve state, up regulates the promoter driving GFP expression, at both the DelTel7 allele and at an interstitial control transgene. My results show that, as previously described in a yeast model of TPE, the silencing imposed by proximity to the telomeres can be at least partially overcome by increased transcription. A parent-of-origin effect at the GFP reporter of the DelTel7 allele was revealed in vivo, but TPE does not spread to nearby imprinted genes in ESCs. Finally, while working on the development of an episomal system to screen for modifiers of TPE, I noted that prolonged zeocin exposure has drastic effects on TPE. My results suggest a previously unappreciated relationship between DSB repair and TPE.
Lay Summary

Telomere position effect (TPE) refers to the negative effect that the ends of chromosomes, called telomeres, can have on the expression of a nearby gene. We have developed a system to study this phenomenon in mouse embryonic stem cells (ESCs) and gain further understanding of the molecular mechanisms involved. First, I showed that TPE is active in all ESCs, irrespective of their developmental state. In the process of developing a screen for genes affecting TPE, I also discovered that exposure to the drug zeocin, which kill cells by inducing breaks in the cell’s DNA, has profound effects on our TPE reporter in ESCs. These results suggest a previously unappreciated relationship between the machinery involved in the repair of DNA breaks and the function of telomeres.
Preface

The candidate (R. Szeto) performed all experiments except for Figure 4-1. All experiments were conducted with the help of A. Bogutz.

FACS was done at the UBC Flow Lab in Life Sciences Institute with sorting completed by J. Wong. DNA Sequencing was done by Genewiz DNA Sequencing and Molecular Biology Services.

The analysis of GFP imprinting *in vivo* from the DelTel7 allele was completed by Dr. T. Gu and A. Bogutz.

DelTel7/+ and +/-DelTel7 ESCs on an F1 (FVB x C57BL/6) background were a gift from Dr. M. Higgins, who established these lines from mice carrying our allele.

Data mining and literature research to choose *Dppa3*, *Nanog*, *Pecam1*, *Zfp42*, and episomal screen was completed by Dr. L. Lefebvre, A. Bogutz, and R. Szeto.

R. Szeto, Dr. L. Lefebvre and A. Bogutz conceived of the study.

Ethics approval was obtained from the Animal Care Committee at the University of British Columbia for mouse work (Protocol number A11-0293).
Table of Contents

Abstract .............................................................................................................................. ii
Lay Summary .................................................................................................................. iii
Preface ............................................................................................................................. iv
Table of Contents .......................................................................................................... v
List of Tables ................................................................................................................... vii
List of Figures ................................................................................................................ viii
List of Abbreviations .................................................................................................... ix
List of Gene Names ........................................................................................................ xi
Acknowledgements ....................................................................................................... xii
Dedication ...................................................................................................................... xiii

Chapter 1: Introduction .................................................................................................. 1

1.1 Telomeres ................................................................................................................. 1
1.2 Position effect variegation (PEV) and telomere position effect (TPE) ............... 2
1.3 Engineering of the DelTel7 allele ............................................................................ 9
1.4 DelTel7 allele and telomere position effect ............................................................ 11
1.5 Thesis theme and objectives .................................................................................... 14

Chapter 2: Materials and Methods .............................................................................. 16

2.1 Tissue Culture .......................................................................................................... 16
2.2 KSR+2i medium ........................................................................................................ 18
2.3 FACS and flow cytometry analysis .......................................................................... 19
2.4 Genomic DNA (gDNA) extraction ......................................................................... 19
2.5 RNA isolation, DNase treatment and cDNA synthesis .......................................... 20
2.6 RT-qPCR. .................................................................................................................. 21
2.7 ESC electroporation ............................................................................................... 22
2.8 pPGKzeo and pMGD20zeo construction ................................................................. 23
2.9 Neomycin selection, zeocin selection, and picking colonies .................................... 23
2.10 Western blotting: SDS-page and immunoblotting ............................................... 24
2.11 Genotyping ............................................................................................................. 25
2.12 Bisulphite sequencing ........................................................................................... 25
2.13 Immunohistochemistry (IHC) ................................................................................ 25

Chapter 3: Relationship between ESC heterogeneity and TPE .................................... 28

3.1 Introduction .............................................................................................................. 28
3.2 Results ...................................................................................................................... 30

Chapter 4: Imprinting effects at the DelTel7 allele ....................................................... 40
4.1 Introduction.................................................................40
4.2 Results.............................................................................41

Chapter 5: Screen for Modifiers of TPE: Development of an Episomal System ..........48
5.1 Introduction........................................................................48
5.2 Results.............................................................................51

Chapter 6: Discussion and Future Directions .................................................................61
6.1 Correlation between heterogeneity in ESCs and GFP variegation in DelTel761
6.2 Interpretation of KSR+2i treatment on DelTel7 ESCs.................................61
6.3 Expression of Igf2 and H19 in DelTel7 sorted population.........................62
6.4 Determination of imprinting at the DelTel7 allele in ESCs.......................63
6.5 Rationale for producing DelTel6.........................................................64
6.6 Verification of episomal system..............................................................67

References..................................................................................70
List of Tables

Table 2-1. Approximate amounts of media used for different sizes of plates used in tissue cultures................................................................. 17

Table 2-2. Percentages of solutions used in different medias for different cell types. .......... 17

Table 2-3. Annealing and reading temperatures for genes used in RT-qPCR experiments .... 22

Table 2-4. List of primers used in this study ................................................................................. 27
List of Figures

Figure 1-1. DelTel7 Production .............................................................................................................. 11
Figure 1-2. GFP expression in Tel7KI and DelTel7 undifferentiated ESCs and embryoid bodies. .............................................................. 12
Figure 1-3. Flow cytometry analysis of sorted DelTel7 ESCs through a 7 day time course. . . . 13
Figure 1-4. Effect of G418 selection on the fraction of GFP+ cells for Tel7KI and two different DelTel7 ESCs .......................................................................................................................... 14
Figure 3-1. RT-qPCR analysis of EGFP and neo on sorted GFP+ and GFP DelTel7 cells . . . . . 31
Figure 3-2. RT-qPCR analysis of Dppa3, Nanog, Pecam1, and Zfp42 on sorted GFP+ and GFP DelTel7 cells. .................................................................................................................................................. 33
Figure 3-3. RT-qPCR analysis of Nanog expression in DelTel7 B3 and C3 and Tel7KI ESCs grown in different media .......................................................................................................................... 36
Figure 3-4. Fluorescent microscopy analysis of DelTel7 ESCs maintained in ESM, KSR, or KSR+2i. ............................................................................................................................................................ 37
Figure 3-5. Effects of KSR + 2i on GFP expression for DelTel7B3, C3, F1, and Tel7KI ESCs 38
Figure 3-6. RT-qPCR of neo expression in DelTel7 B3, C3, F1 and Tel7KI ESCs grown in different media ........................................................................................................................................................................... 39
Figure 4-1. GFP expression of +/-Tel7KI, Tel7KI/+, +/-DelTel7, and DelTel7/+ E7.5 embryos . . 40
Figure 4-2. Illustration of how imprinting of GFP reporter can affect later downstream genes . 42
Figure 4-3. RT-qPCR of H19 mRNA expression on DelTel7 B3, C3, and F1 subclones. . . . . . . . . . 43
Figure 4-4. RT-qPCR of Igf2 mRNA expression on DelTel7 B3, C3, F1 subclones. . . . . . . . . . . . . . . . 43
Figure 4-5. GFP expression of MH-DF and MH-FD DelTel7 ESCs under a microscope. . . . . . 45
Figure 4-6. Flow Analysis of MH-FD and MH-DF line. ........................................................................ 45
Figure 4-7. Ape alignment of bisulphite sequencing at IC2 from the reference sequence, DelTel7, MH-DF and MH-FD ESCs. ........................................................................................................................................................................ 47

Figure 5-1. Strategy for Modifier screen of enhancers and suppressors of TPE using DelTel7 and pMGD20zeo ........................................................................................................................................................................................................................................ 49

Figure 5-2. A modified strategy for the modifier screen in DelTel7 ESCs. ........................................ 50

Figure 5-3. GFP expression of DeTel7-C3 and several DeTel7C3-pMGD20zeo ESCs. ............... 52

Figure 5-4. Percentage of GFP+ DelTel7 and pMGD20zeo ESCs analyzed by flow cytometry. 54

Figure 5-5. Time course done on DelTel7 and selected pMGD20zeo cells. ......................... 55

Figure 5-6. Percentage of GFP expression in pPGKzeo transgenic ESCs. ........................................ 57

Figure 5-7. Time course of DelTel7-pPGKzeo without zeocin selection. ................................. 58

Figure 5-8. Time course of DelTel7-pMGD20zeo without any zeocin selection. .................. 59

Figure 5-9. RT-qPCR for Large T Antigen on 293T cells DelTel7C3 ESCs and DelTel7C3-pMGD20zeo clones. .......................................................................................................................................................................................................................... 60

Figure 6-1. Schematic diagram of the proposed strategy to use MICER techniques to produce DelTel6. .......................................................................................................................................................................................................................... 66

Figure 6-2. Model of the GFP variation seen in DelTel7 ESCs electroporated with Large T antigen and placed under zeo selection. .......................................................................................................................................................................................................................... 69
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1xFM</td>
<td>1x freezing media</td>
</tr>
<tr>
<td>BSC</td>
<td>biosafety cabinet</td>
</tr>
<tr>
<td>ChIP</td>
<td>chromatin immunoprecipitation</td>
</tr>
<tr>
<td>Chr</td>
<td>Chromosome</td>
</tr>
<tr>
<td>CMV-IE</td>
<td>Cytomegalovirus immediate early enhancer</td>
</tr>
<tr>
<td>DEPC</td>
<td>diethylpyrocarbonate</td>
</tr>
<tr>
<td>dH₂O</td>
<td>distilled water</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DSB</td>
<td>double strand break</td>
</tr>
<tr>
<td>ESCs</td>
<td>embryonic stem cells</td>
</tr>
<tr>
<td>ESM</td>
<td>embryonic stem cell medium</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
</tr>
<tr>
<td>FISH</td>
<td>fluorescence in situ hybridization</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>IC1</td>
<td>distal Chr7 imprinting center 1</td>
</tr>
<tr>
<td>IC2</td>
<td>distal Chr7 imprinting center 2</td>
</tr>
<tr>
<td>KSR</td>
<td>knockout serum replacement</td>
</tr>
<tr>
<td>KSR+2i</td>
<td>knockout serum replacement (with) 2 inhibitors</td>
</tr>
<tr>
<td>LIF</td>
<td>leukemia inhibitory factor</td>
</tr>
<tr>
<td>M-MLV</td>
<td>moloney murine leukemia virus</td>
</tr>
<tr>
<td>MPD</td>
<td>mean population doublings</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffer saline</td>
</tr>
<tr>
<td>PBS-T</td>
<td>phosphate buffer saline-Tween 20</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEF</td>
<td>primary embryonic fibroblasts</td>
</tr>
<tr>
<td>PFM</td>
<td>primary embryonic fibroblasts media</td>
</tr>
<tr>
<td>PEV</td>
<td>position effect variegation</td>
</tr>
<tr>
<td>PFA</td>
<td>paraformaldehyde</td>
</tr>
<tr>
<td>PGK</td>
<td>phosphoglycerate kinase</td>
</tr>
<tr>
<td>PI</td>
<td>propidium iodide</td>
</tr>
<tr>
<td>Py</td>
<td>Polyoma virus</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>reverse transcription quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>rpm</td>
<td>rotations per minute</td>
</tr>
<tr>
<td>siRNA</td>
<td>small or short interfering RNA</td>
</tr>
<tr>
<td>SV</td>
<td>Simian virus</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/Borate/EDTA buffer</td>
</tr>
<tr>
<td>TSA</td>
<td>trichostatin A</td>
</tr>
<tr>
<td>TPE</td>
<td>telomere position effect</td>
</tr>
<tr>
<td>TPE-OLD</td>
<td>telomere position effect over long distances</td>
</tr>
</tbody>
</table>
List of Gene Names

ade2  Adenine requiring
Dppa3/Stella  Developmental Pluripotency Associated 3
DsRed2  Discosome red fluorescent protein
EGFP  enhanced Green Fluorescent Protein
GSK  Glycogen synthase kinase
H19  H19 gene, ncRNA
Hp1  Heterochromatin protein 1
 Hp1
Hp1  Heterochromatin protein 1
Hp1t  Hypoxanthine Guanine PhosphoribosylTransferase
(h)TERT  (human) telomerase reverse transcriptase
Igf2  Insulin-like growth factor 2
Ins2  Insulin2
Nanog  NanogHomeobox
Pecam1  Platelet/Endothelial Cell Adhesion Molecule 1
Ppia  PeptidylProlylIsomerase A
Puro  Puromycin resistance gene
neo  Neomycin/G418 resistance gene
Sirt6  Sirtuin-6
 tk promoter  Thymidine kinase promoter
Trf1/2  Telomeric repeating factor 1 or 2
Ura3  Orotidine 5’phosphate decarboxylase
w+  A.1 White gene of Drosophila
wnt  Wingless type
zeo  Zeocin resistance gene Sh ble
Zfp42/Rex1  Zinc finger protein 42
Acknowledgements

First and foremost, I would like to thank Dr. Louis Lefebvre for the opportunity to learn, work, and grow within the lab. Your guidance has taught me to be patient, persistent, and to think positively! Within the Lefebvre Lab, I would also like to thank Aaron Bogutz, Karen Jacob, Ting Gu, and Amanda Ha for all of their amazing support.

