

# Genetic Dissection of the Exit from Pluripotency in Mouse Embryonic Stem Cells by CRISPR-based Screening

Meng Li



Newnham College  
Wellcome Trust Sanger Institute  
University of Cambridge

This dissertation is submitted for the degree of

*Doctor of Philosophy*

September 2017

## **Declaration**

This dissertation describes work carried out from June 2014 to July 2017 under the supervision of Dr Kosuke Yusa at the Wellcome Trust Sanger Institute.

I hereby declare that this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

I state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

This thesis does not exceed the prescribed word limit of 60,000 words as set by the Degree Committee for the Faculty of Biology.

Meng Li

30th September 2017

## **Acknowledgements**

I would like to thank foremost my supervisor Kosuke Yusa, for his guidance and support during the past four years. I consider myself extremely lucky to have a supervisor who cared so much about my work and provided the best supervision he possibly can. His immense knowledge and great passion for science was and will always be a source of inspiration to me.

I would also like to thank the rest of my thesis committee: Prof Allan Bradley and Dr Pentao Liu, for the insightful discussions and encouragement. I am very grateful to my external supervisor Prof Austin Smith, who has been a great source of advice throughout my PhD. I also thank Dr Martin Leeb and Dr Meng Amy Li, for sharing their knowledge on genetic screening and ESC culture.

I am thankful to everybody in the the Yusa lab. I would particularly like to thank Hiroko, who is one of the warmest and kindest person I have ever met. She has given me enormous help and encouragement. This journey would not be as great without her. Jason also deserves special mention for helping me in the lab and always being there to give advice. I thank Kasia, Shannade, Yanzi, Leandro and Swee Hoe for all the help and support they have given me. Thanks also to the Vallier lab members especially Sapna for always being there to cheer me up. I would also like to thank the members of staff at the Flow Cytometry Facility, without whom all the large-scale cell sorting would not be possible.

I am also very glad for having had the luck of being surrounded by great friends. Special thanks to Ziyang, Giulia and Haruna for all the warm encouragement and support. Thanks also to Tom for proof-reading part of this thesis.

Finally, I would like to express my gratitude to my parents, for their constant encouragement and unconditional love. Without their support I would never have gotten to where I am today.

## Abstract

The ground state naive pluripotency is established in the epiblast of the blastocyst and can be captured by culturing mouse embryonic stem cells (mESCs) with MEK and GSK3 inhibitors (2i). The transcription network that maintains pluripotency has been extensively studied with the indispensable core factors being Oct4, Sox2 and Nanog, together with other ancillary factors reinforcing the network. However, how this network is dissolved at the onset of differentiation is still not fully understood. To identify genes required for differentiation in an unbiased fashion, I conducted a genome-wide CRISPR-Cas9-mediated screen in Rex1GFPd2 mESCs. This cell line expresses GFP specifically in the naive state and rapidly down-regulate upon differentiation. I differentiated mutagenised mESCs for two days and sorted mutants that kept higher GFP expression. gRNA representation was subsequently analysed by sequencing. I identified 563 and 8 genes whose mutants showed delayed and accelerated differentiation, respectively, at a false discovery rate (FDR) cut-off of 10%. The majority of the previously known genes were identified in my screen, suggesting faithful representation of genes regulating differentiation. Detailed screening result analysis revealed a comprehensive picture of pathways involved in the dissolution of naive pluripotency. Amongst the genes identified are 19 mTORC1 regulators and components of the mTORC2 complex. Deficiency in the TSC and GATOR complexes resulted in mTORC1 upregulation in consistent with previous studies. However, they showed opposite phenotype during ESC differentiation: TSC complex knockout cells showed delayed differentiation, whereas GATOR1 deficiency accelerated differentiation I found that the pattern of GSK3b phosphorylation is highly correlated with differentiation phenotype. I conclude that mTORC1 is involved in pluripotency maintenance and differentiation through cross-talk with the Wnt signalling pathway. My screen has demonstrated the power of CRISPR-Cas9-mediated screen and provided further insights in biological pathways involved in regulating differentiation. It would be interesting to explore the remaining unstudied genes for better understanding of the mechanisms underlying mESC differentiation.

