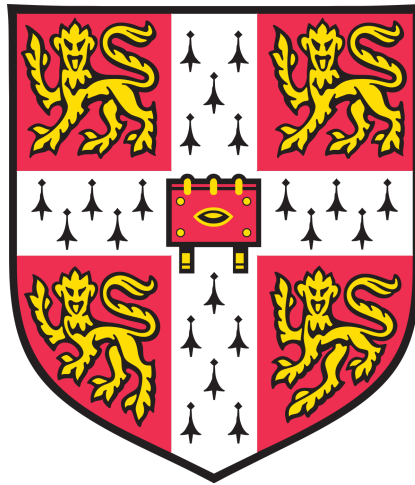


Genome-wide search for regulators of the first cell lineage decision in the mouse: a CRISPR/Cas9 endeavour



Liliana Freitas Antunes

Wellcome Trust Sanger Institute

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

DECLARATION

I hereby declare that the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. 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. This thesis does not exceed the 60.000 word limit set by the Faculty of Biology.

Liliana Antunes
December 2019

ACKNOWLEDGEMENTS

The PhD experience was a key avenue towards both professional and personal growth. I would start by acknowledging my supervisors, Pentao Liu, Emmanouil Metzakopian and Allan Bradley for their complementing and experienced advice throughout this five-year adventure. I am especially thankful to Manos and Allan which embraced my project at a key turning point and have provided guidance and insightful discussions leading to this PhD thesis. Both were not only extremely supportive supervisors, but also thoughtful mentors, allowing me the freedom of thought to shape my own project (while preventing some tempting classical PhD student mistakes!).

I am very grateful to my thesis committee, George Vassiliou and Jennifer Nichols, for their constant encouragement and meaningful critique. In particular to Jenny, for taking the time to share her extensive knowledge in embryo development, both theoretically and experimentally.

Research is ultimately a multi-player effort and therefore I would like to acknowledge the skilled core facilities at Sanger that were a valuable support for the development of this project. Especially the core flow cytometry facility for their countless hours of cell sorting, the molecular cytogenetics core, and the DNA pipelines. Externally, I am very grateful to the Cancer Research UK genome editing core, namely Guocheng Lan, Xiangang Zhu and Alasdair Russell for their collaboration towards the *in vivo* chimera assays.

I thank the Portuguese Foundation for Science and Technology, FCT, as well as the Wellcome Sanger Institute for funding my PhD and allowing me to join a world-leading scientific community. To the Sanger Graduate Office, namely Christina Hedberg-Delouka and Annabel Smith, for all their support. My deepest gratitude towards the Director of Graduate studies, Carl Anderson, for stepping out of his way to genuinely support my professional development.

One of the key take home messages from the PhD experience is that in addition to our own motivation and perseverance, the surrounding environment plays a major role in the way one overcomes challenges. I consider myself incredibly lucky for all the special people I have encountered along the way, who made my life much easier in so many synergistic aspects.

To all the members of past Team 18, that made my first years go smoothly. To David Ryan and Wei Wang for introducing me to both pluripotent and expanded potential stem cells. Jason Tsang for input into single-cell RNAseq. Juexuan Wang and Yong Yu for general help.

A particular thanks to Xuefei Gao and Jian Yang who followed my twists and turns towards this project and provided useful discussions throughout. Furthermore, to Jian, for his prompt help with E6.5 dissections.

To all the members of Teams Bradley and Metzakopian for the supportive and scientifically stimulating environment, which was always mixed with fun. In particular, to Manos Metzakopian and Mathias Friedrich for always having an answer to experimental design and cloning tips. To Michał Kosicki, Kärt Tomberg, Dimitris Garyfallos, Jorge de la Rosa, and Imran Noorani for keen discussions and authentic interest in science. To Ross Cook, Katta Hautaviita, Alex Strong, Frances Law and Haydn Prosser for all the invaluable technical support. From the Metzakopian Team, a big thanks to Nikolaos Patikas for precious help with computational data analysis, and to Sandeep Rajan for assistance with promoter methylation analysis. Together with Stefanie Foskolou and Hugo Fernandes, this could not have been a better crew for the inception of the Metzakopian Lab. Outside research, thanks to everyone who joined the mythical foosball breaks and all the coffee-infused random discussions, a pleasant addition to weekend feeding routines (and everyday life!). It was a privilege to work alongside such inspiring scientists, whom I can also call friends.