Next, I would like to thank my family: Mom, Dad, Reagan and Eurwen. Thank you for giving me the foundation to pursue my dreams and always believing that I will find my way – whether it took 1 year or 5.

Lastly to my amazing other half, Dennis Chen – Thank you for being by my side and calming me down whenever a storm hit.
Dedication

I dedicate this to my mom and dad, Patricia Tham and Raymond Szeto. You two push me to be anything I want to be.
Chapter 1: Introduction

1.1 Telomeres

Telomeres are found at both ends of linear chromosomes in eukaryotes and have a repetitive terminal sequence. Telomeres in humans are generally 10-20 kilobases long while mice telomeres are 20-100 kb long\(^1\). In humans and mice, the repeated telomere sequence is TTAGGG or T\(_2\)AG\(_3\)\(^2,3\). In *Tetrahymena*, a ciliate, the sequence of the tandem repeats is TGGGG\(^2\). Some species have regular tandem repeats while others have irregular repeats. Some examples are in yeast where the telomere contains a T subsequently followed by one to three Gs (TG\(_{1,3}\))\(^1\) and Paramecium where the sequence alternates between TTGGGG and TTTGGG\(^2,4\). The repetitive state, but not the specific nucleotide sequence, of telomeres is highly conserved across species because they share the same functions.

Telomeres have three main functions. The first function is to help maintain the structural integrity of the genome. Because telomeres are at the end of the chromosome, they help to prevent chromosome fusion by protecting otherwise exposed double stranded DNA (dsDNA). Usually, dsDNA is interpreted as a double stranded break (DSB). When DSB is detected by the cell, a cascade of events is triggered called the DNA damage response (DDR). This pathway activates a variety of factors to fix the DSB, which can cause chromosomes to fuse to each other. Another function is to ensure that DNA replication is complete during cell divisions. This ensures essential genes near the end of the chromosomes are replicated and passed down onto sister cells. After many cycles of cell division, the DNA at the end of the chromosomes can be lost. Telomeres at the end of chromosomes act as a buffer. The shelterin complex was found to be involved in protecting the telomeric ends.

The last function that telomeres have been implicated in is subnuclear positioning. Heterochromatin regions generally gravitate more towards the outer areas of the nucleus as
opposed to the center. Heterochromatin and euchromatin are two conformations of DNA. Euchromatin refers to an open state where genes are accessible and transcribed. Euchromatin is generally associated with H3K4 and H3K27 acetylation marks. Heterochromatin is the exact opposite in the closed conformation, which is associated with H3K9 methylated repression marks. Genes in heterochromatin are not accessible and are usually not transcribed. There are two types heterochromatin: facultative and constitutive. Constitutive heterochromatic regions are the same in all cells, like the centromeres and telomeres. Specific areas of some chromosomes have also been found to be constantly compact. Facultative heterochromatin refers to highly compact regions that are not necessarily compact in all cells like the X chromosome in female cells, since there is random X inactivation of one of the two X chromosomes.

1.2 Position effect variegation (PEV) and telomere position effect (TPE)

Position effect variegation (PEV) was discovered by Hermann Muller in 1939 using forward genetics in *Drosophila*. The wild-type phenotype of *Drosophila* eyes is red and through radiation, a patchy red and white eye phenotype was found. In later years, it was determined that an inversion occurred and caused the white (w+) gene to be placed closer to the centromere. As previously mentioned, the centromere is a constitutive heterochromatin or compact region. The stochastic spreading of the heterochromatin region caused the w+ gene to sometimes be expressed and sometimes be repressed. This phenomenon of a gene being repressed or expressed due to proximity to a heterochromatic region is called PEV. One particular type of PEV is Telomere Position Effect (TPE). TPE is the expression or repression of a gene based on its proximity to the telomere. This epigenetic phenomenon was first discovered in yeast but since then has been shown to be widespread in different model organisms.


1.2.1 Drosophila

Drosophila with its fast life cycle, only four chromosomes, and ease of maintenance has made it a great choice to study genetics. Their telomeres contain tandem arrays of two types of non-LTR retrotransposons⁷. These two retrotransposons are Het-A and TART and are orientated to have their oligo-A tails facing the centromere⁷. Many studies have looked at the differences between PEV and TPE using Drosophila as a model organism. They have all concluded that TPE does not have the same suppressors and enhancers of PEV⁸-¹⁰. One exception is Su(z)25 which has been found to be a stronger suppressor¹¹.

1.2.2 Yeast

The yeast genome and the human genome have many differences. The yeast genome contains 16 chromosomes whereas the human genome contains 23 chromosomes. Another key difference is in the structure of their telomeres. Yeast have a simple telomere repeat at the ends of all linear chromosomes (TG₁₃)ₙ¹². Yeast telomeres also contain repetitive telomere-associated sequences which are equivalent to mammalian subtelomeres¹³,₁⁴. Telomeres without any telomere-associated sequences can exert a transcriptional effect on the expression of genes located nearby. For instance when the URA3, TRP1, HIS3 or ADE2 markers were placed near a telomere, the expression of these genes was repressed¹⁵. This led to the concept that telomeres exert a position-dependent effect on transcription, a phenomenon called telomere position effect (TPE).

Many studies with the yeast system take advantage of the negative/positive selectable marker URA3 and the toxic “suicide” drug 5-fluororotic acid (FOA)¹⁵-¹⁷. TPE is able to act on URA3 if the gene is inserted near the telomere. The production of URA3 will render the colonies sensitive to FOA medium whereas URA3 repression leads to FOA₅ cells. Gottschling et al took
advantage of the URA3-FOA relationship and found that when cells are pregrown in medium containing uracil, most cells plated on FOA were resistant, suggesting repression of the URA3 marker\textsuperscript{15}. These FOA\textsuperscript{R} cells were then replated onto media lacking uracil and cells were still able to grow\textsuperscript{15}. These results suggest that URA3 expression was high enough to sustain growth on uracil drop-out medium but also low enough to confer a FOA\textsuperscript{R} phenotype. This unexpected characteristic of FOA\textsuperscript{R} and URA3\textsuperscript{+} cells led the researchers to examine URA3 RNA expression directly. RNA levels were normal for URA3 gene at its normal locus\textsuperscript{15}. URA3 RNA expression was 5-fold reduced under non-selective conditions when the URA3 gene was placed near the telomere\textsuperscript{15}. Under selective conditions, (uracil drop-out medium), URA3 at telomere and URA3 at chromosome had similar RNA expression\textsuperscript{15}. These results show that URA3 RNA is significantly reduced when the gene is placed next to the telomere but that activation of transcription can overcome this effect.

The researchers wanted to determine if these position effects could be generalized with other genes like TRP1, HIS3 and ADE2 when placed near the telomere. Under non-selective conditions, most of the cells with the TRP1 gene near the telomere exhibited very low to no expression of TRP1\textsuperscript{15}. ADE2\textsuperscript{+} colonies are white while ade2\textsuperscript{−} colonies are red when grown on appropriate medium. When ADE2 was placed near the telomere, all the colonies contained red and white sectors\textsuperscript{15}. Reversibility of the phenotypes within these colonies was shown when white and red sectors were separated and replated. Each new growth consisted of red and white colonies\textsuperscript{15}. All these results show that genes close to the telomere are affected by TPE in a reversible way that the phenomenon share similarities in yeast and Drosophila.

In 1994, Sandell \textit{et al} introduced a galactose-inducible promoter in yeast chromosome VII in three different strains\textsuperscript{17}. The inducible promoter contained GAL1, UAS and a TATA element in that order\textsuperscript{17}. The specific order directed the transcription to occur towards the end of
the chromosome. The modified chromosome VII also contains URA3, a portion of ADH4 gene and a small segment of TG DNA. The strain containing all the above was labeled UGPT, while UGXT contained the same components with the elimination of the TATA element. The UT version did not have the GAL1, UAS, and TATA element. On glucose medium, all modified telomeres showed normal telomere lengths. However, in galactose medium, UGXT and UT-modified telomeres exhibited ~25bp reduction of telomere length while the UGPT telomeres were reduced by ~90bp\textsuperscript{17}. These results confirm that transcription through a telomere decreases the telomere length.

To link TPE to telomere transcription, all strains were grown in glucose or galactose medium and then plated on FOA medium\textsuperscript{17}. All strains grown in glucose had mostly cells that were FOA resistant, which means the \textit{URA3} gene was repressed\textsuperscript{17}. UT-telomere modified cells had mostly FOA resistant cells but UGPT and UGXT-telomere modified cells had a 4-fold reduction of FOA resistant cells when grown on galactose medium compared to glucose\textsuperscript{17}. These results highlight two important characteristics of TPE: \textit{i}) neighbouring genes inserted at telomeres tend to be co-regulated by TPE; \textit{ii}) activation of transcription can compete with TPE and allow greater expression of telomeric genes.

Mishra \textit{et al} used the negative selection between URA3 and FOA to demonstrate that Ku likely recruits or activates Sir4 proteins but more importantly, Ku is required to overcome the inhibitory effects of Rif proteins to restore telomere silencing\textsuperscript{18}. \textit{Yku70} mutants lack Ku protein and are defective in telomeric silencing\textsuperscript{18}. Rif1 was found to co-localize with Rap1 at telomeres\textsuperscript{18}, which is expected as Rif and Sir proteins have been shown to compete with the Rap1p carboxyl terminus\textsuperscript{19}. With a telomeric-URA3 gene, the triple mutants \textit{rif1 rif2} and \textit{yku70} yeast phenotype in FOA media is indistinguishable from WT colonies\textsuperscript{18}. This result confirms that it is essential for Ku to overcome Rif protein’s inhibitory effect for telomere silencing. A
substantial increase in telomere length would also restore TPE but only if the telomere is increased by 400-500bp\textsuperscript{18}. The presence of Sir4 also greatly improved TPE\textsuperscript{20}. This important work demonstrated that components of non-homologous end joining DNA repair pathway are also implicated in telomere maintenance and TPE\textsuperscript{16}.

1.2.3 Mammalian systems and TPE

In 1996, the first study was published that investigated gene expression and TPE in mammalian cells. Sprung et al used human fibroblasts transfected with a plasmid containing a neo selectable marker. Subclones were isolated that contained neo on chromosome 13 with varying lengths of telomere repeat sequences\textsuperscript{21}. The telomere sequences were approximately 0.5, 1, 3, 6, and 25 kb from the neo gene\textsuperscript{21}. Through several experiments like colony forming ability, growth rates, and RNA blot analysis, the group found that TPE had no effect on neo expression\textsuperscript{21}.

Another study examined TPE using a linear plasmid containing luciferase alongside a 1.6 kb telomere repeat transfected into telomerase-positive HeLa cells\textsuperscript{22}. Results showed that cells with a telomeric reporter produced 10-fold less expression than the interstitial clones\textsuperscript{22}. Luciferase expression was increased by 51 ± 37-fold when treated with trichostatin A (TSA), which is an inhibitor of histone deacetylases\textsuperscript{22}. More importantly, luciferase expression was returned back to normal within 72 hours when TSA treatment was withdrawn\textsuperscript{22}. These results demonstrate reversibility of the luciferase expression. HeLa cells were also infected with human telomerase reverse transcriptase (hTERT) to determine the length dependence of the silencing effect\textsuperscript{22}. A 2- to 10-fold decrease of luciferase was seen when telomerase extended the telomeress\textsuperscript{22}. These results suggest that TPE in mammalian cell is affected by telomere length.
Koering et al. used C33-A cells, a human carcinoma cervical cell line, to study the influence of telomere proximity and composition on EGFP expression\textsuperscript{23}. They transfected two plasmids; one containing only CMV-EGFP (pCMV) and the other containing CMV-EGFP and 1.6 kb of telomere repeats (pCMVTelo) \textsuperscript{23}. EGFP expression was found to be significantly higher in cells that were transfected with pCMVTelo than pCMV\textsuperscript{23}. After 10 days of hygromycin selection, there was less expression of EGFP in cells transfected with the telomere array\textsuperscript{23}. The expression plateaus at 10% GFP positive cells and the selection seems to be reversing the activating effect seen earlier. Sorted populations of mostly GFP\textsuperscript{+} and GFP\textsuperscript{−} cells were grown and cultured for 20 mean populations doublings (MPDs) \textsuperscript{23}. The GFP profiles remained relatively stable independent of hygromycin selection. Increased Trf1 gene dosage led to a large increase of EGFP-positive cells in two lines just after 7 MPDs; however no detected change in telomere length was observed after 20 MPDs\textsuperscript{23}. EGFP expression was increased when treated with TSA and indicates TSA effect is dependent on telomere location since control cells did not have any change\textsuperscript{23}. Focusing on the localization of three isoforms of heterochromatin protein, HP1α, β and γ, it was shown that the HP1 did not change FISH patterns during first TSA treatment. However, during the second TSA treatment, HP1 α and HP1β were released but HP1 γ was not\textsuperscript{23}.

