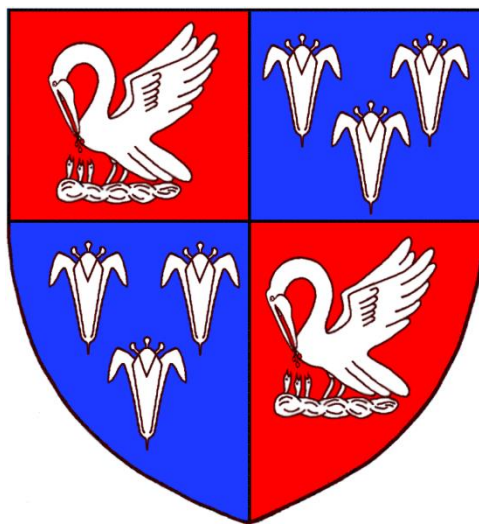


*DEFINING THE ESSENTIAL
REGULATORS OF NAÏVE HUMAN
PLURIPOTENT STEM CELL
REPROGRAMMING*

Adam Bendall

Corpus Christi College



University of Cambridge

Babraham Institute

This dissertation is submitted for the degree of Doctor of Philosophy

July 2022

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed the prescribed word limit for the Faculty of Biology Degree Committee of 60,000 words.

Abstract

Human pluripotent stem cells (PSCs) exist in two distinct states – naïve and primed, which are interconvertible. Naïve PSCs uniquely harbour several desirable properties that primed PSCs do not, including their ability to model human pre-implantation biology and developmentally-regulated epigenetic phenomena *in vitro*. Naïve PSCs also have an expanded differentiation potential to produce cell types with possible therapeutic applications.

Naïve PSCs are predominantly generated by reprogramming primed PSCs. This project aimed to address a significant gap in our knowledge about the molecular events that underpin reprogramming, which could assist in unlocking the full potential of naïve PSCs. A genome-wide screen identified novel regulators of naïve cell reprogramming, unexpectedly highlighting the Polycomb Repressive Complex 1 (PRC1) subtype PRC1.3. In this project, I sought to understand the role of PRC1.3 in human pluripotency and reprogramming, and determine the PRC1.3-dependent mechanisms that modulate the acquisition of naïve pluripotency.

I used CRISPR-Cas9 to delete the core PRC1.3 complex component *PCGF3* in primed PSCs and ascertained that PRC1.3 is dispensable for human primed pluripotency but critical for naïve cell reprogramming. I used ChIP-sequencing to identify genes that gain PRC1.3 occupancy during reprogramming and these genes become transcriptionally repressed as cells successfully reprogramme, suggesting a key gene silencing role for PRC1.3. Deletion of PRC1.3 reduces the global abundance of the PRC1-deposited H2AK119ub1 modification, and levels of PRC2-deposited H3K27me3 also, in primed PSCs.

Using a proteomics approach, I identified a novel, naïve-specific interaction between PRC1.3 and PRDM14, which cooperate in transcriptional repression of PRC1.3 target genes, I also discovered a developmentally-regulated composition shift in the PRC1.3 complex between human naïve and primed PSCs, which was corroborated in early human embryo datasets. This change in PRC1.3 composition does not appear to alter the enzymatic activity of PRC1.3 *in vitro*.

Overall, I identified that the PRC1.3 complex plays a critical role in human naïve cell reprogramming. The PRC1.3 complex influences the human primed PSC epigenome and is associated with transcriptional repression, with the complex undergoing both intrinsic and exogenous regulation across the cell state transition. These findings advance our knowledge of molecular events that underpin successful naïve cell reprogramming, and this progress could optimise our ability to produce human naïve PSCs, as an *in vitro* developmental model and a cell type with considerable therapeutic potential.

Acknowledgements

Throughout my PhD, I have received so much warmth and kindness, all of which has enhanced this wonderful experience immeasurably. None of this work would have been possible without funding from the BBSRC DTP, the fantastic core facilities at the Babraham Institute and collaborators, both in Cambridge and internationally. Everyone in the Epigenetics ISP has been nothing but delightful, and I will particularly miss my chats in tissue culture with Amy, Aled, Chris, Brendan and Jasmin. I would also like to thank my internship colleagues at Page Medical for taking me under their wing, making me a valued member of their team and providing me with a new perspective on science and learning.

I would particularly highlight everyone in the Rugg-Gunn lab. They have been a pleasure to work with, without exception, and have made coming to work every day (even at weekends!), a joy. The intimidatingly impressive work ethic and intense patience of Mandy and Charlene were instrumental in my training, and I am thankful to them for the skills that I have acquired. I also thank Claudi and Matteo for their dedicated mentorship and my fellow PhD companions Andrew and Kate for their friendship and humour.

I must also thank my first-class supervisor, Peter. His infectious enthusiasm for science was the reason I applied for this project and was a consistent source of motivation throughout the years, even during the bleakest points of lockdown and when I was facing my biggest experimental obstacles. His balance of ambition, approachability, and encouragement is something that I hope to emulate one day.

As well as professionally, I have also grown so much personally these past four years. For this, I must thank my friends, and especially my girlfriend, Genevieve, for their love and loyalty, despite the years of cancelled plans and very late replies to messages! Last but certainly not least, my amazing family. I am particularly thankful to my parents for instilling within me the value of education, as well as their support, sacrifice, and love.

Table of Acknowledgements

Initial training in techniques and laboratory practice and subsequent mentoring:

Aled Parry, Amanda Collier, Amy Wilkinson, Andrew Malcolm, Charlene Fabian, Christine Hall, Clara Novo, Claudia Semprich, Irene Zorzan, Katarzyna Wojdyla, Kate Maskalenka, Matteo Molè, Peter Rugg-Gunn, Stephen Bevan and Yang Wang

Data obtained from a technical service provider:

Histone post-translational modification proteomics – Sigrid Verhelst and Maarten Dhaenens at the Laboratory of Pharmaceutical Biotechnology, University of Ghent

qPLEX-RIME sample processing, mass spectrometry and analysis – Valar Nila Roamio Franklin, Kamal Kishore and Clive D’Santos – Cancer Research UK Cambridge Institute, Proteomics and Bioinformatics Facility

Imaging facility training – Simon Walker

Flow cytometry facility and sorting – Rebecca Roberts and Aleksandra Lazowska-Addy

Next-Generation Sequencing facility – Nicole Forrester and Paula Kokko-Gonzalez

Data produced jointly:

Generation of CRISPR knockout cell lines – Amanda Collier and Charlene Fabian

Generation of rescue clone – Charlene Fabian

Generation of RNA- and ChIP-sequencing libraries – Claudia Semprich

Calibrated CUT&Tag data – Yang Wang

ChIP-sequencing data analysis – Peter Rugg-Gunn

Data or materials provided by someone else:

Genome-wide CRISPR-Cas9 screen – Amanda Collier

ChIP-sequencing and CUT&Tag plots – Andrew Malcolm

Table of Contents

Declaration.....	2
Abstract.....	3
Acknowledgements	4
Table of Acknowledgements	5
Table of Contents.....	6
List of Figures	12
List of Tables	20
List of Abbreviations.....	22
1 – Introduction	24
1.1 – Pluripotency and pluripotent stem cells.....	25
1.1.1 – Human development and embryonic pluripotency	25
1.1.2 – Mouse embryonic stem cells	27
1.1.3 – Human pluripotent stem cells	27
1.1.4 – Human induced pluripotent stem cells	28
1.2 – Ground state pluripotency	29
1.2.1 – Mouse ground state pluripotency	29
1.2.2 – Primed human PSCs.....	33
1.2.3 – Human naïve pluripotency.....	34
1.2.4 – Signalling requirements of human naïve PSCs	35
1.3 – Differences between human primed and naïve PSCs	37
1.3.1 – Pluripotency transcriptional networks	38
1.3.2 – DNA methylation	39
1.3.3 – X chromosome activation status.....	39
1.3.4 – Enhancer usage.....	40
1.3.5 – Metabolism.....	41
1.4 – Generation of naïve PSCs by primed to naïve reprogramming (PTNR)	42
1.4.1 – 5i/L/A reprogramming.....	43
1.4.2 – t2iL + PKCi reprogramming and Chemical Resetting.....	43
1.4.3 – Comparison of PTNR methods.....	44
1.5 – Unique functions of human naïve PSCs.....	46
1.5.1 – Generation of extra-embryonic lineages.....	46
1.5.2 – Generation of human blastocyst-like ‘blastoid’ structures	47
1.5.3 – Generation of early amniotic cell types	48
1.5.4 – Defining the regulators of human PTNR could unlock the full potential of human naïve pluripotency	49

1.5.5 – Using cell surface marker proteins to identify human naïve and primed PSCs .	50
1.5.6 – Using genome-wide CRISPR-Cas9 screening to identify key regulators of human PTNR	52
1.5.6 – PRC1.3 is one of the most critical epigenetic complexes for human PTNR	53
1.6 – Polycomb Repressive Complexes	57
1.6.1 – Histone post-translational modifications and epigenetics.....	57
1.6.2 – Identification of Polycomb group genes in <i>Drosophila</i>	58
1.6.3 – Homology between <i>Drosophila</i> and mammalian PRC1 complexes	59
1.6.4 – Polycomb Repressive Complex 2.....	59
1.6.5 – Polycomb Repressive Complex 1.....	60
1.6.6 – Diversity in Polycomb Repressive Complex 1 composition.....	61
1.7 – Cooperativity between PRC1 and PRC2 complexes.....	64
1.7.1 – Canonical and variant PRC1 complexes	64
1.7.2 – H3K27me3 independent recruitment of variant PRC1 complexes	64
1.7.3 – H2AK119Ub1-dependent PRC2 complex recruitment	65
1.8 – Polycomb Repressive Complex 1 function	68
1.8.1 – Regulation of transcription by H2AK119ub1	68
1.8.2 – Developmental importance of PRC1 complexes.....	68
1.8.3 – Catalysis-dependent Polycomb Repressive Complex 1 functions.....	69
1.8.4 – Catalysis-independent Polycomb Repressive Complex 1 functions.....	70
1.9 – PRC1.3 and PRC1.5 complexes.....	72
1.9.1 – PRC1.3 and PRC1.5 complex function.....	72
1.9.2 – X-chromosome inactivation	72
1.9.3 – Transcriptional repression	73
1.9.4 – Transcriptional activation.....	74
1.10 – Thesis aims	77
2 – Materials and Methods.....	79
2.1 – Cell Lines.....	80
2.1.1 – Cell line characterisation	80
2.2 – Cell Culture Techniques	81
2.2.1 – Generation of Mouse Embryonic Fibroblasts (MEF)	81
2.2.2 – Primed human PSC culture	82
2.2.3 – Naïve human PSC culture.....	83
2.2.4 – Cryopreservation of cells.....	85
2.2.5 – Primed to naïve reprogramming (PTNR)	85
2.2.6 – <i>Drosophila</i> Schneider 2 (S2) culture.....	88
2.3 – Flow Cytometry.....	89
2.3.1 – Flow Cytometry Analysis	89

2.3.2 – Fluorescence Activated Cell Sorting (FACS)	90
2.4 – Transfections	92
2.5 – Molecular Biology	93
2.5.1 – Plasmid Cloning	93
2.5.2 – Genotyping and Polymerase Chain Reaction	98
2.5.3 – RNA extraction and cDNA synthesis	99
2.5.4 – Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR) ...	99
2.5.5 – Western Blotting.....	100
2.5.6 – Immunofluorescence	102
2.5.7 – Co-immunoprecipitation	103
2.6 – Proteomics	106
2.6.1 – Quantitative Tandem Mass Tag-Pro 16plex Rapid Immunoprecipitation Mass Spectrometry of Endogenous Protein (qPLEX-RIME)	106
2.6.2 – qPLEX-RIME Analysis.....	109
2.6.3 – Histone proteomics sample preparation	110
2.6.4 – Histone proteomics analysis.....	110
2.6.5 – <i>In vitro</i> histone H2A Lysine 119 ubiquitylation assay	110
2.7 – Next Generation Sequencing.....	112
2.7.1 – Cross-linked chromatin immunoprecipitation (ChIP)-sequencing.....	112
2.7.2 – ChIP-sequencing analysis.....	115
2.7.3 – Calibrated Cleavage Under Targets and Tagmentation (CUT&Tag).....	116
2.7.4 – Calibrated CUT&Tag analysis	119
2.7.5 – RNA-sequencing	120
2.7.6 – RNA-sequencing analysis	120
2.7.7 – Gene ontology analysis	121
2.7.8 – Motif analysis	121
2.7.9 – Gene set enrichment analysis (GSEA)	121
3 – PRC1.3 is essential for human naïve pluripotent stem cell reprogramming	123
3.1 – Introduction.....	124
3.1.1 – Characterisation of PRC1.3 complex action in human pluripotency and reprogramming	124
3.1.2 – Aims.....	125
3.2 – Investigating the role of PRC1.3 in human primed pluripotency	126
3.2.1 – Analysing the specific requirement for PRC1.3 from the genome-wide screen of human PTNR	126
3.2.2 – Assessing PRC1.3 component gene expression in human embryos and PSCs	128
3.2.3 – Generating PRC1.3-deficient human primed PSCs	130

3.2.4 – Testing the impact of <i>PCGF3</i> deletion on the association of the PRC1.3 complex	133
3.2.5 – Characterising the primed state of pluripotency in PRC1.3-deficient PSCs	134
3.2.6 - <i>PCGF3</i> deletion causes very few transcriptional changes in primed PSCs	139
3.3 – The PRC1.3 complex is essential for successful human PTNR.....	141
3.3.1 – PRC1.3-deficient primed PSCs fail to produce KLF17-expressing colonies during 5i/L/A reprogramming	141
3.3.2 – PRC1.3-deficient primed PSCs fail to produce upregulate naïve-specific cell-surface markers during 5i/L/A reprogramming	144
3.3.3 – PRC1.3 is also essential for PTNR by chemical resetting.....	148
3.4 – Confirming that PRC1.3 is essential for PTNR using <i>FBRS</i> KO primed human PSCs	154
3.5 – Generating PRC1.3 knockdown naïve PSCs	161
3.6 – Chapter Discussion.....	167
3.6.1 – PRC1.3 is dispensable for the analysed characteristics of human primed PSCs	167
3.6.2 – The absence of PRC1.3 causes cells to fail to reprogramme to naïve PSCs ..	168
3.6.3 – It is not known whether PRC1.3 specifically regulates PTNR or pluripotency cell state transitions more broadly	169
3.6.4 – A naïve PRC1.3 knockdown system could address a number of outstanding questions	170
4 – PRC1.3-dependent transcriptional repression is associated with successful human PTNR.....	174
4.1 – Introduction.....	175
4.1.1 – Studying the molecular function of the PRC1.3 complex in human PTNR	175
4.1.2 – Studying the epigenetic function of PRC1.3 complexes in human primed PSCs	177
4.1.3 – Aims.....	178
4.2 – <i>PCGF3</i> KO primed PSCs fail to upregulate naïve pluripotency genes	179
4.2.1 - WT and <i>PCGF3</i> KO cells transcriptionally diverge during reprogramming.....	179
4.2.2 - <i>PCGF3</i> KO cells upregulate neural marker genes during reprogramming	181
4.3 – PRC1.3 target genes were identified by ChIP-sequencing.....	186
4.3.1 - Generating PRC1.3 and PRC1.5-deficient (DKO) human primed PSCs.....	186
4.3.2 – ChIP-sequencing identifies PRC1.3 target genes.....	187
4.4 – PRC1.3 target genes are downregulated during reprogramming	193
4.4.1 - PRC1.3 occupancy is associated with gene downregulation during PTNR.....	193
4.4.2 - Failure to repress PRC1.3 targets is associated with a failure to reprogramme	196
4.5 – PRC1.3 influences the epigenetic landscape of primed PSCs	200
4.5.1 - H2AK119ub1 and H3K27me3 are more abundant in human naïve PSCs	200
4.5.2 - DKO primed PSCs show reduced levels of H2AK119ub1 and H3K27me3	201

4.5.3 - PRC1.3-deficiency influences the histone H2A variants present within nucleosomes.....	208
4.6 – Discussion	210
4.6.1 – WT and <i>PCGF3</i> KO cells transcriptionally diverge during PTNR	210
4.6.2 – PRC1.3 naïve target genes are repressed during reprogramming.....	210
4.6.3 – The PRC1.3 complex influences the primed PSC epigenome	212
5 – PRC1.3 undergoes pluripotency state-specific regulation between human naïve and primed PSCs.....	214
5.1 – Background	215
5.1.1 – Investigating the PRC1.3 protein interaction network in human PSCs	215
5.1.2 – Investigating the PRC1.3 complex catalytic activity in primed and naïve PSCs	216
5.1.3 – Aims.....	217
5.2 – Using qPLEX-RIME to identify PRC1.3 complex interactors	218
5.2.1 – Optimisation of the qPLEX-RIME methodology in human PSCs.....	218
5.2.2 – qPLEX-RIME can identify the PRC1.3 complex.....	221
5.2.3 - qPLEX-RIME identifies PRDM14 as a naïve-specific PRC1.3 interactor	225
5.3 – PRC1.3 and PRDM14 cooperate to repress transcription during human PTNR	229
5.3.1 – PRC1.3 and PRDM14 interact in naïve PSCs	229
5.3.2 – PRDM14 binds with PRC1.3 at naïve target genes	231
5.3.3 – <i>PRDM14</i> is also essential for human PTNR	234
5.3.4 – PRDM14 and PRC1.3 cooperate to repress PRC1.3 target genes	238
5.4 – The PRC1.3 complex undergoes a developmentally-regulated composition switch	243
5.4.1 – <i>FBRS</i> and <i>AUTS2</i> are enriched in human naïve and primed PSCs respectively	243
5.4.2 – <i>FBRS</i> and <i>AUTS2</i> expression is mutually exclusive in human PSCs and embryos.....	249
5.5 – Interrogating PRC1.3 enzymatic activity in primed and naïve PSCs	255
5.6 – Discussion	260
5.6.1 – PRC1.3 and PRDM14 cooperate in transcriptional repression during reprogramming	260
5.6.2 – <i>AUTS2</i> and <i>FBRS</i> undergo a paralogue shift in human PSCs and development	261
5.6.3 – The PRC1.3 complex does not appear to differ in catalytic activity between the primed and naïve state of pluripotency.....	264
6 – Summary and Conclusions	267
6.1 – Reprogramming remains an enigmatic process	268
6.2 – PRC1.3 is essential for human PTNR	269
6.3 – PRC1.3 is associated with transcriptional repression during reprogramming	272

6.4 – PRC1.3 undergoes pluripotency state-specific regulation	273
6.5 – Final conclusions	276
Bibliography	278

List of Figures

Figure 1.1.1 – The first ten days of human embryonic development

Figure 1.1.2 – Lineage specification and differentiation of human epiblast cells

Figure 1.2.1 – LIF signalling in PSCs

Figure 1.2.2 – BMP4 signalling in PSCs

Figure 1.2.3 – MAPK/ERK signalling in PSCs

Figure 1.2.4 – Wnt signalling in PSCs

Figure 1.2.5 – TGF- β signalling in PSCs

Figure 1.2.6 – Developmental progression of human and mouse epiblast derived PSCs

Figure 1.3.1 – Epigenetic and molecular differences between human naïve and primed PSCs

Figure 1.5.1 – The applications of the extended differentiation potential of naïve PSCs

Figure 1.5.2 – Naïve and primed human PSC specific cell-surface markers

Figure 1.5.3 – Screening strategy to identify regulators of naïve PSC reprogramming

Figure 1.5.4 – Ranked differential-enrichment (DE) plots comparing successfully reprogrammed (nascent naïve) versus failed to reprogramme (refractory) populations

Figure 1.5.5 – Gene ontology analysis of the ‘essential’ regulators of naïve PSC reprogramming

Figure 1.6.1 – The mammalian ubiquitin transfer system

Figure 1.6.2 – Diversity of mammalian PRC1 complexes

Figure 1.7.1 – Interdependence of PRC1 and PRC2 complexes in chromatin recruitment

Figure 1.9.1 – Proposed modes of transcriptional regulation by AUTS2-PRC1.3/5 complexes

Figure 3.2.1 – Chart shows the significance of all PRC1 complex component genes in human PTNR

Figure 3.2.2 – Charts show the p-values of PRC2 complex components as essential regulators of human PTNR

Figure 3.2.3 – Violin plots show the expression level distribution of PRC1.3/5 complex components in single cells throughout human preimplantation embryo development

Figure 3.2.4 – Charts show *PCGF3* and *PCGF5* expression levels in log₂ RPKM (reads per kilobase transcript per million reads), in primed and naïve PSCs

Figure 3.2.5 – Schematic of the CRISPR-Cas9 strategy used to create *PCGF3* KO primed PSCs

Figure 3.2.6 – Electrophoresis image of PCR-based genotyping of the *PCGF3* gene

Figure 3.2.7 – Multiple sequence alignment of Sanger sequencing identifies successful exon deletion within *PCGF3*

Figure 3.2.8 – Western blotting shows the absence of PCGF3 protein in two KO primed PSC lines

Figure 3.2.9 – Co-immunoprecipitation with RING1B antibody shows PRC1.3 complex dissociation in *PCGF3* KO primed PSCs

Figure 3.2.10 – Phase contrast images show the morphology of WT and *PCGF3* KO primed PSCs

Figure 3.2.11 – Immunofluorescence microscopy shows the uniform expression of OCT4 and NANOG in WT and both *PCGF3* KO primed PSC lines

Figure 3.2.12 – SSEA-4 flow cytometry in WT and *PCGF3* KO primed PSC lines

Figure 3.2.13 – Charts shows proliferation rates of WT and both *PCGF3* KO primed PSCs

Figure 3.2.14 – Karyotyping of WT and two *PCGF3* KO primed PSC clones

Figure 3.2.15 – Gene expression is nearly identical between WT and *PCGF3* KO primed PSCs

Figure 3.2.16 – Chart shows gene expression similarity between WT and *PCGF3* KO primed PSCs

Figure 3.3.1 – Phase contrast images of parental WT and *PCGF3* KO PSCs at day 2 and day 10 of PTNR in 5i/L/A conditions

Figure 3.3.2 – Immunofluorescent microscopy of WT and *PCGF3* KO cells at day 10 of PTNR (5i/L/A conditions)

Figure 3.3.3 – Cell colony classification from Fig. 3.3.2 based on OCT4 and KLF17 expression

Figure 3.3.4 – Flow cytometry gating strategy to identify nascent naïve cell populations

Figure 3.3.5A – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 14 of PTNR in 5i/L/A conditions

Figure 3.3.5B – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 14 of PTNR in 5i/L/A conditions

Figure 3.3.5C – Chart showing the quantification of flow cytometry data in Fig. 3.3.5B

Figure 3.3.6A – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 24 of PTNR in 5i/L/A conditions

Figure 3.3.6B – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 24 of PTNR in 5i/L/A conditions

Figure 3.3.11 – Western blotting shows the production of a *PCGF3* rescue clone

Figure 3.3.12 – Electrophoresis image of PCR-based genotyping of the *PCGF3* gene

Figure 3.3.13 – Phase contrast images of parental WT, *PCGF3* KO, *PCGF3* Het and *PCGF3* rescue PSCs at day 2 and day 14 of PTNR in chemical resetting conditions

Figure 3.3.14A – Flow cytometry contour plots of parental WT, *PCGF3* heterozygous (Het), *PCGF3* KO and *PCGF3* rescue PSCs at day 10 of PTNR in chemical resetting (CR) conditions

Figure 3.3.14B – Flow cytometry contour plots of parental WT, *PCGF3* heterozygous (Het), *PCGF3* KO and *PCGF3* rescue PSCs at day 10 of PTNR in CR conditions

Figure 3.3.14C – Chart showing the quantification of flow cytometry data shown in Fig. 3.4.14B

Figure 3.4.1 – Schematic of the CRISPR-Cas9 strategy for the creation of *FBRS* KO primed PSCs

Figure 3.4.2 – Western blotting does not detect *FBRS* protein in the KO primed PSC line

Figure 3.4.3 – Phase contrast images showing the morphology of WT and *FBRS* KO primed PSCs

Figure 3.4.4 – Immunofluorescence microscopy shows the uniform expression of OCT4 and NANOG in WT and *FBRS* KO primed PSC lines

Figure 3.4.5 – SSEA-4 flow cytometry in WT and *FBRS* KO primed PSC lines

Figure 3.4.6 – Karyotyping of the *FBRS* KO primed PSC line

Figure 3.4.7 – Phase contrast images of parental WT and *FBRS* KO PSCs at day 2 and day 10 of PTNR in CR conditions

Figure 3.4.8A – Flow cytometry contour plots of parental WT and *FBRS* KO PSCs on day 12 of PTNR in CR conditions

Figure 3.4.8B – Flow cytometry contour plots of parental WT and *FBRS* KO PSCs on day 12 of PTNR in CR conditions

Figure 3.4.8C – Chart showing the quantification of flow cytometry data shown in Fig. 3.4.8B

Figure 3.5.1 – Schematic of CRISPRi-mediated inducible *PCGF3* knockdown

Figure 3.5.2 – mCherry flow cytometry in Doxycycline treated and untreated CRISPRi primed iPSCs

Figure 3.5.3 – Mapping of the guide RNAs predicted to successfully knockdown *PCGF3* expression

Figure 3.5.4 – Fluorescence microscopy confirms successful vector transfection

Figure 3.5.5 – RT-qPCR showing *PCGF3* expression across a Doxycycline addition time-course

Figure 3.5.6 – RT-qPCR showing *PCGF3* expression across eight CRISPRi clones

Figure 4.2.1 – Correlation matrix of RNA-sequencing dataset at days 0 and 10 of PTNR in 5i/L/A conditions

Figure 4.2.2 – Principal component analysis (PCA) of RNA-sequencing data at days 0 and 10 of PTNR in 5i/L/A conditions

Figure 4.2.3 – MA plot comparing WT and *PCGF3* KO cells on day 10 of PTNR in 5i/L/A conditions

Figure 4.2.4 – Charts show gene expression levels in log₂ RPKM (reads per kilobase transcript per million reads), in WT and *PCGF3* KO cells at days 0 and 10 of PTNR in 5i/L/A conditions

Figure 4.2.5 – RT-qPCR showing gene expression levels in WT and *PCGF3* KO cells at days 0 and 10 of PTNR in 5i/L/A conditions

Figure 4.2.6 – Clustering of RNA-sequencing within differentially expressed genes

Figure 4.3.1 – Schematic for the creation of *PCGF3/5* DKO primed PSCs by *PCGF5* gene deletion

Figure 4.3.2 – Electrophoresis image of PCR-based genotyping of the *PCGF5* gene

Figure 4.3.3 – Metaplot of PCGF3 ChIP-seq. signal over naïve-specific and primed-specific target genes in naïve PSCs, WT and DKO primed PSCs

Figure 4.3.4 – Heatmaps of normalised PCGF3 and RING1B ChIP-sequencing read counts within a 20 kb transcription start site (TSS)-centred window in naïve and primed PSCs

Figure 4.3.5 – Metaplot of H2AK119ub1 CUT&Tag signal over naïve-specific PRC1.3 target genes in naïve PSCs, WT and DKO primed PSCs

Figure 4.3.6 – Heatmaps of normalised PCGF3 and RING1B ChIP-seq., and H2AK119ub1 and IgG CUT&Tag, read counts over a 20 kb TSS-centred window in naïve, WT and DKO primed PSCs

Figure 4.3.7 – Chromatin profiling shows *LETM1* is a naïve-specific PRC1.3 target gene

Figure 4.4.1 – Violin plot showing the fold-change in the expression of PRC1.3 target genes in naïve compared to primed PSCs

Figure 4.4.2 – Chart shows the gene expression change of PRC1.3 target genes between naïve and primed PSCs

Figure 4.4.3 – Violin plots showing the fold-change in expression of PRC1.3 target genes in WT and *PCGF3* KO cells at day 0 compared to day 10 of PTNR (5i/L/A conditions)

Figure 4.4.4 – Heatmap showing the row normalised expression levels for a subset of naïve PRC1.3 target genes in WT and *PCGF3* KO cells at days 0 and 10 of PTNR (5i/L/A conditions)

Figure 4.4.5 – Violin plots showing the fold-change in the expression of PRC1.3 target genes in cell-sorted populations at day 0 (primed) and day 10 (nascent naïve and refractory) of PTNR (5i/L/A conditions)

Figure 4.5.1 – Western blotting shows the abundance of H2AK119ub1 in naïve and primed PSCs

Figure 4.5.2 – Histone proteomics shows the relative abundance of the H3K27me3 modification in naïve and primed PSCs

Figure 4.5.3 – Western blotting shows the abundance of H2AK119ub1, with and without PRC2 inhibition, in primed PSCs

Figure 4.5.4 – Western blotting shows the abundance of H2AK119ub1 in WT and *PCGF3/5* DKO primed PSCs

Figure 4.5.5 – Kernel density estimate of normalised H2AK119ub1 CUT&Tag reads in naïve, WT and DKO primed PSCs

Fig. 4.5.6 – A density plot showing H2AK119ub1 CUT&Tag reads across chromosome 18

Figure 4.5.7 – Histone proteomics show the relative abundance of the H3K27me1, H3K27me2, H3K27me3 and H3K27ac modifications in WT and DKO primed PSCs

Figure 4.5.8 – Histone proteomics shows the relative abundance of all detected histone H3 and H4 modifications detected in WT and DKO primed PSCs

Figure 4.5.9 – Histone proteomics shows the relative abundance of histone H2A variants in WT and DKO primed PSCs

Figure 5.2.1 – Amino acid multiple sequence alignment shows similarity between human and mouse PCGF3

Figure 5.2.2 – RIME shows PCGF3 peptides can be detected in *PCGF3/5* DKO primed PSCs grown on MEFs

Figure 5.2.3 – Immunofluorescence detects PCGF3/5 in MEFs and human primed PSCs

Figure 5.2.4 – Flow cytometry confirms the MACS-dependent depletion of Cd90.2 positive MEFs and enrichment of CD147 positive human cells

Figure 5.2.5 – Correlation matrix of qPLEX-RIME dataset

Figure 5.2.6 – Box and whisker plot shows the peptide abundance intensity in each biological replicate

Figure 5.2.7 – Principal Components Analysis of the 14-plex qPLEX-RIME data

Figure 5.2.8 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change (FC) in abundance and significance of PCGF3-interacting proteins between WT and DKO in primed PSCs

Figure 5.2.9 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change and the significance of proteins identified as key for human PTNR when comparing primed and 48 h 5i/L/A PSCs

Figure 5.2.10 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change and the significance of proteins identified as key for human PTNR when comparing primed and naïve PSCs

Figure 5.3.1 – Co-immunoprecipitation-Western blotting confirms an interaction between PRDM14 and PCGF3 in naïve PSCs

Figure 5.3.2 – Motif enrichment analysis for PRC1.3 target sites in naïve PSCs

Figure 5.3.3 – Normalised PCGF3, RING1B and PRDM14-AID-VENUS ChIP-sequencing read counts, displayed in a heatmap, within a 20 kb peak-centered window in naïve PSCs

Figure 5.3.4 – The distribution of the PRDM14-AID-VENUS ChIP-sequencing normalised read counts are shown in a scatter plot, with and without IAA-induced PRDM14 degradation

Figure 5.3.5 – Flow cytometry histogram of SSEA4 expression in PRDM14-AID-VENUS primed PSCs

Figure 5.3.6 – Immunofluorescence microscopy images show the uniform expression of OCT4 and NANOG in PRDM14-AID-VENUS primed PSCs

Figure 5.3.7 – Karyotyping of the PRDM14-AID-VENUS primed PSC line

Figure 5.3.8 – Immunofluorescence in PRDM14-AID-VENUS primed PSCs in the presence and absence of IAA treatment, confirming PRDM14 depletion

Figure 5.3.9 – Flow cytometry contour plots of PRDM14-AID-VENUS PSCs after 10 days of chemical resetting, in the absence (left) or presence (right) of IAA-induced PRDM14 degradation

Figure 5.3.10 – Phase contrast images of PRDM14-AID-VENUS PSCs at day 2 and day 10 of naïve PSC reprogramming, with and without IAA (CR conditions)

Figure 5.3.11 – Immunofluorescence microscopy shows the uniform expression of OCT4, NANOG and KLF17 in PRDM14-AID-VENUS naïve PSCs

Figure 5.3.12 – Immunofluorescence microscopy in PRDM14-AID-VENUS naïve PSCs in the presence and absence of IAA treatment, confirming PRDM14 depletion