To the PhD gang - Masha, Veli, Alice, Dimitris, GM and Fernando -, for their unconditional support through the good, the bad and the uncanny! To Fernando Aguilera for the great (and dark) sense of humour, precisely tailored to need. To Gianmarco Raddi for his passion for politics and “socialism”. To Dimitris Garyfallos for the insight into Greek mythology and all sorts of super heroes. To Alice Mann for the Johnian-Sanger adventures (together with Masha!). To Velislava Petrova for shared views of the world. And to Masha Levitin, for simply holding us all together. Sharing the highs and lows of PhD surrounded by true friendship has made it possible to overcome all challenges and make the most of Cambridge daily life. The PhD experience would otherwise be incomplete.

To the lifelong friends, that always find the time to catch-up in my short-notice and rather brief visits, Nuno Santos, Daniela Araújo, Daniela Silva, Tânia Rodrigues and Cátia Sousa. To Cátia, for being the bravest person I've had the luck to come across.

Finally, to my sister and parents, for all their support and encouragement throughout the different stages of my life that lead me to this amazing opportunity. In particular to our parents for prioritising our education despite all challenges. To Miguel, for his understanding of the demands of stem cell culture and PhD in general, and for truly believing in my potential.

ABSTRACT

In vitro cell lines derived from the early embryo retain the identity and lineage restrictions of their embryonic counterparts: embryonic stem cells (ESCs) are pluripotent and similarly to ICM contribute to all tissues in the embryo proper, being generally excluded from the extraembryonic layer. On the other hand, trophoblast stem cells (TSCs) resemble TE and are restricted to the extra-embryonic lineage. Despite the efforts of several groups, it is still not possible to completely convert ESCs to TSCs. The molecular players that underpin this restriction in ESCs remain largely unknown. I therefore proposed to identify suppressors of ESC differentiation to trophoblast in an unbiased approach, using CRISPR/Cas9 genome-wide loss of function screening.

I developed a reporter cell line that constitutively expressed *Cas9* from the *Rosa26* locus, and was engineered with *knock-in* of *T2A-Venus* into the *Elf5* locus (*Elf5::Venus* ESCs). *Elf5* is a stringent TSC marker that fails to be activated in most published methods for ESC conversion to TSC. Proof-of-concept experiments used *Oct4* knockout via CRISPR/Cas9 to validate this line for both efficient gene editing and faithful tracking of TSC differentiation. *Elf5::Venus* ESCs were then transduced with a genome-wide lentiviral gRNA library for CRISPR/Cas9 genetic screen. Mutant cells were differentiated in TSC media for fourteen days, at which point Venus positive and negative populations were sorted. gRNA representation was subsequently analysed by next generation sequencing. At a false discovery rate of 10%, I could identify 42 genes whose loss of function allowed ESC conversion to TSC. I could validate 22 genes with different phenotype strength, implicated in the maintenance of the first cell lineage decision. Importantly, the screen identified four components of the non-canonical Polycomb Repressive Complex 1.1 as strong lineage regulators. The role of two of these genes, *Bcor* and *Rnf2*, was further explored in ESCs for their differentiation profile as well as transcriptional changes that underline their mechanism. Preliminary data for *in vivo* chimera assays showed that derivatives of ESCs deficient for *Bcor* or *Rnf2* could be found within extraembryonic tissues in chimeric embryos. Further studies however will be necessary to assess both reproducibility and cell identity, and evaluate if this lineage restriction is indeed impaired *in vivo* in knockout ESCs.

Overall, my studies demonstrate the power of CRISPR/Cas9 screen, highlighting new layers of regulation for the first cell lineage decision and its maintenance in ESCs. I believe further characterisation of remaining screen hits will continue to elucidate the molecular mechanisms of TE and ICM segregation in early embryos.