## Abbreviation

2i	2 inhibitors
AP	Alkaline phosphate
<i>Blm</i>	Bloom's syndrome
BMP	Bone morphogenetic protein
cDNA	complementary DNA
ChIP	Chromatin immunoprecipitation
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
CRISPRi	CRISPR interference
CRISPRa	CRISPR-mediated gene activation
dCas9	nuclease-deficient Cas9
DE score	Depletion/Enrichment score
DSB	Double stranded break
dsRNA	double stranded RNA
DT-A	Diphtheria toxin fragment-A
EC	Embryonic carcinoma
EGC	Embryonic germ cells
ENU	N-ethyl-N-nitrosourea
EpiSC	Epiblast stem cell
ESC	Embryonic stem cell
ESCRT	Endosomal sorting complexes required for transport
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FDR	False discovery rate
GAP	GTPase-activating protein
GATOR	Gap activity toward rags
GeCKO	CRISPR-Cas9 Knockout
gRNA	guide RNA
GSEA	Gene set enrichment analysis
HDR	Homologous directed repair
HOPS	Homotypic fusion and protein sorting
HR	Homologous recombination
ICM	Inner cell mass
indels	Insertion or deletions
iPSC	Induced pluripotent stem cell

---

KSR	Knockout serum replacement
LIF	Leukemia inhibitory factor
LncRNA	Long non-coding RNA
LTR	Long terminal repeat
M15L	Knockout DMEM supplemented with 15%FBS and LIF
m <sup>6</sup> A	N <sup>6</sup> A-methyladenosine
MAGeCK	Model-based analysis of genome-wide CRISPR/Cas9 knockout
MMR	Mismatch repair
MOI	Multiplicity of infection
mTOR	Mammalian target of rapamycin
mTORC1/2	mTOR complex 1/2
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NMD	Non-sense mediated decay
NuRd	Nucleosome remodeling deacetylase
OSN	Oct4/Sox2/Nanog
RA	Retinoid acid
RIGER	RNAi gene enrichment ranking
RNP	Ribonucleoproteins
RSA	Redundant siRNA activity
PAM	Protospacer adjacent motif
PB	piggyBac
PGC	Primordial germ cells
PRC	Polycomb repressive complex
RISC	RNA-inducing silencing complex
RNAi	RNA interference
RVDs	Repeat variable diresidues
SAM	synergistic activation mediator
SB	Sleeping beauty
siRNA	short interfering RNA
shRNA	short hairpin RNA
ssODNs	single-strand oligodeoxynucleotides
TALEN	Transcription activator-like effector nuclease
tracrRNA	trans-encoded smallRNA
TKO	Toronto knockOut
TSC	Tuberous sclerosis complex
XEN	Extra-embryonic endoderm
ZFN	Zinc finger nuclease

# Contents

<b>1</b>	<b>Introduction</b>	<b>14</b>
1.1	Reverse Genetics . . . . .	15
1.1.1	Gene targeting . . . . .	15
1.1.2	RNA interference (RNAi) . . . . .	16
1.1.3	Genome engineering with programmable nucleases . . . . .	17
1.1.3.1	The repair pathways and applications . . . . .	18
1.1.3.2	Programmable nucleases before the CRISPR era . . . . .	19
1.1.3.2.1	ZFN . . . . .	19
1.1.3.2.2	TALEN . . . . .	20
1.1.3.3	CRISPR-Cas systems . . . . .	21
1.1.3.3.1	The discovery of CRISPR-Cas systems . . . . .	21
1.1.3.3.2	The diversity of CRISPR-Cas systems . . . . .	21
1.1.3.3.3	CRISPR-Cas9 system as a genome-editing tool . . . . .	23
1.1.3.3.4	Structure and working mechanism of Cas9 . . . . .	25
1.1.3.3.5	Off-target effect . . . . .	26
1.1.3.3.6	Applications of CRISPR-Cas9 system . . . . .	26
1.2	Forward genetics . . . . .	29
1.2.1	Mutagenesis using chemical and physical agents . . . . .	29
1.2.1.1	N-ethyl-N-nitrosourea (ENU) . . . . .	29
1.2.1.2	Irradiation . . . . .	30
1.2.2	Insertional mutagenesis . . . . .	30
1.2.2.1	Retroviral-mediated mutagenesis . . . . .	31
1.2.2.2	DNA transposon-mediated mutagenesis . . . . .	31
1.2.3	The use of <i>Blm</i> -deficient and haploid cell lines in genetic screens . . . . .	32
1.2.3.1	<i>Blm</i> -deficient ESC systems . . . . .	32
1.2.3.2	Haploid cell lines . . . . .	32