Figure 5.3.13 – Gene set enrichment analysis (GSEA) of naïve PRC1.3 target genes, ranked according fold change in transcription, following IAA-induced PRDM14 depletion

Figure 5.3.14 – Charts show log₂ RPKM of gene expression in PRDM14-AID-VENUS naïve PSCs, in the presence (no IAA) and absence (plus IAA) of PRDM14

Figure 5.4.1 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change in abundance and statistical significance of PCGF3-interacting proteins in naïve and primed PSCs

Figure 5.4.2 – Chart shows the fold change in abundance on PRC1.3 complex components using chromatin enrichment for proteomics (ChEP) between naïve and primed PSCs

Figure 5.4.3 – Heatmap shows PRC1.3/5 component expression levels during PTNR, on day 0 (primed PSCs), day 10 (nascent naïve PSCs) and day 34 (established naïve PSC lines)

Figure 5.4.4 – Heatmap shows *FBRS* and *AUTS2* expression levels over a naïve (day 0) to primed (day 10) capacitation time course

Figure 5.4.5 – Western blotting shows protein levels of PRC1.3 components in naïve and primed PSCs

Figure 5.4.6 – Chart shows the *FBRS* and *AUTS2* expression levels in primed PSCs, nascent naïve PSCs (day 10, reprog.), refractory cells (day 10, refract.) and established naïve PSC lines

Figure 5.4.7 – Plot shows the expression levels of *FBRS* and *AUTS2* in single cells in primed and naïve PSCs

Figure 5.4.8 – Heatmap shows *FBRS* and *AUTS2* expression levels in early human embryos

Figure 5.4.9 – Plot shows the *FBRS* and *AUTS2* expression levels in single cells in early human embryos

Figure 5.4.10 – Violin plots show the expression levels of *FBRS* and *AUTS2* in single cells during cynomolgus monkey embryo development

Figure 5.4.11 – Schematic showing the developmentally-regulated paralogue switch of *FBRS* and *AUTS2* within the PRC1.3 complex

Figure 5.5.1 – Schematic showing the method used for the PRC1.3 *in vitro* ubiquitylation assay

Figure 5.5.2 – Endogenous PRC1 complexes can catalyse H2AK119ub1 on *in vitro* polynucleosomes

Figure 5.5.3 – H2AK119ub1 deposited by PRC1.3 complexes can be detected *in vitro*

Figure 5.5.4 – H2AK119ub1 deposited by PRC1.3 complexes can be detected *in vitro*

Figure 5.6.1 – Schematic showing the homology between *AUTS2* and *FBRS* proteins

List of Tables

Table 2.2.1 – Composition of MEF media

Table 2.2.2 – Composition of KSR+FGF2 media

Table 2.2.3 – Composition of t2iL+PKCi media

Table 2.2.4 – Composition of PXGL media

Table 2.2.5 – Composition of 5i/L/A media

Table 2.2.6 – Composition of Chemical Reset Media 1

Table 2.3.1 – Antibodies used for flow cytometry and Fluorescence Activated Cell Sorting (FACS)

Table 2.3.2 – Flow cytometry machine settings used for analyser acquired data

Table 2.5.1 – Plasmids used in this study

Table 2.5.2 – gRNA sequences used for individual gene knockouts, with 3' protospacer adjacent motif emboldened

Table 2.5.3 – Primer sequences for *PCGF3* amplification and restriction site creation

Table 2.5.4 – Primer sequence for Sanger sequencing of the pCAG-PCGF3 vector

Table 2.5.5 – Guide RNA sequences with overhangs for *PCGF3* CRISPRi knockdown

Table 2.5.6 – Primer sequence for Sanger sequencing of the pgRNA-CKB vector

Table 2.5.7 – Primer sequences for CRISPR gene knockout genotyping

Table 2.5.8 – Primer sequences for Sanger sequencing of genotyping

Table 2.5.9 – Primer sequences used for RT-qPCR experiments

Table 2.5.10 – Primary antibodies used for Western blotting

Table 2.5.11 – Secondary antibodies used for Western Blotting

Table 2.5.12 – Primary antibodies used for immunofluorescence microscopy

Table 2.5.13 – Secondary antibodies used for immunofluorescence microscopy

Table 2.5.14 – Composition of Co-IP Buffer 1

Table 2.5.15 – Composition of Co-IP Buffer 2

Table 2.5.16 – Composition of Co-IP Wash Buffer

Table 2.6.1 – Composition of Wash Buffer 1

Table 2.6.2 – Composition of Wash Buffer 2

Table 2.6.3 – Composition of Lysis and Sonication Buffer

Table 2.6.4 – Composition of RIME RIPA Buffer

Table 2.6.5 – Composition of ubiquitylation buffer

Table 2.7.1 – Composition of ChIP Dilution Buffer

Table 2.7.2 – Composition of Wash Buffer A

Table 2.7.3 – Composition of Wash Buffer B

Table 2.7.4 – Composition of Wash Buffer C

Table 2.7.5 – Composition of ChIP Elution Buffer

Table 2.7.6 – Antibodies used for chromatin immunoprecipitation-sequencing

Table 2.7.7 – Buffers used for CUT&Tag

Table 2.7.8 – Antibodies used for CUT&Tag

List of Abbreviations

AID	Auxin Inducible Degron	NHSM	Naïve human stem cell media
AUTS2	Autism susceptibility candidate 2	NK2	NANOG and KLF2
BAP1	BRCA1 associated protein-1	PCA	Principal Components Analysis
(d)CAS9	(Dead) CRISPR-associated protein 9	PCGF	Polycomb Group Ring Finger
CBX	Chromobox	(RT-q)PCR	(Reverse Transcription-Quantitative) Polymerase Chain Reaction
CD	Cluster of Differentiation	PRC	Polycomb Repressive Complex
ChIP	Chromatin-immunoprecipitation	PRDM14	PR/SET Domain 14
CKII/CSNK2	Casein Kinase II	(i)PSC	(Induced) Pluripotent Stem Cell
CR(M1)	Chemical resetting (media 1)	PTNR	Primed-to-naïve reprogramming
CRISPR(i)	Clustered Regularly Interspaced Short Palindromic Repeats (Interference)	PXGL	PD03, XAV939, Gö6983 and LIF
CUT&RUN	Cleavage Under Targets and Release Using Nuclease	qPLEX-RIME	Quantitative Multiplexed Rapid Immunoprecipitation Mass Spectrometry of Endogenous Protein
CUT&Tag	Cleavage Under Targets and Tagmentation	RING	Really Interesting New Gene
DNA	Deoxyribonucleic Acid	RNA	Ribonucleic acid
DNMT	DNA methyltransferase	ROCK	Rho-associated protein kinase
DPPA	Developmental pluripotency associated	RP(K)M	Reads (per kilobase transcript) per million reads
FACS/MACS	Fluorescence Activated Cell Sorting/Magnetically Activated Cell Sorting	RYBP	RING1B and YY1 binding protein
FBRS	Fibrosin	SAGA	Spt-Ada-Gcn5 acetyltransferase
FGF2/4	Fibroblast Growth Factor 2/4	SSEA-4	Stage-specific embryonic antigen 4
GFP	Green Fluorescence Protein	SUSD2	Sushi Domain Containing 2
GSEA	Gene Set Enrichment Analysis	TGF- β	Transforming Growth Factor- β
GSK3	Glycogen Synthase Kinase 3	TSC	Trophoblast Stem Cell
KLF	Krüppel-like factor	TSS	Transcription Start Site
KO/DKO	Knockout/Double Knockout	t2iL+PKCi	Titrated 2 inhibitors, LIF and Protein Kinase C inhibitor
LIF	Leukaemia Inhibitory Factor	VPA	Valproic Acid
MaGeCK	Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout	Wnt	Wingless/integrated
MAPK/ERK	Mitogen-associated protein kinase/extracellular regulated kinase	WT	Wild-type
MEF	Mouse Embryonic Fibroblast	XACT	X-activating transcript
MEK	Mitogen-activated protein kinase kinase	XIST	X-inactive specific transcript
mEpiSC	Mouse Epiblast Stem Cell	YAF2	YY1 associated factor 2
mESC	Mouse Embryonic Stem Cell	5i/L/A	5 inhibitors, LIF and Activin A

List of Publications

Bendall, A., Semprich, C.I., 2022. Chromatin Profiling of Human Naïve Pluripotent Stem Cells, in: Rugg-Gunn, P. (Ed.), Human Naïve Pluripotent Stem Cells, Methods in Molecular Biology. Springer US, New York, NY, pp. 181–200. https://doi.org/10.1007/978-1-0716-1908-7_12

Collier, A.J.*, **Bendall, A.***, Fabian, C., Malcolm, A.A., Tilgner, K., Semprich, C.I., Wojdyla, K., Nisi, P.S., Kishore, K., Roamio Franklin, V.N., Mirshekar-Syahkal, B., D'Santos, C., Plath, K., Yusa, K., Rugg-Gunn, P.J., 2022. Genome-wide screening identifies Polycomb repressive complex 1.3 as an essential regulator of human naïve pluripotent cell reprogramming. Sci. Adv. 8, eabk0013. <https://doi.org/10.1126/sciadv.abk0013>

Zijlmans, D.W.*, Talon, I.*, Verhelst, S.*, **Bendall, A.***, Van Nerum, K., Javali, A., Malcolm, A.A., van Knippenberg, S.S.F.A., Biggins, L., To, S.K., Janiszewski, A., Admiraal, D., Knops, R., Corthout, N., Balaton, B.P., Georgolopoulos, G., Panda, A., Bhanu, N.V., Collier, A.J., Fabian, C., Allsop, R.N., Chappell, J., Pham, T.X.A., Oberhuemer, M., Ertekin, C., Vanheer, L., Athanasouli, P., Lluís, F., Deforce, D., Jansen, J.H., Garcia, B.A., Vermeulen, M., Rivron, N., Dhaenens, M., Marks, H., Rugg-Gunn, P.J., Pasque, V., 2022. Integrated multi-omics reveal polycomb repressive complex 2 restricts human trophoblast induction. Nat. Cell Biol. <https://doi.org/10.1038/s41556-022-00932-w>

Asterisk (*) denotes joint co-first authorship

1 – Introduction

1.1 – Pluripotency and pluripotent stem cells

1.1.1 – Human development and embryonic pluripotency

Pluripotency is characterised by two key properties. Firstly, pluripotent cells can self-renew and pass-through multiple cell divisions to produce identical pluripotent daughter cells. Secondly, pluripotent cells retain the ability to give rise to all of the specialised cell types of an adult mature organism through differentiation (Hackett *and* Surani, 2014). The state of pluripotency is fundamental to development, but the process remains poorly understood.

Pluripotent cells are found in early embryos prior to lineage specification and cell fate commitment. Mammalian development begins with the fusion of haploid gametes, with the fertilisation of an oocyte with a sperm to produce a diploid zygote. The zygote is totipotent, meaning that it can form all of the differentiated cells of an organism, including extra-embryonic tissues (Seydoux *and* Braun, 2006).

The first lineage specification event occurs approximately five days post-fertilisation, to specify the trophectoderm and inner cell mass. The trophectoderm subsequently develops into extra-embryonic tissue required for successful uterine invasion implantation, and formation of the maternal-placental interface. The inner cell mass subsequently generates pluripotent epiblast and primitive endoderm cells (Fig. 1.1.1) (Knöfler *et al.*, 2019).

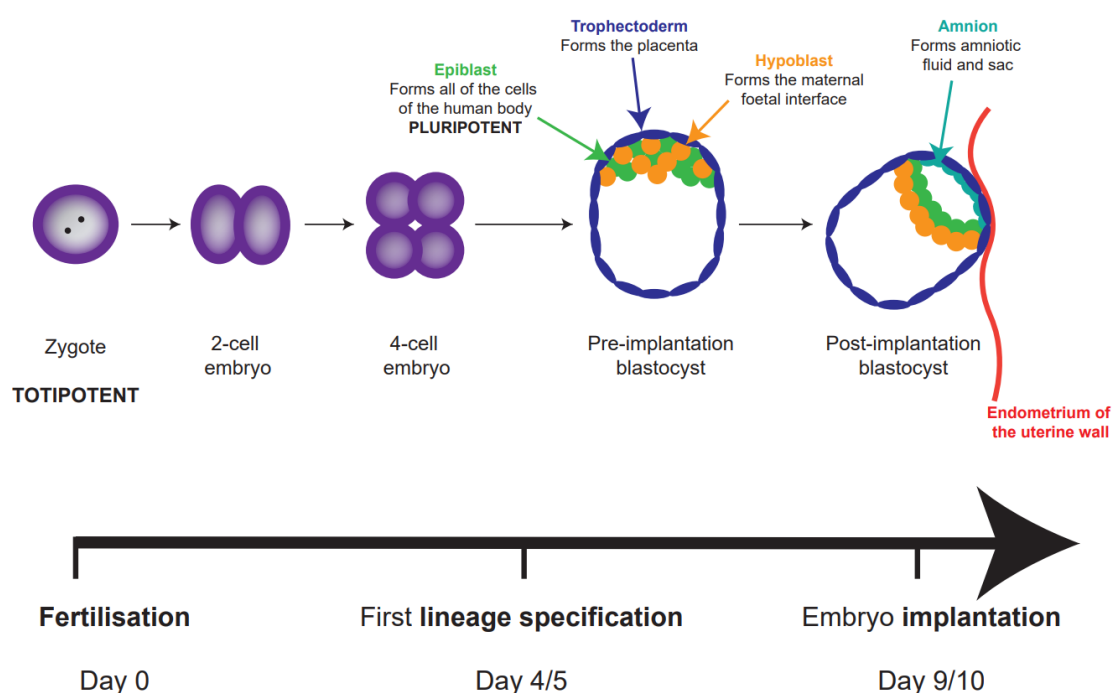
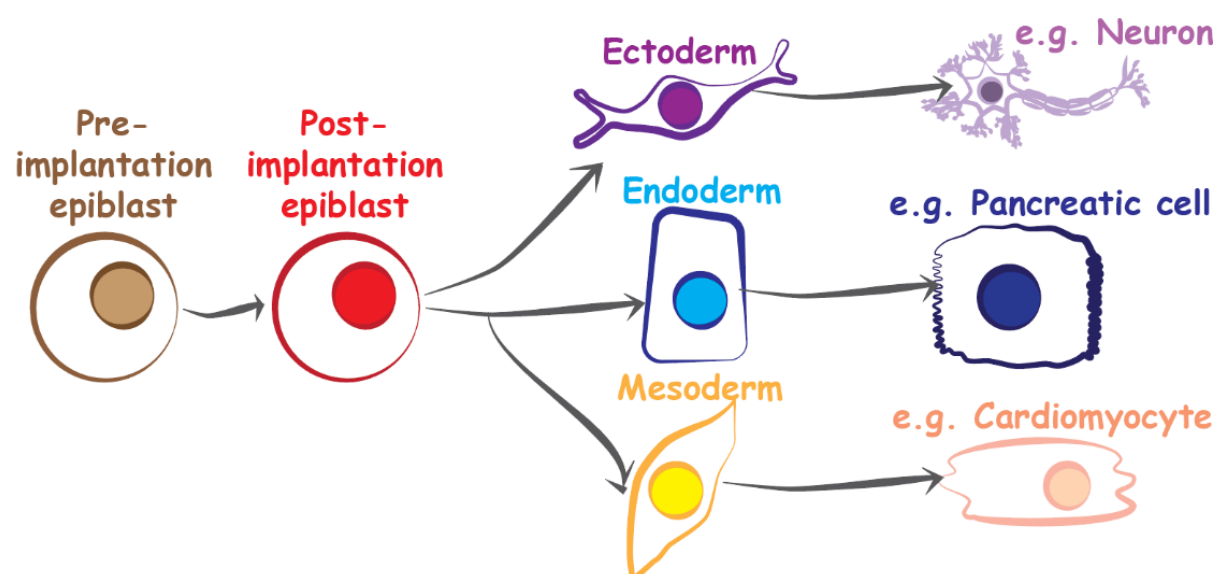


Figure 1.1.1 – The first ten days of human embryonic development

Early human development summarised from oocyte fertilisation to form a diploid zygote, the first lineage specification event segregating embryonic and extra-embryonic tissues, spatiotemporal organisation into blastocysts and implantation into the maternal uterine wall. Adapted from (Ortega *et al.*, 2018).

Pluripotency is retained in unspecialised epiblast cells and lasts for an approximately two-week window in human development, throughout implantation until the embryo undergoes gastrulation. At gastrulation, the endoderm, mesoderm and ectoderm lineages are specified (Fig. 1.1.2), with these cells still possessing the ability to form an array of cell types, but no longer fully satisfying the definition of pluripotency (Rossant *and* Tam, 2017). Primate embryo gastrulation has recently been spatially modelled *in utero*, showing the specification of early post-implantation cell lineages (Bergmann *et al.*, 2022).

**Figure 1.1.2 – Lineage specification and differentiation of human epiblast cells**

Embryo-forming epiblast becomes specified into three distinct germ layers; ectoderm, mesoderm and endoderm during gastrulation. Each lineage has restricted multipotency to form terminally differentiated cell types but cannot form cell types of another germ lineage (Ghimire *et al.*, 2021).

1.1.2 – Mouse embryonic stem cells

In order to study embryonic pluripotency more accessibly, pluripotent stem cell lines were derived from mouse embryos, known as mouse embryonic stem cells (mESCs) (Evans *and* Kaufman, 1981; Martin, 1981). The advantage of pluripotent stem cells is that they continuously self-renew and provide an indefinite supply of pluripotent cells in the correct cell culture conditions, whereas embryonic pluripotency is transient (Wobus *and* Boheler, 2005).

To benchmark the pluripotency of mESCs, two key assays were performed. Firstly, it was established that injecting mESCs into a mouse blastocyst could produce chimeric mice, with the pluripotency of mESCs enabling their contribution across multiple specialised cell types of the adult mouse (Beddington *and* Robertson, 1989). Secondly, mESCs were able to contribute exclusively to the embryonic tissue and ultimately the full adult mouse using a tetraploid complementation assay (Nagy *et al.*, 1993). This involves mESC aggregation with tetraploid cells, formed by the fusion of both cells of a two-cell embryo, meaning that only mESCs can form the inner cell mass, as the tetraploid cells form the extra-embryonic tissue.

1.1.3 – Human pluripotent stem cells

Following stem cell derivation from mice, pluripotent stem cells were also derived from human blastocyst-stage embryos to generate human pluripotent stem cells (PSCs). Human PSCs self-renew continuously in the correct culture conditions, which will be discussed in section 1.2.2, and retain the developmental potential of their embryonic counterparts (Thomson *et al.*, 1998). The first derived human PSCs are known as conventional or ‘primed’ human PSCs, with the cells preserved in a state that can readily differentiate in response to extracellular cues (Schuldiner *et al.*, 2000).

The developmental potential of human PSCs cannot be assessed using human chimera or tetraploid complementation assays due to obvious ethical implications, and so different methods must be used. Human primed PSCs are able to form teratomas, which can form cells of all three germ lineages when injected into immunodeficient mice (Prokhorova *et al.*, 2009; Przyborski, 2005). Human PSCs can also form human embryoid bodies (Warmflash *et al.*, 2014) or gastruloids (Moris *et al.*, 2020), where all three germ lineages can be generated from primed PSCs *in vitro*. Bioinformatics analysis of transcriptomic data can also be used to assess the human undifferentiated state (Müller *et al.*, 2011; The International Stem Cell Initiative, 2018).

1.1.4 – Human induced pluripotent stem cells

Human PSCs are not exclusively derived from human embryos. Somatic cell nuclear transfer experiments, where the nuclei of terminally differentiated somatic cells are transplanted into the enucleated unfertilised oocytes (Briggs *and* King, 1952; Gurdon, 1960), confirmed that terminally differentiated cells retain all of the genes required to form all of the other specialised cell types of the adult organism. This was first extended to mammals with the live birth of Dolly the Sheep (Campbell *et al.*, 1996). These experiments were performed by generating pluripotent cells from somatic cells. Reprogramming of differentiated somatic cells to form PSCs (known as induced PSCs, iPSCs) requires the overexpression of only four transcription factors; *OCT4*, *SOX2*, *KLF4* and *c-MYC* (known collectively as OSKM factors). These genes were able to erase the differentiated cell fate and generate pluripotent cells that are largely indistinguishable from embryo-derived human PSCs (Park *et al.*, 2008; Yu *et al.*, 2007).

1.2 – Ground state pluripotency

1.2.1 – Mouse ground state pluripotency

Pluripotent cells possess specific growth requirements. mESC derivation required growth on a layer of mouse embryonic fibroblasts (MEFs) to provide supplementary growth factors (Martin, 1981). Mitotically-inactivated MEFs are utilised to secrete pluripotency-sustaining factors, without themselves proliferating. Efforts to define the factors secreted from MEFs identified Leukaemia Inhibitory Factor (LIF). Exogenous LIF supplementation allows the growth of mESCs in serum-containing medium in the absence of MEFs (Williams *et al.*, 1988).

Exogenous LIF signals through its receptor, promoting Janus kinase (JAK)-dependent phosphorylation of Signal Transducer and Activator of Transcription 3 (STAT3). Phosphorylated STAT3 translocates to the nucleus and activates pluripotency gene, and represses differentiation gene, expression (Fig. 1.2.1) (Niwa *et al.*, 1998).

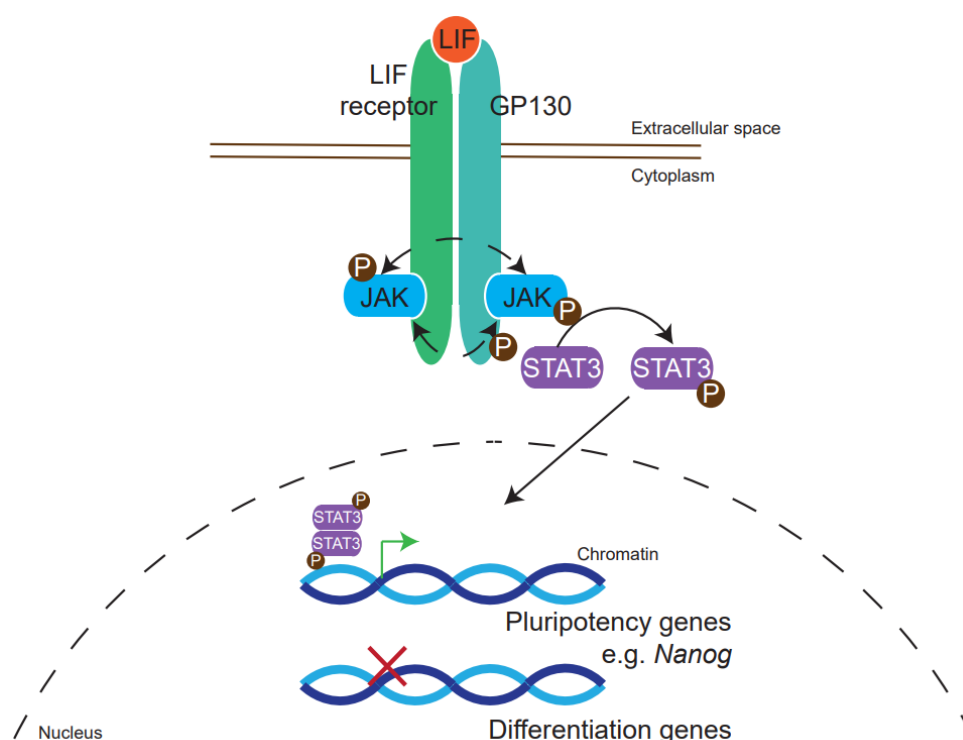


Figure 1.2.1 – LIF signalling in PSCs

LIF ligand binding promotes dimerisation of the LIF receptor to GP130, bringing JAKs into close proximity. The kinases are able to transphosphorylate one another and the receptor itself. The

phosphorylated residues within the receptor are able to recruit SH2 (Src homology 2)-domain proteins such as STAT3. JAK phosphorylates the STAT3, which then homodimerises and translocates into the nucleus. Nuclear STAT3 acts alongside transcription factors and the basal transcriptional machinery to regulate gene transcription (Hirai *et al.*, 2011).

LIF alone is not sufficient to inhibit spontaneous mESC differentiation and additional factors present in serum are required for pluripotency. Serum-free growth of mESCs could be achieved with the additional supplementation of bone morphogenetic protein 4 (BMP4), to produce a more defined growth medium. BMP4 signalling induces the expression of inhibitor of differentiation (ID) genes to maintain the pluripotent state, via a SMAD-dependent signalling pathway (Fig. 1.2.2) (Ying *et al.*, 2003).

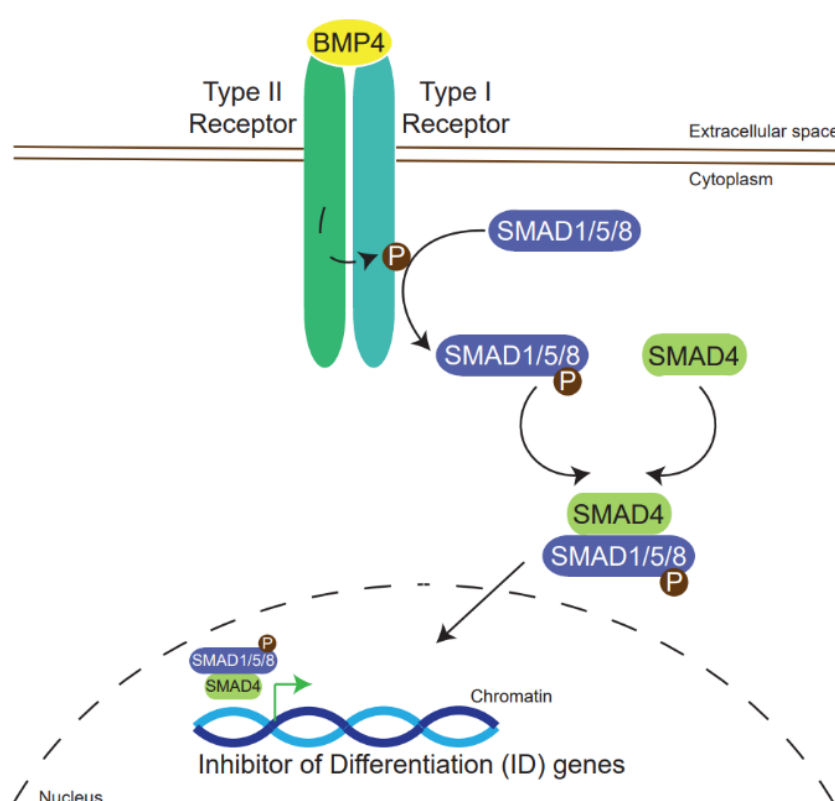


Figure 1.2.2 – BMP4 signalling in PSCs

Type II transforming growth factor- β (TGF- β) family BMP4 receptors form heterodimers with type I receptors and activate type I receptor kinase activity by serine/threonine phosphorylation. Active type I receptors phosphorylate SMAD1/5/8 proteins, which form complexes with SMAD4, translocate into the nucleus and activate expression of ID genes (Morikawa *et al.*, 2016; Ying *et al.*, 2003).

Further interrogation of the signalling requirements of mESCs identified that inhibition of the mitogen activated protein kinase/extracellular regulated kinase (MAPK/ERK) signalling pathway could promote mESC self-renewal (Burdon *et al.*, 1999). PD0325901 acts as a potent inhibitor of ERK phosphorylation by MAPK/ERK kinase (MEK), interrupting the signalling cascade that regulates gene expression (Fig. 1.2.3) (Sebolt-Leopold *and* Herrera, 2004).

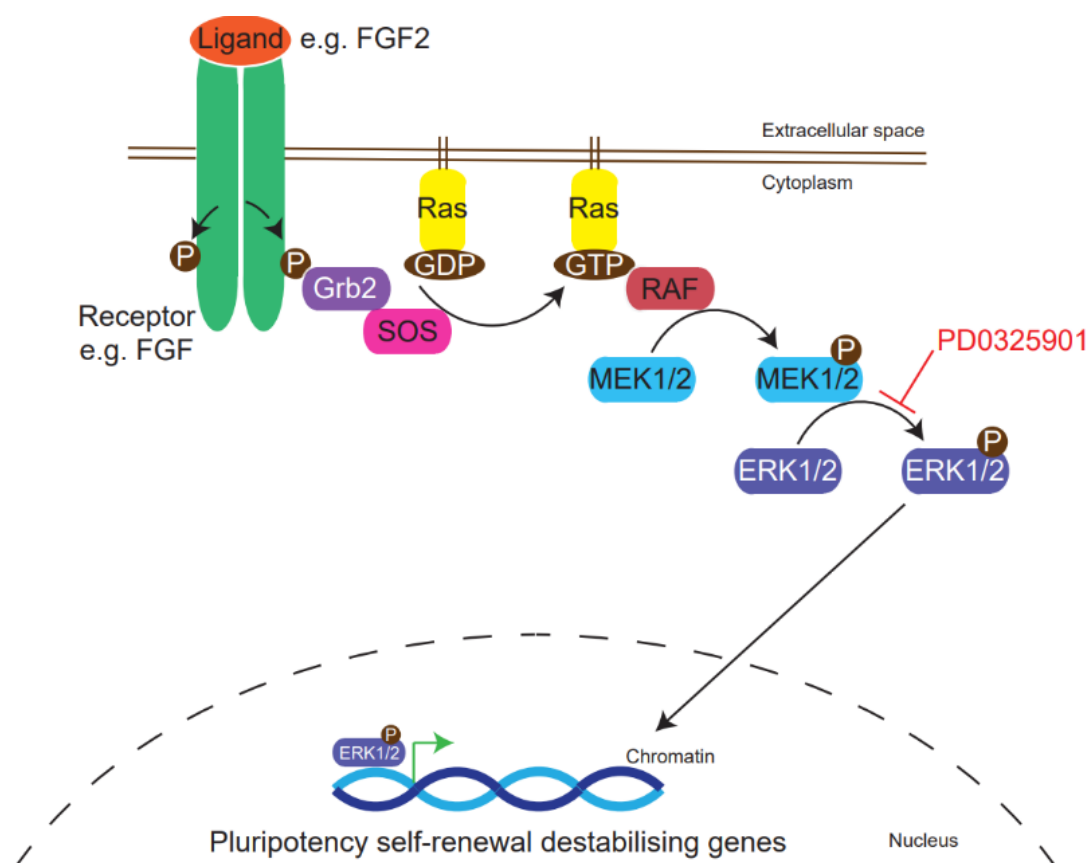


Figure 1.2.3 – MAPK/ERK signalling in PSCs

Ligands, such as Fibroblast Growth Factors (FGF), bind and promote the autophosphorylation activity of receptor tyrosine kinases. The phosphorylated receptor recruits the SH2 domain-containing Growth factor receptor-bound protein 2 (Grb2) and the Guanine nucleotide exchange factor Son of Sevenless (SOS). SOS ensures that Ras (Rat sarcoma virus) kinase is in an active, GTP-bound state. Ras begins a phosphorylation cascade, through Raf (Rapidly Accelerated Fibrosarcoma), MEK and ERK. Each kinase phosphorylates multiple substrates, activating their kinase activity. Phosphorylated ERK1/2 (extracellular regulate kinase) translocates into the nucleus and influences gene expression. MEK inhibition using PD0325901 stabilises self-renewal of pluripotency (Burdon *et al.*, 1999; Ma *et al.*, 2016).

Co-treatment of mESCs with the MEK inhibitor PD0325901 and an activator of the canonical wingless/integrated (Wnt) signalling pathway, using the Glycogen Synthase Kinase 3 β (GSK-3 β) inhibitor CHIR99021, further stabilises the pluripotent state (Fig. 1.2.4). Active MEK signalling, stimulated by fibroblast growth factor 4 (FGF4) stimulates mESC differentiation (Kunath *et al.*, 2007) and active Wnt signalling promotes the expression of the core pluripotency factors *Oct4* and *Nanog* (Sato *et al.*, 2004; Ying *et al.*, 2008). The transcriptional network underpinning pluripotency will be discussed further in section 1.3.1.