TABLE OF CONTENTS

1	INTRODUCTION	1
1.1	Totipotency and preimplantation mouse embryo development	1
1.1.1	Totipotency and epigenetic dynamics in early preimplantation	2
1.1.2	The first cell lineage decision: inner cell mass versus trophoctoderm specification	4
1.1.3	The second cell lineage decision: Epiblast versus Primitive Endoderm	9
1.2	Stem cell lines derived from preimplantation mouse embryos	14
1.2.1	Embryonic Stem Cells (ESCs)	14
1.2.2	Trophoblast Stem Cells (TSCs).....	17
1.2.3	Extraembryonic Endoderm Stem Cells (XEN cells)	18
1.3	The trophoblast lineage and early placenta development	19
1.3.1	Early placenta development and trophoblast differentiation	19
1.3.2	Trophoblast Stem Cell Regulatory Networks	22
1.3.3	<i>In vitro</i> Trophoblast Stem Cell Lines – Identity and Potency.....	28
1.4	Mammalian Genome Engineering	31
1.4.1	Gene targeting and Homologous Recombination	31
1.4.2	Genome Editing and the use of engineered chimeric nucleases	32
1.4.3	CRISPR/Cas9 Genome editing	34
1.5	Thesis Aims	37
2	MATERIALS AND METHODS	39
2.1	Cell culture	39
2.1.1	Embryonic (ESC) and Trophoblast (TSC) stem cell lines	39
2.1.2	Cell culture conditions	39
2.1.3	ESC and TSC passaging and cryopreservation.....	41
2.1.4	<i>In vitro</i> differentiation towards trophoblast lineage	41
2.1.5	ESC transfection	41

2.2	Molecular Biology	42
2.2.1	Molecular cloning	42
2.2.2	Nucleic Acid Extraction.....	44
2.2.3	Genotyping.....	45
2.2.4	cDNA preparation.....	46
2.2.5	RT-qPCR.....	46
2.2.6	<i>Elf5</i> promoter DMR analysis.....	47
2.3	Flow cytometry and fluorescent-activated cell sorting (FACS)	47
2.4	Lentivirus methods and genome-wide CRISPR/Cas9 screening	48
2.4.1	Lentivirus Production.....	48
2.4.2	ESC transduction and Lentivirus titration	48
2.4.3	Genome-wide mouse gRNA lentiviral library	49
2.4.4	Genome-wide ESC mutant libraries and screening	49
2.4.5	Illumina sequencing of gRNAs	50
2.5	Immunostaining	50
2.5.1	Immunocytochemistry	50
2.5.2	Embryo Immunostaining.....	51
2.5.3	Image acquisition	51
2.5.4	Cell surface Immunostaining for flow cytometry analysis.....	51
2.5.5	Intracellular Immunostaining for flow cytometry analysis.....	52
2.5.6	Western Blot.....	52
2.6	RNAseq library preparation and NGS sequencing	53
2.7	Embryo work	53
2.7.1	Preparing embryos and foster recipients	54
2.7.2	ESC microinjection into 8-cell embryos for chimera assays	54
2.8	Computational methods	54
2.8.1	CRISPR/Cas9 screening analysis.....	54
2.8.2	RNAseq analysis.....	55
2.8.3	ChIPseq data analysis (Published dataset).....	55
2.9	Methods Tables	56