Baur et al. studied telomere position effect in human cells using a single-cell fluorescent reporter similar to ours (see below). This group used a linearized plasmid to integrate a DsRed2 reporter into the genome, beside a newly formed telomere or randomly at an interstitial position. Subcloning and growth of colonies from individual positive or negative cells was used to determine the reversibility of the DsRed phenotype. The presence of any cells of opposite phenotype from the original single cell would clearly show that a transcriptional switching event has occurred at the reporter. Switching from positive to negative appeared to be easier than switching from negative to positive as more occurrences were noticed\textsuperscript{24}. Using 96-well plates to
separate the single cells during growth, cells initially negative were expressing DsRed2 within 1-2 weeks, and weakly positive cells gave rise to strongly positive and negative cells also within 1-2 weeks.\textsuperscript{24}

The first group to study TPE in mouse ESCs was Pedram \textit{et al} in 2006. Clones were obtained using transfection into mouse ESCs and a plasmid that contained neo and HSV \textit{tk}. These two selectable markers were transcribed outwards from each other with neo being transcribed using a HSV \textit{tk} promoter and polyoma virus enhancer and \textit{tk} was driven by the mouse \textit{Pgkl} promoter.\textsuperscript{25} Clones were isolated by Southern blot to determine if the plasmid was integrated near a telomere. Two types of clones were isolated: telomeric clones where the plasmid integrated near the telomere and interstitial clones where the plasmid was not near the telomere. Some of the interstitial clones lost some or most of the telomere repeat sequence but the telomeric clones did not.\textsuperscript{25}

Expression level of neo was measured using RT-qPCR and interstitial clones showed variable expression of neo with one clone having very high expression levels.\textsuperscript{25} This one clone also had the least amount of interstitial telomere repeat sequence.\textsuperscript{25} Telomeric clones had 4.5-fold lower expression of neo compared to the interstitial clones.\textsuperscript{25} Removal of G418 selection also decreased neo expression in telomeric clones, which again suggests that TPE in ESCs globally reduces transgene expression.\textsuperscript{25}

Another group (Tennan \textit{et al}) in 2011 used HeLa cells and determined that SIRT6 was essential for telomere position effect.\textsuperscript{26} Previously it was shown that overexpression of telomerase decreases telomeric luciferase expression.\textsuperscript{22} SIRT6 depletion was accomplished using siRNA and luciferase expression was increased at the telomere. Furthermore, this effect was also found to be reversible.\textsuperscript{26} To establish a link between chromatin structure and TPE, Tennan \textit{et al} also looked directly at H3K9ac active marks and H3K9me3 repressive marks. ChIP analysis
showed that hTERT overexpression lead to less active marks and significantly higher repressive marks\textsuperscript{26}. Knockdown of SIRT6 had opposite results of more active marks and less repressive marks\textsuperscript{26}. These series of results point to the fact that SIRT6 is necessary for telomere chromatin structure maintenance, which is implicated in TPE and the silencing of telomeric genes.

In 2014, Robin et al discovered that telomeres not only have silencing abilities on genes close by but also on genes over longer distances away from the telomere\textsuperscript{27}. They coined this process TPE-OLD, for telomere position effect over long distances. Using TERT cDNA flanked by loxP sites, this group was able to remove TERT at any time point to normalize the telomere lengths between different chromosomes\textsuperscript{27}. The group focused on 1p, 6p, and 12p and also long and short telomeres depending on when TERT was excised\textsuperscript{27}. Several genes on 12p were differentially expressed in these cells. These genes were located within 10Mb of the telomere and also 10Mb away\textsuperscript{27}. This result could mean that differential expression based on telomere length could affect genes further away from the telomere. Expression of BMP6 and DSP in myoblasts and fibroblasts were found to be significantly different between short telomeres and long telomeres\textsuperscript{27}. Using 3D-FISH, probes were used to look at subtelomeric ends of 1p and 12p and also C1S on 12p and ISG15 on 1p\textsuperscript{27}. Cells with longer telomeres had 85-92\% of the probes adjacent whereas shorter telomere cells had 76-96\% separation\textsuperscript{27}. All the above results indicate that shortening of telomere length distorts chromosomal looping and causes differential expression of genes located away from the telomere.

1.3 Engineering of the DelTel7 allele

At the distal end of mouse chromosome 7, there are two imprinting centers (IC) regulating imprinted gene expression over a 1 Mb domain. IC1 and IC2 are cis-acting differentially DNA methylated regions (DMRs)\textsuperscript{28}. These DNA methylated marks at those DMRs
are directly inherited from the gametes: the sperm in the case of IC1 at the \(H19-Igf2\) region, and the oocyte for the IC2 \(Kcnq1ot1\) promoter. These two ICs are inherited together on Chr 7 in mouse and the Deletion of Telomeric end of chromosome 7 (DelTel7) allele was created to study the relationship between these two imprinted domains\(^{28}\). Separating IC1 and IC2 would allow for further understanding of how they interact in the context of Chr7 and whether it is essential for them to be inherited together.

To generate the DelTel7 allele (official name Del\(^{(7}Ins2-Tel\)1LoxP, MGI ID: 3772922), a \(loxp\) docking site within a promoter-less \(loxp\)-neo-pA cassette was first targeted at the distal end of mouse chromosome 7 in ESCs, 2.6 kb downstream of \(Ins2\)^{28}. The \(Ins2\) gene is within the last cytogenic band on chromosome 7 (band 7F5) and the targeted allele \(I2loxP\) (official name \(Ins2\) \(tm1.1Nagy\), MGI ID: 3807183), is located immediately telomeric of \(Ins2\), in between the IC1- and IC2-regulated imprinted domains. The targeted allele is quite close to the end of the chromosome and led to the hypothesis that a Chr7 telomeric truncation could be introduced at this location. To achieve this, a linear vector containing a terminal array of telomeric repeat, a “ubiquitous” EGFP reporter (pCAGGS-\(EGFP\))\(^{29}\) and a \(Pgk\) promoter-\(loxp\) fusion to activate the targeted neo at \(Ins2\) was electroporated in \(I2loxP/+\) ESCs along with a Cre-expressing vector\(^{28}\). This particular technique is known as recombinase-mediated chromosome truncation (RCMT). The vector, when linear, produced the DelTel7 truncation allele, whereas if the vector is circular, Tel7KI was produced\(^{28}\). In Tel7KI, the vector is simply inserted at \(I2loxP\) (pop-in reaction), with no deletion of Chr7 sequences and provides an interstitial EGFP reporter used as a control line for most of my experiments. From the distal end, DelTel7 contains a new telomere, extended by telomerase in ESCs, the CAG-EGFP reporter, and a reformed functional \(Pgk-loxp\)-neo-pA selectable marker. Those same elements are all present at the Tel7KI insertional allele (Figure 1-1). Both of
these alleles were taken to the germline through transmitting male chimeras and our laboratory maintains colonies of Tel7KI and DelTel7 mice.

Figure 1-1. DelTel7 Production.

Recombinase Mediated Chromosome Truncation (RCMT) was used to truncate chromosome 7 with an EGFP reporter inserted upstream of Ins2 at the I2loxP allele (A). Tel7KI was produced because there was some circular vector in the mix and the vector was popped into I2loxP site by Cre (B). Both DelTel7 and Tel7KI were recovered from the same experiment in ESCs (C). Figure adapted from Jones et al (2009)30.

1.4 DelTel7 allele and telomere position effect

Several preliminary experiments were conducted on DelTel7/+ ESCs (referred to below as DelTel7 ESCs) and showed that those ESCs can be used to study telomere position effect. The
first experiments seen in Figure 1-2 showed that there is variegated expression of GFP in undifferentiated ESCs as well as embryoid bodies. This variegation is indicative that there is some similarity with the patchy red and white *Drosophila* eyes seen in PEV mutants or in yeast ADE2 TPE reporter strains. The control cell line, Tel7KI, exhibited ubiquitous expression of GFP as expected from an interstitial GFP reporter.

**Figure 1-2. GFP expression in Tel7KI and DelTel7 undifferentiated ESCs and embryoid bodies.**

Tel7KI/+ ESCs are always fluorescing green whereas the DelTel7/+ ESCs have variegated GFP expression. These GFP profiles are seen in undifferentiated ESCs and embryoid bodies.

The second factor discovered is that the state of EGFP expression is reversible. DelTel7 ESCs were sorted by flow cytometry into GFP$^+$ and GFP$^-$ populations. Both populations were then placed back into culture and analyzed for seven days. We can easily see in Figure 1-3 that at the top, the sorted population started off as mostly GFP$^-$ and within four days, GFP$^+$ cells were already seen. Finally, at day seven the distribution of GFP expression has reverted back to the original profile. For the sorted population containing mostly GFP$^+$ cells, we can see that throughout the course of seven days, there is a slow progression towards a 20:80 percentage of GFP$^+$ to GFP$^-$ ESCs. This reversibility is seen in many TPE and PEV alleles because the
telomere and centromere are thought to fluctuate between open and closed configuration, or to nucleate variable spreading of heterochromatin to the reporter. The open and closed state causes genes nearby to be expressed or repressed.

Figure 1-3. Flow cytometry analysis of sorted DelTel7 ESCs through a 7 day time course.

After DelTel7 ESCs are sorted, they are placed back into culture and we can see that the GFP profile reverts back to its original unsorted population of 10-25% GFP+ population and 75-90% GFP− population.

The final key aspect of DelTel7 is that the expression of GFP and neo were found to be linked. This was assessed through G418 selection, which gradually eliminates DelTel7 cells in which the neo marker is silenced. Without G418 selection, Tel7KI cells remain ~98% GFP+ whereas DelTel7 ESCs are usually between 10-25% GFP+. With G418 selection, Tel7KI cells also remain highly GFP+ whereas 50-58% of DelTel7 ESCs are now GFP+. This 5-fold increase in the proportion of GFP− expressing DelTel7 cells between no G418 selection and G418
selection can be seen in Figure 1-4. The results suggest that the *neo* marker is also subject to TPE, and that cells expressing *neo* are also more likely to express GFP. Looking back at Figure 1-1, the GFP and *neo* are very close to the telomere in DelTel7, suggesting that the state of *neo* expression is closely correlated with that of the GFP reporter.

**No G418 Selection**

![Bar graph showing % of GFP+ and GFP- cells for Tel7KI, DelTel7B3, and DelTel7F1 under No G418 Selection.]

**G418 Selection**

![Bar graph showing % of GFP+ and GFP- cells for Tel7KI, DelTel7B3, and DelTel7F1 under G418 Selection.]

**Figure 1-4. Effect of G418 selection on the fraction of GFP+ cells for Tel7KI and two different DelTel7 ESCs.**

1.5 **Thesis theme and objectives**

The broad objective of my project is to identify factors that are correlated or associated with telomere position effect in our DelTel7 ESCs system. In this thesis, I will be describing the techniques, results, and interpretation of data in our efforts to learn more about telomere position...
effect in ESCs using the DelTel7 and Tel7KI cell lines. Several main topics I will study further include the relationship between TPE and the heterogeneity of gene expression seen in ESCs, the imprinting occurring at DelTel7, and ongoing efforts to develop a system for a modifier screen to identify novel enhancers and suppressors of TPE.
Chapter 2: Materials and Methods

2.1 Tissue Culture

All tissue culture procedures were performed in a Nuaire biosafety cabinet class II type A2 (BSC). All equipment going into the BSC was sprayed with 70% ethanol and wiped down before being used. Pipette tips and Eppendorf tubes were autoclaved and designated for tissue culture. These materials were not used previously in other areas of the lab. Cell cultures were incubated in a Nuaire DH Autoflow CO\textsubscript{2} air-jacketed incubator at 5% CO\textsubscript{2} and 37°C.

2.1.1 Maintenance and culture of embryonic stem cells (ESCs)

ESCs were maintained in embryonic stem cell medium (ESM; Gibco). ESM contained 15% FBS (Hyclone, Premium Multicell), 5% sodium pyruvate (Gibco), 5% MEM non-essential amino acids (Gibco), 5% 2-mercaptoethanol (Sigma), 5% L-Glutamate (Gibco), 5% Pen-Strep (Gibco), and 0.01% LIF (A. Bogutz, Lefebvre Lab).

ESCs were generally thawed from -80°C freezer or liquid nitrogen tank onto a feeder layer of mitomycin C treated primary embryonic fibroblasts (PEFs). Depending on the time of day that the ESCs were thawed, a same day media change could be possible. To equally spread out the cells, plates were moved horizontally from East to West several times and then vertically from North to South several times. Circular motions were discouraged.

ESCs were passaged 1:6 every two days or when the cells’ confluence reached ~70%.

The protocol for passaging is as follows. First, two washes of appropriate amount of PBS to remove any residual FBS found in ESM. Then an approximate volume of trypsin, found in Table 2-1, is added and plates are incubated at 37°C for 5 minutes. The plates are then aggravated a little by physically tapping. Afterwards, depending on the plate size, the cells were collected with adequate amount of ESM in 14 mL tubes. These tubes were then spun at 5000 rpm for 5 minutes.
and the supernatant was aspirated. Pellets were then resuspended in the desired amount of ESM for passaging.