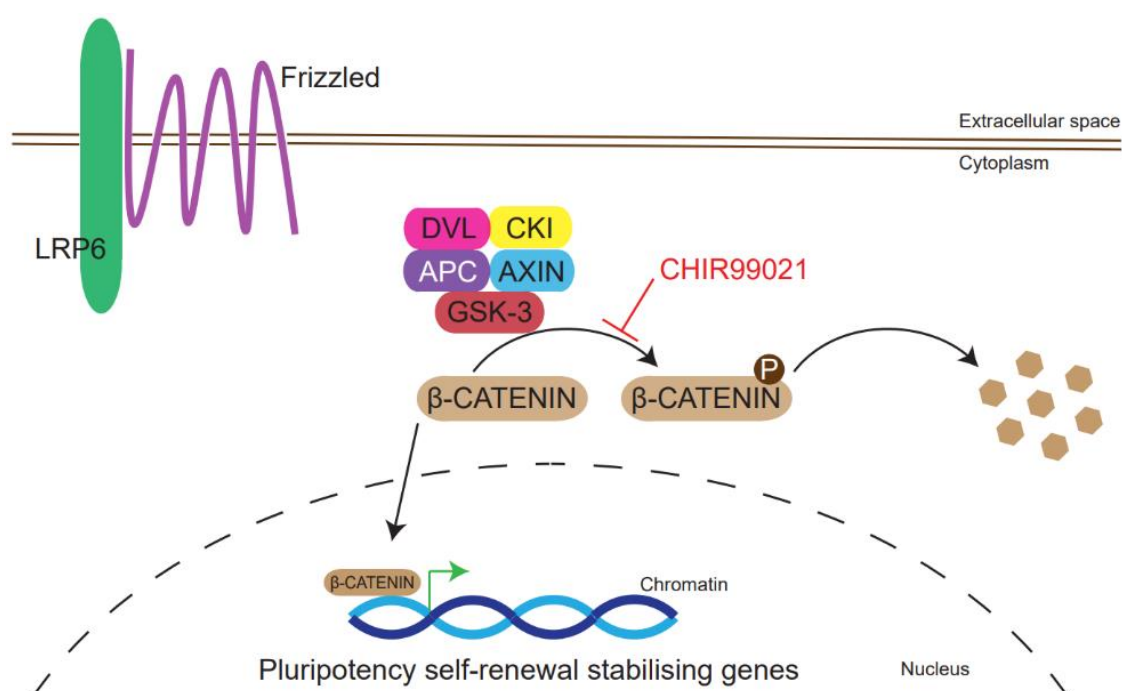


Figure 1.2.4 – Wnt signalling in PSCs

Activation of the Wnt receptor Frizzled and the co-receptor lipoprotein receptor-related protein (LRP6) cause a direct interaction between Frizzled and Dishevelled (DVL). This interaction sequesters the β -catenin phosphorylation and degradation complex, composed of DVL, Axin, Adenomatosis Polyposis Coli (APC), Casein Kinase I (CKI) and Glycogen Synthase Kinase-3 β (GSK-3). GSK-3 inhibition by CHIR99021 prevents β -catenin phosphorylation and degradation, allowing β -catenin translocation into the nucleus and activation of pluripotency-promoting genes (de Jaime-Soguero *et al.*, 2018).

Culturing mESCs in PD0325901, CHIR99021 and LIF, or 2i/LIF, allowed the propagation of pluripotent cells from mouse pre-implantation blastocysts. These mESCs have a pre-implantation epiblast signature (Boroviak *et al.*, 2014) and are described as the ‘ground’ or ‘naïve’ state of pluripotency (Silva *et al.*, 2008; Ying *et al.*, 2008). Unlike serum/LIF grown mESCs, 2i/LIF grown mESCs are highly homogeneous (Godwin *et al.*, 2017), and the ground

state of pluripotency is characterised by minimal epigenetic restriction and minimal requirement for extrinsic signalling stimuli (Wray *et al.*, 2010).

1.2.2 – Primed human PSCs

Human PSCs initially derived from human pre-implantation embryos (Thomson *et al.*, 1998) have very different signalling requirements from mESCs. LIF signalling is dispensable for maintenance of primed human PSCs (Dahéron *et al.*, 2004), which instead rely on both FGF (Greber *et al.*, 2007) and transforming growth factor- β (TGF- β) (James *et al.*, 2005; Vallier *et al.*, 2009a) signalling (Fig. 1.2.5) for the self-renewal of the pluripotent state.

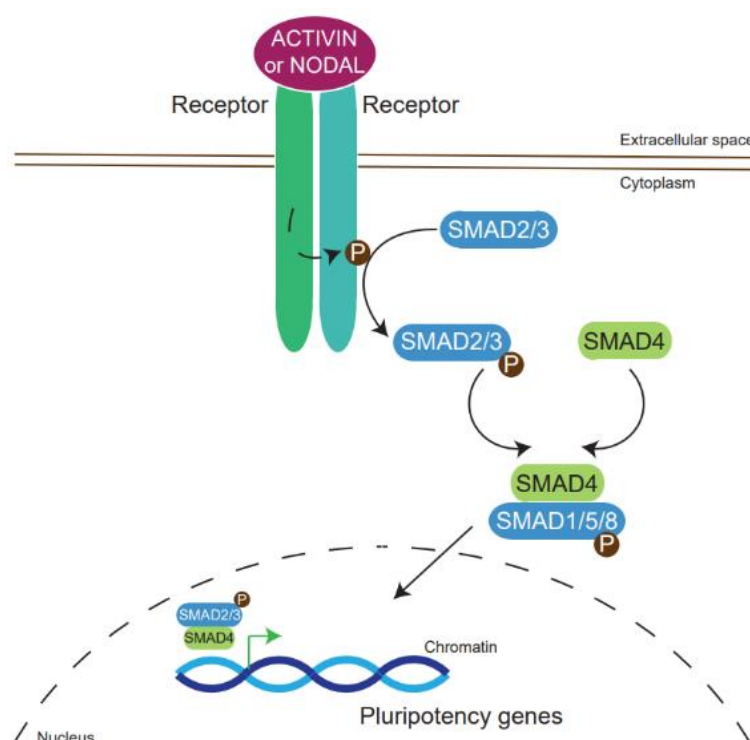


Figure 1.2.5 – TGF- β signalling in PSCs

Binding of TGF- β ligands such as ACTIVIN A or NODAL to their transmembrane receptor tyrosine kinases causes receptor dimerisation and activation of kinase activity. The substrates SMAD2 and 3 become phosphorylated by this kinase activity, and phosphorylated SMAD2/3 proteins heterodimerise with SMAD4, translocate into the nucleus and activate expression of pluripotency genes such as *NANOG* (Brown *et al.*, 2011; Osnato *et al.*, 2021; Pauklin and Vallier, 2015; Vallier *et al.*, 2009a).

Inhibition of FGF signalling results in neuroectodermal differentiation (Greber *et al.*, 2011; Vallier *et al.*, 2009b). Secreted or exogenously supplied FGF, often FGF2, acts as a ligand that binds to an FGF receptor tyrosine kinase, activating the kinase activity and promoting MAPK/ERK signalling cascades (Greber *et al.*, 2007), which contradicts the action of PD0325901 in mESCs (see Fig. 1.2.3).

1.2.3 – Human naïve pluripotency

The signalling requirements of mESCs and primed human PSCs, both derived from pre-implantation blastocysts, are profoundly different. In fact, the growth conditions of primed human PSCs can instead be used to support the growth of mouse post-implantation epiblast cells, called epiblast stem cells (EpiSCs) (Tesar *et al.*, 2007), which are more developmentally advanced and are poor at producing chimeras (Brons *et al.*, 2007). EpiSCs, like primed human PSCs, cannot be maintained under 2i/LIF conditions (Guo *et al.*, 2009).

The similarities between primed PSCs and EpiSCs suggests that after derivation, human PSCs developmentally advance to represent a post-implantation state of human pluripotency. Single-cell RNA-sequencing of human embryos identified a developmental progression from human pre-implantation epiblast cells to primed human PSCs (Tyser *et al.*, 2021; Yan *et al.*, 2013). Therefore, endeavours were made to generate and sustain the human naïve state of pluripotency.

2i/LIF conditions, used to stabilise the naïve state of pluripotency in mice, were used as a starting point in attempting to produce human naïve PSCs and the ‘reprogramming’ conversion of serum/LIF mESCs or mEpiSCs into naïve mESCs is a fairly straightforward process under 2i/LIF conditions (Silva *et al.*, 2008). However, 2i/LIF conditions alone are not sufficient to reprogramme human primed PSCs to a naïve state (Hanna *et al.*, 2010). Therefore 2i/LIF conditions were expanded upon in a number of different ways in order to try to ‘reprogramme’ primed human PSCs to a naïve state (Chan *et al.*, 2013b; Gafni *et al.*, 2013; Guo *et al.*, 2017; Hanna *et al.*, 2010; Takashima *et al.*, 2014; Theunissen *et al.*, 2014, 2016; Ware *et al.*, 2014). These methods will be summarised more extensively in section 1.4.

These experiments confirmed that human PSCs, as in mice, exist in two different states, termed naïve and primed (Fig. 1.2.6) (Davidson *et al.*, 2015; Weinberger *et al.*, 2016). Both primed and naïve PSCs can continuously self-renew and undergo multi-lineage differentiation, but are distinct in a number of molecular characteristics, as reviewed in section 1.3. Naïve PSCs largely recapitulate the transcriptome of pre-implantation embryos, and primed PSCs

are instead similar to early post-implantation embryos (Blakeley *et al.*, 2015; Huang *et al.*, 2014; Nakamura *et al.*, 2016; Petropoulos *et al.*, 2016; Stirparo *et al.*, 2018; Tesar *et al.*, 2007; Xiang *et al.*, 2020).

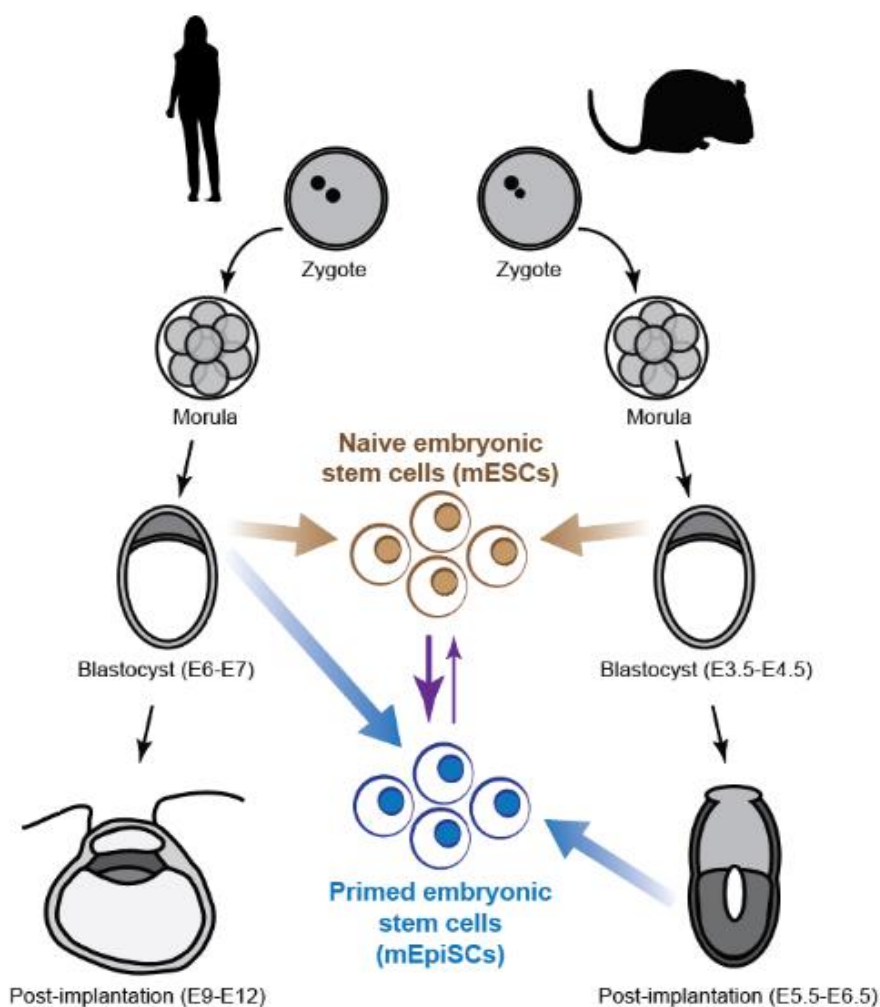


Figure 1.2.6 – Developmental progression of human and mouse epiblast derived PSCs

Pre-implantation epiblast derived primed human PSCs capture the properties of the post-implantation epiblast derived mouse EpiSCs, rather than the mouse ground state mESCs. Specific signalling modulation, mimicking those used to grow mESCs was required to culture human naïve PSCs or derive them from human pre-implantation embryos (Guo *et al.*, 2016a; Weinberger *et al.*, 2016).

1.2.4 – Signalling requirements of human naïve PSCs

The signalling requirements of human naïve PSCs resemble the conditions used to derive mESCs, including a need for LIF signalling. Two of the most stringent and commonly used naïve PSC sustaining media are described as t2iL+PKCi (Takashima *et al.*, 2014) and PXGL (Bredenkamp *et al.*, 2019b; Rostovskaya *et al.*, 2019). t2iL+PKCi medium is composed of titrated concentrations of the MEK inhibitor PD0325901 and the Wnt activator CHIR99021, combined with LIF and the Protein Kinase C (PKC) inhibitor Gö6983. PXGL is composed of the same components, except that the Wnt activator CHIR99021 is exchanged for the tankyrase 1 and 2 inhibitors XAV939. While CHIR99021 inhibits the activity of GSK-3 β , XAV939 stabilises the AXIN-GSK-3 β complex and inhibits Wnt signalling (Huang *et al.*, 2009), and is proposed to suppress differentiation in naïve PSCs cultures (Rostovskaya *et al.*, 2019). The use of a Protein Kinase C inhibitor Gö6983 in naïve PSC cultures is corroborated with naïve mESC culture, where Gö6983 can promote pluripotency without STAT3 activation (Dutta *et al.*, 2011), and is proposed to inhibit epithelial polarity (Takashima *et al.*, 2014). Inhibitors of other components of the MAPK/ERK signalling pathway can also stabilise human naïve PSCs (Khan *et al.*, 2021). Like primed PSCs, naïve PSCs also rely on TGF- β signalling for their maintenance. Inhibition of SMAD2 phosphorylation during TGF- β signalling, using the inhibitor SB-431542, causes reduced expression of pluripotency factors including *NANOG* and *LEFTY1/2* and disruption to the morphology of naïve PSCs (Osnato *et al.*, 2021).

As naïve PSCs model an earlier stage of human development than primed PSCs, they are less likely to possess biased differentiation propensities, which has been observed in primed PSC cell lines (Osafune *et al.*, 2008), although naïve PSCs are not epigenetically unrestricted (Kumar *et al.*, 2022; Zijlmans *et al.*, 2022). Naïve PSCs are incompetent to differentiation cues, with acquisition of ‘competence’ required for multi-lineage differentiation. Treatment of naïve PSCs with XAV939 is proposed to generate a poorly understood intermediate state of pluripotency between the naïve and the primed state, known as formative pluripotency, which can then respond to differentiation signals (Kinoshita *et al.*, 2021; Rostovskaya *et al.*, 2019).

1.3 – Differences between human primed and naïve PSCs

Primed and naïve PSCs have the common feature of pluripotency, meaning both can continuously self-renew in culture and are capable of multi-lineage differentiation into a number of specialised cell types. However, naïve and primed PSCs also possess fundamental differences (Fig. 1.3.1) (Takahashi *et al.*, 2018).

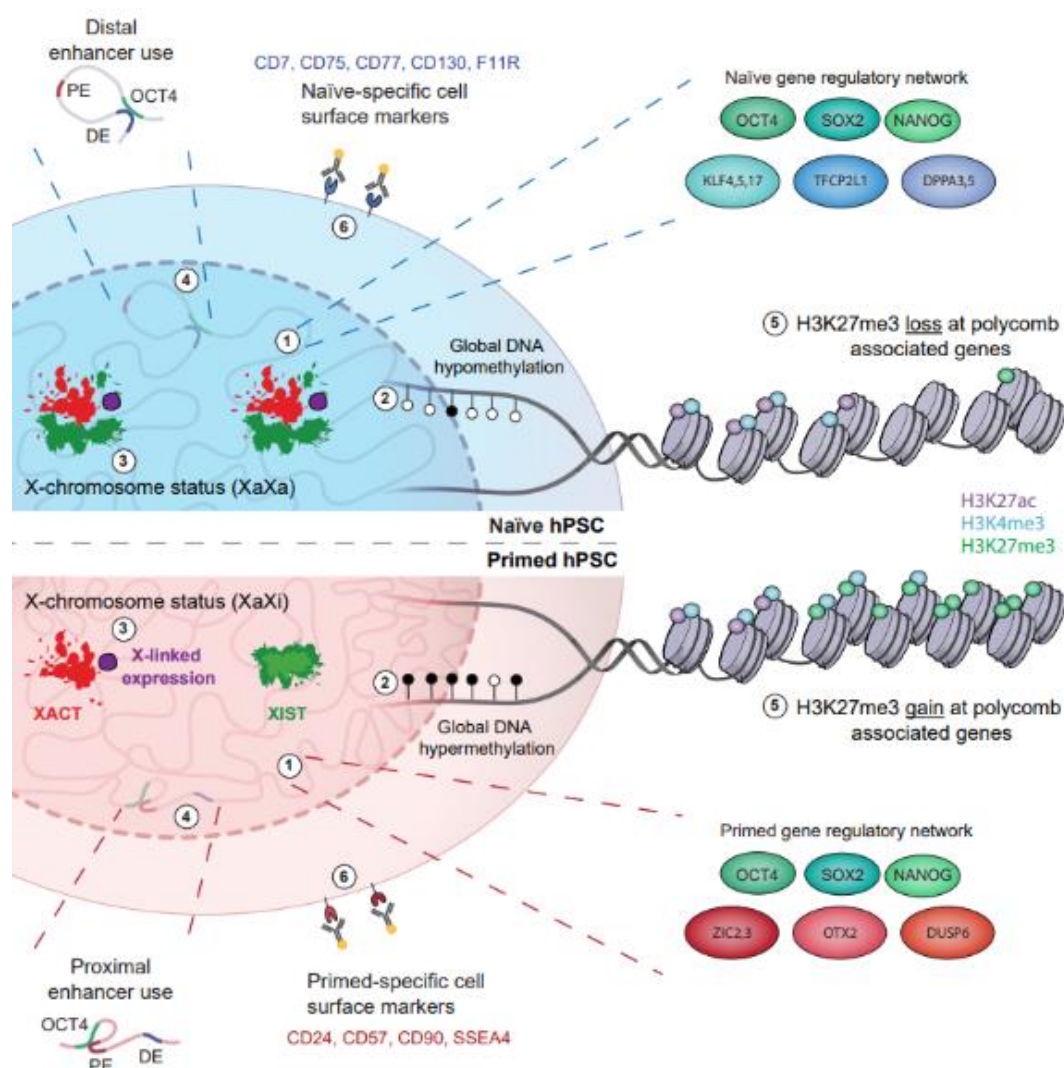


Figure 1.3.1 – Epigenetic and molecular differences between human naïve and primed PSCs

Differences are seen between primed and naïve PSCs in their gene regulatory networks, DNA methylation state, X-chromosome activation state, developmental enhancer usage, epigenetic landscape and cell-surface marker proteins, which will be discussed in section 1.5.5. Obtained from (Collier *and* Rugg-Gunn, 2018).

Naïve and primed PSCs are very morphologically different. Naïve PSCs form smaller, rounded, dome-like colonies, with high levels of clonogenicity and single-cell survival (Warrier *et al.*, 2017). Conversely, primed PSCs form larger, more dispersed and flattened colonies and have lower single-cell survival (Healy and Ruban, 2015). Theunissen and colleagues, 2016, generated a set of molecular criteria that can be used to evaluate naïve PSCs, by comparing their properties to those of human pre-implantation epiblast cells, and this has been subsequently expanded (Boroviak and Nichols, 2017).

1.3.1 – Pluripotency transcriptional networks

Human primed and naïve PSCs share a large degree of their core pluripotency transcriptional networks. The components of this core pluripotency transcription network are also conserved with mouse pluripotency, composed of *OCT4*, *SOX2* and *NANOG* (Avilion *et al.*, 2003; Chambers *et al.*, 2003; Nichols *et al.*, 1998). *OCT4* deletion in human primed PSCs results in endoderm and trophoblast differentiation (Hay *et al.*, 2004), and in human embryos causes destabilisation of pluripotency factor expression within the inner cell mass (Fogarty *et al.*, 2017). In human PSCs *SOX2* knockdown causes primitive streak marker upregulation and *NANOG* knockdown instead promotes differentiation (Hyslop *et al.*, 2005), particularly towards neuroectoderm and neural crest lineages (Wang *et al.*, 2012). *OCT4*, *SOX2* and *NANOG* all show strong co-localisation on chromatin, reinforcing one another in the regulation of the pluripotency transcriptional network (Boyer *et al.*, 2005), but these factors co-localise at different sites in chromatin in primed and naïve PSCs (Chovanec *et al.*, 2021).

As well as core pluripotency factors, primed and naïve PSCs also possess distinct transcriptional networks, underpinning their distinct cell identities. Primed pluripotency transcription networks include *ZIC2*, *OTX2* and *DUSP6*, while naïve-specific transcription networks instead contain *KLF4*, *KLF17*, *FGF4*, *DPPA3*, *DPPA5*, *DNMT3L* and *TFCP2L1* (Messmer *et al.*, 2019; Theunissen *et al.*, 2014). *KLF17* in particular is an important transcription factor in human naïve pluripotency, but reports are conflicting about whether *KLF17* is important for the acquisition of naïve pluripotency by reprogramming (Lea *et al.*, 2021; Wang *et al.*, 2022). Single-cell RNA-sequencing demonstrated a large degree of transcriptional homogeneity within primed PSCs and within naïve PSCs (Messmer *et al.*, 2019). The transposable element transcriptome also differ between human naïve and primed PSCs (Theunissen *et al.*, 2016).

1.3.2 – DNA methylation

DNA methylation levels show conserved fluctuations in early mammalian embryogenesis (Okabe *et al.*, 2014; White *et al.*, 2016). Global DNA methylation levels are high in gametes and in zygotes, before reducing such that the genome of inner mass cells is hypomethylated, before increasing to a hypermethylated state in post-implantation epiblast cells (Guo *et al.*, 2014; Smith and Meissner, 2013; Zhu *et al.*, 2018). This difference in embryonic DNA methylation levels is recapitulated *in vitro*, with naïve PSCs being globally DNA hypomethylated at CpG sequences and primed PSCs instead hypermethylated (Pastor *et al.*, 2016; Takashima *et al.*, 2014). DNA methylation state is useful barometer to assess the quality of naïve PSC lines compared to their embryonic counterparts (Theunissen *et al.*, 2016), and acts as a useful system to study the developmental regulation of DNA methylation.

1.3.3 – X chromosome activation status

Another epigenetic difference between naïve and primed PSCs relates to the transcriptional silencing of one of the two X chromosomes in female cells (XX) as part of a dosage compensation mechanism, providing the same level of X-linked gene expression in female cells as male cells (XY) (Disteche, 2012; Lyon, 1961). Naïve PSCs possess two active X chromosomes while primed PSCs have an inactivated X chromosome, in both mouse (Minkovsky *et al.*, 2012) and human (Sahakyan *et al.*, 2017) pluripotency. This is consistent with the random inactivation of an X chromosome prior to implantation of mouse (Okamoto *et al.*, 2004) and human (Petropoulos *et al.*, 2016) embryos.

X chromosome inactivation mechanisms are not fully conserved between human and mouse. In mouse, X chromosome silencing occurs upon induction of the long non-coding RNA *Xist* (X-inactive specific transcript) from the inactive X chromosome (Xi) (Brockdorff *et al.*, 1992, 1991; Penny *et al.*, 1996), which coats the length of the chromosome and initiates silencing of gene expression (Clemson *et al.*, 1996; Lee and Jaenisch, 1997; Marahrens *et al.*, 1998). In humans, *XIST* is expressed from both X chromosomes in the inner cell mass and *XIST* expression alone is not sufficient for transcriptional silencing. (Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016; Vallot *et al.*, 2017). An additional, primate-specific, long non-coding

RNA *XACT* (X-activating transcript) exists, which coats the transcriptionally active X chromosome (Xa) (Petropoulos *et al.*, 2016; Vallot *et al.*, 2017, 2013).

Although mESCs have provided valuable fundamental insight into mammalian X chromosome inactivation, the presence of distinct regulation within humans makes naïve PSCs a stronger *in vitro* model to study human X-chromosome inactivation. In human naïve PSCs, *XIST* and *XACT* can be expressed from the same chromosome and both X chromosomes remain transcriptionally active. Upon transition of naïve PSCs to the primed state, one of the two X chromosomes becomes transcriptionally silenced, expressing only *XIST*, while the active X chromosome retains *XACT* expression (Sahakyan *et al.*, 2017; Vallot *et al.*, 2017). However, the inactive X chromosome in human primed PSCs is unstable and can become partially 'eroded', causing *XACT* re-expression and gene re-activation, prior to *XIST* silencing (Bar *et al.*, 2019; Vallot *et al.*, 2015).

During human primed to naïve reprogramming, the inactivated X chromosome becomes largely transcriptionally re-activated. This occurs at a late stage during the generation of naïve PSCs (Collier *et al.*, 2017) in a manner that is thought to be heterogeneous (Panova *et al.*, 2018) and non-uniform (Sarel-Gallily and Benvenisty, 2022). The X-chromosome activation state of female human naïve PSCs also acts a useful rubric to assess the properties of cell types and different culture or reprogramming methods (Theunissen *et al.*, 2016).

1.3.4 – Enhancer usage

The *Oct4* gene in mouse contains two cis-regulatory elements, acting as enhancers to interact with the *Oct4* promoter and initiate gene expression. These are named the distal and proximal enhancer. *Oct4* is preferentially expressed from the distal enhancer in mESCs and the proximal promoter in mEpiSCs, showing pluripotency state-specific regulation of the *Oct4* expression as a core pluripotency factor (Choi *et al.*, 2016; Tesar *et al.*, 2007; Yeom *et al.*, 1996).

The preferential expression of *Oct4* from its distal enhancer was assumed to be conserved in human naïve pluripotency and was used as the basis of a reporter assay to screen for chemical inhibitors that support the growth of naïve PSCs (Theunissen *et al.*, 2014), as described in section 1.4.1. This suggests that *OCT4* enhancer usage may also be developmentally regulated between human naïve and primed pluripotency.

1.3.5 – Metabolism

Finally, naïve PSCs obtain most of their cellular energy from oxidative phosphorylation in comparison to primed PSCs, which are predominantly glycolytic (Diamante *and* Martello, 2022; Takashima *et al.*, 2014; Zhou *et al.*, 2012). The metabolic pathways and their fluxes in the different pluripotent states influences the abundance of metabolic intermediates, which act as substrates for epigenetic modifying enzymes, and could therefore influence the underlying epigenetic landscape (Betto *et al.*, 2021).

1.4 – Generation of naïve PSCs by primed to naïve reprogramming (PTNR)

Although naïve PSCs can be directly derived from the inner cell mass of human pre-implantation embryos (Guo *et al.*, 2016a) and by somatic cell induced pluripotent stem cell reprogramming (Giulitti *et al.*, 2019; Kilens *et al.*, 2018), the vast majority of naïve PSCs in culture are generated by the reprogramming of primed PSCs. Exposing primed PSCs to a variety of different signalling conditions, both with and without the overexpression of transgenes, are able to induce this cell state conversion and can also be used to maintain naïve PSCs for long-term culture.

Initial attempts to derive human naïve PSCs were informed by knowledge of ground state mESCs, which are cultured in 2i/LIF conditions. LIF signalling is not required for culturing human primed PSCs (Dahéron *et al.*, 2004), but is important in mESCs to promote the expression of ground-state specific transcription factors such as *TFCP2L1* (Martello *et al.*, 2013). The earliest attempts to produce naïve PSCs relied upon transgene induction-based approaches as 2i/LIF alone is not sufficient to reprogramme or sustain human naïve pluripotency. The combination of 2i/LIF conditions and inducible overexpression of *OCT4*, *KLF2* and *KLF4* generated PSCs with some properties of the naïve state (Hanna *et al.*, 2010). Following on from this, it was discovered that primed PSCs could be reprogrammed to naïve pluripotency in 2i/LIF conditions in the presence of an inducible overexpression of two transcription factors, the pan-pluripotency marker gene *NANOG* and the naïve-specific marker *KLF2*, given the shorthand NK2 (Takashima *et al.*, 2014; Theunissen *et al.*, 2014).

Attempts were also made to improve the culture of naïve PSCs in a transgene-independent manner by chemical modulation. These conditions were first obtained under NHSM (naïve human stem cell medium) or 6i conditions, containing 2i/LIF, as well as an additional four inhibitors; p38 inhibitor BIRB796, JNK inhibitor SP600125, Protein Kinase C inhibitor Gö6983 and the ROCK (Rho-associated protein kinase) inhibitor Y-27632, as well as FGF2 and TGF- β . NHSM (naïve human stem cell medium) forms the basis of the commercial 'RSeT' naïve PSC medium (Gafni *et al.*, 2013).

A number of other transgene-independent formulations have also been developed (Chan *et al.*, 2013; Chen *et al.*, 2015; Duggal *et al.*, 2015; Guo *et al.*, 2017; Qin *et al.*, 2016; Szczerbinska *et al.*, 2019; Takashima *et al.*, 2014; Kilens *et al.*, 2018; Theunissen *et al.*, 2014; Ware *et al.*, 2014; Zimmerlin *et al.*, 2016), most of which are based on a common logic of 2i/LIF conditions with added supplementation. Some studies replaced CHIR99021 with other GSK-3 β inhibitors including BIO (Chan *et al.*, 2013b) or IM-12 (Theunissen *et al.*, 2014), or substituted Wnt activation with Wnt inhibition using the tankyrase inhibitor XAV939

(Bredenkamp *et al.*, 2019b; Rostovskaya *et al.*, 2019). MEK inhibition independent naïve PSCs culture has been reported using inhibitors against different components of the MAPK/ERK signalling pathway (Khan *et al.*, 2021), or alternatively using inhibitors of either cyclin dependent kinase 8/19 mediator (Lynch *et al.*, 2020) or NOTCH signalling (Bayerl *et al.*, 2021), with the latter two approaches showing varying degrees of retention of a DNA hypermethylated state.

Two particular signalling cocktails that derive naïve PSCs, with the ability to recapitulate the properties of the human pre-implantation epiblast, are 5i/L/A (Theunissen *et al.*, 2014) and t2iL + PKCi conditions (Takashima *et al.*, 2014).

1.4.1 – 5i/L/A reprogramming

By using a reporter system, in which the *OCT4* proximal enhancer (PE) is deleted (*OCT4-ΔPE-GFP*), the *OCT4-GFP* reporter, which indicates *OCT4* expression from its distal enhancer, can be used as a reporter of human naïve pluripotency. This system was used to establish which inhibitors can be used in combination with 2i/LIF to allow transgene-independent reprogramming of primed PSCs, to eliminate the need for genetic manipulation of human PSCs. High-throughput screening of kinase inhibitors was used to ascertain the conditions required to stabilise *OCT4-GFP* naïve PSCs.

This led to the formulation of 5i/L/A conditions, composed of five inhibitors; PD0325901 (MEK inhibitor), IM-12 (an alternative, more potent GSK-3 β than CHIR99021), (Schmöle *et al.*, 2010), WH-4-023 (Lymphocyte-specific protein tyrosine kinase/sarcoma (LCK/SRC) inhibitor), SB590885 (a BRAF inhibitor that like PD0325901 suppresses the MEK signalling pathway) and Y-27632 (Rho-associated protein kinase (ROCK) inhibitor), in combination with LIF and the TGF- β family ligand Activin A. The addition of FGF2 (5i/L/A/F) could also stabilise the naïve state of pluripotency (Theunissen *et al.*, 2014).

1.4.2 – t2iL + PKCi reprogramming and Chemical Resetting

As another method to stabilise 2i/LIF-grown naïve PSCs when NK2 overexpression is withdrawn, titrated concentrations of the 2i inhibitors were combined with LIF and protein kinase C inhibitor Gö6983 (t2iL+PKCi) (Takashima *et al.*, 2014), which stabilizes mouse

ground state of pluripotency in mESCs (Dutta *et al.*, 2011). t2iL+PKCi conditions, with the addition of Y-27632 (ROCK inhibitor) to promote cell survival (Watanabe *et al.*, 2007) (t2iL+PKCi+Y), can be used to derive naïve PSC lines, such as the HNES (human naïve embryonic stem cell) directly from the inner mass cell of human pre-implantation embryos (Guo *et al.*, 2016a).