3	PROOF-OF-CONCEPT STUDIES FOR THE USE OF CRISPR/CAS9 GENETIC SCREENS IN THE DIFFERENTIATION OF ESCS TO TSCS	65
3.1	Introduction	65
3.1.1	Bypassing the first cell lineage decision: conversion of ESCs to TSC fate	66
3.1.2	Epigenetic memory of the first cell lineage	69
3.1.3	Role of <i>Elf5</i> in the trophoblast compartment	72
3.1.4	Scope of this chapter	73
3.2	Results	74
3.2.1	Generation of an <i>Elf5::Venus</i> ESC reporter line for trackable detection of trophoblast differentiation	74
3.2.2	Proof-of-concept for the use of CRISPR/Cas9 in genome-wide knockout screening for TSC differentiation: <i>Oct4</i> knockout and optimisation experiments	82
3.2.3	Using <i>Dnmt1</i> knockout as a second case-study for TSC differentiation	92
3.3	Discussion	97
3.3.1	Development of <i>Elf5::Venus</i> ESCs that faithfully report TSC differentiation	97
3.3.2	Optimisation of experimental setup for CRISPR/Cas9 genome-wide screening using <i>Oct4</i> knockout model	98
3.3.3	Limitations in CRISPR/Cas9 screening: the <i>Dnmt1</i> knockout scenario	101
3.4	Supplementary Figures	103
4	CRISPR/CAS9 GENOME-WIDE RECESSIVE SCREENING IDENTIFIES NEW GENES PREVENTING ESC DIFFERENTIATION TO TROPHOBLAST	105
4.1	Introduction	105
4.1.1	High-throughput loss-of-function genetic screens	106
4.1.2	CRISPR/Cas9 in genome-wide screenings	107
4.1.3	Arrayed and Pooled screening approaches	108
4.2	Scope of this chapter	109
4.3	Results	110
4.3.1	High-throughput CRISPR/Cas9 recessive screening uncovers genes regulating the first cell lineage barrier in ESCs	110
4.3.2	Validation of candidate genes	121

4.4	Discussion	136
4.4.1	CRISPR/Cas9 genome-wide screen for <i>Elf5</i> activation: strengths and pitfalls	136
4.4.2	A renewed look at regulators of ESC to TSC lineage barrier	138
5	ROLE OF NCPRC1.1 MEMBERS <i>BCOR</i> AND <i>RNF2 (RING1B)</i> IN THE MAINTENANCE OF THE FIRST CELL LINEAGE DECISION.....	143
5.1	Introduction.....	143
5.1.1	Polycomb Group (PcG) Proteins.....	144
5.1.2	Models for PRC1 and PRC2 recruitment to chromatin	146
5.1.3	The role of PRC1 and PRC2 in ESC self-renewal and differentiation.....	147
5.1.4	The function of <i>Bcor</i> in mouse embryo development and ESCs.....	149
5.1.5	The role of <i>Rnf2</i> in mouse embryo development and ESCs	150
5.1.6	Scope of this chapter	151
5.2	Results.....	152
5.2.1	The role of <i>Bcor</i> in the maintenance of ESC and TSC lineage barrier.....	152
5.2.2	The role of <i>Rnf2</i> in the restriction of ESC to TSC differentiation.....	171
5.2.3	Transcriptomic characterisation of <i>Bcor-KO</i> and <i>Rnf2-KO</i> ESCs	181
5.2.4	Surveying <i>Bcor</i> - and <i>Rnf2</i> -mediated lineage restriction <i>in vivo</i>	191
5.3	Discussion	198
5.3.1	<i>Bcor</i> and ESC conversion to TSC.....	198
5.3.2	<i>Rnf2</i> and ESC conversion to TSC.....	200
5.3.3	Role of <i>Bcor</i> and <i>Rnf2</i> in the maintenance of the first cell lineage decision in ESCs	201
5.3.4	Future Perspectives	202
6	DISCUSSION AND FUTURE DIRECTIONS.....	207
6.1	Summary	207
6.1.1	Proof-of-concept studies	207
6.1.2	Genome-wide CRISPR/Cas9 loss-of-function screen identifies new suppressors of ESC differentiation to trophoblast	208
6.1.3	<i>Bcor</i> and <i>Rnf2</i> in the lineage barrier preventing ESC differentiation to TSC	209
6.2	Future Perspectives	210

6.2.1	The potential of genome-wide CRISPR/Cas9 screens as tools for dissection of lineage specification and restriction	210
6.2.2	ncPRC1.1 as key player in the maintenance of the first cell lineage decision	212

