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 Ziyan, 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
Blm	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	Fluoresence-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^6A	N^6A-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-indicing 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

List of Figures