Primed to naïve reprogramming can be initiated by transient broad-spectrum inhibition of histone deacetylases using valproic acid (VPA) and subsequently stabilisation of naïve PSCs using t2iL+PKCi conditions – known as chemical resetting (Guo *et al.*, 2017). Naïve PSCs can be further stabilised by the tankyrase inhibitor XAV939 – t2iL+PKCi +XAV939. The replacement of CHIR99021 with XAV939, generates the naïve PSC growth condition called PXGL, which can be used to stabilise 5i/L/A and chemically reset reprogramming cultures (Bredenkamp *et al.*, 2019b; Rostovskaya *et al.*, 2019).

1.4.3 – Comparison of PTNR methods

There is a large degree of variability in the quality of naïve PSCs produced by PTNR, when compared to the human pre-implantation epiblast, which naïve PSCs model *in vitro*. This has comparison become easier to assess as resources such as single-cell sequencing datasets increase our knowledge of human pre-implantation embryo biology (Blakeley *et al.*, 2015; Petropoulos *et al.*, 2016; Stirparo *et al.*, 2018; Xiang *et al.*, 2020).

Comparing reprogramming protocols suggests that the 5i/L/A (Theunissen *et al.*, 2014) and t2iL+PKCi (Takashima *et al.*, 2014) reprogrammed naïve PSCs had the closest transcriptional profile to human blastocysts (Huang *et al.*, 2014; Stirparo *et al.*, 2018). Other methods showed a reduced transcriptional drift towards a human pre-implantation epiblast signature (Chan *et al.*, 2013b; Gafni *et al.*, 2013; Ware *et al.*, 2014). 5i/L/A and t2iL+PKCi reprogrammed PSCs also generate DNA hypomethylated naïve PSCs with two active X chromosomes in female cells (Theunissen *et al.*, 2016), showing that the epigenetic properties of these naïve PSCs in these conditions have stronger developmental relevance to the *in vivo* state.

There does however remain scope for improvement in the generation and culture of naïve PSCs. For example, the DNA hypomethylated genome in PSCs, but not in embryos, causes loss of inherited imprints (Pastor *et al.*, 2016). Alternative reprogramming methods where MEK inhibitors have been omitted causes protection of crucial DNA imprints, but at the expense of a retention of a partially DNA hypermethylated state (Bayerl *et al.*, 2021; Lynch *et al.*, 2020).

The accessibility, scalability and ability to genetically modify naïve PSCs makes them a strong *in vitro* tool for the study of human embryology and developmentally regulated phenomena, such as human X chromosome inactivation. Naïve PSCs also possess a number of additional, powerful properties that will be discussed in section 1.5.

1.5 – Unique functions of human naïve PSCs

By reprogramming primed PSCs to the naïve state, the level of epigenetic restriction decreases and the differentiation potential of pluripotent cells expands, endowing naïve PSCs with desirable properties that are not possessed by primed PSCs.

1.5.1 – Generation of extra-embryonic lineages

As summarised in section 1.1, the first lineage specification decision allocates totipotent cells of the early human embryo to form either embryonic or extra-embryonic cell types. These extra-embryonic lineage cells are called trophectoderm, and they ultimately develop into the trophoblast cells of the placenta. Several different trophoblast cell types exist within the human placenta. Syncytiotrophoblast acts as a maternal-foetal interface for the exchange of gas and nutrients and is also an endocrine tissue, responsible for the release of pregnancy-supporting hormones (Fournier, 2020). Extravillous trophoblast instead acts as an invasive tissue for uterine implantation and vascularisation by the embryo (Pollheimer *et al.*, 2018).

Human trophoblast stem cells (TSCs) can be derived and cultured *in vitro* from human blastocysts and placenta (Okoe *et al.*, 2018). The ability of TSCs to continuously self-renew in culture provides a sustainable reservoir of placental cells for research, rather than relying on scarce human embryo and placental samples, can be useful for the study of human reproduction and complications such as miscarriage (Dong *et al.*, 2020).

It was subsequently discovered that naïve PSCs can be differentiated into TSCs and subsequently matured into both extravillous and syncytiotrophoblast (Dong *et al.*, 2020). Using different combinations of signalling modulators, this finding has since been corroborated a number of times (Castel *et al.*, 2020; Cinkornpumin *et al.*, 2020; Guo *et al.*, 2020b; Io *et al.*, 2021; Osnato *et al.*, 2021). Emerging evidence suggests that primed PSCs may also be able to produce TSCs, with similar molecular characteristics to naïve PSC derived TSCs (Soncin *et al.*, 2022; Wei *et al.*, 2021), corroborated by other reports (currently available as pre-prints (Viukov *et al.*, 2022; Zorzan *et al.*, 2022)). Naïve PSCs, however, remain the most efficient way in which TSCs can be generated *in vitro*.

Embryonic hypoblast cells are specified from the inner cell mass prior to embryo implantation. Naïve PSCs are also able to differentiate into primitive endoderm cells, which represent the

embryonic hypoblast *in vitro* (Linneberg-Agerholm *et al.*, 2019). In embryos, the hypoblast ultimately develops into the yolk sac, which forms structures such as the umbilical cord for embryo sustenance and survival (Mackinlay *et al.*, 2021; Ross and Boroviak, 2020). The broad differentiation potential of naïve PSCs is likely derived from the prolonged developmental plasticity of human epiblast cells (Guo *et al.*, 2020a).

1.5.2 – Generation of human blastocyst-like ‘blastoid’ structures

As human naïve PSCs recapitulate pre-implantation epiblast cells and retain the ability to generate both trophoblast and primitive endoderm cell states, it was speculated that they could generate human blastocyst models *in vitro*. mESCs are able to generate blastoids when mESCs were aggregated with mouse TSCs (Rivron *et al.*, 2018).

Indeed, a number of studies have confirmed that naïve PSCs can generate three-dimensional human blastocyst-like structures *in vitro* named ‘blastoids’, without the need for aggregation of different stem cell types. Blastoid formation requires the specification of the three defined lineages of the human blastocyst from stem cells; epiblast, trophectoderm and primitive endoderm, as well as their specific three-dimensional spatial organisation by cavitation to envelope the epiblast (Kagawa *et al.*, 2022; Yanagida *et al.*, 2021; Yu *et al.*, 2021). It was also reported that blastoids can be generated during the induced PSC reprogramming of fibroblasts, known as iBlastoids (Liu *et al.*, 2021), however re-analysis and re-evaluation of the transcriptomic data used to benchmark the identity of iBlastoids suggests that these blastoids do not generate any trophectoderm, but instead produce amniotic ectoderm (Zhao *et al.*, 2021). Primed PSCs have not been reported to produce blastoids, and therefore only naïve PSCs can generate authentic human blastocyst models *in vitro*.

In Kagawa *et al.*, 2022, blastoids were generated with a very high efficiency by subjecting naïve PSCs to the presence of triple inhibition of HIPPO (lysophosphatidic acid, LPA), TGF- β (A83-01) and MAPK/ERK (PD0325901) signalling. HIPPO signalling inhibition is important for directing peripheral cells to become trophectoderm (Gerri *et al.*, 2020). Going one step further, human blastoids were able to implant onto a hormonally-stimulated receptive endometrial organoid monolayer, to provide a model for human embryo implantation (Kagawa *et al.*, 2022).

Naïve PSC-derived blastoids are a powerful tool for research. Due to the self-renewal capacity of naïve PSCs, contrasted with the scarcity of human embryos for research, blastoids can provide a sustainable *in vitro* model of human pre-implantation development and reproduction.

Additional advantages of blastoids system include their ability to be generated in only three to four days and their ability to be genetically modified without complex ethical approval and restrictions, allowing study of the effects of genetic perturbations on the model system, albeit there are a number of caveats of blastoids compared to the bona fide study of human embryos.

1.5.3 – Generation of early amniotic cell types

Pre-implantation epiblast cells, which naïve PSCs closely resemble, provide a suitable source material for generating an additional extra-embryonic tissue, the early amnion (Rostovskaya *et al.*, 2022). The amnion forms a membrane, containing fluid that protects the embryo, and acts as a signalling environment to support embryo development (Mamede *et al.*, 2012).

Amniotic tissue arises from epiblast cells around the time of embryo implantation. Capacitation of naïve to primed PSCs represents a cell state transition that mimics molecular events across embryo implantation *in vitro*, and generates human peri-implantation epiblast-like cells. Partially capacitated, peri-implantation epiblast-like, cells can differentiate into early amniotic cells *in vitro* and generate epithelial spheres, which closely resemble primary early amnion cells. Primed PSCs are unable to achieve this outcome under the same conditions (Rostovskaya *et al.*, 2022).

The ability of naïve PSCs to produce amniotic tissue highlights a therapeutic opportunity. Amniotic tissue has powerful wound healing properties, with pro-epithelialisation (Ruiz-Cañada *et al.*, 2021) and pro-angiogenic properties (Duan-Arnold *et al.*, 2015; Maan *et al.*, 2015). Amniotic tissue is able to achieve medically favourable outcomes without immune rejection or the requirement for immunosuppressive prophylaxis (Parolini *et al.*, 2009). As PSC derived amniotic cells do not rely on primary tissue, they could provide an avenue for the high-throughput *in vitro* production of amniotic derived medicines, either as individual factors or using a more complete amniotic tissue formulation, as well as the other applications previously explained throughout section 1.5 (Fig. 1.5.1).

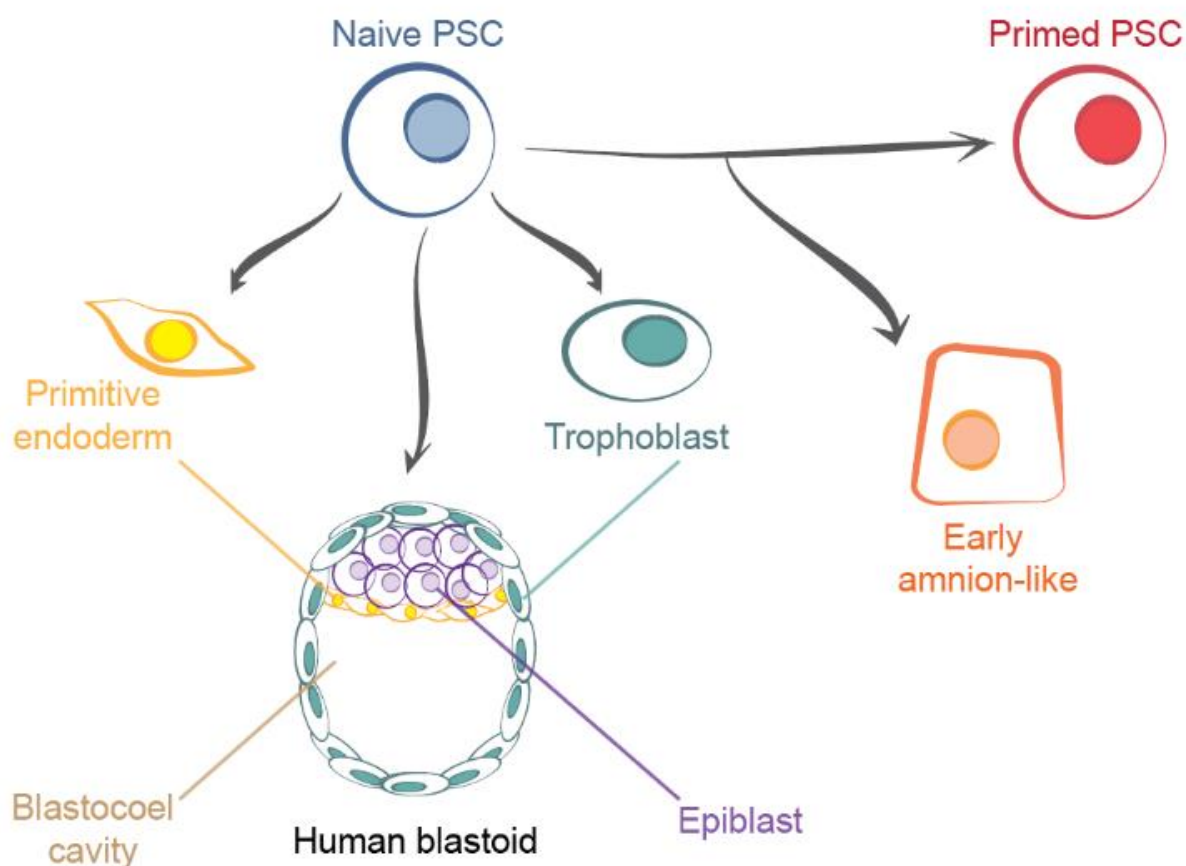


Figure 1.5.1 – The applications of the extended differentiation potential of naïve PSCs

Naïve PSCs are able to generate trophoblast, primitive endoderm and blastocyst-modelling ‘blastoids’ *in vitro*, and partially capacitated cells derived from naïve PSCs can generate early amnion-like cells.

1.5.4 – Defining the regulators of human PTNR could unlock the full potential of human naïve pluripotency

The ability of naïve PSCs to form extra-embryonic and amniotic lineages, as well as blastoids, shows the power of naïve PSCs compared to their primed counterparts, as well as their ability to study earlier stages of human development *in vitro*. Recently, it has even been reported that a sub-population of naïve PSCs capture the properties of the human 8-cell embryo (Taubenschmid-Stowers *et al.*, 2022), meaning that naïve PSCs could model even earlier stages of human development than the pre-implantation epiblast state.

The vast majority of naïve PSCs in culture are derived by human primed-to-naïve reprogramming (PTNR) due to the scarcity of human embryos. The reprogramming process remains poorly understood and is strongly handicapped by its low efficiency, high levels of cell death, production of a heterogeneous mixture of cells and time-consuming nature (Collier *et al.*, 2017; Liu *et al.*, 2020, 2017). This limits our ability to generate high quality naïve PSCs in culture for study and means that human naïve PSCs are comparatively understudied compared to their primed counterpart, despite the additional benefits of naïve PSCs.

Studies have been performed in order to understand transcriptional changes that occur in cells during PTNR (Bi *et al.*, 2022; Collier *et al.*, 2017; Liu *et al.*, 2020), but a huge amount remains unknown about this cell state transition. For example, the factors and pathways needed for cells to enter into the human naïve state by reprogramming are not known. Current reprogramming protocols are crude, with modulators often identified from high-throughput screens, with a limited understanding of the underlying biology. Further knowledge of PTNR could enable optimisation of reprogramming protocols, and allow us to unlock and harness the full potential of human naïve PSCs. In addition, some molecular events that occur in the establishment of naïve PSCs during reprogramming also occur during developmental establishment of pre-implantation epiblast cells, such as global DNA demethylation. Therefore, better understanding the critical steps underpinning human PTNR could also reveal molecular information about cell state transitions in early human development.

1.5.5 – Using cell surface marker proteins to identify human naïve and primed PSCs

In order to study human PTNR, particularly the early stages of this process, a method for isolating nascent naïve PSCs from a heterogeneous mixture of cells is required. Methods of enriching and isolating naïve PSCs are reviewed in (Collier *and* Rugg-Gunn, 2018). Naïve PSCs were traditionally enriched in culture by continuous single-cell passaging, which selectively advantages naïve PSCs due to their strong clonogenicity, but this does not allow isolation of naïve PSCs. To isolate nascent naïve PSCs, early experiments used genetically-modified reporter PSC lines, such as the OCT4- Δ PE-GFP system described previously, where the human *OCT4* coding sequence is tagged with a GFP reporter when the proximal *OCT4* enhancer has been genetically deleted (Theunissen *et al.*, 2014). This approach allows flow-sorting of live OCT4-GFP positive cells as a read-out for human naïve pluripotency. However, the stringency and fidelity of the OCT4 enhancer reporter system is not well understood in

human pluripotency and the ability to isolate naïve PSCs without the requirement for genetic manipulation would be ideal.

Human naïve and primed PSCs possess different cell-surface marker protein signatures (Goodwin *et al.*, 2020). Exploiting these differences allows the segregation and isolation of genetically unmodified PSCs by flow cytometry, allowing isolation of live nascent naïve PSCs during reprogramming. Cell-surface markers identified on primed PSCs include as SSEA-4 (Henderson *et al.*, 2002; The International Stem Cell Initiative, 2007) and CD24 (Shakiba *et al.*, 2015). However, identifying naïve PSCs by simply excluding cells that express cell-surface markers seen on primed PSCs is unsatisfactory for the characterisation of naïve PSCs. Primed PSC markers have been shown to be heterogeneously expressed in naïve PSCs (O'Brien *et al.*, 2017; Pastor *et al.*, 2016) and so an understanding of naïve-specific cell-surface markers is needed.

In Collier *et al.*, 2017, an antibody library screening approach identified a number naïve-specific, primed-specific and shared cell-surface markers. Naïve-specific cell-surface markers include CD75, CD77, and CD130, which are also detected on the cell surface of human pre-implantation blastocysts. Primed PSC-specific cell-surface marker proteins include CD24, CD57 and CD90 (Collier *et al.*, 2017). Subsequent experiments have identified SUSD2 (Sushi Containing Domain 2) was identified as robust cell-surface marker of naïve PSCs and human pre-implantation blastocysts (Bredenkamp *et al.*, 2019a; Wojdyla *et al.*, 2020). Unbiased plasma membrane proteome profiling identified the complete cell-surface marker differences between human primed and naïve PSCs, broadening the repertoire of known cell-surface proteins. Many of these proteins are related to signalling pathways, allowing naïve and primed PSCs to communicate and interact with their environment in different ways (Wojdyla *et al.*, 2020). A summary of naïve and primed-specific cell-surface markers are shown in Fig. 1.5.2.

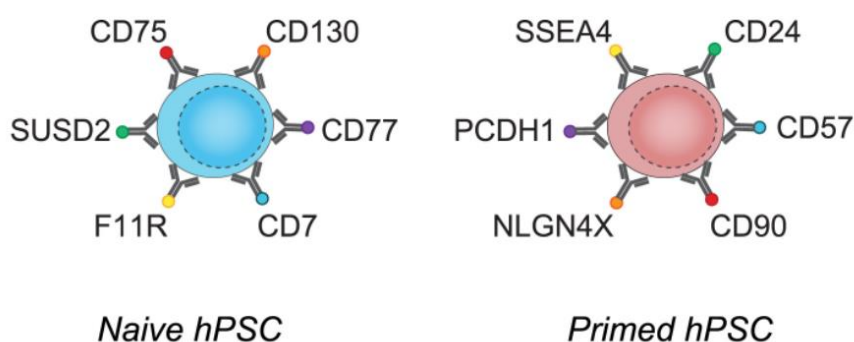


Figure 1.5.2 – Naïve and primed human PSC specific cell-surface markers

CD7, CD75, CD77, CD130, SUSD2 and F11R are naïve enriched cell-surface marker proteins. CD24, CD57, CD90, SSEA-4, PCDH1 and NLGN4X are primed enriched cell-surface marker proteins. Obtained from (Goodwin *et al.*, 2020).

1.5.6 – Using genome-wide CRISPR-Cas9 screening to identify key regulators of human PTNR

We have a substantial gap in our knowledge about PTNR, as the cell state transition that governs entry of cells into human naïve pluripotency. To better understand this process, including the key driving factors and molecular pathways that underpin this process, a previous member of the lab, Amanda Collier, used a genome-wide CRISPR-Cas9 screen to identify the critical regulators of PTNR in an unbiased manner (Collier *et al.*, 2022). This approach was facilitated by rapid developments in genome-wide CRISPR-Cas9 screening technology (Yu and Yusa, 2019) and advances in our understanding of the cell-surface markers of human naïve and primed PSCs.

This genome-wide screen used a fluorescently-conjugated antibody panel, using antibodies against naïve-specific and primed-specific PSCs, that can stringently phenotype live naïve and primed PSCs at single-cell resolution by flow cytometry. This panel used a viability dye to identify live cells and a MEF-specific cell-surface marker Cd90.2 (Ishihara *et al.*, 1987) to allow gating specifically to identify human cells from the MEFs on which they are grown. The panel was then composed of antibodies against the naïve-specific cell surface markers CD75, CD77 and CD130 and the primed-specific cell surface markers CD24 and CD57 (Collier *et al.*, 2017). CAS9-expressing primed PSCs were transduced with an optimised human v3 CRISPR-based loss-of-function mutant guide RNA library, using a sufficient multiplicity of infection to ensure that the guide RNA library is fully represented in the population and six guide RNAs for every protein-encoding gene to ensure successful knockout generation for every protein-encoding gene. Cells were then subject to 5i/L/A reprogramming (Theunissen *et al.*, 2014) and after 10 days, cells were flow-sorted using the fluorescently-conjugated antibody panel, into either nascent naïve PSCs or unsuccessfully reprogrammed 'refractory' cells (Fig. 1.5.3).

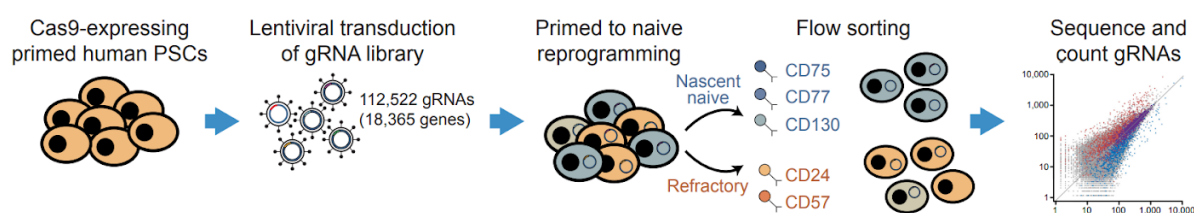


Figure 1.5.3 – Screening strategy to identify regulators of naïve PSC reprogramming

Primed FSPS13B iPSCs, expressing CAS9 constitutively from the AAVS1 (Adeno-Associated Virus Integration Site 1) safe harbour site (Hayashi *et al.*, 2020), were transduced with a genome-wide guide RNA library and selected for transduced cells. 5i/L/A reprogramming was initiated for 10 days and the cell population was cell-sorted using a panel of cell-surface markers into nascent naïve cells (successfully reprogrammed – CD75+, CD77+, CD130+, CD24– and CD57–) and refractory cells (not reprogrammed – CD75–, CD77– and CD130–, irrespective of CD24 and CD57). High-throughput sequencing counted gRNAs in the sorted cell populations.

After cell sorting, the genomic DNA was extracted and sequenced in order to detect the abundance of each guide RNA in the two cell populations. The data was analysed using the MAGeCK algorithm, which corrects for multiple guide RNAs per protein-encoding gene (Li *et al.*, 2014). Guide RNAs over-represented in nascent naïve PSCs indicates that deletion of this gene improves reprogramming, as an impediment gene. Of more interest for my project, guide RNAs over-represented in the refractory cells indicates that deletion of this gene impairs reprogramming, implicating these genes as being ‘essential’ for human PTNR.

Performing the screening study at an early time-point in reprogramming (Day 10) meant that primary, direct regulators of the cell state transition could be identified. This time-point also allowed the capture of enough cells in both sorted populations to achieve statistically significant findings, as cell death occurs at very high levels during PTNR and the nascent naïve PSC population makes up only a small proportion of the live human cells.

1.5.6 – PRC1.3 is one of the most critical epigenetic complexes for human PTNR

Within the scope of this project, I was most interested in complexes that are needed for the key molecular events that reconfigure cell identity across this cell state transition to generate naïve PSCs. The most statistically significant genes identified in this category were largely components of two specific epigenetic complexes, Spt-Ada-Gcn5 acetyltransferase (SAGA) complex and Polycomb Repressive Complex PRC1.3 (Fig. 1.5.4).

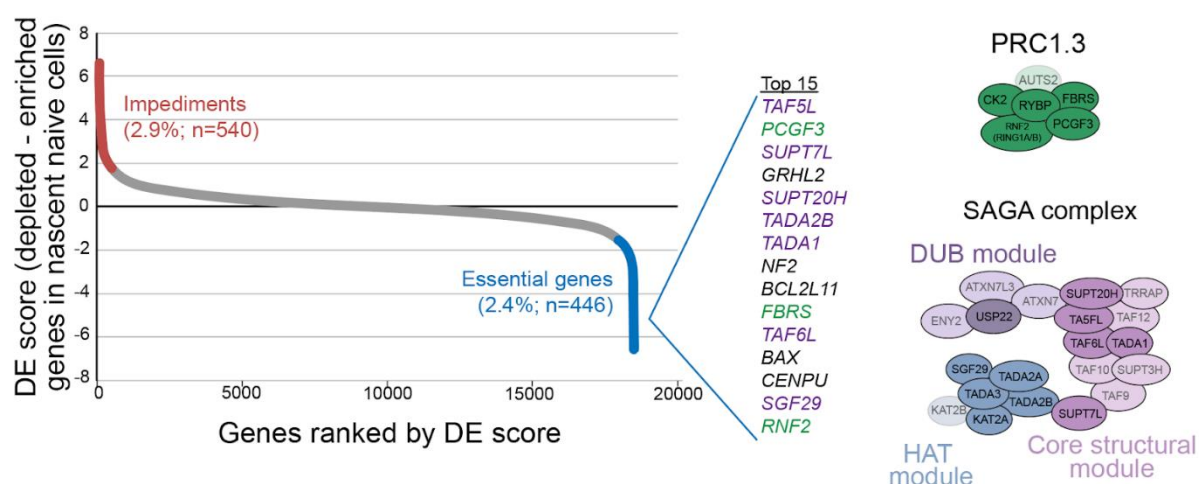


Figure 1.5.4 – Ranked differential-enrichment (DE) plots comparing successfully reprogrammed (nascent naïve) versus failed to reprogramme (refractory) populations

446 ‘essential’ genes and 540 ‘impediment’ genes were identified when applying a $p < 0.02$ threshold using a permutation test. The top fifteen essential genes are shown, with PRC1.3 components labelled in green and SAGA components labelled in purple. The PRC1.3 and SAGA complexes are shown, with genes with significant p -values in the screen highlighted with black outlines.

The identification of multiple components of the same protein complex provides compelling evidence that the entire complex, rather than specific components, may be important for PTNR. Due to the epigenetic differences between human primed and naïve PSCs, a strong degree of epigenetic remodelling must occur as cells reprogramme to the naïve state. We therefore expected to identify epigenetic complexes, but the screen powerfully identified exactly which epigenetic regulators appear to be involved. For example, the chemical resetting protocol (Guo *et al.*, 2017) uses the broad-spectrum histone deacetylase inhibitor VPA acid early in the establishment phase of reprogramming. One could therefore speculate that histone acetylation is important for reprogramming, but the screen can specifically identify the histone acetyltransferase SAGA as a strong candidate for critical histone acetylation.

The identification of almost all of the components of the PRC1.3 complex was one of the most surprising findings of the screen. PRC1.3 has never been studied in the context of human pluripotency or reprogramming. Amongst the genes with the highest level of statistical significance for acting as essential factors in PTNR, three of them were components of the PRC1.3 complex. These are; *RING1B* (#15 factor), which is common to the catalytic core of all PRC1 complexes, *FBRS* (#10 factor), which is unique to PRC1.3 and PRC1.5 complexes,

and *PCGF3* (#2 factor), which is the core, uniquely identity determining component of the PRC1.3 complex. Other identified PRC1.3 genes include *RYBP* (#398 factor), which is common to non-canonical PRC1 complexes such as PRC1.3 and *CSNK2A3* (#88 factor), which might be present within the Casein Kinase II complex, found within PRC1.3 and PRC1.5 complexes (Gao *et al.*, 2012; Hauri *et al.*, 2016).

One of the key, previously reported functions of PRC1.3 and PRC1.5 complexes in pluripotency, is during X chromosome inactivation in mESCs (Almeida *et al.*, 2017; Bousard *et al.*, 2019; Bowness *et al.*, 2022; Nesterova *et al.*, 2019; Pintacuda *et al.*, 2017). As the X chromosome activation status alters between primed and naïve PSCs, with the inactive X chromosome becoming reactivated during reprogramming (Sahakyan *et al.*, 2017), a function for the PRC1.3 complex in reprogramming by regulating of X chromosome activation status made an attractive hypothesis. However, the genome-wide screen was performed in a male (XY) FSPS13B iPSC line, meaning that any candidates identified by this experiment must be X chromosome dosage compensation independent.

Gene ontology analysis of the ‘essential’ genes identified by the genome-wide screen, within the biological processes category, identified the most highly-ranked terms as including ‘regulation of transcription’, ‘histone H3 acetylation’ and ‘histone H2A-K119 monoubiquitination’. These categories suggest that transcriptional and epigenetic regulators underpin the reprogramming process and corroborates the importance of both the SAGA (H3 acetylation) and PRC1.3 complexes (histone H2A lysine 119 monoubiquitination) (Fig. 1.5.5).

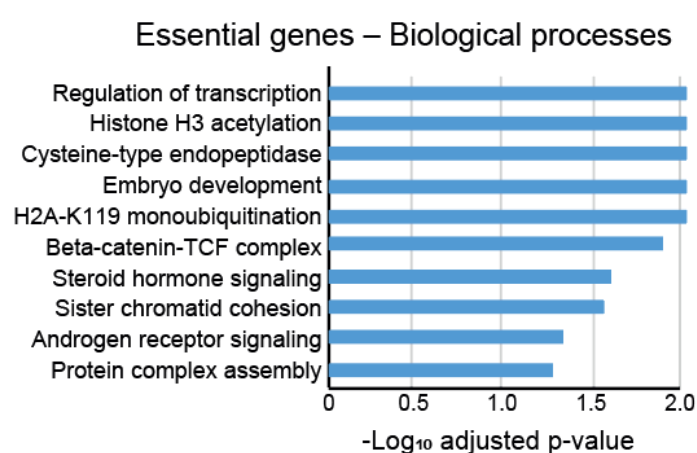


Figure 1.5.5 – Gene ontology analysis of the ‘essential’ regulators of naïve PSC reprogramming

Charts show the biological processes gene ontology category and adjusted p-values, determined by a Fisher exact test.

In summary, genome-wide screening identified PRC1.3 as a key regulator of the formation of the desirable naïve state of pluripotency. An additional screen (Collier *et al.*, 2022) was used to identify genes that are important for human primed PSC proliferation, and unlike the reprogramming screen, this did not identify the core, identity-determining component of the PRC1.3 complex, *PCGF3*.

1.6 – Polycomb Repressive Complexes

1.6.1 – Histone post-translational modifications and epigenetics

Eukaryotic DNA possesses basic units, composed of spherical structures that compact DNA filaments (Oudet *et al.*, 1975). DNA is extensively packaged and compacted to fit into the cell nucleus by histone proteins, which form an octamer on DNA, known as a nucleosome (Arents *and* Moudrianakis, 1993; Oudet *et al.*, 1975). This protein-bound DNA macromolecule is named chromatin and can form higher order structures such as mitotic chromosomes (Paulson *et al.*, 2021). Chromatin dynamics can be regulated by covalent post-translational modifications to the N and C terminal regions of histone octamer proteins on DNA (Bowman *and* Poirier, 2015). These modifications form ‘epigenetic’ regulation within cells, which describes changes to the genome that occur independently of changes in the underlying DNA sequence (Noble, 2015).

A number of different types of covalent modification can occur on defined histone residues, including acetylation, methylation, phosphorylation, ubiquitylation, citrullination, crotonylation and SUMOylation (Bannister *and* Kouzarides, 2011). Each epigenetic modification has enzymatic ‘writers’ that deposit, and ‘erasers’ that undo the modification, as well as ‘readers’ that recognise the modification (Biswas *and* Rao, 2018). How histone modifications are read, often described as the ‘histone code’, influences their functional impact within cells. The same type of modification can have drastically different effects depending on the residue modified, and a number of histone modifications are associated with specific gene regulatory regions. For example, the presence of trimethylation of lysine 4 of histone H3 (H3K4me3) is associated with active gene promoters, trimethylation of lysine 27 of histone H3 (H3K27me3) is associated with repressed gene promoters, and acetylation of lysine 27 and monomethylation of lysine 4 on histone H3 (H3K27ac and H3K4me1) are associated with active gene enhancers (Kimura, 2013).

In certain instances, both active and repressed histone modifications will co-occur at the same regions on chromatin, with gene promoters regions modified with both H3K4me3 and H3K27me3 being described as ‘bivalent’ (Azuara *et al.*, 2006; Bernstein *et al.*, 2006). A number of developmental genes possess bivalent epigenetic modifications and are not expressed but are thought to be ‘poised’ for activation or repression depending upon the developmental signals encountered. Specific histone modifications also play a role in chromatin associated processes such as DNA-damage repair and replication (Karakaidos *et al.*, 2020; Stewart-

Morgan *et al.*, 2020) and can be propagated across cell divisions (Kheir *and* Lund, 2010) and inherited trans-generationally (Fitz-James *and* Cavalli, 2022).