For freezing cells, passaging was completed as above. Pellets were then either resuspended in ESM or 1x freezing medium (FM) depending on the portion of the cells to freeze down. 1xFM was in a ratio of 1:1:8 of DMSO (Sigma): FBS: ESM. 2xFM was in a ratio is 1:1:3 of DMSO: FBS: ESM. Cells being frozen down were then placed in cryovials (Corning Incorporated) and stored in a freezing container with controlled cooling rate (Mr. Frosty, Nalgene) or a Styrofoam box at -80°C to slow the decrease in temperature.

Table 2-1. Approximate amounts of media used for different sizes of plates used in tissue cultures.
The lower amount is used for cell lines that are less demanding like PEFs and recently passaged ESCs whereas the higher amount of media is used when ESCs are close to 50-70% confluence.

<table>
<thead>
<tr>
<th>Plate name (diameter in cm)</th>
<th>Amount of media (mL)</th>
<th>Amount of Trypsin (mL)</th>
<th>Approximate Growth Area (cm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-well/24-well (1.6)</td>
<td>0.5 – 0.8</td>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>Small (3.5)</td>
<td>1.5 - 2</td>
<td>0.5</td>
<td>10</td>
</tr>
<tr>
<td>Medium (6.0)</td>
<td>3 - 4.5</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Large (10.0)</td>
<td>8 - 12</td>
<td>2</td>
<td>60</td>
</tr>
<tr>
<td>Extra Large (15.0)</td>
<td>18 - 30</td>
<td>4</td>
<td>160</td>
</tr>
</tbody>
</table>

Table 2-2. Percentages of solutions used in different medias for different cell types.

<table>
<thead>
<tr>
<th></th>
<th>PEFM (%)</th>
<th>ESM</th>
<th>KSR</th>
<th>KSR+2i</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBS (Hyclone)</td>
<td>15</td>
<td>15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KSR</td>
<td>-</td>
<td>-</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Sodium pyruvate</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MEM non-essential amino acids</td>
<td>-</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>L-Glutamate</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pen-strep</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2-mercaptoethanol</td>
<td>-</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LIF</td>
<td>-</td>
<td>0.01</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>PEFM (%)</td>
<td>ESM</td>
<td>KSR</td>
<td>KSR+2i</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>Insulin</td>
<td>-</td>
<td>-</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>PD0325901</td>
<td>-</td>
<td>-</td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td>CHIR99021</td>
<td>-</td>
<td>-</td>
<td></td>
<td>0.002</td>
</tr>
</tbody>
</table>

### 2.1.2 Maintenance of PEFs

PEFS were maintained in primary embryonic fibroblasts media (PEFM; Gibco) which contains 10% FBS (Hyclone, High Quality Multicell), 5% sodium pyruvate (Gibco), 5% L-Glutamate (Gibco), and 5% Pen-Strep (Gibco). These PEFs were established from E13.5 embryos by Aaron Bogutz in the Lefebvre Lab. PEFs were thawed from liquid nitrogen onto large plates. Passaging PEFs is just like passaging ESCs with the exception that PEFs generally do not grow as fast and so usually a passage of 1:3 is sufficient. A higher ratio of passaging could prevent the PEFs from growing further. PEFs must be treated with 5% mitomycin C (Sigma) to prevent further growing of PEFs and establish ESC feeder layers. Once the PEFs have been inactivated from growing, they can be passaged onto different sized plates if needed and then ESCs can be thawed on them. For mitotic inactivation, mitomycin C inactivation is added to the plates at 5% and the plates are incubated at 37°C for 1.5 to 3 hours followed by several washes in PBS. The plates can be maintained in PEFM or directly in ESM if ESCs are plated immediately.

### 2.2 KSR+2i medium

KSR medium stands for Knock-out Serum Replacement. KSR allows the elimination of Fetal Bovine Serum added to ESM. KSR is used to provide cells with just the minimum amount of materials. KSR+2i replaces FBS with KSR and includes two inhibitors, PD0325901 and CHIR99021. The addition of these two inhibitors has been shown to maintain ESCs.
undifferentiated and push them towards an earlier stage of development like the ICM (called naïve state) as opposed to later development as the pre-epiblast.

### 2.3 FACS and flow cytometry analysis

For FACS and flow cytometry analysis, cells were passaged using the above protocol. A second spin to remove the ESM was performed. Pellets were then resuspended in FACS buffer with 1xPI (Roche). Samples were then placed in a single-cell strainer (Falcon Corning) and given to Justin Wong to sort on a BD FACS Aria or analyzed by the candidate, R. Szeto, by flow cytometry on a BD LSRII-561 analyzer.

After FACS or flow cytometry analysis, the samples were moved to 1.5 mL tubes, spun down at 5000 rpm for 5 minutes, and the supernatant was aspirated. The samples were then snap frozen on dry ice at -80°C or with liquid nitrogen depending on which was available. Tubes were stored in -80°C for further RNA extraction.

### 2.4 Genomic DNA (gDNA) extraction

Lysis buffer can be prepared as a stock beforehand and contained 100mM Tris-HCl at pH 8.5, 5mM EDTA, 200mM NaCl and 0.2% SDS. Proteinase K lysis solution must be prepared fresh and have final concentration of 100μg/mL of proteinase K.

gDNA was prepared from 24-well dishes. Two PBS washes were used to remove residual ESM. This can be done in the BSC or at the lab bench. Then 500μL of proteinase K lysis solution was added for each well. The dish was shaken at room temperature for at least 15 minutes. After 15 minutes, 500μL of isopropanol was added and remained shaking for several hours. The longer the dish shook, the easier it was to see the precipitated gDNA. The white gDNA can directly be picked out using a Pipetman and placed into a 1.5 mL Eppendorf tube. 1mL of 70% ethanol was added for washing and then spun at maximum speed for 10 minutes.
The supernatant was aspirated and 1mL of 70% ethanol was added and spun again. Tubes were left in room temperature to air dry and then 100 – 200 μL of TE was added to resuspend the gDNA. To help with resuspension, the tubes were placed at 50°C for several hours.

2.5 RNA isolation, DNase treatment and cDNA synthesis

Cells were sorted off a large dish (60cm²) on a FACS machine. For RNA isolation, 1 mL (500 μL at a time) of TRIZOL reagent was added carefully in the fumehood. Resuspension of the cells in the solution was done by pipetting up and down vigorously. Once all cells were in suspension, 200 μL of chloroform was added and the tube vortexed for 5 seconds. The cells were then spun at 10 000 rpm for 10 minutes at 4°C. The supernatant was moved to a newly labeled tube and 500 μL of isopropanol was added and inverted to mix. Cells were spun using previously mentioned speed and temperature for 15 minutes and the supernatant was aspirated. Next, 1 mL of 75% ethanol was added and the tube centrifuged at 4°C at 10000 rpm for 5 minutes. The supernatant was aspirated and the samples were left to dry at room temperature for roughly 10 to 15 minutes. Dried samples were resuspended in 50 μL of DEPC-treated dH₂O.

DNase treatment started with adding 10μL of RNA sample to each tube along with a DNase treatment solution. The solution contained 2.5 μL of 10x RQI buffer (Promega), 1 μL of RNase inhibitor (Promega), 1 μL of DNaseI, and 10.5 μL of dH₂O. All tubes were incubated at 37°C for 30 minutes and then 65°C for 15 minutes.

For cDNA synthesis by reverse transcription, M-MLV RT (Invitrogen) was used. First a master mix contained 2 μL of N15 primers, 5 μL of dH₂O, and 2 μL of 10uM dNTPs. 10μL of DNase treated RNA was added to the odd-numbered tubes and were incubated in the PCR machine at 65°C for 10 minutes. Upon completion, the tubes were removed from the PCR machine and chilled on ice. The second master mix contained 8μL of 5x Buffer (Invitrogen), 4
μL 0.1M DTT (Invitrogen), 0.5 μL RNase inhibitor (Promega), and 6.5 μL dH₂O. 19μL of this mix was added to the odd numbered tubes, pipetted up and down to ensure homogenous solution and then 19μL was moved to the even numbered tubes. These tubes are then placed at 42°C for 2 minutes and finally the odd numbered tubes would have 1μL of Reverse Transcriptase (Invitrogen, SuperScript II 200U/μ) added to it.

2.6 RT-qPCR.

Primers for the control gene Ppia were previously described. The primers used for RT-qPCR were designed by the candidate (R. J. Szeto) using the online Primer Express 3.0 program. All primers used in this study are listed in Table 2-4 and were ordered from IDT. RT-qPCR results were analyzed using the LinReg PCR software and 2-ΔΔCT method of determining relative gene expression. Significance was determined by the Student’s t-test (threshold set to p<0.05)

RT-qPCR requires precision and cleanliness so a different lab bench and different set of pipettes were used for these particular experiments. Master mixes included 10X Buffer, 10% Magnesium, 10% dNTPS, 10% EvaGreen, 0.2μL of Tsg (5μ/μL, Bio Basic Canada Inc) per tube and the remainder was dH₂O bringing the volume up to 23μL. 1μL was set aside for the reverse and forward primers and 1 μL was set aside for the cDNA.

Sometimes if the samples were precious or many runs were needed from a particular set of samples, the cDNA was diluted with TE. To decrease primer dimers, the master mix was made without Tsg, tubes with reverse and forward primers were prepared in advance, and everything was placed onto ice. If primer dimers were still a problem, a different approach was used. 5μL of diluted cDNA was added to the strip tubes first and placed on an ice block in an ice
bucket. The master mix with primers would only be a multiple of 20 and a repeat pipetor was used to decrease the amount of time primer dimers could form.

**Table 2-3. Annealing and reading temperatures for genes used in RT-qPCR experiments**

<table>
<thead>
<tr>
<th>Target</th>
<th>Annealing Temperature (°C)</th>
<th>Reading Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ppi</em></td>
<td>58 – 62</td>
<td>81</td>
</tr>
<tr>
<td><em>Dppa3</em></td>
<td>60 / 62</td>
<td>79</td>
</tr>
<tr>
<td><em>Nanog</em></td>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td><em>Pecam</em></td>
<td>60 / 62</td>
<td>81</td>
</tr>
<tr>
<td><em>Zfp42</em></td>
<td>60</td>
<td>79</td>
</tr>
<tr>
<td><em>neo</em></td>
<td>58</td>
<td>82</td>
</tr>
<tr>
<td><em>EGFP</em></td>
<td>58</td>
<td>86</td>
</tr>
</tbody>
</table>

### 2.7 ESC electroporation

ESCs were thawed from -80°C or liquid nitrogen onto mitomycin C-treated PEFs. After a minimum of two passages on gelatin-coated plates to eliminate most of the feeders, the ESCs were ready for electroporation. Approximately 40μg of DNA was used for each electroporation. ESCs were passaged as mentioned previously and resuspended in 500μL of Embryomax electroporation buffer (Millipore). Dilutions of 1:100 and 1:1000 and counting on a hemacytometer were used to determine the concentration of cells from the plates. A final concentration of 7.6x10^6 cells/mL was required. Each electroporation cuvette (0.4 electrode gap size) held 800μL of cells to which the DNA was added. A Gene Pulser X cell electroporator (Biorad) was used, set at a voltage of 250 V and capacitance of 500μF. Cuvettes were flicked and wiped down prior to electroporation. Time constants were written down to determine salt content of mixture. Cuvettes were then placed on ice for roughly 30 minutes to allow the cells to recover.
To place electroporated cells into two large dishes evenly, drops were counted and mimicked onto the other plate. Alternating plates periodically ensured even distribution.

2.8 pPGKzeo and pMGD20zeo construction

The gene for zeocin resistance (Sh ble gene called zeo below) was PCR amplified from pFpuro.1zeopA using primer pairs ZeoF1 and ZeoR1 (Table 2-4). ZeoF1 introduces an XhoI cut site whereas ZeoR1 introduces an EcoRV cut site. The amplified product is roughly 375 bp long. The pPGKPuro plasmid (Addgene #11349) was digested with XhoI and SmaI to remove the Puro marker. The zeo PCR product was cut with XhoI and EcoRV and then ligated into pPGKPuro to produce pPGKzeo. Directional cloning was then used to produce pMGD20zeo.

pPGKZeo and pMGD20neo were both cut with SpeI and SalI to obtain pMGD20zeo, in which the neo marker has been replaced with the zeo marker.

2.9 Neomycin selection, zeocin selection, and picking colonies

Neomycin (G418, Sigma) selection was accomplished over 10 days and at a concentration of 150-200 μg/mL. Each day the media was changed to remove dead cells and other debris. Daily media change also ensured cells had access to LIF to remain undifferentiated. Sometimes one to two PBS washes were used to completely remove the excessive number of dead cells.