LIST OF FIGURES

Figure 1.1 Preimplantation mouse embryo development.	2
Figure 1.2 Models for the first cell lineage decision.	6
Figure 1.3 Major signalling and transcriptional networks controlling TE and ICM segregation.....	9
Figure 1.4 The second cell lineage specification: Epiblast (Epi) versus Primitive Endoderm (PrE).	13
Figure 1.5 Stem cell lines can be established from blastocyst stage mouse embryos and cultured indefinitely <i>in vitro</i>	15
Figure 1.6 Early placenta development and trophoblast differentiation.	21
Figure 1.7 Signalling and transcriptional regulation of self-renewal in TSCs.	24
Figure 1.8 Epigenetic regulation of TSCs.	27
Figure 1.9 Schematic representation of type II CRISPR/Cas9 mediated genome editing.	36
Figure 3.1 Epigenetic restriction of the first cell lineage decision.	71
Figure 3.2 Schematic representation of the strategy used to generate the <i>Elf5::Venus</i> reporter ESC line.	75
Figure 3.3 Generation of the <i>Elf5::Venus</i> knock-in ESC line.	77
Figure 3.4 Absence of Cas9-mediated indels in the three predicted exonic off-targets of gRNA #1.	78
Figure 3.5 Analysis of <i>Elf5</i> promoter DMR methylation status upon gene targeting demonstrates it remains highly methylated.....	79
Figure 3.6 Validation of the <i>Elf5::Venus</i> reporter ESC line, through CRISPR/Cas9-mediated <i>Oct4</i> knockout followed by induction of trophoblast differentiation.	81
Figure 3.7 Selection of gRNAs to efficiently target the <i>Oct4</i> locus in ESCs.	85
Figure 3.8 Using lentiviral delivery of gRNA against <i>Oct4</i> allows for differentiation towards trophoblast lineage.	87
Figure 3.9 Initiating differentiation 3 days post-transduction captures a higher percentage of frame-shifting mutations of <i>Oct4</i> but does not increase trophoblast differentiation efficiency.	89
Figure 3.10 Immunostaining of differentiated cells confirms co-expression of Venus and TSC markers <i>Elf5</i> and <i>Cdx2</i>	91
Figure 3.11 Selection of gRNAs to efficiently target the <i>Dnmt1</i> locus in ESCs.	93
Figure 3.12 Knockout of <i>Dnmt1</i> through CRISPR/Cas9 does not allow for differentiation towards trophoblast lineage in the same setting as <i>Oct4</i> knockout.	95

Figure 3.13 Knocking out <i>Dnmt1</i> with a different range of gRNAs does not result in improved differentiation capacity.....	96
Figure 4.1 Genome-wide CRISPR/Cas9 knockout screening outline.....	111
Figure 4.2 Summary of the different screening replicates.....	113
Figure 4.3 CRISPR/Cas9 screenings could be performed in a reproducible manner.....	114
Figure 4.4 CRISPR/Cas9 knockout screening uncovers 42 genes involved on the differentiation of ESCs to TSCs.....	116
Figure 4.5 Venn diagram depicting enriched genes detected with screen v1 or v2, and the added effect of the four replicates.....	117
Figure 4.6 Gene set enrichment analysis of the 42 hits from the screening.....	119
Figure 4.7 STRING Network analysis of screening hits.....	121
Figure 4.8 Shortlisting candidate genes for validation.....	122
Figure 4.9 Validation of candidate genes through lentiviral delivery of gRNAs.....	123
Figure 4.10 TSC differentiation time-course for candidate genes validated through low MOI lentiviral transductions.....	125
Figure 4.11 Stable integration of multiple copies of gRNAs via PBase transposition results in increased differentiation efficiency compared to the use of lentivirus at an MOI 0.3.....	126
Figure 4.12 Validation of candidate genes using PBase to mediate integration of two gRNAs per gene.....	128
Figure 4.13 TSC differentiation time-course for candidate genes with strong phenotypes.....	129
Figure 4.14 Kinetics of TSC differentiation for candidate genes with an intermediate phenotype.....	130
Figure 4.15 Representative live microscopy of Venus expressing cells by day 14 of TSC differentiation in the indicated knockouts.....	132
Figure 4.16 STRING network analysis of validated hits with expanded interactors.....	133
Figure 4.17 Expression profiles for selected genes in preimplantation mouse embryo development.....	135
Figure 5.1 Composition of mammalian PcG complexes.....	146
Figure 5.2 Expression profiling of <i>Bcor</i> knockout and its differentiation to TSC lineage.....	154
Figure 5.3 Expression profiling of TSC differentiation upon <i>Oct4</i> or <i>Bcor</i> knockout.....	156
Figure 5.4 A stable TSC-like line could be obtained upon <i>Bcor</i> knockout in ESCs followed by differentiation to TSC.....	158
Figure 5.5 Characterisation of the TSC-like line established from ESCs with <i>Bcor</i> knockout demonstrates a similar profile to wild-type TSCs.....	160
Figure 5.6 Transcriptomics analysis of TSC-like cells shows they resemble bona-fide TSCs.....	162
Figure 5.7 Terminal differentiation of TSC-like cells indicate they can differentiate, but with lower efficiency compared to TSCwt.....	163
Figure 5.8 Generation of clonal <i>Bcor-KO Eif5::Venus</i> ESCs.....	165