1.6.2 – Identification of Polycomb group genes in *Drosophila*

Genetic screens in *Drosophila melanogaster* were used to identify a number of developmental regulatory genes, whose deletion caused developmentally-defective flies (Lewis, 1947). *Drosophila* body plan development, including axes formation, is largely established by spatially restricted expression of genes within the multinucleated syncytium of embryos (Schuettengruber *et al.*, 2017).

Polycomb group (PcG) genes were identified using a genetic screen (Lewis, 1947). Polycomb acts by developmentally repressing the transcription of Homeotic (Hox) genes, such as Bithorax, thus restricting Bithorax expression to posterior segments of the *Drosophila* embryo (Bajusz *et al.*, 2018; Jürgens, 1985; Lewis, 1978; Sánchez-Herrero *et al.*, 1985). PcG gene mutations cause anterior segments to gain a more posterior fate (Bajusz *et al.*, 2018; Denell *and* Frederick, 1983; Lewis, 1978). The name Polycomb is derived from this developmental aberration seen in the mutant phenotype, whereby additional sex combs appeared aberrantly on the legs of the flies. Transcriptional repression by the *Drosophila* Polycomb protein was confirmed using a GAL4 (Galactose-induced 4)/UAS (upstream activator sequence) system, where artificial recruitment of Polycomb causes heritable silencing of a reporter gene (Müller, 1995). This transcriptional repression finding was corroborated by a mammalian PcG protein named BM1-1 (Bunker *and* Kingston, 1994).

While PcG genes suppress posterior cell fate genes in anterior segments in *Drosophila* embryos, this action is antagonised by Trithorax group (TrxG) genes, with TrxG gene mutations causing posterior segments to become more anterior in character (Ingham, 1985; Schuettengruber *et al.*, 2017; Struhl *and* Akam, 1985). Antagonism is the result of TrxG proteins activating the expression of Hox genes like Bithorax, which are repressed by PcG proteins (Schuettengruber *et al.*, 2017). TrxG proteins are now known to catalyse the H3K4me3 epigenetic modification, associated with active gene promoters (Miller *et al.*, 2001; Mohan *et al.*, 2011).

1.6.3 – Homology between *Drosophila* and mammalian PRC1 complexes

Purification of Polycomb protein from *Drosophila* embryos identified a complex composed of Polycomb, Posterior Sex Combs, Polyhomeotic, Sex Combs on Midleg and *Drosophila* RING1 (dRING1) (Satijn *et al.*, 1997; Saurin *et al.*, 2001; Shao *et al.*, 1999). Human-specific interactors of the *Drosophila* Polycomb protein, determined using a two-hybrid screen, identified RING1B (also known as RING2/RNF2, homologue of dRING1) and BMI-1 (also known as PCGF4, homologue of Posterior Sex Combs) (Reijnen *et al.*, 1995; Satijn *et al.*, 1997). Ectopic targeting of RING1B to genes identified that the human complex is also capable of transcriptional repression (Shao *et al.*, 1999). Purification. This complex was named Polycomb Repressive Complex 1 (PRC1).

Homology comparisons show that mammalian PRC1 complexes have an increased degree of complexity compared to the *Drosophila* complex, which will be discussed in section 1.6.5. Functional homology is also seen in the repression of ATP-dependent chromatin remodelling by the SWI/SNF (Switch/Sucrose non-fermenting) complex, an early characterised function (Francis *et al.*, 2001; Levine *et al.*, 2002; Shao *et al.*, 1999). The presence of a PRC1 complex is conserved from *Drosophila melanogaster* to humans, but is absent in lower eukaryotes such as *Saccharomyces cerevisiae*, *Dictyostelium discoideum*, *Arabidopsis thaliana* and *Caenorhabditis elegans* (de Napoles *et al.*, 2004; Robzyk, 2000).

1.6.4 – Polycomb Repressive Complex 2

Drosophila PcG proteins Extra Sex Combs (ESC) and Enhancer of Zeste (EZ), and their mammalian homologues EED and EZH2 respectively, were shown to interact with one another, and not with known components of PRC1 complexes. This shows that there are two distinct Polycomb Repressive Complexes in both *Drosophila* and mammalian cells, with this second complex discovered named Polycomb Repressive Complex 2 (PRC2) (Ng *et al.*, 2000; Sewalt *et al.*, 1998).

Purified *Drosophila* PRC2 complexes, composed of ESC, EZ and additional PcG proteins SUZ12 and NURF-55, possess methyltransferase enzymatic activity against lysine 27 of histone H3, forming mono-(H3K27me1), di-(H3K27me2) and trimethylated (H3K27me3) modifications to this residue. The methyltransferase activity was contained within the SET domain of the EZ protein, with enzymatic activity strongly activated by its accessory proteins,

and SET domain mutations resulted in derepression of Hox genes in *Drosophila* embryos. This finding linked the function of PRC1 as depositing the H3K27me3 modification for developmental gene silencing (Czermin *et al.*, 2002; Müller *et al.*, 2002).

Mammalian PRC2 complexes, composed of a core of EED, EZH2 and SUZ12 were also confirmed to catalyse H3K27me3 (Cao *et al.*, 2002; Pasini *et al.*, 2004), and the transcriptional repressive action of the PRC2 complex is functionally conserved from flies to mammals (Cao and Zhang, 2004). Histone editing to substitute all lysine 27 residues on all histone H3 isoforms for arginine (K27R) in mESCs showed derepression of a large number of genes and defective gene expression patterns during *in vitro* gastrulation in mouse embryoid body formation (Sankar *et al.*, 2022). With the same common core, two different forms of mammalian PRC2 complex exist (Hauri *et al.*, 2016; Healy *et al.*, 2019; Holoch and Margueron, 2017; Youmans *et al.*, 2021). PRC2.1 complexes consist of EPOC (Beringer *et al.*, 2016) and MTF2 (van Mierlo *et al.*, 2019b). PRC2.2 complexes instead contain JARID2 and AEBP2 (Grijzenhout *et al.*, 2016).

PRC2 plays a key role in mammalian development and homeostasis (Piunti and Shilatifard, 2022) and developmentally-regulated epigenetic phenomena such as X chromosome inactivation (Brockdorff, 2017; Wang *et al.*, 2001). PRC2 has also been shown to play an important role in the *in vitro* differentiation of human primed PSCs, with *EZH2* deletion causing a loss of H3K27me3-dependent repression of developmental regulatory genes and a failure of differentiation into mature specialised cell types, showing the importance of PRC2 and H3K27me3 in pluripotency cell state transitions (Collinson *et al.*, 2016; Shan *et al.*, 2017).

1.6.5 – Polycomb Repressive Complex 1

Although first discovered in *Drosophila melanogaster*, mammals have a much broader repertoire of PRC1 complex genes. In mammals, dRING1 has two homologues, *RING1A* and *RING1B*, Polyhomeotic has three homologues, *PHC1-3*, Posterior Sex Combs has six homologues, *PCGF1-6* (*PCGF4* is also known as *BMI-1*), as well as various homologues of the Polycomb gene, named Chromobox (*CBX*).

Human *RING1B* was characterised as an E3 ubiquitin ligase enzyme, capable of covalent transfer of a ubiquitin moiety on lysine 119 of histone H2A (H2AK119ub1) (H. Wang *et al.*, 2004), with histone ubiquitination acting as a conserved eukaryotic epigenetic modification, predominantly on histones H2A and H2B (Nickel and Davie, 1989). The ubiquitin ligase activity

of RING1B, and its homologue dRING1, is consistent with the presence of a RING finger domain, which is usually found across ubiquitin and small ubiquitin-ligase modifier (SUMO) E3 ligase enzymes (Fig. 1.6.1) (Pickart, 2001).

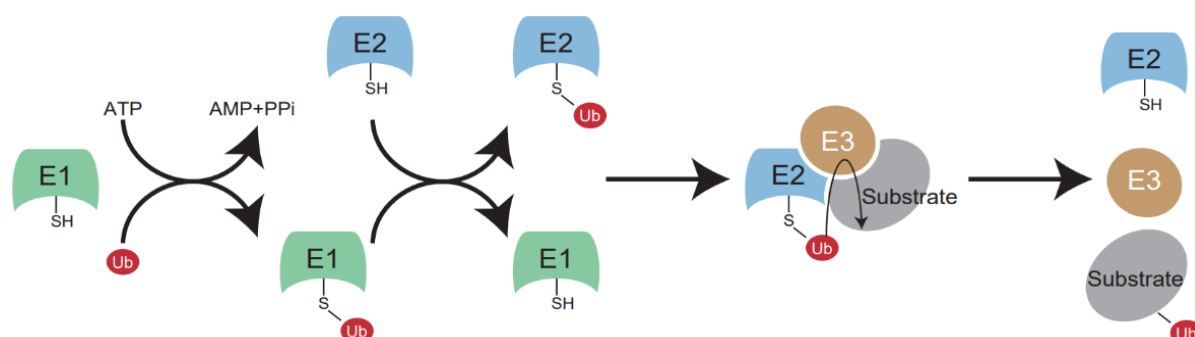


Figure 1.6.1 – The mammalian ubiquitin transfer system

E1 ubiquitin-activating enzymes capture and activate free ubiquitin, which is then transferred to an E2 ubiquitin conjugating-enzyme. The E3 ubiquitin ligase, e.g., RING1B, then catalyses the transfer of ubiquitin to its specific substrate. PPi, pyrophosphate; –SH, cysteine residue side-chain; Ub, free or conjugated ubiquitin group. Adapted from (Gong *et al.*, 2020).

RING finger loss-of-function mutations in human RING1B or *Drosophila* dRING1 caused a loss of H2AK119ub1 (H2AK118ub1 in *Drosophila*) and loss of H2AK118ub1 at the Ultrabithorax gene in *Drosophila* embryos caused gene derepression (Fritsch *et al.*, 2003; H. Wang *et al.*, 2004). Like PRC2 complexes, PRC1 complexes also possess intrinsic enzymatic activity, catalysing an epigenetic modification that is important for developmental gene silencing (Cao *et al.*, 2005; L. Wang *et al.*, 2004).

1.6.6 – Diversity in Polycomb Repressive Complex 1 composition

Due to the increased number of homologues compared to the *Drosophila* PRC1 complex (Levine *et al.*, 2002) and the detection of novel proteins, such as BCOR within PRC1 complexes (Gearhart *et al.*, 2006), it was unclear how mammalian PRC1 complexes are composed. The larger number of proteins expands the complexity of PRC1 complexes and creates mammalian-specific forms of regulation.

In mammals, two different broad classes of PRC1 complexes exist – ‘canonical’ complexes and ‘non-canonical’ or ‘variant’ PRC1 complexes. Our understanding of mammalian PRC1 complexes has been shaped by elucidating their composition. PRC1 complexes were subsequently discovered in *Drosophila* that did not contain the chromodomain-containing protein Polycomb (Lagarou *et al.*, 2008; Lee *et al.*, 2015). Indeed, only some mammalian PRC1 complexes contain Polycomb homologues, named Chromobox (CBX) 2, 4, 6, 7 and 8, and these are named canonical PRC1 complexes, based on the canonical understanding of PRC1 complexes from *Drosophila*. PRC1 complexes that do not contain CBX proteins instead possess either RYBP (RING1B and YY1 binding protein) or its homologue YAF2 (YY1 associated factor 2), and these are variant PRC1 complexes (Gao *et al.*, 2012).

The catalytic components of the complex RING1A and RING1B were detected within all PRC1 complexes, meaning that they form a core module of the complexes and that all PRC1 complexes possess enzymatic potential. The six Polycomb group ring finger (PCGF) proteins, PCGF1-6, are present mutually exclusively within all PRC1 complexes, showing that one PCGF protein accompanies RING1A/B in the core of PRC1 complexes. The identity of PRC1 complex is determined by which PCGF protein is present at its core, for example PCGF3-containing PRC1 complexes are named PRC1.3 complexes (Gao *et al.*, 2012).

Canonical, CBX protein-containing PRC1 complexes, include some PRC1.2 and PRC1.4 complexes. The remaining PRC1.2 and PRC1.4 complexes, and all PRC1.1, PRC1.3, PRC1.5 and PRC1.6 complexes contain RYBP/YAF2 and are variant complexes (Gao *et al.*, 2012). RYBP and CBX proteins share a common binding site on RING1B and so bind mutually exclusively (Chittock *et al.*, 2017; Wang *et al.*, 2010).

Each PRC1 complex subtype possesses a distinct set of accessory proteins, which can provide them with unique functionality. For example, FBRS (Fibrosin), AUTS2 (Autism susceptibility candidate 2) and the Casein Kinase II complex are present within PRC1.3 and PRC1.5 complexes, but absent from all other sub-types (Gao *et al.*, 2012; Hauri *et al.*, 2016). The known composition of all PRC1 complexes is summarised in Fig. 1.6.2.

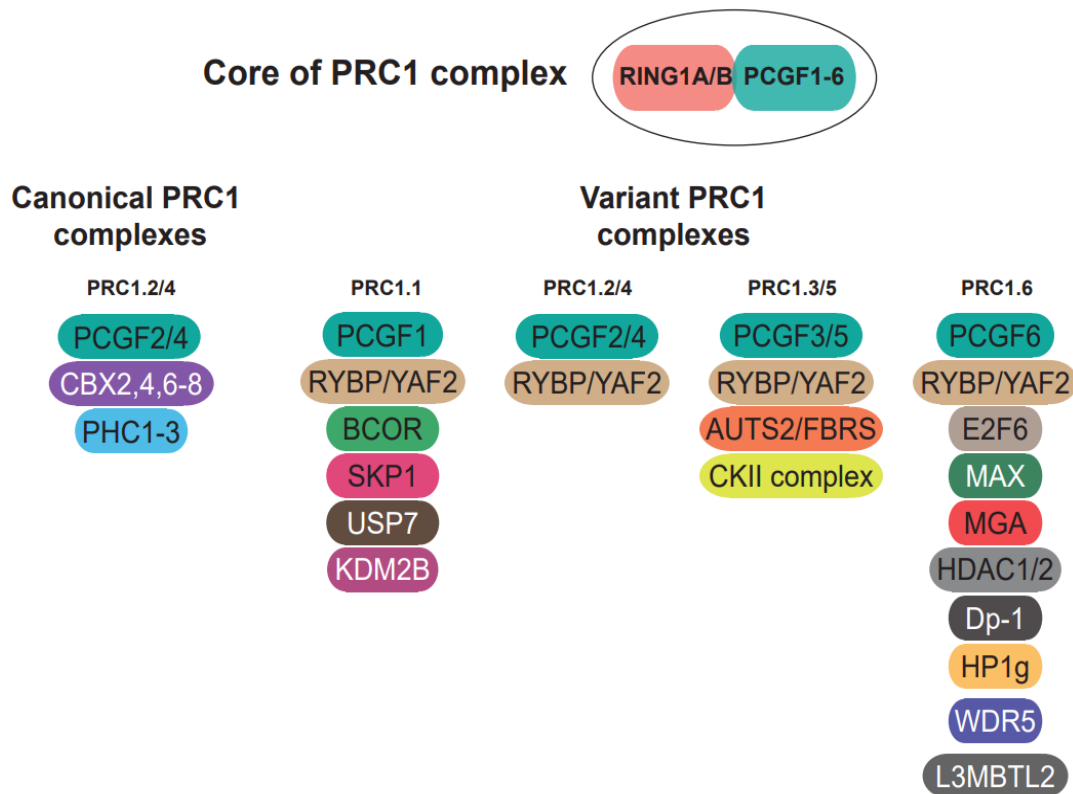


Figure 1.6.2 – Diversity of mammalian PRC1 complexes

Significant diversity exists within PRC1 complexes, with a common core composed of RING1A/RING1B and a single PCGF protein. Each PRC1 complex sub-type contains distinct accessory proteins, and can either be categorised as canonical, by the presence of CBX proteins, or variant, by the presence of RYBP/YAF2. Adapted from (Aranda *et al.*, 2015).

1.7 – Cooperativity between PRC1 and PRC2 complexes

1.7.1 – Canonical and variant PRC1 complexes

Canonical and variant PRC1 complexes differ in their composition and their chromatin recruitment mechanisms. Chromodomains bind to chromatin (Messmer *et al.*, 1992), and the chromodomain in the *Drosophila* Polycomb and mammalian CBX proteins is homologous to the chromodomain within *Drosophila* heterochromatin-binding protein HP1, which binds H3K9me3 (Fischle, 2003; Paro and Hogness, 1991). The Polycomb chromodomain specifically recognises the H3K27me3 epigenetic modification deposited by PRC2 complexes (Cao *et al.*, 2002; Fischle *et al.*, 2003). The specificity of H3K27me3 binding was confirmed by swapping the chromodomain of Polycomb and HP1, which reversed their specificities and caused HP1 recruitment to H3K27me3 instead (Fischle *et al.*, 2003). All *Drosophila* PRC1 complexes contain the chromodomain-containing Polycomb protein and are recruited to chromatin dependent upon H3K27me3 (Cao *et al.*, 2002; Fischle, 2003; Gao *et al.*, 2012; L. Wang *et al.*, 2004).

As chromodomains with canonical PRC1 complexes bind H3K27me3, these complexes in both *Drosophila* and mammals are hierarchically recruited to chromatin in an H3K27me3-dependent manner (Cao *et al.*, 2002; Czermin *et al.*, 2002; Fischle *et al.*, 2003; Min *et al.*, 2003; L. Wang *et al.*, 2004), with PRC2 deletion and loss of H3K27me3 in mESCs resulting in reduced canonical PRC1 component binding to chromatin (Tavares *et al.*, 2012). In *Drosophila*, PcG proteins are targeted to specific elements within the genome, named Polycomb response elements (PREs) (Czermin *et al.*, 2002; Kassis and Brown, 2013). In mammals, no PREs or equivalent have been discovered and targeting is instead often enriched at hypomethylated CpG islands (Schuettengruber *et al.*, 2017; Tanay *et al.*, 2007).

1.7.2 – H3K27me3 independent recruitment of variant PRC1 complexes

Variant, RYBP-containing PRC1 complexes are able to bind chromatin even in the absence of H3K27me3 in mESCs (Schoeftner *et al.*, 2006; Tavares *et al.*, 2012). Therefore, an H3K27me3-independent recruitment mechanism to chromatin for PRC1 complexes must also exist. As a factor common to variant PRC1 complexes, RYBP was an attractive candidate for the recruitment of these complexes to chromatin. However, *RYBP* knockout in mESCs was

able to retain RING1B occupancy on chromatin, making recruitment dependent upon RYBP unlikely (Tavares *et al.*, 2012). It is poorly understood how exactly variant PRC1 complexes become recruited to chromatin, but experiments in PRC2-deficient cells treated with the reversible inhibitor MG132, which completely abolishes H2AK119ub1 from cells, confirms that variant PRC1 complexes can establish H2AK119ub1 in the absence of any pre-existing H2AK119ub1 and H3K27me3 (Tavares *et al.*, 2012). It is therefore thought that accessory proteins within variant PRC1 complexes might play a role. For example, the KDM2B component of PRC1.1 complexes is able to directly bind chromatin at CpG island regions independently of H3K27me3 (Blackledge *et al.*, 2014; Farcas *et al.*, 2012). The PRC1.6 complex contains E2F6 and MGA-MAX dimers that may directly bind DNA (Gao *et al.*, 2012) and while PRC1.3/5 subunits lack an auxiliary component that directly binds to DNA, proteins including USF1 and the RNA binding protein hnRNPK have been identified as recruiting factors of PRC1.3/5 to chromatin (Pintacuda *et al.*, 2017; Scelfo *et al.*, 2019).

Although direct *de novo* recruitment by RYBP is unlikely, RYBP contains a NZF (Npl4 zinc finger) domain, which is able to bind PRC1 deposited H2AK119ub1 (Arrigoni *et al.*, 2006). This is proposed to act as a positive feedback loop to allow variant PRC1 complexes to propagate H2AK119ub1 and promote stable gene silencing (Zhao *et al.*, 2020), supported by the binding of variant PRC1 complexes at sites on chromatin with higher levels of H2AK119ub1 than canonical PRC1 complexes (Tavares *et al.*, 2012).

1.7.3 – H2AK119Ub1-dependent PRC2 complex recruitment

JARID2 and AEBP2 were both detected as interactors of H2AK119ub1 on chromatin, and both are present within the mammalian PRC2.2 complex subtype (Hauri *et al.*, 2016; Healy *et al.*, 2019; Holoch *and* Margueron, 2017; Kalb *et al.*, 2014). This raised the possibility that PRC2 complexes might be able to bind chromatin downstream of PRC1-deposited H2AK119ub1. Forcing ectopic H2AK119ub1 deposition on chromatin is able to recruit PRC2 and promote H3K27me3 deposition, dependent upon JARID2 (Cooper *et al.*, 2016, 2014). JARID2 contains a ubiquitin interacting motif that allows H2AK119ub1-dependent recruitment of JARID2-containing PRC2 complexes (Cooper *et al.*, 2016, 2014; Fisher *et al.*, 2003). This H2AK119ub1-dependent PRC2 recruitment is consistent with depletion of H2AK119ub1 in mESCs causing a reduction in H3K27me3 levels (Endoh *et al.*, 2012).

Using an ectopic targeting system to recruit PRC1 and PRC2 complexes to an inert, unmodified region of DNA added into the mESC genome demonstrated that PRC1 complexes

are capable of H2AK119ub1 catalysis in the absence of pre-existing H3K27me3, and that H2AK119ub1 alone is able to recruit PRC2 complexes that establish H3K27me3 (Blackledge *et al.*, 2014), likely dependent upon the presence of JARID2 (Cooper *et al.*, 2016). The recruitment of PRC2 downstream of PRC1-deposited H2AK119ub1 has been observed in the earliest stages of zebrafish and mouse embryos post-fertilisation (Chen *et al.*, 2021; Hickey *et al.*, 2022; Mei *et al.*, 2021). Likewise, PRC2 complexes are capable of H3K27me3 catalysis in the absence of pre-existing H2AK119ub1, and H3K27me3 alone is able to recruit canonical PRC1 complexes. This system creates a positive feedback loop to reinforce H2AK119ub1 and H3K27me3 deposition at common sites (Blackledge *et al.*, 2014) (Fig. 1.7.1).

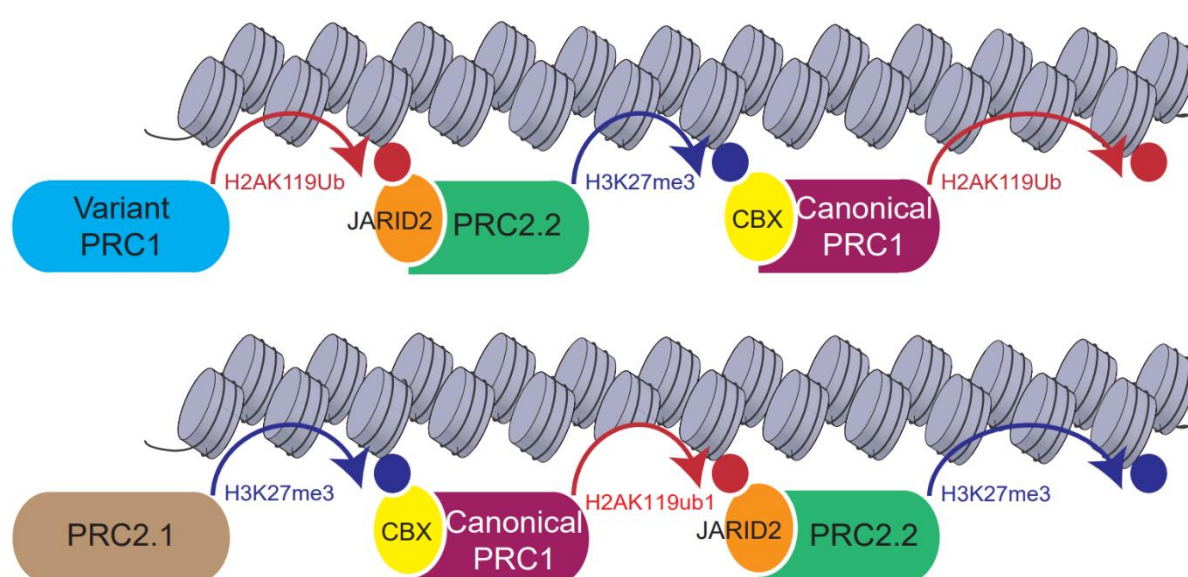


Figure 1.7.1 – Interdependence of PRC1 and PRC2 complexes in chromatin recruitment

PRC1 and PRC2 complexes can both be recruited *de novo* to unmodified chromatin and deposit H2AK119ub1 and H3K27me3 respectively, which can subsequently recruit PRC2 and PRC1 respectively and promote co-localisation of the two modifications (Blackledge *et al.*, 2015, 2014).

Notably, canonical PRC1 complexes recruited downstream of H3K27me3 in the targeting system did not catalyse large amounts of H2AK119ub1 (Blackledge *et al.*, 2014). This is because canonical PRC1 complexes are less enzymatically active than variant PRC1 complexes (Taherbhoy *et al.*, 2015) and a large proportion of cellular H2AK119ub1 is catalysed in an H3K27me3-independent manner (Blackledge *et al.*, 2014). RYBP has also been shown to stimulate the enzymatic activity of a reconstituted PRC1.1 complex by stabilising a catalytically active conformation of the complex (Rose *et al.*, 2016).

In summary, canonical PRC1 complexes are recruited to chromatin dependent upon H3K27me₃, while variant PRC1 complexes are recruited to chromatin independently of H3K27me₃, predominantly through sub-type specific accessory proteins, and catalyse higher levels of the H2AK119ub1 modification. PRC2.2 complexes are recruited to chromatin dependent upon H2AK119ub1. These interlaced mechanisms reinforce the co-localisation of both of these modifications at 'Polycomb domains' (Blackledge *et al.*, 2014).

1.8 – Polycomb Repressive Complex 1 function

1.8.1 – Regulation of transcription by H2AK119ub1

PRC1-catalysed H2AK119ub1 is associated with developmental gene silencing, supported by the exclusion of H2AK119ub1 from DAPI-dense regions, which indicates that the modification is more strongly associated with facultative heterochromatin, such as at silenced genes, rather than constitutive heterochromatin (de Napoles *et al.*, 2004). Both RING1B and the H2AK119ub1 modification are present at a number of 'bivalent' gene promoters, which are marked with both the active epigenetic modification H3K4me3 and the repressive modification H3K27me3 (Stock *et al.*, 2007). H2AK119ub1 appears to play a role in the inhibition of transcriptional elongation, showing a co-localisation with promoter-proximal, non-elongating RNA polymerase II (Brookes *et al.*, 2012; Stock *et al.*, 2007; Zhou *et al.*, 2008). Conditional PRC1 deletion resulted in stalled RNA polymerase II release and productive transcriptional elongation at these genes, which are considered 'poised' for expression (Brookes *et al.*, 2012).

An inducible depletion for the catalytic component of PRC1 complexes RING1B in mESCs was able to rapidly ablate genome-wide H2AK119ub1 levels and upregulate Polycomb gene targets, independently of H3K27me3 dynamics (Dobrinić *et al.*, 2021), with H3K27me3 turning over on chromatin slowly and diluting out with each DNA replication and cell division (Jadhav *et al.*, 2020). This suggests that H2AK119ub1 may be more critical for repression of transcription than H3K27me3 in certain contexts. H2AK119ub1 removal allows the RNA polymerase II engagement on chromatin, suggesting that H2AK119ub1 might in fact regulate transcription initiation rather than elongation (Dobrinić *et al.*, 2021). H2AK119ub1 is proposed to constrain expression of genes expressed at low levels, which is alleviated by RING1B depletion (Dobrinić *et al.*, 2021), and exacerbated by depletion of the H2AK119ub1 eraser or deubiquitinase BAP1 (Fursova *et al.*, 2021).

1.8.2 – Developmental importance of PRC1 complexes

PRC1 and PRC2 complexes deletions result in developmental delay in mouse embryos (O'Carroll *et al.*, 2001; Schumacher *et al.*, 1996; Voncken *et al.*, 2003), with *Ring1b* deletion causing arrest at gastrulation and embryonic lethality (Voncken *et al.*, 2003). Experiments in oocytes and early zebrafish and mouse embryos suggest that H2AK119ub1 plays a critical

role in the silencing of premature zygotic genome activation immediately post-fertilisation, and in the recruitment of PRC2 to establish *de novo* embryonic H3K27me3 deposition (Chen *et al.*, 2021; Hickey *et al.*, 2022; Mei *et al.*, 2021). This shows the importance of PRC1 complexes and H2AK119ub1 in vertebrate, and particularly, mammalian development.

In mESCs, conditional deletion of *Ring1b* strongly impairs the morphology and self-renewal capacity over long-term culture, with *Ring1A* only able to contribute to a small amount of H2AK119ub1 *in vivo* (Endoh *et al.*, 2008). Knockdown of *Ring1A* alone in mouse embryos showed only a mild phenotype, showing that RING1B is the dominant homologue with the complex (del Mar Lorente *et al.*, 2000; Voncken *et al.*, 2003). In human PSCs, *RING1B* (Shan *et al.*, 2020) and *RING1A* knockouts (Zhang *et al.*, 2022) have been generated with little effect on the pluripotent state, but chemical inhibition of both RING1A and RING1B causes biased differentiation dynamics towards the neuroectoderm state in embryoid body formation (Desai *et al.*, 2020; Ismail *et al.*, 2013). PRC1 complexes and H2AK119ub1 appear therefore to be important for mammalian developmental progression and multi-lineage differentiation.

1.8.3 – Catalysis-dependent Polycomb Repressive Complex 1 functions

To study the global effects of the catalytic activity of PRC1 complexes, structural analyses were performed to identify specific residues within RING1B that could be mutated to obliterate the catalytic activity of the enzyme. This approach has the advantage that it does not require gene deletions, which may have secondary effects, and allows H2AK119ub1 to be studied in isolation, separably from PRC1. The substitution of RING1B isoleucine 53 (e.g. to alanine, I53A) was used to catalytically inactivate PRC1 complexes (Buchwald *et al.*, 2006; Eskeland *et al.*, 2010; Illingworth *et al.*, 2015). However, an additional aspartic acid 56 to lysine (D56K) mutation is required to fully disrupt the E2 to E3 ubiquitin transfer interface and fully abolish RING1B enzymatic activity (Blackledge *et al.*, 2020; Tsuboi *et al.*, 2018). As RING1A also possesses some enzymatic activity (de Napoles *et al.*, 2004; Endoh *et al.*, 2008), RING1A catalytic inactivation mutations (I50A and D53K) are also required to completely eliminate all H2AK119ub1 from cells (Blackledge *et al.*, 2020).

Catalytic inactivation of RING1B alone (I53S) resulted in gene derepression (Tamburri *et al.*, 2020), corroborated by the conditional catalytic inactivation of both RING1A and RING1B in mESCs, which derepresses ~3,000 Polycomb target genes while not affecting PRC1 binding to chromatin (Blackledge *et al.*, 2020). This evidences that the H2AK119ub1 modification is important for regulation of gene expression in the pluripotent state. In addition, attempts to

create catalytically inactivating RING1A and RING1B mutations in mESCs non-conditionally was lethal, demonstrating that H2AK119ub1 is likely essential for the survival of pluripotent mESCs (Blackledge *et al.*, 2020).

PRC1 complex subtypes have their own distinct protein interaction networks (Gao *et al.*, 2012; Hauri *et al.*, 2016) and individual PRC1 complex subtypes contribute differently towards H2AK119ub1 catalysis and gene regulation. The combined deletion of both *Pcgf2* and *Pcgf4/Bmi-1* in mESCs showed almost no effect on either H2AK119ub1 levels or gene expression, consistent with low rates of catalytic activity of canonical PRC1 complexes (Fursova *et al.*, 2019). PRC1.1 complexes are associated with Polycomb target gene promoter enriched H2AK119ub1, while the entire genome is blanketed with low levels of PRC1.3/5 dependent H2AK119ub1 (Almeida *et al.*, 2017; Fursova *et al.*, 2019; Scelfo *et al.*, 2019).

In spite of the individual contributions to H2AK119ub1 levels and distribution, a large degree of synergy between variant PRC1 complexes has been seen in mESCs with respect to regulation of gene expression (Fursova *et al.*, 2019; Zhu *et al.*, 2022). A quadruple deletion of all of the variant specific PCGF genes, *Pcgf1*, *Pcgf3*, *Pcgf5* and *Pcgf6* was required to cause a large magnitude of gene derepression. The existence of multiple variant PRC1 subtypes likely buffers regulation of gene expression and can compensate the effect of individual deletions (Fursova *et al.*, 2019). However, other studies have shown instead that PCGF proteins are distinct and that the distribution of these proteins on chromatin did not alter in the presence of deletions in other PCGF proteins, arguing against synergy (Scelfo *et al.*, 2019).