Zeocin (Invitrogen) selection was performed over a span of 10 days and at a 25μg/mL concentration where each day, the media was changed. PBS washes were used if large amounts of dead cells were present. At the 9th or 10th day mark, colonies were selected for picking and cultured individually. Colonies were picked using a Leica microscope, trypsinized with 10 μL of trypsin (1% trypsin) in trypsin buffer at 37°C, and placed onto 4- or 24-well dishes containing ESM. Since the ratio of trypsin to ESM was quite small, a spin to remove trypsin was not
necessary. At this point, antibiotic selection was stopped to allow the cells to grow sufficiently.

Next, each line was passaged onto small dishes and frozen down.

### 2.10 Western blotting: SDS-page and immunoblotting

All SDS-page equipment was from BioRad. Cell pellets were collected from a small plate with 80-90% confluence. Samples were boiled twice with roughly 200μL of SDS loading buffer. Between the two boiling incubations, samples were physically disaggregated via pipetting up and down as well as pressing the pellet along the edge of the tube. The MW ladder used contained 5μL of PageRuler pre-stained protein ladder (Thermo) and 10μL of loading buffer. 8% resolving gel and 5% stacking gel was optimal for the large T Ag protein. All solutions were made following standard instructions. Gels were run for approximately 3 hours at 100V and then transferred to a membrane at 4°C at 100V for one hour. 10x Transfer buffer was premade and stored and 1X transfer buffer was made fresh. Membranes (Hypond-P, Amersham Biosciences) were then immediately dried on top of clean Whatman paper, sealed, and stored at 4°C.

The second identical membrane was stained with Coomassie blue stain for 5 minutes at 20% power in the microwave and shaken for 10 minutes. Destaining buffer was then used to remove excess Coomassie blue by shaking for 10-20 minutes two times and then further destained overnight. The next day, the destained gel was immersed in dH2O and shaken for at least an hour before being imaged. AlphaImager was used to produce images to ensure consistency in gel lanes and protein amount.

For immunoblotting, the membrane was removed from 4°C and blocked with 15 mL of blocking solution which contained 5% non-fat milk powder (Safeway) in PBS-T. PBS-T was prepared by the addition of 1mL of Tween20 (Sigma) for every 500mL of PBS. Blocking occurred on a rotator at room temperature for one hour. Washes were completed three times with 5mL of PBS-T at 5 minutes and rotating. All washes will be in the above conditions. Primary
antibody (mAb anti-Polyoma Virus, Large T Antigen Antibody PyLT; Novus) was diluted 1:1000 in 5 mL of 5% non-fat milk powder PBS-T and rotated at 4°C overnight.

The next day, three PBS-T washes were used to remove any unattached primary antibody. Secondary antibody (Anti-mouse Ig horseradish peroxidase linked whole antibody from sheep; GE healthcare) was also diluted 1:1000 and rotated at room temperature for 1 hour followed by three PBS-T washes. Finally, ECL solution (ECL plus Western Blotting Detection System; RPN2132) was added following the manufacturer’s instructions. Membranes were then exposed to film (Thermo Scientific) in a dark room.

2.11 Genotyping

DelTel7-pMGD20zeo ESCs were genotyped using primers used to genotype the DelTel7 mouse. These primer names are I2E, I2A.2 and pAF2 (Table 2-4).^{28}

2.12 Bisulphite sequencing

Bisulphite conversion was completed on gDNA. 0.2N of NaOH was added to gDNA and suspended at 42°C water bath for 30 minutes. Samples were then placed on a heat block set at 100°C for one minute. The working solution of sodium bisulphite solution was made of 1.64 g of NaHSO₃ with 2 mL of dH₂O, followed by 90 μL of 10N NaOH, 200μL of hydroquinone, finally 1.4 mL of dH₂O. 400μL of sodium bisulphite solution was added and vortexed. The samples were placed in a 55°C incubator overnight.

2.13 Immunohistochemistry (IHC)

IHC for PECAM1 was done on 4-well plates and circular microscope slides. The primary antibody was a rat anti-mouse PECAM1/CD31 and the secondary antibody was biotin
sp. goat anti-rat IgG. Cells were fixed with 4% PFA for 30 minutes. Plates were stored at 4°C with 70% ethanol and sealed with Parafilm for long-term storage.

All steps were done at room temperature and three washes of PBS were used between steps and to remove any excess 70% ethanol. The blocking step was accomplished with 0.3% H₂O₂ for 30 minutes followed by PBS-T washing. The second blocking was done with 5% goat normal serum in 0.5% BSA in PBS-T for 30 minutes and also followed with PBS-T washing. Coverslips were flipped upside down and incubated with primary antibody overnight. Parafilm was used as a seal.

The next day, washing with PBS-T was used to remove unattached primary antibodies. Secondary antibody was diluted 1:250 and incubated for 30 minutes. After washing with PBS-T, ABC was added using manufacturer’s instructions (Amersham Biosciences) for 30 minute incubation. PBS-T wash three times and then adding 1% DAB and 3%H₂O₂ into PBS for roughly three minutes. Wash with dH₂O immediately and then staining with Mayer’s hematoxylin. Hematoxylin was prepared with 5% Aluminum Potassium Sulfate (Fisher Scientific), 0.1% Hematoxylin (Fisher Scientific), 0.02% Sodium Iodate (Fisher Scientific) and 2% Glacial Acetic Acid (Fisher Scientific). Wash with three times dH₂O and wash with Scotts Tap Water. Final steps are to dehydrate the coverslip via five minute incubations in 50%, 70%, 90%, 2x 100% Ethanol and Xylene. Entellan (Harleco) was used to adhere coverslip to microscope slide. Slides are stored at room temperature in a cool dark drawer.
Table 2-4. List of primers used in this study

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Sequence</th>
<th>Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning (pMGD20zeo)</td>
<td>GCAGCTCGAGATGGCAAGTTGACCAGTGCCG</td>
<td>ZeoF1</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (Ppia)</td>
<td>GCCGTCTCCTTCGAGCTGTGTTTG</td>
<td>PPIA</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>TGTTAATGCACCACCTTGCGCAGATC</td>
<td>PPIA</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td>RT-qPCR (Nanog)</td>
<td>TCGCCATACACTGACATGAG</td>
<td>Nanog F</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>CAAGAATAGTCTCCGGATGAAAAC</td>
<td>Nanog R</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (Pecam1)</td>
<td>GAGGAAGCAGGCGGCAAAC</td>
<td>Pecam1F</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>TTTCACTAGGCTCAGGATCTTC</td>
<td>Pecam1R</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (Zfp42/Rex1)</td>
<td>TGCTGCCACAGTGGTTC</td>
<td>Rex1F</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>TGCTGCCACAGTGGTTC</td>
<td>Rex1R</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (Dppa3/Stella)</td>
<td>TCTGGTTGTAGGAGCTGATGC</td>
<td>Stella F</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>CAAGAATAGTCTCCGGATGAAAAC</td>
<td>Stella R</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (neo)</td>
<td>GCAAAGCCTCCTCAGCAATATC</td>
<td>NeoqR1</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (EGFP)</td>
<td>AAGCAGAGAGACGGAATCACG</td>
<td>EGFPqF1</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>GGACTGGGTGCTGGATTGAGT</td>
<td>EGFPqR1</td>
<td>This study</td>
</tr>
<tr>
<td>RT-qPCR (large T)</td>
<td>GCCTTAATGCAAGATTGAAAG</td>
<td>LargeT-Ex1F</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>CTCCGTGGCGCGTTCTAG</td>
<td>LargeT-Ex2R</td>
<td>This study</td>
</tr>
<tr>
<td>Bisulphite (KvDMR1)</td>
<td>AAAACCTTTTTCTATCACTATTTTATACAC</td>
<td>Kcnq1ot1 OR</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>GTTTTTAAGATTATTTTTTGTGGATTG</td>
<td>Kcnq1ot1 IF</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>AATTCTCTAAATATAATTTTTTCTCAAC</td>
<td>Kcnq1ot1 IR</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td>Genotyping (DelTel7)</td>
<td>TGTCCCACTGGCAGCGTCACT</td>
<td>I2E</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>GTCAAGGCTTTAGGTGCTTAGG</td>
<td>I2A2</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>GGAGGATTGGGAAGAAATAGC</td>
<td>pAF2</td>
<td>Oh et al&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
Chapter 3: **Relationship Between ESC Heterogeneity and TPE**

### 3.1 Introduction

Many studies have shown that ESCs grown in serum show heterogenous expression of several genes like *Dppa3*, *Rex1*, and *Oct4*. There have been several theories as to why this heterogenous expression occurs. The first theory is that the heterogeneity stems from the fact that ESCs fluctuate between different states and range from the 2-cell stage to the morula stage to the pre-epiblast stage. All these cells stages are considered ESCs because they are pluripotent and are able to make an entire organism. With this theory in mind, it means that the heterogeneity seen in gene expression could be correlated to slight differences in specific stages of the embryo. Chambers *et al* (2007) showed that *Nanog* positive cells are biased towards producing undifferentiated populations whereas *Nanog* negative cells are more likely to yield differentiated populations. This observation leads to the theory that to maintain ESCs, heterogeneity is required. Heterogeneity is important to maintain the balance between self-renewal and differentiation.

The heterogenous expression of key developmental genes seen in undifferentiated ESCs looks similar to the stochastic expression of GFP observed in DelTel7 cells. In light of this similar variegated expression, I wanted to see if there was any correlation between the heterogenous expression of genes and the stochastic expression of GFP in DelTel7.

The four genes I focused on are *Dppa3*, *Nanog*, *Pecam1*, and *Zfp42*. These genes were chosen because they have all been shown to be involved in early development, pluripotency, and are highly characterized. These four genes are not found on mouse chromosome 7 so they have no physical linkage to the GFP reporter on the DelTel7 allele. *Dppa3/Pgc7/Stella* is found on mouse chromosome 6 and is highly expressed in the early embryo. *Dppa3* is associated with
pluripotency and germ cells. Dppa3-positive ESCs have been observed to resemble ICM compared to Dppa3-negative cells resembling epiblast-like cells\textsuperscript{35,37}. As a marker of germ line lineage, Dppa3 is first detected in preimplantation embryos\textsuperscript{37}. Dppa3 is an essential gene as a maternal factor for the mouse cleavage stages in embryogenesis as well as ensuring the maternal genome does not oxidate from methylated cytosines (5mC) to hydroxymethylated cytosines (5hmC)\textsuperscript{38}. This gene is poorly conserved across mammalian species as there is only 32% similarity between mice, cows, and humans\textsuperscript{39}. In bovine studies, it has also been shown that Dppa3 is associated with preventing cytosine hydroxymethylation of the maternal genome\textsuperscript{39}.

Nanog is also found on mouse chromosome 6 and is a marker for pluripotency\textsuperscript{40,41}. It has been shown to be highly expressed in the inner cell mass (ICM) compared to the epiblast\textsuperscript{42}. Therefore, Nanog expression is required in early development compared to later development. Nanog null embryos do not develop pass the preimplantation stage, however it has been shown that Nanog is not required to maintain pluripotency for ESCs in an ex vivo environment\textsuperscript{43,44}.

Pecam1 or Cd31 is found on mouse chromosome 11 and codes for a 130-120 kD integral membrane glycoprotein\textsuperscript{43}. This protein can be found on the surfaces of platelets, endothelial cell boundaries, leukocytes, and myeloid lineage cells\textsuperscript{45,46,47}. Using confocal microscopy, PECAM1 is first detected in the ICM of the mouse blastocyst\textsuperscript{48}. The function for Pecam1 prior to vascularization is unclear. PECAM-1 deficient mice exhibited defects in leukocyte trans-basement membrane migration but they were viable\textsuperscript{49}. Pecam1 expression shows heterogeneity in undifferentiated ESCs\textsuperscript{49}.

Zfp42/Rex1 is found on mouse chromosome 8 and is a common gene used to determine pluripotency\textsuperscript{50,51}. Zfp42 is strongly expressed in the ICM and has significantly decreased expression in the epiblast\textsuperscript{50,51}. Consequently, the observation that Zfp42 is heterogeneously
expressed in ESCs reflects the hypothesis that ESCs fluctuate from a more naïve to a more
developmentally advanced state\textsuperscript{52}. \textit{Ppia} is a housekeeping gene that was identified as one of the
most stably expressed reference genes for expression studies by RT-qPCR during mouse
development\textsuperscript{31}. It is used here as a control to normalize our measurements for experimental
genres.

KSR+2i medium includes two inhibitors to push ESCs into a more ICM-like stage as opposed to the pre-epiblast stage. The two inhibitors are PD0325901 and CHIR99021.

PD0325901 is derived from a MEK inhibitor CI-1040. This organic molecule binds to MEK and inhibits the phosphorylation and activation of MAPK/ERK pathway\textsuperscript{53}. CHIR99021 is an aminopyrimidine derivative capable of activating the WNT pathway by inhibiting glycogen synthase kinase (GSK3α/β), thus increasing cytosolic β-catenin levels\textsuperscript{54}.