1.2.4	RNAi-mediated screens . . . . .	33
1.2.5	CRISPR-Cas9-mediated screens . . . . .	34
1.2.5.1	The establishment of CRISPR-Cas9 screening technology . . . . .	34
1.2.5.2	Screening format . . . . .	35
1.2.5.2.1	Arrayed screening . . . . .	35
1.2.5.2.2	Pooled screening . . . . .	35
1.2.5.3	Applications of CRISPR-Cas9-mediated screens . . . . .	36
1.2.5.4	Experimental design of CRISPR-Cas9-mediated screens . . . . .	38
1.2.5.4.1	gRNA design . . . . .	38
1.2.5.4.2	Cas9 and gRNA delivery . . . . .	39
1.2.5.4.3	Gene identification and data analysis . . . . .	40
1.3	Embryonic stem cells (ESCs) . . . . .	42
1.3.1	Early development of mouse embryo . . . . .	42
1.3.2	Derivation of mouse ESCs . . . . .	43
1.3.3	Regulation of the pluripotency state . . . . .	45
1.3.3.1	Extrinsic signalling pathways . . . . .	46
1.3.3.1.1	LIF-mediated signalling pathway . . . . .	46
1.3.3.1.2	TGF- $\beta$ -mediated pathway . . . . .	48
1.3.3.1.3	FGF/MAPK pathway . . . . .	48
1.3.3.1.4	Wnt signalling pathway . . . . .	49
1.3.3.1.5	Serum/LIF culture and 2i culture . . . . .	51
1.3.3.2	Transcription factor network . . . . .	52
1.3.3.2.1	Core pluripotency factors . . . . .	52
1.3.3.2.2	Ancillary pluripotency regulators . . . . .	55
1.3.3.3	Epigenetic regulation . . . . .	56
1.3.3.3.1	DNA methylation . . . . .	56
1.3.3.3.2	Bivalent domains . . . . .	57
1.3.3.3.3	Heterochromatin organisation . . . . .	58
<b>2</b>	<b>Materials and Methods</b>	<b>59</b>
2.1	Cell Culture . . . . .	60
2.1.1	Materials . . . . .	60
2.1.1.1	Cell lines . . . . .	60
2.1.1.2	Media components, inhibitors and other reagents used in cell culture . . . . .	60



---

2.1.1.3	Dissociation agents . . . . .	61
2.1.1.4	Other chemicals and kits . . . . .	61
2.1.1.5	Media . . . . .	61
2.1.2	Methods . . . . .	62
2.1.2.1	Routine culture and maintenance mESCs . . . . .	62
2.1.2.2	Differentiation condition . . . . .	62
2.1.2.3	Transfection of mESCs . . . . .	62
2.1.2.4	Lentivirus production and transduction . . . . .	63
2.1.2.5	Genetic manipulation of cell lines . . . . .	63
2.1.2.6	Commitment assay . . . . .	64
2.1.2.7	Flow cytometry and cell sorting . . . . .	65
2.2	Molecular Biology . . . . .	66
2.2.1	Materials . . . . .	66
2.2.1.1	Molecular chemicals and Kits . . . . .	66
2.2.1.2	Immunoblotting . . . . .	67
2.2.2	Methods . . . . .	69
2.2.2.1	Isolation of nucleic acids . . . . .	69
2.2.2.2	gRNA cloning . . . . .	69
2.2.2.3	Lysate PCR . . . . .	69
2.2.2.4	Illumina library preparation . . . . .	70
2.2.2.5	Western blotting . . . . .	71
2.3	Bioinformatics analysis . . . . .	72
2.3.0.1	RNA-Seq analysis . . . . .	72
2.3.0.2	gRNA sequencing result analysis . . . . .	72
2.3.0.3	Gene Set Enrichment Analysis (GSEA) analysis . . . . .	72
<b>3</b>	<b>Screening Preparation and Condition Optimisation</b>	<b>73</b>
3.1	Introduction . . . . .	74
3.2	Results . . . . .	75
3.2.1	Analysis of self-renewal and differentiation conditions . . . . .	75
3.2.2	Establishment of Cas9 expression in Rex1:GFPd2 cell line . . . . .	78
3.2.3	Proof of principle studies . . . . .	80
3.2.4	A Preliminary screen . . . . .	82
3.3	Discussion and Conclusion . . . . .	85