1.8.4 – Catalysis-independent Polycomb Repressive Complex 1 functions

The assumption that the action of PRC1 complexes in developmental gene silencing is mediated solely through catalysis of the H2AK119ub1 modification has been challenged. *Drosophila* PRC1 complexes can compact chromatin *in vitro* (Francis *et al.*, 2004) and within mammalian cells (Eskeland *et al.*, 2010). A catalytically inactivating RING1B point mutation (I53A) (Buchwald *et al.*, 2006) was able to sustain chromatin compaction and *Hox* gene repression – suggesting that regulation of transcription by PRC1 complexes does not act solely through H2AK119ub1 (Eskeland *et al.*, 2010). However, RING1A was not catalytically inactivated or deleted, meaning that some H2AK119ub1 was likely retained that could have sustained this gene repression.

Homozygous RING1B I53A mouse embryos could progress through gastrulation as far as embryonic day 15.5, suggesting that catalytically active PRC1 complexes could facilitate a

substantial degree of transcriptional repression during lineage specification via chromatin compaction. However, the retained presence and activity of RING1A could provide a likely explanation for this sustained developmentally-regulated transcriptional repression. RING1B I53A embryos did not produce live pups, indicating that the RING1B catalytic activity is ultimately important for complete developmental progression (Illingworth *et al.*, 2015). The corresponding mutation in *Drosophila*, where only one catalytically active protein is present, also had a limited effect on *Hox* gene expression in *Drosophila* embryos, which were able to complete all stages of larval development with few observable defects (Pengelly *et al.*, 2015). The I53A RING1B mutation has minimal impact on chromatin compaction of the *HoxA* gene cluster when directly visualised by super-resolution microscopy (Kundu *et al.*, 2017).

Canonical PRC1 complex components CBX2 and PHC2 were confirmed to strongly compact chromatin in mice. For CBX2, chromatin compaction was facilitated by a positively charged region (Grau *et al.*, 2011), while PHC2 possesses a sterile alpha motif that causes sequential polymerisation of PRC1 complexes on chromatin into clusters. Disruption of the CBX2 positive region or the PHC2 sterile alpha motif resulted in mouse posteriorisation transformations, consistent with defective *Hox* gene silencing and mammalian body patterning (Isono *et al.*, 2013; Lau *et al.*, 2017). As well as chromatin compaction, PRC1 complexes also appear to regulate chromatin on a finer scale. PRC1 complex proteins playing a role in nucleosome enrichment and spacing over gene promoters, in a manner that was not observed for PRC2 proteins, perhaps regulating gene transcription with more complexity than simply occluding promoters from the transcriptional apparatus by compaction (King *et al.*, 2018).

The catalysis independent functions of PRC1 complexes are predominantly attributed to canonical PRC1 complexes, which have a lower enzymatic activity than variant PRC1 complexes (Blackledge *et al.*, 2014; Gao *et al.*, 2012; Rose *et al.*, 2016; Taherbhoy *et al.*, 2015). Canonical PRC1 complexes are proposed to strongly influence genome organisation and mediate long-range chromatin interactions (Boyle *et al.*, 2020; Kundu *et al.*, 2017; Ogiyama *et al.*, 2018; Schoenfelder *et al.*, 2015), with self-interacting Polycomb-associated genomic compartments anchored by PRC1 and PRC2 present in mouse oocytes and embryos (Du *et al.*, 2020).

In summary, both catalysis of H2AK119ub1, principally by variant PRC1 complexes, and H2AK119ub1-independent regulation, predominantly by canonical PRC1 complexes, are both important for the regulation of pluripotency and development.

1.9 – PRC1.3 and PRC1.5 complexes

1.9.1 – PRC1.3 and PRC1.5 complex function

The variant PRC1.3 and PRC1.5 subtypes of PRC1 complexes are highly understudied and their functions are not well understood. PRC1.3 and PRC1.5 share a common protein interactome, with PCGF3 or PCGF5 binding the same accessory proteins; RING1A/B, RYBP/YAF2, the Casein Kinase II complex, AUTS2 and FBRS (Gao *et al.*, 2012; Hauri *et al.*, 2016), and are largely assumed to have shared functions.

In mESCs, PRC1.3 and PRC1.5 contribute to around half of global H2AK119ub1 levels, with a low-level genome-wide blanketing of the modification, rather than an enrichment at gene promoters. PRC1.3/5 dependent H2AK119ub1 however shows very little regulation of gene expression in steady-state conditions (Fursova *et al.*, 2019). The abundance of PRC1.3/5-dependent H2AK119ub1 suggests that these complexes have a very high catalytic activity in cells. Single-molecule live-cell imaging estimated that each of the ~6,000 PRC1.3 and PRC1.5 complexes in mESCs deposit an H2AK119ub1 modification every 16 seconds (Huseyin *and* Klose, 2020). The H2AK119ub1 modification is rapidly turned over, unlike H3K27me3 (Dobrinić *et al.*, 2021) and is removed by the H2AK119 deubiquitinase BAP1, which is proposed to counteract genome-wide PRC1.3/5-dependent H2AK119ub1 and prevent more widespread transcription repression (Conway *et al.*, 2021; Fursova *et al.*, 2021).

Attempts to profile PRC1.3 binding to chromatin, which might be difficult due to its rapid chromatin binding kinetics (Huseyin *and* Klose, 2020), identified a very small number of target genes, 80 % of which did not overlap with RING1A/B binding sites, and were not strongly enriched for the H2AK119ub1 modification. The heterodimer USF1/2 was also identified as a recruitment mechanism to target the PRC1.3 complex to chromatin (Scelfo *et al.*, 2019). PRC1.3 (and PRC1.5) complexes have been specifically implicated in a number of context-specific roles with varying functional effects that I discuss in the following sections.

1.9.2 – X-chromosome inactivation

The Polycomb catalysed H2AK119ub1 and H3K27me3 epigenetic modifications are abundant along the transcriptionally silenced X chromosome (Xi) in female cells (de Napoles *et al.*, 2004; Plath *et al.*, 2003; Silva *et al.*, 2003). It was discovered that variant PRC1.3/5 complexes are

specifically critical for the initiation of X chromosome inactivation (Almeida *et al.*, 2017). PRC1.3/5 complexes are recruited to the inactive X chromosome by the B/C repeat domains of the *Xist* long non-coding RNA, with an interaction that is bridged by Heterogeneous Nuclear Ribonucleoprotein K (hnRNPK) (Pintacuda *et al.*, 2017). Deletion of either *Pcgf3/Pcgf5* or the Polycomb interaction domain within *Xist* results in loss of H2AK119ub1 and H3K27me3 epigenetic modifications, loss of gene silencing, increased chromatin accessibility across the X chromosome and female-specific embryo degradation (Almeida *et al.*, 2017; Pintacuda *et al.*, 2017).

The ectopic integration of an *Xist* transgene into an autosomal chromosome and induction of *Xist* expression is sufficient for PRC1.3/5 complex recruitment and deposition of H2AK119ub1 along the length of the chromosome (Fursova *et al.*, 2019; Nesterova *et al.*, 2019). PRC1.3/5 deposited H2AK119ub1 acts upstream of PRC2.2 and other variant PRC1 complexes (Almeida *et al.*, 2017; Grijzenhout *et al.*, 2016; Nesterova *et al.*, 2019), and acts in parallel alongside other pathways such as SPEN-mediated X chromosome inactivation (Bowness *et al.*, 2022; Nesterova *et al.*, 2019). Although initiation of transcriptional silencing of the inactive X chromosome has been reported in the absence of PRC1.3/5, the complexes were still reported to be important in the stabilisation of early gene silencing (Bousard *et al.*, 2019), with differentiation apparently required for the complete transcriptional silencing of the inactive X chromosome (Bowness *et al.*, 2022).

1.9.3 – Transcriptional repression

A small number of instances exist in which PRC1.3 complexes have been seen to specifically repress transcription, in a way that cannot be compensated by other variant PRC1 complexes. The gene *Meis2* is important in mouse limb development, being expressed specifically in the proximal forelimb bud, activated by retinoic acid signalling (Yakushiji-Kaminatsui *et al.*, 2018, 2016). PRC1.3/5 complexes specifically repress expression of *Meis2* in the mouse distal forelimb bud, with deletion of *Pcgf3* and *Pcgf5* resulted in gene derepression and loss of H3K27me3 in tissues (Yakushiji-Kaminatsui *et al.*, 2018). PCGF5 binding at specific gene promoters was also seen in mESCs, with many of these genes becoming derepressed when *Pcgf5* was deleted during neuronal differentiation (Yao *et al.*, 2018).

1.9.4 – Transcriptional activation

Although they are named Polycomb *Repressive* complexes, emerging evidence opines that variant PRC1 complexes may actually be able to perform the reverse function and activate transcription. This phenomenon was first described for PRC1.3 and PRC1.5 complexes.

AUTS2 and FBRS are components of mammalian PRC1.3 and PRC1.5 complexes (Gao *et al.*, 2012; Hauri *et al.*, 2016). *AUTS2* is a paralogue of *FBRS* and *FBRS1*, arising from a genome duplication event and possessing evolutionary homology and sequence similarity (Kondrychyn *et al.*, 2017; Sellers *et al.*, 2020; Singh *et al.*, 2015). *AUTS2* has been studied in the field of neurobiology, with mutations associated with autism spectrum disorders (Sultana *et al.*, 2002), delayed development, microcephaly and epilepsy (Beunders *et al.*, 2013; Oksenberg *and* Ahituv, 2013).

Auts2 has been shown to play an important role in neurodevelopment in zebrafish (Beunders *et al.*, 2013; Oksenberg *et al.*, 2013) and mice (Bedogni *et al.*, 2010; Hori *et al.*, 2014), and this is proposed to be as a transcriptional activator within PRC1.3/5 complexes. Genes bound by AUTS2 at their promoters were highly expressed in the developing mouse brain (Gao *et al.*, 2014; Oksenberg *et al.*, 2014). Artificial recruitment of AUTS2 to a reporter gene resulted in PRC1 complex recruitment and reporter gene activation, dependent upon the presence of the histone acetyltransferase P300 (Gao *et al.*, 2014), with this finding subsequently corroborated for the proteins PCGF3 and PCGF5 (Scelfo *et al.*, 2019). These findings suggest that AUTS2-PRC1.3/5 complexes can activate transcription in specific contexts.

As H2AK119ub1 represses transcription, as previously described, it was unclear how enzymatically active PRC1.3/5 complexes might instead activate transcription. The casein kinase II complex within AUTS2-PRC1.3/5 complexes can phosphorylate RING1B on serine 41 and 168, which inhibits H2AK119ub1 catalysis by RING1B. The ability of AUTS2-PRC1.3/5 complexes to interact with P300 provides a model for how gene transcription is activated by these complexes, with inhibition of transcriptionally repressive H2AK119ub1 and deposition of transcriptionally permissive histone acetylation modifications (Fig. 1.9.1) (Gao *et al.*, 2014).

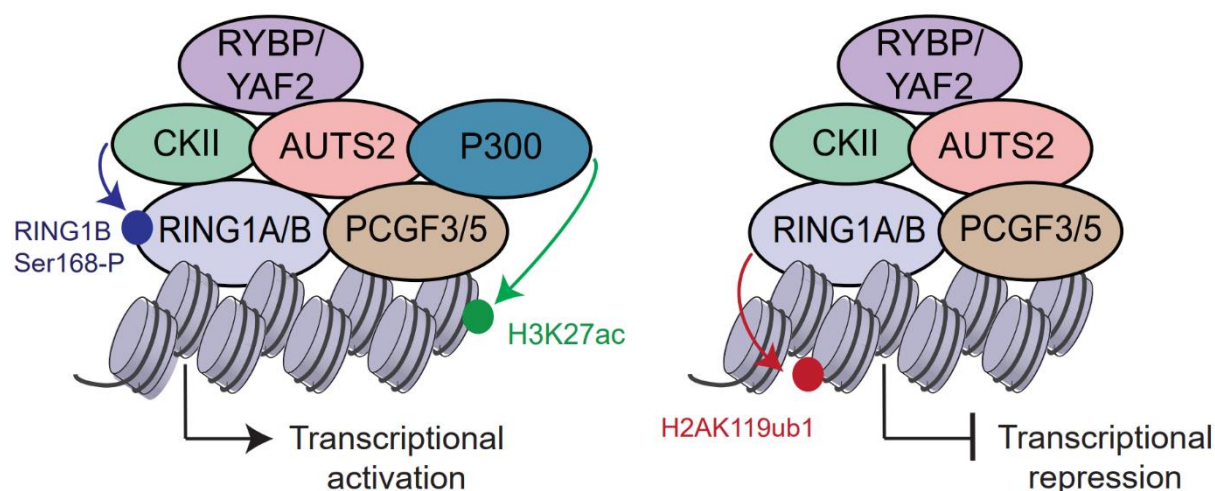


Figure 1.9.1 – Proposed modes of transcriptional regulation by AUTS2-PRC1.3/5 complexes

Left – AUTS2-PRC1.3/5 component Casein Kinase II becomes activated and phosphorylates fellow complex member RING1B on serine 41 and 168, inactivating the E3 ubiquitin ligase enzymatic activity of the complex. AUTS2 also recruits the histone acetyltransferase P300, which deposits the transcriptionally active H3K27ac epigenetic modification (Gao *et al.*, 2014). Right – PRC1.3/5 complexes catalyse the transcriptionally repressive H2AK119ub1 modification.

Activation of neuronal differentiation genes by the PRC1.3/5 complexes is important for central nervous system development (Gao *et al.*, 2014) and requires the central scaffold protein DCAF7/WDR68 (Hauri *et al.*, 2016; Wang *et al.*, 2018). Reinforcing this finding, *Pcgf5* deletion in mESCs resulted in a failure to activate transcription of neuronal differentiation genes, which could interestingly be partially rescued by chemical inhibition of TGF- β signalling (Yao *et al.*, 2018). PRC1.3/5 complexes have also been reported to activate transcription of mesoderm lineage marker genes during mouse embryoid body differentiation, acting alongside P300 and the pluripotency factor TEX10 (Zhao *et al.*, 2017). PRC1.3 has also been implicated in transient Vascular Endothelial Growth Factor (VEGF)-dependent transcriptional activation of angiogenesis transcription factors in mouse post-natal endothelial cells (Kanki *et al.*, 2022).

Other variant PRC1 complexes besides PRC1.3/5 have also been reported to activate transcription in some contexts. For example, the PRC1.1 complex is important for mesoderm and ectoderm differentiation gene activation in mESC embryoid body differentiation (Yan *et al.*, 2017), RING1B within variant PRC1 complexes activates transcription of the key spermatogenesis regulator *Sall4* in a manner essential for fertility (Maezawa *et al.*, 2017) and variant PRC1 complexes activate transcription of epidermal regulatory genes in mice to

maintain the integrity of the mouse skin (Cohen *et al.*, 2019). This body of evidence opines that PRC1 is not always repressive in its function (Wang *and* Yi, 2019).

PRC1.3-dependent, context-specific functions are not currently well understood and the majority of our understanding of the function and regulation of gene expression by PRC1 complexes has been derived from the mouse system. The identification of PRC1.3 as a specific functional regulator of human PTNR is the first time that a specific function for this complex has been identified in human pluripotency and reprogramming. This finding offers a strong opportunity to develop our mechanistic understanding of the key events that underpin the reprogramming pluripotency cell state transition.

1.10 – Thesis aims

Due to the profound epigenetic differences between primed and naïve PSCs, the process of primed-to-naïve human pluripotent stem cell reprogramming (PTNR) must involve a large amount of epigenetic remodelling and erasure within cells. This includes reactivation of an inactive X chromosome and global DNA hypomethylation in response to different cocktails of intracellular signalling inhibitors. As such, epigenetic complexes would be expected to be central to the regulation of the reprogramming process.

By analysing the molecular events underpinning the poorly understood human PTNR process, opportunities may arise for optimisation of the reprogramming process. This could allow many classic primed PSC lines to be readily reprogrammed to generate naïve PSCs, unlocking the full potential of naïve PSCs. The reduced epigenetic restriction and increased differentiation potential of naïve PSCs could offer exciting opportunities in stem cell therapeutics with the production of desirable cell types or high quality or efficiency stable tissue production. The *in vitro* system of naïve PSCs could also be a powerful system to study key human pre-implantation developmental and epigenetic processes such as lineage specification, implantation and human-specific X chromosome inactivation.

Genome-wide CRISPR-Cas9 screening allowed an unbiased exploration of the key essential regulators that govern the entry of human primed PSCs into the naïve state. Amongst the key regulators identified included epigenetic complexes, and surprisingly included almost all of the components of the Polycomb Repressive Complex PRC1.3, which has never been examined previously in the context of cell reprogramming or human pluripotency. This finding led me to pose the central question: how does the PRC1.3 complex modulate the reprogramming of primed PSCs to the naïve state? Addressing this gap in our knowledge should prove informative about the molecular events underpinning the reprogramming process that control cell identity and how they are governed.

Aims:

1. What is the phenotype of PRC1.3-deficient human primed PSCs?
2. Is the PRC1.3 complex indeed essential for human PTNR?
3. Where does the PRC1.3 complex bind on chromatin in primed and naïve PSCs?
4. How does the PRC1.3 complex influence gene transcription during reprogramming?
5. How does the PRC1.3 complex influence the epigenetic landscape of human PSCs?
6. How is the PRC1.3 complex recruited to chromatin?
7. Does the PRC1.3 complex undergo pluripotency-state specific regulation?

2 – Materials and Methods

2.1 – Cell Lines

All work has been conducted with appropriate approval from the UK Stem Cell Bank Steering Committee.

- H9 (WA09) primed PSCs (Thomson, 1998) were obtained with permission from WiCell and H9 naïve PSCs were obtained by chemically resetting (Guo *et al.*, 2017).
- H9 NK2 naïve PSCs (Takashima *et al.*, 2014) were obtained from Austin Smith (Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute) with permission from WiCell. H9 NK2 primed cells were derived from H9 primed PSCs by transfection of piggyBac vectors carrying a doxycycline-inducible *NANOG* and *KLF2* overexpression construct, coupled to a *VENUS* reporter. H9 NK2 naïve PSCs were produced by doxycycline induced transgene-driven resetting of H9 NK2 primed PSCs.
- H9 *PRDM14-AID-VENUS* primed PSCs (Sybirna *et al.*, 2020), produced by CRISPR-Cas9 mediated knock-in of an auxin inducible degron (AID) tag and the *VENUS* reporter to the *PRDM14* locus, were obtained from Azim Surani (Wellcome Trust-Cancer Research UK Gurdon Institute).
- *WTB CRISPRi Gen1B* primed iPSCs (Mandegar *et al.*, 2016), produced by a TALENS dCas9-KRAB donor knock-in at the *AAVS1* locus, were obtained from Bruce R. Conklin.

2.1.1 – Cell line characterisation

WT cells described in this study correspond to untargeted parental WA09/H9 cells. All unmodified and CRISPR-modified cell lines used in this study were confirmed to be mycoplasma-negative using the LookOut® Mycoplasma PCR Detection Kit (Sigma-Aldrich, MP0035) (Dobrovolsky and Bess, 2011). Karyotype analysis of G-banded chromosomes (carried out by Cell Guidance Systems) confirmed that all cell lines have the normal complement of chromosomes, and twenty cells were analysed per cell line.

2.2 – Cell Culture Techniques

All cells were cultured at 37 °C in humidified and hypoxic incubators at 5 % CO₂ and 5 % O₂.

2.2.1 – Generation of Mouse Embryonic Fibroblasts (MEF)

Active MF1 or DR4 mouse embryonic fibroblasts (MEFs) (Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute) were plated in Nunc™ EasYFlask™ 175 cm² flasks (Thermo Scientific™, 159910) and grown in MEF media.

Table 2.2.1 – Composition of MEF media

MEF media		
Component	Concentration	Company and Catalogue
DMEM	90 % v/v	Gibco™, 41965039
Foetal Bovine Serum (FBS)	10 % v/v	Sigma-Aldrich, F7524
L-Glutamine	2 mM	Gibco™, 25030024
Penicillin-Streptomycin	50 U/ml	Gibco™, 15140122
Sodium Pyruvate	1 mM	Gibco™, 11360070
MEM Non-Essential Amino Acids	1x	Gibco™, 11140035
β mercaptoethanol	0.1 mM	Gibco™, 31350010

Active MEFs were passaged every 2 to 3 days. Cell culture media was removed and flasks were gently washed with sterile Dulbecco's Phosphate Buffered Saline (D-PBS) without calcium or magnesium (Sigma-Aldrich, D8537). MEFs were incubated for 5 min with TrypLE™ Express (Gibco™, 12604021) at 37 °C. TrypLE™ Express was quenched at least 1:1 with MEF media and cells were collected by centrifugation at 300 x g for 3 min. The supernatant was removed and cell pellets were re-suspended in fresh MEF media and passaged at a 1:3 ratio into fresh flasks. Once expanded in number, MEFs were harvested as described above,

and irradiated with 30 Grays of gamma irradiation to mitotically inactivate the cells. The cells were cryopreserved and stored at -80 °C.

2.2.2 – Primed human PSC culture

Primed human PSCs were maintained either on mitotically inactivated MEFs (feeders) or in feeder-independent conditions.

Feeder-cultured conditions: Sterile tissue culture plates were coated for 1 h at 37 °C with 0.1 % gelatin (Sigma-Aldrich, G1890) dissolved in sterile water. Excess gelatin was aspirated from plates and MEFs were plated in MEF media, at a density of 2×10^6 cells per 6 well plate or 100 mm dish. The next day, MEF media was removed and primed PSCs were seeded onto irradiated MEFs in Knockout Serum Replacement and Fibroblast Growth Factor 2 (KSR+FGF2) media, which was changed daily.

Table 2.2.2 – Composition of KSR+FGF2 media

KSR+FGF2 media		
Component	Concentration	Company and Catalogue
Advanced DMEM	80 % v/v	Gibco™, 12491015
Knockout Serum Replacement (KSR)	20 % v/v	Gibco™, 10828028
L-Glutamine	2 mM	Gibco™, 25030024
Penicillin-Streptomycin	50 U/ml	Gibco™, 15140122
β mercaptoethanol	0.1 mM	Gibco™, 31350010
Fibroblast Growth Factor 2 (FGF2)	4 ng/ml	Cambridge Stem Cell Institute

Feeder-dependent primed PSCs were passaged at a 1:3 – 1:6 ratio every 4-5 days. During passaging, cell culture media was removed and plates were washed gently with D-PBS. Cells were incubated for 5 min with 200 U/ml Collagenase type IV (Gibco™, 17104019) at 37 °C. Collagenase was aspirated and KSR+FGF2 media was added. Cells were manually dissociated from tissue culture plates, gently triturated using a 5 ml serological pipette (Sarstedt, 86.1253.001) and passaged onto MEF-coated plates in KSR+FGF2 media.

Feeder-free culture conditions: Sterile tissue culture plates were coated with 5 µg/ml Vitronectin (Gibco™, A14700) for 1 h at room temperature (RT). Excess vitronectin was aspirated and primed PSCs were seeded in complete TeSR-E8™ (Stem Cell Technologies, 05991/05990) or complete mTeSR™-Plus media (Stem Cell Technologies, 05826/05827). Cells were passaged at a 1:5 – 1:20 ratio every 3-5 days.

To passage feeder-independent primed PSCs, cell culture media was removed and the cells were washed gently with D-PBS. Cells were incubated for 5 min with Gentle Cell Dissociation Reagent (GCDR), consisting of 0.5 mM EDTA pH 8.0 (Invitrogen™, AM9260G) at RT. GCDR was aspirated and TeSR E8™ or mTeSR™-Plus media was added. Cells were manually dissociated from tissue culture plates and passaged to freshly prepared Vitronectin plates. Media was changed daily. For PRC2 inhibition in primed PSCs, UNC1999 (Sigma-Aldrich, SML0778 or Abcam, ab146152) was applied at 2.5 µM for 4 d. For CRISPRi knockdown in primed PSCs, Doxycycline hyclate (Tocris, 4090) was applied at 1 µM for 6 d, unless specified otherwise. For PRDM14-AID-VENUS depletion in primed PSCs, Indole-3-acetic acid sodium salt (IAA; Cayman Chemical, CAY16954) was applied at 100 µM for 24 hours.

2.2.3 – Naïve human PSC culture

Naïve PSCs were either maintained in feeder-dependent or feeder-independent conditions, and cultured in either t2iL + PKCi media (Takashima *et al.*, 2014) or PXGL media (Bredenkamp *et al.*, 2019b; Rostovskaya *et al.*, 2019), with media changed daily.

Table 2.2.3 – Composition of t2iL+PKCi media

t2iL+PKCi media		
Component	Concentration	Company and Catalogue
DMEM/F12	50 % v/v	Gibco™, 21331020
Neurobasal	50 % v/v	Gibco™, 21103049
N-2 supplement	0.5x	Gibco™, 17502048
B-27 supplement	0.5x	Gibco™, 17504044
L-Glutamine	2 mM	Gibco™, 25030024

Penicillin-Streptomycin	50 U/ml	Gibco™, 15140122
β mercaptoethanol	0.1 mM	Gibco™, 31350010
PD0325901	1 μM	Cambridge Stem Cell Institute
CHIR99021	1 μM	
Recombinant human LIF	20 ng/ml	
Gö6983	2 μM	Tocris, 2285

Table 2.2.4 – Composition of PXGL media

PXGL media		
Component	Concentration	Company and Catalogue
DMEM/F12	50 % v/v	Gibco™, 21331020
Neurobasal	50 % v/v	Gibco™, 21103049
N-2 supplement	0.5x	Gibco™, 17502048
B-27 supplement	0.5x	Gibco™, 17504044
L-Glutamine	2 mM	Gibco™, 25030024
Penicillin-Streptomycin	50 U/ml	Gibco™, 15140122
β mercaptoethanol	0.1 mM	Gibco™, 31350010
PD0325901	1 μM	Cambridge Stem Cell Institute
XAV939	2 μM	Sigma-Aldrich, X3004
Recombinant human LIF	20 ng/ml	Cambridge Stem Cell Institute
Gö6983	2 μM	Tocris, 2285

Naïve PSCs were passaged at a 1:3 ratio every 3 days. For passaging, cell culture media was removed and the cells were washed gently with D-PBS, followed by a 5 min incubation with Accutase (BioLegend®, 423201) at 37 °C. Accutase was quenched at least 1:1 with N2B27 basal media containing 0.1 % Bovine Serum Albumin (BSA) (Gibco™, 15260037) and cells

were collected and centrifuged at 300 x g for 3 min. Media was aspirated and cell pellets were resuspended in complete t2iL + PKCi or PXGL media and passaged onto fresh MEF-coated plates or feeder-independent plates.

Feeder-independent plates were either pre-coated with Corning® Matrigel® Growth Factor Reduced Basement Membrane Matrix LDEV-Free (Corning Inc., 354230), diluted 1:100 in cold DMEM/F12 for 1 h at RT, or coated immediately after cell plating with Geltrex™ LDEV-Free Reduced Growth Factor Basement Membrane Matrix (Gibco™, A1413202), diluted 1:300 in complete media. Feeder-independent naïve PSCs were passaged from feeder-cultured conditions and harvested for experimental work only after two passages in feeder-free conditions. For PRDM14-AID-VENUS depletion in naïve PSCs, IAA was applied at 100 µM for 24 hours.

2.2.4 – Cryopreservation of cells

Primed PSCs cultured on MEFs were harvested with Collagenase type IV and cryopreserved in 45 % KSR+FGF2 media, 45 % KSR and 10 % DMSO (Sigma-Aldrich, D2650), containing 10 µM Y-27632 (Cell Guidance Systems, SM02). Primed PSCs cultured in feeder-free conditions were harvested using GCDR and cryopreserved in 45 % TeSR-E8™/mTeSR™-Plus, 45 % KSR and 10 % DMSO, containing 10 µM Y-27632. Naïve PSCs were harvested with Accutase and cryopreserved in 45 % complete t2iL + PKCi or PXGL media, 45 % FBS and 10 % DMSO, containing 10 µM Y-27632. For all lines, cells were cryopreserved at –80 °C overnight in a Nalgene® Mr. Frosty freezing container (Sigma-Aldrich, C1562) and stored long term in liquid nitrogen at below –80 °C.

2.2.5 – Primed to naïve reprogramming (PTNR)

Primed PSCs were reprogrammed to a naïve state using either 5i/L/A (Theunissen *et al.*, 2014) or chemical resetting conditions (Guo *et al.*, 2017).

2.2.5.1 – 5i/L/A reprogramming

On Day 0 of 5i/L/A reprogramming, feeder-dependent primed PSCs were dissociated into single cells with Accutase and counted. Cells were seeded in KSR+FGF2 media containing 10 μ M Y-27632 at 1.2×10^6 cells per 100 mm plate or 6 well plate that were pre-coated with irradiated MEF at twice the normal density (4×10^6 MEFs per dish). The following day (day 1), media was changed to complete 5i/L/A media.

Table 2.2.5 – Composition of 5i/L/A media

5i/L/A media		
Component	Concentration	Company and Catalogue
DMEM/F12	50 % v/v	Gibco™, 21331020
Neurobasal	50 % v/v	Gibco™, 21103049
N-2 supplement	0.5x	Gibco™, 17502048
B-27 supplement	0.5x	Gibco™, 17504044
L-Glutamine	2 mM	Gibco™, 25030024
Penicillin-Streptomycin	50 U/ml	Gibco™, 15140122
β mercaptoethanol	0.1 mM	Gibco™, 31350010
Bovine Serum Albumin (BSA)	50 μ g/ml	Gibco™, 15260037
KSR	0.5 % v/v	Gibco™, 10828028
PD0325901	PD0325901	Cambridge Stem Cell Institute
IM-12	1 μ M	Cell Guidance Systems, SM04
SB590885	0.5 μ M	Cell Guidance Systems, SM48
WH-4-023	1 μ M	Cell Guidance Systems, SM16
Y-27632	10 μ M	Cell Guidance Systems, SM02
Recombinant human LIF	20 ng/ml	Cambridge Stem Cell Institute
Activin A	20 ng/ml	Cambridge Stem Cell Institute

Media was changed daily and cells were passaged on day 5 and day 10 with Accutase. Cells (1.2×10^6 cells per 10 cm dish) were seeded at each passage onto fresh twice-density MEF plates. In wild-type (WT) PSCs, dome-shaped naïve colonies appear in the first 10 days of 5i/L/A reprogramming. After 10 days, cells were passaged with Accutase every 4 days at a 1:1 – 1:3 ratio and on Day 18, cells were transferred into PXGL media to promote stabilisation of naïve colonies.

2.2.5.2 – Chemical resetting

On Day 0 of chemical resetting (CR), feeder-dependent primed PSCs were dissociated into single cells with Accutase and 5.7×10^5 primed PSCs were seeded per 100 mm dish in KSR+FGF2 media containing 10 μ M Y-27632, onto a freshly prepared irradiated MEF layer. The following day (day 1), media was changed to KSR+FGF2 media without Y-27632. On day 2, media was changed to Chemical Reset Media 1 (CRM1).

Table 2.2.6 – Composition of Chemical Reset Media 1

Chemical Reset media 1 (CRM1)		
Component	Concentration	Company and Catalogue
DMEM/F12	50 % v/v	Gibco™, 21331020
Neurobasal	50 % v/v	Gibco™, 21103049
N-2 supplement	0.5x	Gibco™, 17502048
B-27 supplement	0.5x	Gibco™, 17504044
L-Glutamine	2 mM	Gibco™, 25030024
Penicillin-Streptomycin	50 U/ml	Gibco™, 15140122
β mercaptoethanol	0.1 mM	Gibco™, 31350010
PD0325901	1 μ M	Cambridge Stem Cell Institute
Valproic acid sodium salt (VPA)	1 mM	Sigma-Aldrich, P4543
Recombinant human LIF	10 ng/ml	Cambridge Stem Cell Institute

Media was not changed on day 3. On day 4, half of the media was changed for fresh CRM1. On Day 5, media was changed to PXGL media and refreshed daily thereafter. On day 9-10, cells were passaged with Accutase at a 1:1 – 1:3 ratio and cells were subsequently passaged with Accutase every 4 days at a 1:1 – 1:3 ratio, for several passages until stable naïve colonies form. For PRDM14-AID-VENUS depletion chemical resetting, IAA was applied at 100 μ M from 48 h of reprogramming onwards.