3.2 Results

3.2.1 Purification of ESC populations by FACS

After sorting DelTel7 cells into GFP\textsuperscript{+} and GFP\textsuperscript{−} populations, the first experiment I conducted was a RT-qPCR for the EGFP reporter itself to determine the purity of the sorted ESC populations. From Figure 3-1, it is clear that the sorting went well and generated very pure populations of GFP\textsuperscript{+} and GFP\textsuperscript{−} ESCs on the basis of both GFP fluorescence and \textit{EGFP} mRNA levels. Next, I looked at expression of the \textit{neo} marker in the same two populations. Previously it was shown that GFP expression follows \textit{neo} expression in the DelTel7 and upon G418 selection to enrich for neo+ cells. For the first time, I was able to ask here whether \textit{neomycin} and GFP expression are similarly correlated, without selection, in sorted cells. The results show that \textit{neo} expression does follow GFP expression since the GFP\textsuperscript{+} population has a higher \textit{neo} expression
than the GFP− population. In other words, when GFP is transcribed and turned on, neo is also transcribed and turned on from the DelTel7 allele in the same cell population. The distance between these two genes is only 2.6 kb (Figure 1-1), suggesting that they are similarly affected by TPE.

**Figure 3-1. RT-qPCR analysis of EGFP and neo on sorted GFP+ and GFP DelTel7 cells.**

Expression data is presented for the individual subclones (left) and for the average, comparing GFP+ and GFP− populations. Values presented are mean±SD; *p<0.05. Relative expression level normalized to the house keeping gene Ppia.

### 3.2.2 Dppa3, Nanog, Pecam1 and Zfp42 expression in sorted DelTel7 ESC populations

RT-qPCR was used to determine if there is any correlation between Dppa3, Nanog, Pecam1, and Zfp42 expression in relation to GFP expression from the DelTel7 allele. At the individual subclone levels, the results for Dppa3 suggest that the GFP+ population expresses
more *Dppa3* than the GFP population (Figure 3-2). However, looking at the average of the three subclones, there is an increase from 1 to 1.5 in *Dppa3* expression between GFP and GFP$^+$ populations, but this difference is not significant ($p = 0.05$). Looking into *Nanog* expression, at the individual clone level it looks like B3 and F1 have slightly lower expression in GFP$^+$ population while C3 has a higher level of *Nanog* expression in GFP$^+$ population. This data counteracts each other such that the average expression of *Nanog* is not significantly different in the two populations. For *Pecam1* and *Zfp42* expression, subclones B3 and C3 appear to have more expression in the GFP$^+$ population compared to GFP$^-$ cells, but in subclone F1, the data is the exact opposite. Here again, the differences in both *Pecam1* and *Zfp42* expression levels between average subclones are insignificant. Based on these results, I have observed no relationship between these heterogeneously expressed genes and the stochastic expression of GFP at the DelTel7 allele.
Figure 3-2. RT-qPCR analysis *Dppa3*, *Nanog*, *Pecam1*, and *Zfp42* on sorted GFP*+* and GFP*−* DelTel7 cells.
On the left, the data presented is for the individual subclones and the data shown on the right is the average for the three subclones, comparing GFP*+* and GFP*−* populations. Values presented are mean±SD. Expression levels are normalized to *Ppia*, a house keeping control gene.
3.2.3 PECAM1 expression

Another approach to assess the relationship between TPE and heterogeneity is using IHC or flow analysis on DelTel7 ESCs using anti-PECAM1 antibodies. IHC would allow us to see visually any stochastic expression of PECAM1. However, IHC does not allow us to document GFP expression simultaneously with quantitative precision. This is where flow analysis can be used to determine any association between GFP expression and PECAM1 levels. I first completed some IHC for PECAM1 on DelTel7 ESCs and determined a good dilution of antibody to use. However, when I attempted to use flow cytometry to analyze whether GFP expression and PECAM1 expression were correlated, I obtained inconclusive results. There were not a lot of PECAM1+ cells. So far, I have concluded that this PECAM1 staining did not work due to a variety of reasons, for instance problems with our antibody.

3.2.4 KSR+2i treatment on DelTel7/+ and Tel7KI/+ ESCs

Previously, it has been observed that ESCs grown in serum represent a heterogeneous mixture of pluripotent cells ranging from the 2-cell stage to the ICM of the blastocyst at one end to an early and epiblast-like cell at the other. Based on these observations, other studies have identified specific small molecules that force ESCs more towards an ICM- state found in the blastocyst. The ICM state, referred to as naïve or ground state, is different from the epiblast in the sense that it represents an earlier stage in development. ESCs in this naïve state were shown to express more Oct4, Nanog, and Zpf42. On the other hand, the epiblast has no expression of Oct4, and reduced expression of Nanog and Zfp42. Conversely, lineage commitment markers such as Sox1 and T/Brachyury are expressed in the epiblast. The KSR+2i growth medium is based on a serum replacement minimal medium containing LIF and the two inhibitors described
above (KSR+2i). Growth of ESCs in this medium keeps ESCs undifferentiated while driving them towards the ground state.

CHIR99021, a GSK3α/β inhibitor, increases ESC density in culture and also arrests their neural differentiation\(^55\). A double knock-out of Gsk3a and Gsk3b has a similar phenotype as the addition of CHIR99021\(^55\). The addition of PD0325901 inhibits differentiation and allows cells to grow continuously\(^54\). In the context of DelTel7, I wanted to see whether there would be a change in GFP expression by forcing cells to become more homogenous and “naïve”. 
3.2.4.1 Confirming ESC ground state using Nanog as a marker

To confirm that our KSR+2i culture condition is indeed pushing cells towards an earlier ICM-like state, I completed a quick RT-qPCR experiment on Nanog expression. Nanog is highly expressed in ICM-like state cells compared to epiblast cells or later developmental stages. We can see from Figure 3-3 that the Nanog expression levels in cells grown in KSR+2i are higher than those seen for cells grown in regular ESM with serum.

![Figure 3-3. RT-qPCR analysis of Nanog expression in DelTel7 B3 and C3 and Tel7KI ESCs grown in different media.](image)

Relative expression of Nanog in DelTel7 B3, DelTel7 C3 and Tel7KI grow in ESM is 0.175 and in KSR is 0.16. Upon KSR+2i treatment, the expression level is almost increased by 3-fold compared to normal ESM conditions. RT-qPCR conditions specifications can be found in Section 2.6 in Materials and Methods. Values presented are mean±SD. Relative expression is normalized to the house keeping gene Ppia.
3.2.4.2 GFP expression using flow cytometry analysis

First, observation of GFP expression by fluorescence microscopy shows that growth in the KSR+2i medium leads to more GFP-expressing DelTel7 ESCs compared to cells grown in ESM.

<table>
<thead>
<tr>
<th></th>
<th>ESM</th>
<th>KSR</th>
<th>KSR + 2i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlay</td>
<td>![Overlay Image]</td>
<td>![Overlay Image]</td>
<td>![Overlay Image]</td>
</tr>
<tr>
<td>Fluorescent</td>
<td>![Fluorescent Image]</td>
<td>![Fluorescent Image]</td>
<td>![Fluorescent Image]</td>
</tr>
</tbody>
</table>

Figure 3-4. Fluorescent microscopy analysis of DelTel7 ESCs maintained in ESM, KSR, or KSR+2i.

The same cell populations of ESCs were also analyzed by flow cytometry (Figure 3-5). When DelTel7 ESCs were grown in KSR+2i medium, both the percentage of GFP-expressing cells as well as the GFP levels increased. This means that as ESCs are forced to be more ICM-like, there is an increase in the GFP⁺ population. Interestingly, when the control cell line Tel7KI with the interstitial CAG-GFP reporter was grown in KSR+2i, we also saw a striking increase in GFP expression levels, by approximately 10-fold. We already know that Tel7KI ESCs always
show robust and ubiquitous GFP fluorescence so the increase in GFP levels suggest KSR+2i has an effect on the transcriptional activity of the CAG promoter/enhancer combination driving GFP expression in both the Tel7KI and DelTel7 alleles. This hypothesis is consistent with the observation that for Tel7KI growth in KSR+2i does not change the fraction of GFP+ cells but rather leads to a higher expression level, as seen by the shift of the GFP peak to the right (Figure 3-5). For DelTel7, this result leads us to believe that through its effect on the expression level of the GFP reporter, growth in KSR+2i essentially acts as a suppressor of TPE by allowing the GFP reporter to compete with telomeric silencing.

Figure 3-5. Effects of KSR + 2i on GFP expression for DelTel7B3, C3, F1, and Tel7KI ESCs.
This data was generated using flow cytometry analysis on DelTel7 B3, C3, F1 and Tel7KI ESCs after growth in ESM, KSR, or KSR+2i. * refers to increase of GFP levels in individual cells.
3.2.4.3  neo expression in KSR+2i treated ESCs

To determine if the GFP expression is a promoter/enhancer effect, I also looked at the expression of the neighbouring neo marker. Looking at the expression of neo is important because the GFP is driven by a CAG promoter/enhancer whereas neo is expressed under the endogenous mouse Pgk1 promoter. Upon growth in KSR+2i, we do not see an increase in neo expression in the DelTel7 clones (Figure 3-6). These results support the idea that the increase seen in GFP expression levels in the KSR+2i treatment is due to a promoter/enhancer effect acting on the CAG-GFP reporter.

Figure 3-6. RT-qPCR of neo expression in DelTel7 B3, C3, F1 and Tel7KI ESCs grown in different media.

Image on the left is the original data and image on the right is scaled down to 0.1 to better visualize the data for the DelTel7 clones. Values presented are mean±SD. Relative expression is normalized to the house keeping gene Ppia.
Chapter 4: **Imprinting Effects at the DelTel7 Allele**

4.1 **Introduction**

DelTel7 was originally produced to study the relationship between the IC1 and IC2 imprinted domains. Imprinted genes are characterized by expression differences between the paternal and maternal allele. To analyze imprinting effects at DelTel7 allele, breeding pairs were set up to determine whether there was a difference in GFP expression in embryos upon paternal and maternal transmission of DelTel7. Looking at E7.5 embryos, we noticed that GFP expression is high when Tel7KI is inherited maternally, but silenced when paternally transmitted, as we previously reported\(^56\). However, the more interesting result is that maternally inherited DelTel7 shows more GFP expression than paternally inherited DelTel7. This would suggest that *in vivo,*

---

**Figure 4-1.** GFP expression of +/Tel7KI, Tel7KI/+, +/DelTel7, and DelTel7/+ E7.5 embryos.
DelTel7 is also imprinted and maternally expressed, which is the same as Tel7KI. Alternatively, this result could suggest that TPE is more pronounced on a paternally inherited allele.

4.2 Results

4.2.1 Is imprinting due to inheritance of DelTel7 allele (maternally or paternally) or germline differences in TPE?

DelTel7 was originally designed to assess the relationship between the epigenetic functions of the imprinting centers IC1 and IC2 on chromosome 7. With that in mind, there were several questions raised when DelTel7 mice was successfully derived from the mutant ESCs. One of these questions is whether GFP was imprinted since it is now in an imprinting region. Embryos were observed for their GFP expression when DelTel7 was inherited paternally and maternally. As shown in Figure 4-1, it is clear that GFP is imprinted in vivo. When DelTel7 is inherited paternally, GFP is generally turned off whereas maternally inherited DelTel7 shows more GFP expression. To be clear, not all GFP expression is gone in paternally inherited DelTel7. There are two areas that have GFP expression: the heart and gonads. Why this expression pattern occurs has yet to be researched thoroughly. Several theories as to why this imprinting occurs have emerged. The first theory is that there is a difference in paternally and maternally inherited DelTel7 (imprinting effect) and the second theory is that there is a difference between the maternal and paternal germlines with respect to telomere structure and TPE (TPE effect).
4.2.2 The H19 gene adjacent to DelTel7 is not affected by TPE

Since GFP expression from the DelTel7 allele is affected by TPE, I wanted to determine how far this silencing can spread from the DelTel7 allele. This was assessed using RT-qPCR to measure the expression of the neighbouring Igf2 and H19 genes in sorted GFP+ and GFP- DelTel7 ESCs (Figure 1-1). Ins2 expression was not detected because it is not expressed in ESCs (data not shown). Since the DelTel7 ESC line that I am working with has the DelTel7 allele on the maternal allele, we could only expect certain results. The predicted model can be found in Figure 4-2.

![Figure 4-2. Illustration of how imprinting of GFP reporter can affect later downstream genes.](image-url)
$H19$ is expressed from the maternal allele while $Igf2$ is expressed from the paternal allele. If DelTel7 and TPE affect those genes in cis, $Igf2$ will have regular expression from the paternal allele. However for $H19$, TPE could affect the expression from the maternal allele since DelTel7 is present on the maternal chromosome 7 in our ESCs. If the TPE effect has spread from DelTel7 to $H19$, then we expect positively correlated expression from GFP and $H19$, as we saw for the $neo$ marker. However, my RT-qPCR results clearly showed that $H19$ is not affected by TPE (Figure 4-3 and Figure 4-4).

![Figure 4-3. RT-qPCR of H19 mRNA expression on DelTel7 B3, C3, and F1 subclones.](image1)

Individual subclones can be seen on the left and average subclones can be seen on the right. Differences in $H19$ levels between GFP$^+$ and GFP$^-$ cells are not statistically significantly. Values presented are mean±SD. Relative expression levels normalized to the housekeeping gene $Ppia$.