<b>4</b>	<b>Screening Result and Pathway Analysis</b>	<b>88</b>
4.1	Introduction . . . . .	89
4.2	Results . . . . .	92
4.2.1	Screening Strategy and Result Analysis . . . . .	92
4.2.2	Screening Result Validation . . . . .	94
4.2.3	Pathways Analysis . . . . .	96
4.2.3.1	Signalling pathways . . . . .	96
4.2.3.2	mRNA related pathways . . . . .	98
4.2.3.3	Chromatin modifiers . . . . .	99
4.2.3.4	Transcription factors and other transcriptional regulatory proteins . . . . .	101
4.2.3.5	Mitochondria-related pathways . . . . .	102
4.2.3.6	Endocytosis and vesicle trafficking . . . . .	103
4.2.4	Discussion and Conclusion . . . . .	105
4.2.4.1	Functional genomics with CRISPR-Cas9 . . . . .	105
4.2.4.2	Integration of the regulatory network . . . . .	106
4.2.4.3	Metabolic regulation . . . . .	107
<b>5</b>	<b>The Role of mTOR-related Pathways in Pluripotency/Differentiation Regulation</b>	<b>108</b>
5.1	Introduction . . . . .	109
5.2	Results . . . . .	114
5.2.1	Generation of knockout cell lines and phenotype validation . . . . .	114
5.2.2	Disruption of GATOR1 complex induces mTORC1-activated nega- tive feedback loop . . . . .	117
5.2.3	<i>Tsc2</i> knockout demonstrated unconventional cross-talk between mTOR and PI(3)K-Akt pathways . . . . .	120
5.2.4	Transcriptional changes in <i>Nprl2</i> and <i>Tsc2</i> knockouts . . . . .	122
5.3	Discussion and conclusion . . . . .	126
<b>6</b>	<b>Conclusions and Future Perspectives</b>	<b>128</b>
6.1	The immeasurable potential of CRISPR-Cas9-mediated genetic screen in stem cell biology . . . . .	129
6.2	The complex regulatory network that drives dissolution of pluripotency . .	132

6.3 The prominent role of mTOR and its related proteins in pluripotency and differentiation regulation . . . . .	134
--	-----

# List of Figures

1.1	CRISPR-mediated DNA interference in microbial adaptive immunity. . . .	22
1.2	The mechanism of CRISPR-mediated genome engineering. . . . .	24
1.3	Overall structure of <i>Streptococcus pyogenes</i> Cas9 and gRNA. . . . .	25
1.4	Lineage segregation in mouse blastocyst. . . . .	44
1.5	Extrinsic signalling pathways that regulate pluripotency. . . . .	46
3.1	Analysis of Rex1GFP profile under maintenance conditions. . . . .	76
3.2	Analysis of the Rex1GFP profile under differentiation conditions. . . . .	77
3.3	Cas9 knockin . . . . .	78
3.4	Cas9 function assay. . . . .	79
3.5	Positive control study with stable knockout clones. . . . .	81
3.6	Positive control study with single-gRNA knockouts. . . . .	83
3.7	Preliminary screen result analysis. . . . .	86
4.1	Screening Strategy and Result Analysis. . . . .	93
4.2	Screening Result Validation. . . . .	95
4.3	Cellular pathways underpinning the initiation of differentiation. . . . .	97
4.4	Rex1GFP differentiation profile of Oct4 knockout mESCs. . . . .	101
4.5	Validation of genes in mitochondria-related pathways. . . . .	104
5.1	Structure of mTORC1 and mTORC2. . . . .	109
5.2	Structure of mTOR regulating complexes and mechanistic target of mTORC1 amino acid sensing pathway. . . . .	111
5.3	Generation of stable KO cell lines for mTORC1 regulating genes. . . . .	115
5.4	mTOR-related gene knockout phenotype validation. . . . .	116
5.5	GATOR1 KO phenotype resulted from activated GSK3. . . . .	119

5.6	The unconventional cross-talk between mTOR and PI(3)K-Akt pathways in <i>Tsc2</i> KO. . . . .	121
5.7	Transcriptional changes in <i>Nprl2</i> and <i>Tsc2</i> knockouts. . . . .	123
5.8	Pathway analysis of differentially expressed gene sets. . . . .	125
5.9	Summary of major relevant signalling components upstream and down- stream of mTORC1 and mTORC2. . . . .	126