2.2.6 – *Drosophila* Schneider 2 (S2) culture

Drosophila melanogaster S2 cells, derived from late-stage *Drosophila melanogaster* embryos (Schneider, 1972), were grown in a non-humidified incubator at 28 °C without additional CO₂ and at 20 % O₂. *Drosophila* S2 cells were maintained in Nunc™ EasYFlask™ 75 cm² flasks (Thermo Scientific™, 156499) and grown in Schneider's *Drosophila* Media (Gibco™, 21720001), supplemented with 10 % v/v heat inactivated FBS. *Drosophila* S2 cells were passaged every 3 days by gently tapping and washing flasks with growth media, as well as pipetting up and down to break up clumps, and transferring cells at a 1:3 ratio into fresh flasks.

2.3 – Flow Cytometry

2.3.1 – Flow Cytometry Analysis

Cells for flow cytometry were dissociated into single cells using Accutase for 5 minutes at 37 °C and quenched at least 1:1 with DMEM/F12 containing 0.1 % BSA. Cells were collected by centrifugation at 300 x g for 3 minutes. Media was aspirated and cell pellets were washed with PBS containing 2 % FBS (flow buffer) and passed through 50 µm cell strainers (Sysmex, 04-0042-2317).

Conjugated antibodies and Fixable Viability Dye-eF780 (eBioscience, 65-0865-18) were mixed in 50 µl Brilliant Stain Buffer (BD Biosciences, 563794) and added to 50 µl of cell suspension (500,000 cells per reaction). Cells were incubated for 30 min at 4 °C in the dark and washed once with flow buffer and centrifuged at 300 x g for 3 minutes. Cell pellets were resuspended in 300 µl of flow buffer for analysis. Controls for flow cytometry include single antibody-stained cells were used for compensation controls, using the cell types with the strongest signal for each antigen. Unstained and fluorescence minus one (FMO) controls were used in cytometer gating and setup. Flow cytometry for discrimination of primed versus naïve PSCs was performed as previously described (Rugg-Gunn, 2022a).

Table 2.3.1 - Antibodies used for flow cytometry and Fluorescence Activated Cell Sorting (FACS)

Antibody target	Conjugation	Source	Optimised dilution
CD24	BUV395	BD Biosciences, 563818	1:40
CD75	eF660	eBioscience, 50-0759-42	1:20
Cd90.2	APC-Cy7	BioLegend®, 105328	1:20
CD147	BV421	BD Biosciences, 562583	1:40
SSEA4	BV605	BD Biosciences, 563119	1:40
SUSD2	PE	Miltenyi Biotec, 130-111-641	1:100
Fixable Viability Dye	eF780	eBioscience, 65-0865-18	1:166

Flow cytometry analysis was performed using a BD LSR Fortessa™ cell analyser (BD Biosciences) at the Babraham Institute Flow Cytometry Facility.

Table 2.3.2 - Flow cytometry machine parameters used for analyser acquired data

Laser	Filter(s)	Antibody and fluorophore
335 nm	379/28	CD24-BUV395
405 nm	450/50	CD147-BV421
	610/20	SSEA4-BV605
561 nm	585/15	SUSD2-PE
640 nm	670/14	CD75-eF660
	780/60	Cd90.2-APC-Cy7
		Fixable Viability Dye eF780

2.3.2 – Fluorescence Activated Cell Sorting (FACS)

Primed or reprogramming cells were sorted on a FACS Aria™ Fusion cell sorter (BD Biosciences). Primed cells were sorted for CAS9-GFP positivity 48 h after transfection and collected into 15 ml falcon tubes containing KSR+FGF2 medium with 10 μ M Y-27632. Cells were washed and transferred onto irradiated MEF-coated plates in KSR+FGF2 containing 10 μ M Y-27632. After 24 hours, medium was replaced with KSR+FGF2 medium without Y-27632.

Reprogramming cells were sorted to enrich and stabilise naïve PSCs in a mixed population. Naïve PSCs were sorted for CD75 and SUSD2 positive cells and CD24 negative cells (Brendenkamp *et al.*, 2019a; Collier *et al.*, 2017; Shakiba *et al.*, 2015; Wojdyla *et al.*, 2020). A viability dye and the MEF marker Cd90.2 (Ishihara *et al.*, 1987), were also included to specifically identify the live, human population. Sorted cells were collected into 15 ml falcon tubes containing PXGL medium with 10 μ M Y-27632. Cells were washed and transferred onto

irradiated MEF-coated plates in PXGL containing 10 μ M Y-27632. After 24 hours, medium was replaced with PXGL media without Y-27632.

Flow cytometry gating and data analysis was performed in FlowJo V10.1 (FlowJo, LLC).

2.4 – Transfections

Nucleofection: Cells were pre-incubated with 10 μ M Y-27632 for 2 h before transfection to improve survival. Cells were dissociated into single cells by incubation with Accutase for 5 min at 37 °C and quenched at least 1:1 with DMEM/F12 containing 0.1 % BSA. Cells were collected and centrifuged at 300 x g for 3 min and media was aspirated. Cell pellets were washed with D-PBS and counted. Two million cells were aliquoted per nucleofection and centrifuged at 300 x g for 3 min and media was aspirated. Cell pellets were resuspended in Nucleofection solution, composed of 82 μ l Nucleofection solution 1 and 18 μ l Supplement 1, from Human Stem Cell Nucleofector™ Kit 2 (Lonza, VPH-5022).

For CRISPR-Cas9 gene deletions: WT H9 primed PSCs were nucleofected with 4 μ g pCas9_GFP (Addgene #44719) and 3 μ g of each of the pCR4-TOPO guide RNA (gRNA) expression vectors, using setting B-016 on the Amaxa 2B Nucleofector™. Cells were quickly transferred to freshly prepared, irradiated MEF-coated plates, containing KSR+FGF2 media supplemented with 10 μ M Y-27632. After 24 h, media was changed to KSR+FGF2 media without Y-27632 and 24 h later cells were flow sorted for the Cas9-GFP positive population.

For CRISPR interference (CRISPRi) gene knockdowns: WTB CRISPRi primed iPSCs were nucleofected with 5 μ g pgRNA-CKB (Addgene #73501), containing the *PCGF3* knockdown gRNA, using setting B-016 on the Amaxa 2B Nucleofector™. Cells were quickly transferred to freshly prepared, vitronectin-coated plates, containing mTeSR™-Plus media supplemented with 10 μ M Y-27632. After 24 h, media was changed to mTeSR™-Plus containing 10 μ M Y-27632 and 10 μ g/ml Blasticidin S hydrochloride (Gibco™, R21001) to select for resistant cells.

Lipofection: *PCGF3*-deficient primed PSCs were pre-incubated with 10 μ M Y-27632 for 2 h before transfection to improve survival. 6 μ l GeneJuice® (Sigma Aldrich, 70967) was added dropwise into 100 μ l Advanced DMEM, mixed by vortexing and incubated at RT for 5 min. 2 μ g of pCAG-PCGF3 plasmid was added to the GeneJuice® mixture, mixed by gentle pipetting and incubated at RT for 15 min. Adherent cells were washed once with D-PBS and KSR+FGF2 media with 10 μ M Y-27632 was added to cells, with KSR was replaced by FBS to improve transfection efficiency. The GeneJuice®-DNA complex mixture was added dropwise to the cells. After 24 h, media was changed to KSR+FGF2 media without 10 μ M Y-27632. After passaging cells onto DR4 (Puromycin-resistant) irradiated MEFs (Tucker *et al.*, 1997), 1 μ g/ml Puromycin (Sigma-Aldrich, P8833) was applied for 48 h to select for resistant cells. Resistant cells were expanded and individual clones were picked and selected based on having the closest to WT expression levels.

2.5 – Molecular Biology

2.5.1 – Plasmid Cloning

Table 2.5.1 - Plasmids used in this study

Plasmid	Reference or Source	Function
pCas9-GFP	Addgene #44719 http://www.addgene.org/44719/	Cas9 transfection for gene knockout (KO)
pCR4-TOPO	Invitrogen™, K457502	Guide RNA cloning for gene KO
pCR2.1-TOPO	Invitrogen™, K450002	PCR product cloning for Sanger sequencing
pCAG-CreGR-ires-puro	(Hitoshi <i>et al.</i> , 1991)	Constitutive <i>PCGF3</i> expression cloning
pgRNA-CKB	Addgene #73501 http://www.addgene.org/73501/	Guide RNA cloning for <i>PCGF3</i> knockdown

2.5.1.1 – Individual gene knockout targeting

To generate individual gene knockouts (KO), pairs of gRNAs were designed using Wellcome Sanger Institute Genome Editing (<https://www.sanger.ac.uk/science/tools/wge>). The plasmids were designed such that an early, critical exon is excised, producing a frameshift that results in a premature stop codon, resulting in nonsense mediated decay of the truncated mRNA.

Table 2.5.2 – gRNA sequences used for individual gene knockouts, with 3' protospacer adjacent motif emboldened

Gene target	Left guide RNA (5'-3')	Right guide RNA (5'-3')
PCGF3	TTAGGAGAGCGTCTAGAGCC AGG	GGCACTCACCCACGTACTGT GG

PCGF5	GTTCTTCTTCAAAAAGTGT AGG	GACAATCCTATGCTTAGAAAT TGG
FBRS	ATAGGCATCCAGGCCCATCT TGG	GTCACCTAAGCAAGTGAACCA AGG

Each gRNA sequence was incorporated into a U6 target gRNA expression vector and synthesised as a g-Block (Integrated DNA Technologies) (Mali *et al.*, 2013). G-Blocks, containing specific gRNAs, were dissolved in 1x TE solution (10 mM Tris, pH 8.0, 1 mM EDTA, pH 8.0; Sigma-Aldrich, 574793) to a final concentration of 10 ng/μl. An adenine overhang is added to the G-block using Qiagen HotStar® Taq Polymerase Kit (Qiagen, 203203) in the following reaction: 5 μl (50 ng) G-block, 1 μl ATP (10 mM, NEB, P0756), 1 μl 10x Buffer and 0.5 μl HotStar® Taq Polymerase in a total reaction volume of 10 μl with nuclease free water (Promega, 7732-18-5). gRNA g-Blocks were sub-cloned into the pCR4-TOPO plasmid using the TOPO™ TA cloning™ kit (Invitrogen™, K457502) in the following reaction: 4 μl PCR product, 1 μl salt solution and 1 μl pCR4-TOPO vector and incubated for 30 min at RT.

pCR4-TOPO g-Block vectors were transformed into sub-cloning efficiency™ DH5α competent bacteria (Invitrogen™, 18265017) by mixing 4 μl of the plasmid mixture with 100 μl bacteria and incubating for 30 min at 4 °C. Bacteria were heat shocked for 30 s at 42 °C and immediately incubated at 4 °C with 200 μl SOC media (Super Optimal broth with Catabolite Repression, Invitrogen™, 15544034). The mixture was incubated at 37 °C with shaking for 1 h. Bacteria were plated onto 100 μg/ml Ampicillin (Thermo Scientific™, J60977-06)-LB (Lysogeny Broth, provided by BI media kitchen) agar plates, with X-Gal (Melford, B71800-0.1) to blue-white screen for PCR product insertion. Bacteria were incubated overnight at 37 °C. The following day, white Ampicillin-resistant colonies were inoculated into 3 ml LB containing 100 μg/ml Ampicillin and grown overnight at 37 °C with shaking. Liquid cultures (1 ml) were subject to Qiagen® Plasmid Miniprep (Qiagen, 12125) as per the manufacturer's protocol, to isolate plasmid DNA for Sanger sequencing (GENEWIZ) of the gRNA sequence. Following gRNA sequencing, bacterial liquid culture was inoculated into 250 ml LB containing 100 μg/ml Ampicillin and grown overnight at 37 °C with shaking. Plasmid DNA was extracted using a Qiagen® Plasmid Maxiprep kit (Qiagen, 12163) as per the manufacturer's protocol.

2.5.1.2 – PCGF3 rescue cDNA cloning

RNA was extracted from WT primed PSCs using the RNeasy® Mini Kit (QIAGEN, 74106) and reverse transcribed to cDNA using SuperScript™ III Reverse Transcriptase (Invitrogen™, 11732020). The full-length *PCGF3* coding sequence was amplified by polymerase chain reaction, using the Qiagen HotStar Taq® Polymerase Kit (Qiagen, 203203) and specifically designed primers.

Table 2.5.3 – Primer sequences for *PCGF3* amplification and restriction site creation

Name	Primer sequence (5' to 3')
PCGF3 Rescue forward	ACGACGCGTGCCACCATGTTGACCAGGAAGATCAAGCTG
PCGF3 Rescue reverse	ATAAGAATGCGGCCGCTCACAGCAAGTCCATCTTGGGT

The pCAG-CreGR-ires-puro plasmid (Hitoshi *et al.*, 1991) and full-length *PCGF3* cDNA were digested using MluI and NotI enzymes in the following reaction: 1 µg pCAG-CreGR-ires-puro vector or 50 µl *PCGF3* cDNA, 5 µl 10x FastDigest buffer (Thermo Scientific™, B64), 1 µl FastDigest MluI (Thermo Scientific™, FD0564) and 1 µl FastDigest NotI (Thermo Scientific™, FD0595) in a 50 µl reaction volume with nuclease-free water. The digestion was incubated at 37 °C for 30 min, followed by heat inactivation at 80 °C for 5 min. Digested vector and cDNA were separately incubated with 6x DNA loading dye (Thermo Scientific™, R1151) and run on 1 % agarose electrophoresis gel containing SYBR™ Safe (Invitrogen, S33102). The linearised plasmid band of 6.4 kb and *PCGF3* cDNA band of 729 bp were excised, and separately purified using a Monarch® DNA Gel Extraction Kit (NEB, T1020).

Linearised plasmid and *PCGF3* cDNA were ligated overnight at 16 °C in the following reaction: 100 ng linearised vector, 300 ng *PCGF3* cDNA, 1 µl T4 DNA ligase (Invitrogen™, 15224090) and 4 µl 5x T4 ligation buffer (Invitrogen™, 15-224-041) in a 20 µl reaction volume with nuclease-free water. The ligated plasmid was transformed into subcloning efficiency™ DH5α competent bacteria as described above, with Ampicillin selection but without blue-white screening of colonies. The following day, Ampicillin resistant colonies were inoculated into liquid cultures as described previously and grown overnight at 37 °C with shaking. Liquid cultures (1 ml) were subject to Qiagen® Plasmid Miniprep as per the manufacturer's protocol, in order to isolate the pCAG-*PCGF3*-ires-puro plasmid for Sanger sequencing (GENEWIZ).

Table 2.5.4 – Primer sequence for Sanger sequencing of the pCAG-PCGF3 vector

Name	Primer sequence (5' to 3')
Rescue cloning sequencing primer	CAACGTGCTGGTTATTGTGC

Following sequencing to confirm the plasmid, bacterial liquid culture was inoculated into 250 ml LB containing 100 µg/ml Ampicillin and grown overnight at 37 °C with shaking. Plasmid DNA was extracted using a Qiagen® Plasmid Maxiprep kit as per the manufacturer's protocol.

2.5.1.3 – CRISPR interference (CRISPRi) cloning

gRNA sequences were selected from (Gilbert *et al.*, 2014) based on their proximity to the *PCGF3* transcription start site. The gRNA sequence and its complementary sequence, with four-nucleotide overhangs of 'TTGG' on the 5' of the gRNA sequence and 'AAAC' on the 5' of the complementary sequence, were synthesised by Sigma-Aldrich.

Table 2.5.5 – Guide RNA sequences with overhangs for *PCGF3* CRISPRi knockdown

gRNA name	Forward oligonucleotide (5'-3')	Reverse oligonucleotide (5'-3')
gRNA 16	TTGGGGGCTCCACCCCGCGGCAA	AAACCCCGAGGGTGGGGCGCCGTT
gRNA 17	TTGGGAGCATGCGCGGTGTGCGTC	AAACCTCGTACGCGCCACACGCAG

Forward and reverse oligonucleotides were phosphorylated and annealed to one another in the following reaction: 1 µl forward oligonucleotide (100 µM), 1 µl reverse oligonucleotide (100 µM), 2 µl 5x T4 ligation buffer (Invitrogen™, 15-224-041), 0.5 µl T4 Polynucleotide Kinase (NEB, M0201S) in a 10 µl reaction volume with nuclease-free water (Promega, 7732-18-5). Samples were incubated at 37 °C for 30 min, heated to 95 °C for 5 min and cooled from 95 °C

to 25 °C by reducing the temperature by 5 °C per min. The phosphorylated annealed guide RNAs were diluted 1:100 in nuclease-free water.

The pgRNA-CKB vector was digested and linearised to allow gRNA insertion in the following reaction: 1 µg pgRNA-CKB vector, 5 µl 10x NEBuffer 3.1 (NEB, B7203S), 1 µl BsmBI (NEB, R0580S) in a 50 µl reaction volume with nuclease-free water. Samples were heated to 55 °C to restriction digest for 15 min. 1 U FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific™, EF0654) was added directly to the digestion reaction and incubated at 37 °C for 10 min to allow dephosphorylation and the phosphatase was heat inactivated at 75 °C for 5 min. The linearised plasmid DNA was mixed with 6x DNA loading dye and run on a 1% agarose electrophoresis gel containing SYBR™ Safe. The 9.6 kb band was excised from the gel and the linearised plasmid DNA was purified using a Monarch® Gel Extraction Kit. Linearised plasmid and diluted phospho-annealed oligonucleotides were ligated overnight at RT in the following reaction: 100 ng linearised vector, 1 µl phospho-annealed oligonucleotides, 1 µl T4 DNA ligase (Invitrogen™, 15224090) and 4 µl 5x T4 ligation buffer in a 20 µl reaction volume with nuclease-free water.

The following day, ligation was stopped by the addition of 1 µl 0.5 M EDTA (Invitrogen™, AM9260G) and the ligated plasmid was transformed into subcloning efficiency™ DH5α competent bacteria as described before, using Ampicillin selection but without blue-white screening of colonies, Ampicillin-resistant colonies were inoculated into liquid cultures and subject to Qiagen® Plasmid Miniprep as per the manufacturer's protocol to isolate plasmid to confirm gRNA insertion by Sanger sequencing (GENEWIZ).

Table 2.5.6 – Primer sequence for Sanger sequencing of the pgRNA-CKB vector

Name	Primer sequence (5' to 3')
CRISPRi cloning sequencing primer	GAGATCCAGTTTGGTTAGTACCGGG

Following sequencing to confirm the plasmid, bacterial liquid culture was inoculated into 250 ml LB containing 100 µg/ml Ampicillin and grown overnight at 37 °C with shaking. Plasmid DNA was extracted using a Qiagen® Plasmid Maxiprep kit as per the manufacturer's protocol.

2.5.2 – Genotyping and Polymerase Chain Reaction

Following transfection and Cas9-GFP sorting, 10,000 cells were seeded onto a 100 mm dish, pre-coated with irradiated MEFs, in KSR+FGF2 media. Cells were expanded and individual colonies were picked under a microscope and transferred into individual wells of an irradiated MEF-coated 96-well plate containing KSR+FGF2 media with 10 μ M Y-27632. Cells were fed with KSR+FGF2 media without Y-27632 and, once confluent, cells were split across two 96-well plates, one to maintain clones in cultures and the other to screen for the desired genotype.

To genotype cells, media was aspirated and cells were washed twice with D-PBS, prior to incubation with 50 μ l Yolk Sac Lysis Buffer (50 mM Tris pH 8.8, 1 mM EDTA, 0.5 % Tween® 20 (Sigma-Aldrich, P1379) containing 2 μ g/ml Proteinase K (Sigma-Aldrich, 39450-01-6) for 5 min at RT. Cells were scraped from the 96-well plate and transferred to a new 96-well plate, where the cells were digested overnight at 65 °C followed by a 95 °C incubation for 15 min.

Extracted DNA was subject to Polymerase Chain Reaction (PCR) using specifically designed primers to confirm exon excision.

Table 2.5.7 – Primer sequences for CRISPR gene knockout genotyping

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
PCGF3	ACAAGGAGGACAGCAGTGAG	GCCCGGCCTACCTAAGTATT
PCGF5	GATGTCCCAATGAGCCTAGAA	GGCCCAAGACAGTTCAGTTG
FBRS	GAGTGGATGCTGTGGAGATG	CCAGGCTCAACTCTTCCTTG

PCR amplification was conducted using the Qiagen HotStar Taq® Polymerase Kit (Qiagen, 203203) in the following reaction: 50 ng extracted DNA, 2.5 μ l 10x Buffer, 0.5 μ l dNTPs (10 mM), 1.25 μ l forward primer (10 μ M), 1.25 μ l reverse primer (10 μ M), 0.2 μ l DMSO and 0.125 μ l HotStar Taq® Polymerase in a total reaction volume of 25 μ l with nuclease free water. PCRs were conducted in a thermocycler, with optimised elongation times and annealing temperatures and KO clones were tested alongside a WT control. PCR products were analysed by 1 % agarose gel electrophoresis, screening for reduced PCR product fragment size in either heterozygous or homozygous KO in modified clones.

PCR products were subsequently transformed into subcloning efficiency™ DH5α competent bacteria as described above, with Ampicillin selection and blue-white screening. The following day, Ampicillin resistant colonies were inoculated into 3 ml LB containing 100 µg/ml Ampicillin and grown overnight at 37 °C with shaking and subject to Qiagen® Plasmid Miniprep as per the manufacturer's protocol, in order to purify the plasmid for Sanger sequencing (GENEWIZ).

Table 2.5.8 – Primer sequences for Sanger sequencing of genotyping

Name	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
M13	GTAAAACGACGGCCAG	CAGGAAACAGCTATGAC

2.5.3 – RNA extraction and cDNA synthesis

RNA for RT-qPCR analysis was extracted using the RNeasy® Mini Kit (QIAGEN, 74106). RNA for sequencing was extracted by lysing cells in TRIzol™ Reagent (Thermo Scientific™, 15596018) for 5 min at RT and shearing genomic DNA by passing the sample twice through a 26-gauge needle. One volume of Chloroform (Thermo Scientific™, 404635000) was added per 5 volumes of TRIzol™ and after a brief incubation, cells were centrifuged at > 21,000 x g for 15 min at 4 °C. The aqueous phase was collected and an equal volume of ice-cold isopropanol was added, prior to overnight incubation at -20 °C. The suspension was again centrifuged at > 21,000 x g for 15 min at 4 °C and the pellet was washed with 75 % ethanol. RNA was again centrifuged at > 21,000 x g for 10 min at 4 °C and the pellet was air dried and resuspended in nuclease free water.

RNA (1 µg) was reverse transcribed using either the QuantiTect® Reverse Transcription Kit (QIAGEN, 205311) or the SuperScript™ III Reverse Transcriptase (Invitrogen™, 11732020).

2.5.4 – Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR)

RT-qPCR was performed using SYBR® Green JumpStart™ ReadyMix™ (Sigma-Aldrich, S4438) in the following reaction: 6 µl SYBR Green, 0.24 µl 10 mM forward primer, 0.24 µl 10 mM reverse primer, 250 ng cDNA, 0.52 µl water to a total volume of 12 µl. RT-qPCR was performed on a CFX96 RT-qPCR (Bio-Rad) machine with the following programme: 94 °C for

2 min, 94 °C for 15 s, 60 °C for 30 s, 72 °C for 30 s – followed by detection of double stranded DNA and repeating steps 2 to 4 for 39 more cycles. RT-qPCR was performed in technical triplicate for three independent biological replicates.

Table 2.5.9 – Primer sequences used for RT-qPCR experiments

Target	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>GAPDH</i>	CGCTGAGTACGTCGTGGAGT	GGGCAGAGATGATGACCCTTT
<i>HMBS</i>	AGGAGTTCAGTGCCATCATCCT	CACAGCATACATGCATTCTCA
<i>DPPA3</i>	AGACCAACAAACAAGGAGCCT	CCCATCCATTAGACACGCAGA
<i>KLF4</i>	GCTGCCGAGGACCTTCTG	GCGAACGTGGAGAAAGATGG
<i>TFCP2L1</i>	TTTGTGGGACCCTGCGAAG	TGCTTAAACGTGTCAATCTGGA
<i>PCGF3</i>	AGGATTGTGATCCACCAGAG	TTTCTCATTTCCGCTTCTTG

2.5.5 – Western Blotting

Whole-cell lysates: Cell pellets were extracted in RIPA (rapid immunoprecipitation assay) buffer, composed of: 150 mM NaCl (Sigma-Aldrich, S6546), 1 % Nonidet™ P-40 (NP40; Sigma-Aldrich, 21-3277), 0.5 % sodium deoxycholate (Sigma Aldrich, 30970), 0.1 % SDS (Sodium Dodecyl Sulphate) (Bio-Rad, 1610418) and 50 mM Tris pH 8.0 (Invitrogen™, 15568025), containing cOmplete™ EDTA-free protease inhibitors (Roche, 11873580001).

Histones: Cells were washed once with D-PBS and dissociated using Accutase. Cell pellets were incubated with PBS containing complete EDTA-free protease inhibitor cocktail for 10 min at 4 °C and centrifuged for 5 min at 300 x g at 4 °C. The cell pellet was resuspended in 1 ml 0.2 M sulphuric acid (Fluka™, 7664-93-9) and incubated for 30 min at 4 °C, before being centrifuged for 2 min at 12,000 x g at 4 °C. The supernatant was transferred to a fresh tube and 8 volumes of acetone (Fisher Scientific, 10244202) were added. Samples were incubated for 10 min at 4 °C and centrifuged at 12,000 x g for 10 min at 4°C. The acetone wash was repeated using 1 ml acetone and the supernatant was removed. Histones were resuspended in 100mM Tris-HCl (pH 8.0) supplemented with protease inhibitor and left to dissolve overnight

at 4 °C with rotation. The dissolved histones were centrifuged for 10 min at 12,000 x g at 4 °C and the supernatant was kept and transferred to a fresh tube.

For whole-cell lysates and histones, protein concentration was determined by Bradford assay, using Bradford reagent (Bio-Rad, 5000006) and a standard curve using known concentrations of BSA protein. Proteins were denatured by boiling at 95 °C for 5 min in water and 1x protein loading dye, composed of: 4 % SDS, 0.25 M Tris pH 6.8, 1 µM Bromophenol Blue (Sigma-Aldrich, B0126), 0.5 mM DTT and 30 % Glycerol (Sigma-Aldrich, G5516).

Protein was separated by electrophoresis on a 12 % (whole-cell lysates) or 15 % (histones) SDS-polyacrylamide gel, using a Mini-PROTEAN® Tetra Vertical Electrophoresis Cell (Bio-Rad, 1658004), at 180 V for 45 minutes, alongside a Precision Plus Protein Dual Colour Standards (Bio-Rad, 1610374) to assess protein molecular weight. Protein was transferred onto a methanol activated 0.45 µm Immobilon®-P PVDF membrane (Millipore®, IPVH00010) by a Trans-Blot® Semi-Dry Transfer Cell (Bio-Rad, 1703848) at 15 V for 1 h. Membranes were blocked for 3 h at RT in TBS-T (1X Tris Buffered Saline and 0.05 % Tween-20) with 5 % dried skimmed milk. Blots were hybridised overnight at 4 °C with primary antibody diluted in TBST-5 % milk. Membranes were washed for 10 min, three times with TBS-T before incubation with an appropriate HRP or fluorophore-conjugated secondary antibody diluted in TBS-T 5 % milk for 1 h at RT. Membranes were washed for 10 min, three times with TBS-T and 10 min once with 1x TBS before detection, which was performed either using Odyssey Imaging Systems (LI-COR Biosciences) or using Clarity Western ECL reagent (Bio-Rad, 1705060)

Table 2.5.10 – Primary antibodies used for Western blotting

Antigen	Host	Source	Concentration
PCGF3+5	Rabbit	Abcam, ab201510	1:500
RING1B	Mouse	(Atsuta <i>et al.</i> , 2001)	1:1,000
AUTS2	Rabbit	Abcam, ab96326	1:500
FBRS	Rabbit	Abcam, ab264404	1:500
RYBP	Rabbit	Abcam, ab185971	1:1,000
PRDM14	Rabbit	Abcam, ab187881	1:500
β-ACTIN	Mouse	Sigma Aldrich, A5441	1:1,000

TUBULIN	Mouse	Sigma Aldrich, T6199	1:1,000
H2AK119ub1	Rabbit	Cell Signalling Technologies, D27C4	1:1,000
Histone H3	Rabbit	Abcam, ab1791	1:1,000
Histone H2B	Mouse	Abcam, ab52484	1:1,000

Table 2.5.11 – Secondary antibodies used for Western Blotting

Antigen and Conjugate	Host	Source	Concentration
Anti-Mouse IgG (H+L)-DyLight-680	Donkey	Invitrogen, SA5-10170	1:10,000
Anti-Rabbit IgG (H+L)-DyLight-800	Donkey	Invitrogen, SA5-10044	1:10,000
Anti-Mouse IgG (H+L)-HRP	Goat	Bio-Rad, 1706516	1:10,000
Anti-Rabbit IgG (H+L)-HRP	Goat	Bio-Rad, 1706515	1:10,000

2.5.6 – Immunofluorescence

Cells were washed with D-PBS and fixed with 4 % methanol-free paraformaldehyde (Agar Scientific, AGR1026) in PBS for 15 min. Cells were permeabilised and blocked overnight in blocking buffer, composed of 0.1 % Triton X-100 (Sigma-Aldrich, T9284) and 5 % FBS in PBS, at 4 °C. Cells were subsequently hybridised overnight with primary antibodies diluted in blocking buffer at 4 °C. Cells were washed three times for 10 min with blocking buffer and secondary antibodies, diluted in blocking buffer, were applied and incubated for 1 h at RT. Cells were again washed three times with blocking buffer, with 0.2 µg/ml DAPI (Sigma-Aldrich, D9542) added in the second wash. Images were collected at the Babraham Institute Imaging Facility on a NIKON A1-R confocal microscope with a 20x oil objective and processed with ImageJ.

Table 2.5.12 – Primary antibodies used for immunofluorescence microscopy

Antigen	Host	Source	Concentration
POU5F1/OCT4	Mouse	Santa Cruz, sc5279	1:300
NANOG	Goat	R&D Systems, 672 AF1997	1:300
PCGF3+5	Rabbit	Abcam, 201510	1:50
KLF17	Rabbit	Atlas Antibodies, HPA024629	1:300
GFP	Goat	Abcam, ab6673	1:300
DAPI	-	Sigma-Aldrich, D9542	1:10,000

Table 2.5.13 – Secondary antibodies used for immunofluorescence microscopy

Antigen and Conjugate	Host	Source	Concentration
Anti-rabbit AF488	Donkey	Invitrogen, A32814	1:400
Anti-goat AF555	Donkey	Invitrogen, A32794	1:400
Anti-mouse AF647	Donkey	Invitrogen, A31571	1:400

2.5.7 – Co-immunoprecipitation

Cells were washed with cold D-PBS and harvested by the addition of 500 µl ice-cold Co-IP buffer 1 containing cOmplete™ EDTA-free protease inhibitor directly onto 15 cm tissue culture plates and manually scraping.

Table 2.5.14 – Composition of Co-IP Buffer 1

Co-IP Buffer 1		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	10 mM	Thermo Scientific™, AAJ60712AK
MgCl ₂	1.5 mM	Invitrogen™, AM9530G

KCl	10 mM	Invitrogen™, AM9640G
DTT	0.5 mM	Promega, P1171
NP-40	0.05 %	Sigma-Aldrich, 21-3277
Benzonase® Nuclease	250 U/μl	Sigma-Aldrich, E1014

The cell suspension was transferred into Falcon tubes and incubated on ice for 10 min. Cells were collected by centrifugation at 1,000 x g for 10 min at 4 °C and the nuclear pellet was resuspended in 376 μl ice-cold Co-IP buffer 2 containing cOmplete™ EDTA-free protease inhibitor, supplemented with 300 mM NaCl to lyse nuclear membranes.

Table 2.5.15 – Composition of Co-IP Buffer 2

Co-IP Buffer 2		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	5 mM	Thermo Scientific™, AAJ60712AK
MgCl ₂	1.5 mM	Invitrogen™, AM9530G
EDTA	0.2 mM	Invitrogen™, AM9260G
DTT	0.5 mM	Promega, P1171
Glycerol	26 %	Sigma-Aldrich, G5516
Benzonase® Nuclease	250 U/μl	Sigma-Aldrich, E1014

Nuclei were further lysed with twenty strokes with a dounce homogeniser and incubated on ice for 30 min. Lysates were centrifuged at 24,000 x g for 20 min at 4 °C and the supernatant was collected and the protein quantified by Bradford assay.