![Figure 4-4. RT-qPCR of Igf2 mRNA expression on DelTel7 B3, C3, F1 subclones.](image2)

The original data is on the left and scaled data is on the right to see the wide variety of expression. Differences in Igf2 levels between GFP$^+$ and GFP$^-$ cells are not statistically significantly. Values presented are mean±SD. Relative expression levels normalized to the housekeeping gene $Ppia$. 

43
4.2.3 Differences in culturing DelTel7/+ and +/DelTel7 ESCs

To answer this question from another angle, we obtained reciprocal ESC lines carrying a maternal or paternal DelTel7 allele. We looked at the expression of GFP in ESCs that inherited DelTel7 maternally (DelTel7/FVB; MH-DF) and paternally (FVB/DelTel7; MH-FD).

Upon receiving the ESCs from our collaborator Dr. Michael Higgins, the morphology of the MH-DF and MH-FD lines were already noted as different. The MH-FD line looked more ESC-like which included round or oval colonies where the individual cells of each colony are not clearly visible. On the other hand, the MH-DF line had more of a spikier or differentiated look where the individual cells of colonies were a little more visible. Growing and expanding the MH-DF line was also a lot more difficult because of the slow growth and differentiated morphology of this particular line.

4.2.4 Flow cytometry analysis and bisulphite sequencing on DelTel7/+ and +/DelTel7 ESCs

Next, we looked at the GFP expression under the microscope and also completed some flow cytometry analysis to quantify the GFP expression. We expected that the MH-DF line would show more GFP expression since the DelTel7/+ embryos exhibit more GFP fluorescence than +/DelTel7 embryos. However, this was not the case. Looking at the microscope pictures in Figure 4-5, it was very obvious that the MH-FD line had more GFP expressing cells. To confirm, I did flow cytometry analysis to look at more cells and to quantify the level of GFP (Figure 4-6). From these results, we can clearly state that MH-FD line had more GFP expressing cells than the MH-DF line, which contradicts the results seen in vivo.
Figure 4-5. GFP expression of MH-DF and MH-FD DelTel7 ESCs under a microscope.
MH-FD has several colonies that are fluorescing green whereas MF-DF only has a one very green colony.

Figure 4-6. Flow Analysis of MH-FD and MH-DF line.
DelTel7 C3 is used as a control. MH-FD has more GFP expressing cells than MH-DF.
With the newly contradicting results, I decided to do some more analysis on these cell lines. The easiest explanation would be that an error occurred in labeling during the derivation, expanding, freezing, and thawing of these two lines. Normal genotyping, using PCR, would not have worked since these two cell lines are genetically identical. The difference in these lines is in their epigenetics since DelTel7 was inherited maternally and paternally. I took advantage of this difference and used bisulphite sequencing at an imprinted location to determine the genotype. We used primers for the *Kcnq1ot1* promoter (defining IC2) which is an imprinted gene DNA methylated on the maternal allele and expressed from the paternal allele. This region is deleted in DelTel7, such that heterozygous cells carry a single epigenetic variant of IC2. We expected DelTel7 and MH-DF to look similar (unmethylated IC2) and MH-FD to look different (methylated IC2). The results from the sequencing showed all cell lines as the same. We concluded that the two MH cell lines have inherited DelTel7 maternally and cannot be used for our studies.
Figure 4-7. Ape alignment of bisulphite sequencing at IC2 from the reference sequence, DelTel7, MH-DF and MH-FD ESCs.

All sequences were placed into an application called Ape and sequences were aligned to the reference genomic IC2 sequence. The first sequence is the reference sequence followed by Ape alignment analysis of the reference sequence to DelTel7, MH-DF and MH-FD. The red highlighted text identifies areas where there is a mismatch with the reference sequence, which all corresponds to unmethylated CpG sites.
Chapter 5: Screen for Modifiers of TPE: Development of an Episomal System

5.1 Introduction

Amongst others, polyoma virus contains a small T, middle T, and a large T antigen gene. The host of polyoma virus (Py) is mouse, monkey for simian virus (SV), and lastly humans for bk\textsuperscript{58}. Py and SV are similar in mRNA alignments\textsuperscript{58}. Some similarities between BK and SV40 are homology in early coding region, similar organization of three late proteins and termini homology\textsuperscript{58}. Some differences include larger gaps and high repeats in the early and late regions of SV40 as well as unique sequences in the middle and Large T antigen of Py\textsuperscript{58}.

There are several Large T antigens of interest in cancer research. The one that I will be using is known as large T. Other labs have used large T in several methods like maintaining vectors extrachromosomally as episomes in pancreatic endothelial cells and mouse embryonic stem cells, as well as synthetic lethality screens in PEFs\textsuperscript{59, 60}. 
5.1.1 Screening for modifiers of TPE

Previous studies have shown that integrating a large T antigen-expressing vector into the genome of ESCs allows other plasmids containing a polyoma origin of replication to be maintained as episomes in those cells\textsuperscript{60,61}. I will be taking advantage of this distinct feature and combining it with the DelTel7 ESCs to screen for modifiers of TPE. The strategy I wanted to develop can be found in Figure 5-1.

![Diagram showing the strategy for modifier screen of enhancers and suppressors of TPE using DelTel7 and pMGD20zeo.]

Figure 5-1. Strategy for Modifier screen of enhancers and suppressors of TPE using DelTel7 and pMGD20zeo.

This is the strategy that will be used to screen for modifiers of TPE. Note that a library construct is the ultimate goal for this particular screen.

5.1.2 Proof of principle

No other lab has attempted to use the large T to study modifiers of TPE so I attempted to produce the new subclone of DelTel7 and conducted control experiments to ensure that this
episomal system can be used to screen for modifiers of TPE. The first goal was to use specific cDNA-expressing plasmids, such as *Trf1* and *Trf2*, to confirm their known effects on TPE.\(^{62}\)

Large T antigen is a gene that has been studied extensively in cancer research. The particular variant that I will be using is the large T from pMGD20neo\(^{63}\.\) This particular plasmid carries a deletion to exclude the small and middle T to render it non-transforming. The particular strategy I will be using as a proof of principle to determine whether a modifier screen is possible can be found in Figure 5-2. This strategy will be a smaller scale screen to show that it is possible to select for cells that have had TPE affected and to recover plasmids that contain genes known to perturb TPE.

![Diagram of Large T Antigen strategy](image)

**Figure 5-2. A modified strategy for the modifier screen in DelTel7 ESCs.**

Known modifiers, *Trf1* and *Trf2*-expressing plasmids, will be used to determine the efficiency of this particular method prior to tackling a library construct as described previously.
5.2 Results

Firstly, I obtained pMGD20neo and electroporated the plasmid into DelTel7C3 ESCs. The electroporated ESCs did not die very fast and after 10 days of G418 selection, there was still a lot of living cells on the plate. The colonies were not pickable as they were touching each other. The \textit{Pgk-neo} marker already on the DelTel7 allele may have been creating an underlying resistance to G418 so I decided to change the antibiotic marker present on the vector. After some careful selection, I decided to use the \textit{zeocin} selection system because of the availability of the gene sequence and the drug in our laboratory.

There were several strategies used to produce pMGD20zeo, but the one that worked included PCR amplification of the \textit{zeo} ORF and adding two restriction enzyme sites at the end. This fragment was approximately 375 bps long and was ligated into pPGK-Puro to generate a \textit{Pgk-zeo} fusion. From here, directional cloning was used to replace the \textit{Pgk-neo} of pMGD20neo with \textit{Pgk-zeo} to produce pMGD20zeo (Section 2.8).

DelTel7C3 ESC line was selected to be electroporated based on its morphology and GFP profile. pMGD20zeo was electroporated into DelTel7C3 ESCs in a circular and linear fashion. Both vectors were used to see if there might be differences between large T being maintained in a circular vector or a linear vector. Gassman \textit{et al} showed that the circular plasmid could potentially be lost after subsequent passages under normal growth conditions without selection\textsuperscript{63}. Because of this possible loss of large T, I decided to maintain all the subclones, circular or linear, in \textit{zeo} selection.
Figure 5-3. GFP expression of DeTel7-C3 and several DelTel7C3-pMGD20zeo ESCs.
The very first control experiment that was conducted on these DelTel7C3-pMGD20zeo clones was to see whether their GFP profile matched DelTel7C3. Microscope pictures of these particular transfected ESCs can be seen in Figure 5-3. Flow cytometry analysis was used for a more quantitative approach. The results of flow cytometry analysis of a few subclones were peculiar so I analyzed a few more subclones at two different time points (Figure 5-4). The results were unsettling because there was quite a bit of variation in the percentage of GFP\(^+\) cells between subclones. Looking at the data, there seems to be three main phenotypes: clones with high, medium, or low numbers of GFP\(^+\) cells. The high GFP expressing line pMD20zeo-A4 ranged from 60-80\% GFP\(^+\) cells, while medium GFP expressing lines ranged from 30-60\% GFP\(^+\). The medium GFP expressing lines were A5, A6, B2, C2, D2, D4, and D5. The low GFP expressing lines had below 5\% GFP\(^+\) and were A1, A2, B1, B5, B6, C1 and D3. Interestingly, there were no cell lines that ranged from 5-30\% GFP\(^+\) cells, as seen for DelTel7.
Figure 5-4. Percentage of GFP⁺ DelTel7 and pMGD20zeo ESCs analyzed by flow cytometry.

DelTel7 has roughly 15% GFP⁺ cells while pMGD20zeo cells have a huge variation in GFP expression. There seems to be three categories: High, medium, and very low percentage of GFP-expressing cells.
I took a closer look at the GFP expression by conducting a time course on five subclones over a span of 12 days. This time course would allow us to see whether the GFP profile of subclones would change overall in 12 days and how much day by day. As seen in Figure 5-5, GFP variation was still present. There was also a trend of obtaining more ESCs expressing GFP throughout the time course under zeocin seledction.

**Figure 5-5. Time course done on DelTel7 and selected pMGD20zeo cells.**

For several of the pMGD20zeo clones, there is a trend towards an increase in GFP-expressing cells.

There are several possible explanations for those particular findings. The first is that the large T has a profound effect on DelTel7 and is somehow affecting GFP expression by perturbing the telomere. If this was true, then the screen for modifiers objective of my project would not work using this system, even if we would have found an aspect that disrupts TPE. Another possibility is that the TPE phenotype of DelTel7 ESCs is lost upon subcloning. The last
possibility is that maybe zeocin itself is having an effect and causing the observed variations in GFP.

To determine whether large T was the culprit, I electroporated pPGKzeo into DelTel7C3 ESCs. As previously mentioned, pPGKzeo was the plasmid used to introduce the zeo gene into pMGD20neo. The main difference between pPGKzeo and pMGD20zeo is that the large T antigen is absent in pPGKzeo. I picked and analyzed several DelTel7C3-pPGKzeo clones using flow cytometry analysis and observed that the variation in GFP was still present (Figure 5-6). This result was unexpected since this was a control plasmid. Looking closer at the time course for pPGKzeo-containing clones, there seems to be three categories that emerge. The first is the medium GFP expressing cells ranging from 40-65%. These cell lines are DelTel7C3-pPGKzeo D1, G2 and I1. Another category is low expressing GFP ranging from 10-30% and are DelTel7C3-pPGKzeo D3, D4, F1, F2, F4, H1, H3, J2, and I3. DelTel7C3-pPGKzeo D2, E1, E2, E3, E4, F3, G1, G3, H2, H4, I2, I4, J1, J3, and J4 are in the last category: very low GFP expressing from 0-5%. Another observation is that a majority of the subclones do not express a lot of GFP and essentially each constitutes a GFP population.
Next, I tested the possibility that zeocin itself was causing the clonal GFP variations observed. I looked at ESCs maintained without zeocin. Through a 12-day time course, the wide range of GFP variation was not seen and DelTel7C3-pPGKzeo ESCs had a GFP profile closer to DelTel7 when cells were grown without zeocin (Figure 5-7).
To confirm that these results would also be applicable to the DelTel7-pMGD20zeo ESCs, I also conducted a 12-day time course on these particular cells (Figure 5-8). I maintained the cells under zeocin selection until I was ready to begin my first day of analysis. With this particular approach, I could also see the effects of removing zeocin and how long it would take for the GFP profile to return back to normal.

**Figure 5-7. Time course of DelTel7-pPGKzeo without zeocin selection.**
Figure 5-8. Time course of DelTel7-pMGD20zeo without any zeocin selection

Western blots were used to determine whether the Large T antigen was present in the electroporated DelTel7 ESCs. Through my many attempts, the results were inconclusive. The first obstacle was ensuring that the protein ran from every sample was similar to each other since confluency of some cultures were not exactly identical to each other. While attempting to optimize the biotinylated reporter, the lab needed to order a new kit which caused a big difference in visualizing the results.