Figure 5.9 Characterisation of the clonal <i>Bcor-KO</i> ESCs generated.	167
Figure 5.10 ESC lineage barrier can be partially rescued upon overexpression of <i>Bcor</i> in <i>Bcor-KO</i> ESCs.	168
Figure 5.11 <i>Bcor</i> maintains trophoblast lineage restriction in preimplantation epiblast-like naïve ESCs.	170
Figure 5.12 Expression profiling of <i>Rnf2</i> knockout and its differentiation to TSC lineage.	172
Figure 5.13 Expression profiling of TSC differentiation upon <i>Rnf2</i> knockout.	173
Figure 5.14 Generation of clonal <i>Rnf2-KO Venus::Elf5</i> ESCs.	175
Figure 5.15 Characterisation of the clonal <i>Rnf2-KO</i> ESCs generated.	177
Figure 5.16 ESC lineage barrier can be fully restored upon overexpression of <i>Rnf2</i> in <i>Rnf2-KO</i> ESCs.	179
Figure 5.17 <i>Rnf2</i> maintains trophoblast lineage restriction in preimplantation epiblast-like naïve ESCs.	180
Figure 5.18 Transcriptome profile for <i>Bcor-KO</i> ESCs.	183
Figure 5.19 Transcriptome profile for <i>Rnf2-KO</i> ESCs.	185
Figure 5.20 Common transcriptional changes in <i>Bcor-KO</i> and <i>Rnf2-KO</i> ESCs.	186
Figure 5.21 Principal Component Analysis (PCA) of differentiating <i>Bcor-KO</i> (A) and <i>Rnf2-KO</i> (B) cells.	187
Figure 5.22 Transcriptomics profiling of differentiated cells by day 18.	189
Figure 5.23 CHIPseq analysis of ncPRC1.1 binding sites in hESCs reveals a direct link to repression of trophoblast lineage markers <i>Eomes</i> , <i>Cdx2</i> and <i>Gata3</i> , but not <i>Elf5</i>	190
Figure 5.24 Preliminary results indicate <i>Bcor-KO</i> ESCs can be found in the extraembryonic compartment in chimeric embryos.	192
Figure 5.25 Live Images from E6.5 Chimeric Embryos resultant from injections of <i>Bcor-KO</i> ESC clone 2 into 8-cell host embryos.	193
Figure 5.26 Live Images from E6.5 Chimeric Embryos resultant from injections of <i>Bcor-KO</i> ESC clone 11 into 8-cell host embryos.	194
Figure 5.27 Preliminary results indicate <i>Rnf2-KO</i> ESCs can be found in the extraembryonic compartment in chimeric embryos.	195
Figure 5.28 Live Images from E6.5 Chimeric Embryos resultant from injections of <i>Rnf2-KO</i> ESC clones into 8-cell host embryos.	196