RT-qPCR was also another technique that could be used to confirm whether the Large T antigen construct was expressed or not in my ESC clones. The PCR product that was amplified is in the Large T Antigen. 293T cells, DelTel7C3-pMGD20zeo-A4, and DelTel7C3-pMGD20zeo-
D2 were used as positive cells since we knew 293T contains Large T Antigen and the other two cell lines most likely did as well. DelTel7C3 was used as a negative control. We can see in Figure 5-9 that the positive control and negative controls followed the expected patterns of Large T Antigen expression and no Large T Antigen expression, respectively. Using primers to amplify for Large T antigen, I was able to see that Large T antigen is indeed found in the DelTel7C3-pMGD20zeo-A4

![Large T Antigen](chart.png)

**Figure 5-9. RT-qPCR for Large T Antigen on 293T cells DelTel7C3 ESCs and DelTel7C3-pMGD20zeo clones.**

293T, DelTel7C3-pMGD20zeo-A4 and DelTel7C3-pMGD20zeo-D2 were used as positive controls. DelTel7C3 was used as a negative control. Values presented are mean±SD. Relative expression levels normalized to the housekeeping gene *Ppia*.

Overall, the data presented in the episomal system shows that the system has normal DelTel7 behaviour and is ready to be tested using known modifiers of TPE. The other aspect that should be validated before moving forward is the plasmid recovery of the episomes which I was fortunate enough to try but was unsuccessful.
Chapter 6: Discussion and Future Directions

6.1 Correlation between heterogeneity in ESCs and GFP variegation in DelTel7

Based on the RT-qPCR results of Nanog, Pecam1, Rex1 and Stella expression levels in GFP+ and GFP- populations, we clearly see that there is no correlation between these four heterogeneously expressed genes and the stochastic expression seen in GFP (Figure 3-2). My results show that the TPE heterogeneity observed in ESCs is not correlated to the heterogeneity existing in undifferentiated ESCs grown in serum. The TPE effect occurs in both cell populations, as an additional layer of epigenetic variability. On the other hand, GFP and neo expression are positively correlated so when one is turned on, the other is most likely turned on as well (Figure 3-1). The same can be said when one is repressed, the other is also repressed. This positive correlation was expected since we know that the GFP reporter and neo marker are very close to one another on the plasmid used to make DelTel7. These results are consistent with what has been observed in the yeast model.

6.2 Interpretation of KSR+2i treatment on DelTel7 ESCs

For the KSR+2i experiments, we were able to see that placing DelTel7 ESCs into KSR+2i caused more cells to fluoresce green. These results were interesting but it was even more interesting when we saw that the control ESCs with the interstitial GFP insertion, Tel7KI also had a similar response to KSR+2i culture conditions. The Tel7KI allele is used here as the control for DelTel7 because there is a normal chromosome 7 with the addition of the GFP and neo constructs inside the chromosome. Tel7KI ESCs are always green but placing them in KSR+2i caused the cells to express the EGFP reporter at higher levels. This can be seen in Figure 3-5 since the peak for the GFP shifts over to the right. Since Tel7KI cells responded
similarly to KSR+2i, we conclude that it is not affecting TPE *per se*, but rather is having an effect on the CAG-EGFP expression cassette instead. Results for a promoter or enhancer effect can be seen in Figure 3-6 since neo is driven by a different promoter than *EGFP*. Our model is that by increasing the activity of CAG promoter, KSR+2i allows DelTel7 to compete with TPE. This competition can also be seen in yeast galactose induction systems where the GAL1-10 promoter is used to induce expression of GAL inducible products^64^.

Bock *et al* showed that different sources of CHIR99021 had varying composition and purity levels^65^. The difference sources of CHIR99021 were able to cause noticeable aneuploidy in the rat ESCs. Upon further investigation, mass spectrometry analysis revealed unexpected small molecules, which may or may not directly influence chromosome stability^65^. Although purity of the 2 inhibitors were not confirmed in my studies, no evidence pointed towards chromosome stability, but a simple metaphase spread on the treated cells could be easily done to determine chromosome stability. A study in porcine induced pluripotent stem cells (iPSCs) treated the cells with the 2 inhibitors and found a decrease in pluripotent markers like OCT4, SOX2, ZFP42 and DPPA3^66^.

Only RNA expression of Nanog was looked at in this experiment, but further analysis of other pluripotency markers would further confirm that KSR+2i treatment is pushing cells towards a “naïve” state.

### 6.3 Expression of *Igf2* and *H19* in DelTel7 sorted population

Imprinted genes undergo a series of events when an embryo is being formed. First there is an erasure process where methylation and histone modification from the parental genome are erased^67^. This step is important to reset the genome to reflect the sex of the germline. Following erasure in the germ line, establishment of the DMR is required. The DMRs mark regions where
imprinted genes are clustered and allow for differential regulation of the imprinted genes on the two parental alleles.\footnote{67}

In our studies, we saw that the \textit{EGFP} reporter and \textit{neo} marker in DelTel7 ESCs cells were affected by TPE (Figure 3-1). We also demonstrated that EGFP reporter is also affected by TPE \textit{in vivo} experiments (Figure 4-1). Even though TPE can affect both the \textit{EGFP} reporter and \textit{neo} marker in DelTel7 ESCs, we were unable to see any effect on the expression of \textit{H19} in ESCs, which is a maternally-expressed gene located proximally of the DelTel7 breakpoint on distal mouse chromosome 7. Although both GFP and \textit{neo} are affected by TPE, my results showed that such an effect does not spread in \textit{cis} to \textit{H19}. With a lengthened telomere, there could also be a chance of TPE-OLD acting in the same area.\footnote{27}

\section*{6.4 Determination of imprinting at the DelTel7 allele in ESCs}

After the production of DelTel7, there were several questions to answer regarding this new ESC line. The first one is whether GFP is imprinted or not and the experiment to determine this is to look at the GFP expression in embryos, when DelTel7 is inherited maternally or paternally. DelTel7/+ embryos had a lot of GFP expression whereas the +/DelTel7 embryos didn’t have any GFP expression with the exception of the heart and gonads. This experiment confirms that GFP expression shows an imprinting-like effect in vivo. The reason behind the differential expression of GFP in DelTel7/+ and +/DelTel7 embryos was unknown so I considered two different explanations.

The first explanation for the differences in GFP expression is caused by the differences in maternally and paternally inherited genes. Since the GFP is in an imprinted region, this is a reasonable explanation since \textit{Igf2} and \textit{H19} expression follows the parent-of-origin-specific
monoallelic expression\textsuperscript{68,69}. A second explanation is due to the fact that males and females have differences in establishing germline cells. Establishment of methylation post erasure is completed earlier in the male germline than the female germline\textsuperscript{70,71}. Male gametes also arrest at a mitotic stage whereas female gametes arrest at a meiotic stage\textsuperscript{70}. Using ESCs provided from the Higgins lab, we were hoping to monitor GFP expression following maternal and paternal transmission. However, both cell lines obtained appear to carry a maternally inherited DelTel7 allele. The only explanation we have for the differences in GFP levels between those two cell lines is that one of them appears less stable and more differentiated in culture.

### 6.5 Rationale for producing DelTel6

Many aspects of determining if there is a mechanism to whether cells are GFP\textsuperscript{+} or GFP\textsuperscript{-} have been negative. So, the biggest question remains: what is causing GFP to be expressed and repressed? The general answer is that the telomere is involved but we still do not know why and how. With that being said, DelTel7 was produced originally to determine if IC1 and IC2 could be separated and still maintain viability since they are always inherited together in mammals. It is clear that the IC1 region is viable when IC2 is not present since this was the structure of DelTel7. To further analyze and confirm the GFP heterogeneity seen in DelTel7, it is essential to produce a similar chromosome truncation on another chromosome. Mouse chromosome 6 has two clusters of imprinted genes in the proximal region, centered on the genes \textit{Mest} and \textit{Peg10}\textsuperscript{56,72}. Therefore, mouse chromosome 6 is a good candidate for a chromosome truncation because the distal end has no known imprinting regions.

The proposed approach would be to use the MICER technique to generate this DelTel6\textsuperscript{73}. MICER takes advantage of a gene named HPRT as a selectable marker. HPRT is normally found
on the X chromosome in human and mice. There is a 3’HPRT and 5’HPRT and when placed together, it is possible to select for cells that are positive and negative for HPRT\textsuperscript{73}. Ideally, we would have a plasmid with some key components to allow for this truncation to occur. The key components are a \textit{loxP} site, fluorescent reporter, 3’HPRT or 5’HPRT and the telomere repeat seed. The fluorescent reporter would allow us to confirm whether the stochastic GFP expression can be repeated. The \textit{loxP} would allow for chromosome truncation to occur and the HPRT would be the selectable marker to screen for cells with the appropriate truncation. Lastly, the telomere repeat seed would allow the telomerase to recognize the telomere and replace it with a lengthier version of itself. This vector would also allow for universal chromosome truncation at any chromosome. A strategy of the MICER technique that can be used to make DelTel6 can be found in Figure 6-1.
Figure 6-1. Schematic diagram of the proposed strategy to use MICER techniques to produce DelTel6.

Mouse chromosome shown contains the centromere on the left and telomere on the right. The first step is to introduce a vector containing 3’HPRT, loxP and puro into the mouse chromosome. The 3’HPRT and loxP will allow for selection and mediated truncation later on. The puro selectable marker allows for the selection of cells with the integrated vector. The second step is to truncate the chromosome using the Cre-loxP system and a pCX-mCherry vector containing 5’HPRT. Cherry fluorescing marker and telomere repeat seed. pCX-mCherry (*) would allow for universal chromosome truncation at any chromosome once the 3’HPRT and loxP were placed near the distal end of targeted chromosome.
6.6 Verification of episomal system

Studies have used Large T in different capacities like maintaining extrachromosomally in pancreatic endothelial cells and mouse embryonic carcinoma cells, as well as synthetic lethality screens in PEFs \(^{59,60}\). Using the Large T antigen and DelTel7 ESCs, we were able to see that an episomal system to screen for new modifiers of TPE is still possible. My experiments focused on the logistics of the episomal system like characterizing DelTel7-pMGD20zeo through its morphology and GFP profile. Through microscope pictures and initial flow cytometry, the GFP profile for DelTel7-pMGD20zeo did not match DelTel7-C3. However, after eliminating a number of possibilities like founder’s effect where we looked at more colonies as well as observing them in a longer capacity, we still saw a big variation in GFP profiles for the transfected ESCs. Another possibility is that Large T is causing the different GFP profiles, which would halt the episomal system but would open up Large T as an avenue of interest. pPGKzeo was transfected into the cells and to our relief, the differential GFP profiles still existed. This new data further supported the possibility of an episomal system but did not help determine what was causing the high variation in GFP expression. The last possibility was that zeo might have a role and looking at the data, it is clear that zeo was causing the high variation in GFP expression (Figure 5-8).

Zeocin resistance is conferred by the Shble gene which inactivates zeocin by binding to the antibiotic. Zeocin is a bleomycin-like compound that introduces lethal double stand breaks (DSB) in DNA \(^{74}\). It would be interesting to see if the cells placed under zeo selection indeed had more DSBs than the ESCs that had no zeo selection. When DSB occur naturally in the genome, the cell uses nonhomologous end joining (NHEJ) and homologous recombination (HR) to repair the break \(^{75}\). Studies in CHO hamster cells demonstrate chronic exposure to 500µg/mL of Zeocin
increases NHEJ products compared to HR products\textsuperscript{76}. Cells are thus placed in an unbalanced environment and chronic DSB stress. Analysis of DelTel7 ESCs under \textit{zeo} selection should also be done to determine if we also see more NHEJ products than HR products.

Studies in yeast have shown that SIR2, SIR3, SIR4 are required for NHEJ in yeast\textsuperscript{16}. Rad50p, Mre11p and Xrs2p are also necessary for NHEJ involving Ku\textsuperscript{16}. Human Ku80 has been seen to co-immunoprecipitate with hamster telomeric DNA, which leads scientists to believe that Ku is also associated with mammalian telomeres\textsuperscript{77}. Scientists have shown that Ku is recruited for NHEJ when DSBs appear and also that Ku is associated with telomere silencing. These research avenues have guided us into thinking that Ku is involved in DSBs and TPE. Thinking along the lines of Ku, DelTel7 ESCs should be looked at to see if Ku is recruited to the DSBs caused by \textit{zeocin}. Also moving forward, experiments could be performed to see if the findings in yeast telomeres are relevant to mouse ESCs telomeres.
Our model is that when cells are placed in the zeocin selectable conditions, the drug induces DSBs. These DSBs would recruit telomeric proteins like Ku to repair the damage and somehow this affects TPE. Indeed, one key observation is that growth in presence of zeocin increases the percentages of GFP<sup>+</sup> cells over time, at least for some clones. This effect is not seen when zeocin selection is not applied. Nevertheless, some clones remain mostly GFP<sup>-</sup> in the same conditions. This would suggest that for these clones, the silent state has been locked in perhaps by a high density of DNA methylation, as has been previously observed for another TPE reporter in ESCs<sup>25</sup>. How zeocin acts as a suppressor of TPE is an interesting research question to address in the future.

![Model of the GFP variation seen in DelTel7 ESCs electroporated with Large T antigen and placed under zeo selection.](image)

Figure 6-2. Model of the GFP variation seen in DelTel7 ESCs electroporated with Large T antigen and placed under zeo selection.
References


49. Furusawa T, Ohkoshi K, Honda C, Takahashi S, Tokunaga T. Embryonic stem cells expressing both platelet endothelial cell adhesion molecule-1 and stage-specific embryonic