For immunoprecipitation, 50 μl of protein A or G Dynabeads (Invitrogen™, 10002D or 10003D) were washed three times with Co-IP wash buffer.

Table 2.5.16 – Composition of Co-IP Wash Buffer

Co-IP Wash Buffer		
Component	Concentration	Company and Catalogue
Tris-HCl pH 7.5	10 mM	Invitrogen™, 15567027
NaCl	150 mM	Sigma-Aldrich, S6546
EDTA	0.5 mM	Invitrogen™, AM9260G

10 µg RING1B (Atsuta *et al.*, 2001) or PCGF3 antibody was immobilised onto the beads for 3 h at 4 °C. Antibody immobilised beads were re-washed three times in Co-IP wash buffer to remove any unbound antibody. 500 µg nuclear protein was added to the antibody-immobilised beads and 50 µg (10 %) was set aside as an input. The immunoprecipitation was incubated overnight at 4 °C with rotation. The beads were washed three times in Co-IP wash buffer and the immunocomplexes were eluted in 1x Protein Loading dye by boiling at 75 °C for 10 min. The inputs were prepared in the same way and immunoprecipitation and input samples were analysed by Western blotting.

2.6 – Proteomics

2.6.1 – Quantitative Tandem Mass Tag-Pro 16plex Rapid Immunoprecipitation Mass Spectrometry of Endogenous Protein (qPLEX-RIME)

qPLEX-RIME was adapted from previously described protocols (Mohammed *et al.*, 2016; Papachristou *et al.*, 2018).

Harvesting, double-fixing and feeder depleting cells: All buffers were pre-chilled to 4 °C and contain cOmplete™ EDTA-free protease inhibitors. Cells were dissociated with Accutase and quenched at least 1:1 with DMEM/F12 containing 0.1 % BSA. Cells were collected in a 50 ml Falcon tube, centrifuged at 300 x g for 5 min at RT, cell pellets were washed with D-PBS and counted. Cells were cross-linked with 2 mM Di-succinimidyl-glutarate (DSG; Sigma-Aldrich, 80424) for 45 min at RT with shaking. Cells were centrifuged at 300 x g for 5 min at RT and the supernatant was removed. The cells were again cross-linked using 1 % methanol-free paraformaldehyde for 12.5 min at RT with shaking. Fixation was quenched with 0.125 M Glycine (Sigma Aldrich, G7126) and incubated for 5 min at RT with shaking. Cells were centrifuged at 300 x g for 5 min at RT and the supernatant was removed. Cells were washed with PBS containing protease inhibitor and cell pellets were flash frozen with liquid nitrogen and stored at –80 °C if required. The fixed cell suspension was subject to feeder depletion using MACS technology as described below.

Magnetically activated cell sorting (MACS) for feeder depletion: Cells grown on irradiated MEF-coated plates underwent selective MEF removal in isolate only human PSCs, exploiting the exclusive expression of the cell-surface marker Cd90.2 on MEFs (Ishihara *et al.*, 1987).

Cells were harvested with Accutase as described in section 2.3 and cells washed in PBS containing 2 % FBS (flow buffer) were passed through a 30 µm strainer (Sysmex, 04-0042-2316). Cells were centrifuged at 300 x g for 3 min and pellets were resuspended in 70 µl flow buffer and incubated with 30 µl Biotin anti-mouse Cd90.2 antibody (BioLegend®, 140313) for 30 min at 4 °C. Cells were subsequently washed with flow buffer, pellets were resuspended in 90 µl flow buffer and incubated with 10 µl Streptavidin microbeads (Miltenyi Biotec, 130-048-102) for 15 min at 4 °C. Cells were again washed with flow buffer and the cell suspension containing both unlabelled human cells of interest and labelled MEFs were passed through flow buffer activated MACS LS columns (Miltenyi Biotec, 130-042-401) on a QuadroMACS magnetic separator (Miltenyi Biotec, 130-091-051). Flow-through from the MACS LS columns containing unlabelled human cells was collected and the column was washed three times with

flow buffer to collect as many human cells remaining on the column. Column-bound cells were detached by removal of the column from the magnet and washing vigorously with flow buffer.

Lysis and Sonication: Twenty million human cells per RIME experiment were resuspended in 10 ml Wash Buffer 1 and incubated for 10 min at 4 °C with rotation.

Table 2.6.1 – Composition of Wash Buffer 1

Wash Buffer 1		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	10 mM	Thermo Scientific™, AAJ60712AK
EDTA pH 8.0	10 mM	Invitrogen™, AM9260G
EGTA	0.5 mM	Millipore, 324626
Triton X-100	0.75 %	Sigma-Aldrich, T9284

After centrifuging at 2,000 x g for 5 min at 4 °C, nuclei were resuspended in 10 ml Wash Buffer 2 and incubated for 10 min at 4 °C with rotation.

Table 2.6.2 – Composition of Wash Buffer 2

Wash Buffer 2		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	10 mM	Thermo Scientific™, AAJ60712AK
EDTA pH 8.0	10 mM	Invitrogen™, AM9260G
EGTA	0.5 mM	Millipore, 324626
NaCl	200 mM	Sigma-Aldrich, S6546

After another 2,000 x g centrifugation for 5 min at 4 °C was performed, the pellet was resuspended in 600 µl Lysis/Sonication Buffer and transferred to protein lobind tubes (Eppendorf, 0030108116).

Table 2.6.3 – Composition of Lysis and Sonication Buffer

Lysis and Sonication Buffer		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	25 mM	Thermo Scientific™, AAJ60712AK
EDTA pH 8.0	5 mM	Invitrogen™, AM9260G
Triton X-100	0.1 %	Sigma-Aldrich, T9284
NaCl	150 mM	Sigma-Aldrich, S6546
SDS	1 %	Bio-Rad, 1610418
Sodium Deoxycholate	0.5 %	Sigma Aldrich, 30970

Lysis was performed for 30 min at 4 °C and chromatin was fragmented by sonication (Microson ultrasonic cell disruptor XL Misonix, output setting 4, 10W) with cycles of 15 seconds on and 30 seconds off. 30 sonication cycles were used to sonicate primed chromatin, 32 cycles for 48 h 5i/L/A reprogramming chromatin and 34 cycles for naive chromatin to produce fragments with a size of around 400-500 bp. Sonication efficiency was tested by reverse cross-linking chromatin at 95 °C for 15 min, DNA purification using the Monarch® PCR & DNA purification kit (NEB, T1030S) and separation of fragments by electrophoresis on a 1 % agarose gel.

Immunoprecipitation: Magnetic Protein A Dynabeads® (Invitrogen™, 10002D), 100 µl per immunoprecipitation in protein lobind tubes, were washed three times with 1 ml cold PBS/0.5 % BSA. Beads were incubated in 250 µl cold PBS/0.5 % BSA containing 5 µg of anti-PCGF3+5 antibody (Abcam, ab201510) for at least 5 h at 4 °C with rotation. Beads were again washed three times with 1 ml cold PBS/0.5 % BSA to remove non-immobilised antibody. Triton-X100 (1 %) was added to the 600 µl fragmented chromatin and the chromatin was incubated with the 100 µl of antibody-immobilised Protein A Dynabeads® overnight at 4 °C with rotation.

Washes: The next day, magnetic beads bound to antibody-chromatin complexes were washed 10 times with RIME RIPA buffer and twice with 100 mM Ammonium Bicarbonate (Sigma-Aldrich, A6141-25G).

Table 2.6.4 – Composition of RIME RIPA Buffer

RIME RIPA Buffer		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	50 mM	Thermo Scientific™, AAJ60712AK
EDTA pH 8.0	1 mM	Invitrogen™, AM9260G
NP-40	1 %	Sigma-Aldrich, 21-3277
LiCl	500 mM	Sigma-Aldrich, L7026
Sodium Deoxycholate	0.7 %	Sigma Aldrich, 30970

All visible liquid was removed from the beads, which were frozen at -80 °C. Samples were prepared as described in (Papachristou *et al.*, 2018) by Valar Nila Roamio Franklin at the Cancer Research UK Cambridge Institute Proteomics facility, who performed on-bead trypsin digestion, sample processing, multiplexing using TMT-16plex reagents (Thermo Scientific™, A44520) and analysed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific™), coupled to a Dionex UltiMate™ 3000 RSLCnano system (Thermo Scientific™)

2.6.2 – qPLEX-RIME Analysis

qPLEX-RIME data was analysed by Kamal Kishore at the Cancer Research UK Cambridge Institute Bioinformatics facility using the Proteome Discoverer 2.1 (ThermoFisher Scientific) and searching spectra against the Uniprot Homo sapiens FASTA database (taxon ID 9606 - Version January 2020). A 1 % false discovery rate was applied and peptides uniquely matched to a protein were used in subsequent quantification and analysis. The qPLEXanalyser (Papachristou *et al.*, 2018) package in Bioconductor was used to perform median scale peptide intensity normalisation and statistical analysis on normalised peptides. Statistical

analysis of differentially-regulated proteins was carried out using the Limma method. Multiple testing correction of p-values was applied using the Benjamini-Hochberg method to control the false discovery rate.

2.6.3 – Histone proteomics sample preparation

Histone proteomics was performed by Sigrid Verhelst and Maarten Dhaenens. Histones were acid extracted from primed and naïve PSCs as described in section 2.5.5. A fraction of histone extract was analysed using one dimensional polyacrylamide gel electrophoresis on an 8–16 % Criterion™ TGX™ pre-cast gel (Bio-Rad, 5671101) to quantify and normalise histone abundances between samples. The remaining histones were propionylated, trypsin digested and re-propionylated to cap all peptide N-termini. Histones were then analysed on a TripleTOF 6600+ mass spectrometer (Sciex) coupled to a nanoLC 425 high-performance liquid chromatography system (Eksigent).

2.6.4 – Histone proteomics analysis

Mass spectrometric data analysis was performed as previously described in (Verhelst *et al.*, 2020) by Sigrid Verhelst and Maarten Dhaenens. Peptides were screened against the Uniprot Homo sapiens Swissprot database, which created a FASTA-database on which histone post-translational modifications were screened, including acetylation, butyrylation, crotonylation and methylation.

2.6.5 – *In vitro* histone H2A Lysine 119 ubiquitylation assay

Cells were harvested and lysed as described in section 2.5.7. Protein levels were quantified by Bradford assay. For immunoprecipitation, 50 µl of protein A or G Dynabeads were washed three times with Co-IP wash buffer. 5 µg RING1B (Atsuta *et al.*, 2001) or PCGF3+5 antibody was immobilised onto the beads for 3 h at 4 °C. Antibody-immobilised beads were washed three times in Co-IP wash buffer. 100 µg nuclear protein was added to the antibody-immobilised beads and the immunoprecipitation was incubated overnight at 4 °C with rotation.

The beads were washed three times in Co-IP wash buffer and ten times in RIME RIPA buffer (Table 2.6.4). After the final wash, the beads were resuspended in ubiquitylation buffer and boiled in 1x Protein Loading dye at 95 °C for 5 min to elute the bound protein.

Table 2.6.5 – Composition of ubiquitylation buffer

RIME RIPA Buffer		
Component	Concentration	Company and Catalogue
HEPES pH 7.5	50 mM	Thermo Scientific™, AAJ60712AK
NaCl	100 mM	Sigma-Aldrich, S6546
MgCl ₂	10 mM	Invitrogen™, AM9530G
Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) pH 7.0	1 mM	Supelco, 646547
ZnCl ₂	10 µM	Sigma-Aldrich, 39059

The protein supernatant was retrieved from the beads and separated on a 12 % SDS-polyacrylamide by electrophoresis using a Mini-PROTEAN® Tetra Vertical Electrophoresis Cell, at 180 V for 45 minutes. Proteins were separated alongside a Precision Plus Protein Dual Colour Standard. The quantity of protein on the gel was detected using the Pierce™ Silver Stain Kit, as per the manufacturer's guidelines, to ensure equal amounts of the PRC1.3 complex are added into the *in vitro* assays.

The *in vitro* ubiquitylation reaction was composed of: 250 ng recombinant polynucleosomes (Active Motif, 31466), 1.5 µM Ubiquitin-conjugated E2 (UBCH5C/UBE2D3) enzyme (R&D Systems, E2-802) and equal amounts of intracellular PRC1.3 E3 ligase from primed or naïve PSCs (or a negative control using *PCGF3* KO primed PSCs) as a titration range, except when any of these components are omitted in controls, and made up to 10 µl with ubiquitylation buffer. The *in vitro* ubiquitylation assay was incubated for 1 h at 37 °C and terminated by boiling in 1x Protein Loading dye at 95 °C for 5 min. The *in vitro* ubiquitylation assay was assessed by Western blotting, as described for histones in section 2.5.5.

2.7 – Next Generation Sequencing

2.7.1 – Cross-linked chromatin immunoprecipitation (ChIP)-sequencing

ChIP-sequencing was performed as described in (Bendall *and* Semprich, 2022). Cells were harvested and double-fixed (Tian *et al.*, 2012), lysed and sonicated (albeit without MACS feeder depletion) as described above for qPLEX-RIME in section 2.6.1.

Immunoprecipitation: Fragmented chromatin was centrifuged at $> 21,000 \times g$ for 15 min at 4 °C and the supernatant was retained and diluted 1:10 with ChIP dilution buffer.

Table 2.7.1 – Composition of ChIP Dilution Buffer

ChIP Dilution Buffer		
Component	Concentration	Company and Catalogue
NaCl	150 mM	Sigma-Aldrich, S6546
EDTA pH 8.0	5 mM	Invitrogen™, AM9260G
Tris-HCl pH 7.5	25 mM	Invitrogen™, 15567027
Triton X-100	1 %	Sigma-Aldrich, T9284
SDS	0.1 %	Bio-Rad, 1610418
Sodium Deoxycholate	0.5 %	Sigma Aldrich, 30970

Diluted chromatin (5 %) was taken as an input for each condition and the remaining diluted supernatant was incubated with a total 10 µg of antibody overnight at 4 °C. Magnetic Protein A Dynabeads® (Invitrogen™, 10002D) or Protein G Dynabeads® (Invitrogen™, 10004D), 60 µl per immunoprecipitation, were washed three times with 1 ml cold Wash Buffer A and blocked for 1 hour with 4 µl yeast tRNA (10 mg/ml) (Invitrogen™, AM7119) and 10 µl BSA (20 mg/ml) (Thermo Scientific™, B14) in 1 ml cold Wash Buffer A at 4 °C with rotation.

Table 2.7.2 – Composition of Wash Buffer A

Wash Buffer A		
Component	Concentration	Company and Catalogue
NaCl	150 mM	Sigma-Aldrich, S6546
EDTA pH 8.0	1 mM	Invitrogen™, AM9260G
Tris pH 8.0	50 mM	Invitrogen™, 15568025
NP-40	1 %	Sigma-Aldrich, 21-3277
SDS	0.1 %	Bio-Rad, 1610418
Sodium Deoxycholate	0.5 %	Sigma Aldrich, 30970

Blocked beads were re-washed three times with 1 ml Wash Buffer A and the 60 µl initial Protein A/G Dynabead volume was made up to 200 µl volume with Wash Buffer A and incubated overnight at 4 °C with rotation. 33.3 µl Protein A/G beads were added to each of the 6 tubes containing antibody-bound chromatin, which was incubated for at least 7 h at 4 °C with rotation to immobilise antibody-bound chromatin on the beads.

Washes, Elution and Library Preparation: The magnetic beads bound to antibody-chromatin complexes were rinsed once with Wash Buffer A and washed twice, for 10 min at 4 °C with rotation, with 1 ml Wash Buffer A and the beads were pooled into a single protein lobind tube across these two washes. The beads were washed once with 1 ml Wash Buffer B, once with 1 ml Wash Buffer C, each with a 10 min incubation at 4 °C with rotation, and rinsed with 1x TE buffer.

Table 2.7.3 – Composition of Wash Buffer B

Wash Buffer B		
Component	Concentration	Company and Catalogue
NaCl	500 mM	Sigma-Aldrich, S6546
EDTA pH 8.0	1 mM	Invitrogen™, AM9260G

Tris pH 8.0	50 mM	Invitrogen™, 15568025
NP-40	1 %	Sigma-Aldrich, 21-3277
SDS	0.1 %	Bio-Rad, 1610418
Sodium Deoxycholate	0.5 %	Sigma Aldrich, 30970

Table 2.7.4 – Composition of Wash Buffer C

Wash Buffer C		
Component	Concentration	Company and Catalogue
LiCl	250 mM	Sigma-Aldrich, L7026
EDTA pH 8.0	1 mM	Invitrogen™, AM9260G
Tris pH 8.0	50 mM	Invitrogen™, 15568025
NP-40	1 %	Sigma-Aldrich, 21-3277
Sodium Deoxycholate	0.5 %	Sigma Aldrich, 30970

Chromatin was eluted from the beads in 450 µl of Elution buffer, including the input samples, and incubated at 37 °C for 2 h, followed by an overnight incubation at 65 °C with shaking to reverse cross-link protein from DNA.

Table 2.7.5 – Composition of ChIP Elution Buffer

ChIP Elution Buffer		
Component	Concentration	Company and Catalogue
SDS	1 %	Bio-Rad, 1610418
Sodium Bicarbonate	0.1 M	Sigma-Aldrich, S6014
Proteinase K	220 µg	Sigma-Aldrich, 39450-01-6

RNase A	20 µg	Promega, A7973

DNA was purified using AMPure XP beads (Beckman Coulter, A63881) into DNA lobind tubes (Eppendorf, 0030108051) and eluted in 50 µl buffer EB (Qiagen, 19086). DNA was quantified using the Qubit fluorometer double stranded DNA high sensitivity assay kit (Thermo Scientific™, Q32854) and libraries were prepared using the NEBNext® Ultra™ II DNA library preparation kit for Illumina® (E7645S) using the manufacturer's protocol, with libraries indexed using NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1 and Set 2) (NEB, E7335 and E7500).

Table 2.7.6 – Antibodies used for chromatin immunoprecipitation-sequencing

Antigen	Host	Protein A/G	Source	µg Antibody per IP
PCGF3+5	Rabbit	Protein A	Abcam, ab201510	10 µg
RING1B	Mouse	Protein G	(Atsuta <i>et al.</i> , 2001)	10 µg

Following library preparation, library fragment size and concentration were determined using the Qubit fluorometer double stranded DNA high sensitivity assay kit, using an Agilent Bioanalyzer 2100 and by KAPA Library Quantification Kit (KAPA Biosystems, KK4824). Samples were sequenced on an Illumina® NextSeq500 instrument as HighOutput 75 bp single-end reads at the Babraham Institute Next Generation Sequencing Facility.

2.7.2 – ChIP-sequencing analysis

Reads were trimmed by the Babraham Institute Bioinformatics facility using trim galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), with default primers to remove the standard Illumina adapter sequence, and mapped to the human genome GRCh38 using Bowtie2 v.2.3.2 (Langmead *and* Salzberg, 2012). BAM files were imported to Seqmonk (<https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/>) and reads were extended by 200 bp at their 5' end to approximate the true insert size.

ChIP-sequencing analysis was performed by Peter Rugg-Gunn. PRC1.3 target promoters were identified by generating 4 kb probes centred on annotated transcriptional start sites. Duplicated reads were removed and the remaining reads were quantified and corrected per million mapped reads. Probes with a $\log_2$ reads per million (RPM) > 3 for RING1B ChIP samples and $\log_2$ RPM > 1 for PCGF3 ChIP samples were retained. Quantitation values were normalised across the RING1B ChIP datasets and percentile normalised (75 %) across the PCGF3 ChIP datasets. A Limma test (after multiple testing correction) was performed between WT and *PCGF3* KO samples and probes were retained if they had a p-value < 0.05 in the PCGF3 ChIP samples. Retained RING1B and PCGF3 probes were overlapped and identified PRC1.3 occupancy on chromatin, and were name-matched to genes.

Peaks were called using a MACS (model-based analysis of ChIP-sequencing) implementation in SeqMonk (Zhang *et al.*, 2008). Each ChIP replicate was analysed using parameters $q < 10^{-12}$ for RING1B and $q < 10^{-5}$ for PCGF3 and a sonicated fragment size of 300 bp. Peaks identified in both replicates were retained if overlapping in at least one 500 bp window where $\log_2$ RPM > 0. PCGF3 peaks were removed if they overlapped with peaks in *PCGF3* KO samples. PRC1.3 peaks were identified by overlapping RING1B and PCGF3 peaks, and visualised using the WashU epigenome browser (Kuhn *et al.*, 2013; Li *et al.*, 2019). Matched control RING1B peaks were identified by filtering peaks further than 10 kb from a PCGF3 peak with a similar distribution to the PRC1.3 peak set. PRDM14 ChIP signal was compared between PRC1.3 peaks and matched RING1B only sites using a two-sided Mann-Whitney test.

Heatmaps were generated by Andrew Malcolm. BAM files were sorted and indexed using SAMTools (Li *et al.*, 2009). BAM files were converted to a bigWig format using Deeptools bamcoverage (Ramírez *et al.*, 2016) with the following parameters: `--binSize 10 --normalizeUsing RPGC --effectiveGenomeSize 2913022398 --extendReads 200 --ignoreForNormalization ChrM ChrY --minMappingQuality 20`. Input files were subtracted from samples using bigWigCompare using `--operation subtract`. Input-subtracted replicates were merged using bigWigCompare using `--operation mean`. To generate aligned probe plots, first computeMatrix was run on each input-normalised coverage file using parameters `--upstream 10000 --downstream10000 --missingDataAsZero --skipZeros`. Plots were generated using plotHeatmap with interpolationMethod and the samples were sorted by PCGF3 signal.

2.7.3 – Calibrated Cleavage Under Targets and Tagmentation (CUT&Tag)

CUT&Tag was performed as described in (Bendall *and* Semprich, 2022; Kaya-Okur *et al.*, 2019).

Table 2.7.7 – Buffers used for CUT&Tag

Nuclear Extraction Buffer	
HEPES-KOH pH 7.9	20 mM
KCl	10 mM
Spermidine	0.5 mM
Triton X-100	0.1 %
Glycerol	20 %

Digitonin Wash Buffer	
Digitonin	0.01 %
NP-40	0.01 %
BSA	0.05 %

Bead Activation Buffer	
HEPES KOH pH 7.9	20 mM
KCl	10 mM
CaCl ₂ (Sigma-Aldrich, C1016)	1 mM
MnCl ₂ (Sigma-Aldrich, 244589)	1 mM

Antibody Buffer – Digitonin Wash Buffer supplemented with:	
EDTA	2 mM

Digitonin 300 Wash Buffer – Digitonin Wash Buffer supplemented with:	
NaCl	300 mM

Tagmentation Buffer – Digitonin 300 Wash Buffer supplemented with:	
MgCl ₂	10 mM

Stop Buffer	
EDTA (0.5 M)	10 µl per sample
SDS (20 %)	1.5 µl per sample
Proteinase K (20 mg/ml)	2.5 µl per sample

Naive or primed PSCs and *Drosophila* S2 cells were dissociated and counted as described previously. Nuclear extraction buffer (500 µl) was added to either naïve or primed PSCs and incubated for 10 min at 4 °C. The suspension was centrifuged at 600 x g for 3 min at 4 °C and the supernatant was carefully discarded. Nuclei were lightly fixed in 0.1 % methanol-free paraformaldehyde at RT for 2 min and quenched with 75 mM Glycine. Fixed nuclei were centrifuged at 600 x g for 3 min at 4 °C and the supernatant was carefully discarded.

Fixed nuclei were re-suspended in Digitonin wash buffer containing cOmplete™ EDTA-free protease inhibitors, cells were counted and 50,000 human cells and 20,000 *Drosophila* cells per reaction were combined and transferred into a fresh protein lo-bind tube or cryovial. At this point, fixed nuclei could be frozen for later use by the addition of 10 % DMSO and slow freezing at –80 °C overnight in a Nalgene® Mr. Frosty freezing container and stored at –80 °C.

Concanavalin A conjugated paramagnetic beads were resuspended and washed twice on a magnetic rack in bead activation buffer. The beads were added to the cell suspension and incubated with rotation for 10 min at RT. Tubes were placed on the magnet and the liquid discarded and replaced with antibody buffer. Antibody (1 µg) was added to beads, which were incubated overnight at 4 °C with rotation.

Table 2.7.8 – Antibodies used for CUT&Tag

Antigen	Host	Source	µg Antibody
H2AK119ub1	Rabbit	Cell Signalling Technologies, D27C4	1 µg
IgG	Rabbit	Invitrogen, 31188	1 µg
Anti- <i>Drosophila</i> Histone H2Av	Mouse	Active Motif, 61686	1 µg

The following day, beads were retained on a magnet and liquid was discarded. Guinea-pig anti-rabbit secondary antibody (1 µg; Fisher Scientific, NBP172763) was added and incubated for 1 h at RT with rotation. Beads were washed three times with Digitonin wash buffer. 50 µl

of Digitonin 300 wash buffer containing a 0.5 µl of pA-Tn5 (Diagenode, C01070001) was added and the beads were incubated for 10 min at RT, before washing three times with Digitonin 300 wash buffer. Tagmentation buffer was added to facilitate DNA tagmentation and incubated for 1 h at 37 °C with mild shaking. Tagmentation was terminated by the addition of stop buffer, and the sample was incubated for 1 h at 55 °C to digest protein and reverse cross-links.

DNA was purified using a 1.2x AMPure XP bead clean-up into DNA lo-bind tubes and eluted in 25 µl 0.1x TE. To generate sequencing libraries, the tagmented DNA was amplified using KAPA HiFi HotStart ReadyMix (2x) (Roche, 7958927001), combining: 25 µl KAPA HiFi 2x PCR master mix, 2 µl i5 primer, 2 µl i7 primer and boiling at 98 °C for 3 min to activate the polymerase. 21 µl tagmented DNA was added and cycled through the following parameters: 72 °C for 5 min, 98 °C for 30 s, 14 cycles of 98 °C for 30 s, 63 °C for 10s, and finally incubated at 72 °C for 1 min. Amplified DNA was purified using a 1.2x AMPure XP bead clean-up into DNA lo-bind tubes and eluted in 20 µl 0.1x TE. Library concentration was determined as for ChIP-sequencing (section 2.7.1) and were sequenced on an Illumina NextSeq500 instrument as HighOutput 75 bp paired-end reads at the Babraham Institute Next Generation Sequencing Facility.

2.7.4 – Calibrated CUT&Tag analysis

Calibrated CUT&Tag analysis was performed by Andrew Malcolm. Raw FastQ files were trimmed by the Babraham Institute Bioinformatics facility using trim galore (v.0.6.6) and aligned to the human GRCh38 genome or the *Drosophila* BDGP6 genome using Bowtie2 v.2.3.2 (Langmead *and* Salzberg, 2012) with the following parameters –very-sensitive -l 10 -X700. High quality reads with a MAPQ value > 20 were retained by filtering with samtools view (v1.11), and the rest were removed (Li *et al.*, 2009). Sorted bam files were then deduplicated using Picard rmDup (v2.26, Broad Institute, 2019). Calibration factors for each sample were determined as the ratio of the sample with the lowest number of unique mapped *Drosophila* spike-in tags within a corresponding replicate over the number of unique mapped *Drosophila* spike-in tags per sample within each replicate. These calibration factors were then used to scale the human genome-mapped reads by random downsampling.

Calibrated bigWig files were produced using deepTools bamCoverage with scaling by the calculated calibration factors (v.3.43) (Ramírez *et al.*, 2016). Replicates were merged and processed using deeptools bigwigCompare with the following parameter --operation mean. Custom R scripts were used to prepare bigwig files for visualisation via the WashU Epigenome

Browser (v.5) (Kuhn *et al.*, 2013; Li *et al.*, 2019). Heatmaps and profiles over naïve-specific and primed-specific promoters were produced using deepTools computeMatrix with the following settings --missingDataAsZero and visualised using plotHeatmap or plotProfile.

Heatmaps were generated over called peaks using deeptools computeMatrix with the following settings --missingDataAsZero and plotted with plotHeatmap or plotProfile. Coverage was analysed within 1 kb genome windows, which were made by sorting and indexing of BAM files by deeptools multiBamSummary and applying calibration factors, with the signal being quantified using multiBigwigSummary. For analysis of 1-kb windows of the genome, the average coverage of 1-kb bins from individual replicate bigwig files was calculated using multiBigwigSummary, coverage was averaged between replicates and then used to generate kernel density estimates of coverage genome-wide. Coverage chromosome-wide was determined using 250-kb windows of the genome, coverage was averaged between replicates and coverage per bin plotted as a line plot from start to end of either chromosome 18.

2.7.5 – RNA-sequencing

Following TRIzol RNA extraction (as described in section 2.5.3), RNA-sequencing libraries were constructed from 500 ng RNA using the NEBNext® Ultra™ RNA Library Preparation Kit for Illumina® (NEB, E7530), with the Poly(A) mRNA Magnetic Isolation Module (NEB, E7490). Libraries were indexed using NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1 and 2) (NEB, E7335 and E7500). Library fragment size and concentration was determined using an Agilent Bioanalyzer 2100 and KAPA Library Quantification Kit (Roche, KK4824). Samples were sequenced on an Illumina® NextSeq500 instrument as 150 bp single-end reads at the Babraham Institute Next Generation Sequencing Facility.

2.7.6 – RNA-sequencing analysis

RNA-sequencing reads were trimmed by the Babraham Bioinformatics facility using trim galore v0.6.5 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with default parameters to remove the standard Illumina adaptor sequence. Reads were mapped to the human GRCh38 human genome using Bowtie2 v.2.3.2 (Langmead *and* Salzberg, 2012). BAM files were imported to Seqmonk

(<https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/>). Raw read counts per transcript were calculated using an RNA-sequencing quantitation pipeline implemented in Seqmonk and data were analysed using DESeq, hierarchical clustering, correlation matrix and principal components analysis in Seqmonk. Differentially expressed genes were identified using DESeq2 (Love *et al.*, 2014), using a Wald test and Benjamini-Hochberg correction with a false discovery rate < 0.05. P-values for fold change in expression comparison of PRC1.3 ChIP targets were calculated using a Kruskal-Wallis test with Dunn's multiple comparison test.

2.7.7 – Gene ontology analysis

The R package enrichR (Kuleshov *et al.*, 2016) was used to analyse gene ontology (GO) data for the GO term: GO_Biological_Process_2018. GO terms were ranked by significance and significance was calculated using a Fisher's exact test.

2.7.8 – Motif analysis

Motif analysis was performed on the central 500 bp region within naïve PRC1.3 peaks using AME (McLeay *and* Bailey, 2010). The HOCOMOCO v11 motif database was used with the core human mono meme format. FIMO (Grant *et al.*, 2011) was used to identify the number of PRDM14 motifs in PRC1.3 peaks versus matched RING1B-only peaks. 53 % of naïve PRC1.3 peaks (201/378) contained at least one PRDM14 motif, compared to around 6 % in the matched RING1B-only peaks (19/295). The significance of PRDM14 motif enrichment in naïve PRC1.3 peaks compared to the control subset was compared with a one-sided Fisher exact test and Bonferroni correction.

2.7.9 – Gene set enrichment analysis (GSEA)

Gene set enrichment analysis (Subramanian *et al.*, 2005) was calculated using the GSEAPreranked tool within GSEA software (www.broadinstitute.org/gsea). RNA-sequencing data of PRDM14-AID-VENUS naïve PSCs were used as an input and the dataset was ranked by fold-change expression between IAA-treated and untreated samples (n=34,624). PRC1.3 target genes (n=267), defined by RING1B and PCGF3 promoter co-localisation, were

integrated to analyse the effect of IAA treatment on expression of these genes. Default settings were used with 1,000 gene set permutations. A positive enrichment (ES) score was calculated with a Kolmogorov-Smirnov test, with a false discovery rate < 0.001 .

3 – PRC1.3 is essential for human naïve pluripotent stem cell reprogramming

3.1 – Introduction

3.1.1 – Characterisation of PRC1.3 complex action in human pluripotency and reprogramming

Naïve PSCs capture the properties of the human embryonic pre-implantation epiblast *in vitro*, with several desirable properties. Naïve PSCs are predominantly produced by reprogramming of primed PSCs (Guo *et al.*, 2017; Takashima *et al.*, 2014; Theunissen *et al.*, 2014). The primed to naïve reprogramming (PTNR) process is poorly understood, and is a low efficiency and time consuming process (Collier *and* Rugg-Gunn, 2018), which limits our ability to produce large numbers of high quality, pure naïve PSCs for research.

Genome-wide CRISPR-Cas9 screening performed by Amanda Collier surprisingly identified almost all of the components of the epigenetic regulator Polycomb Repressive Complex PRC1.3 as key regulators of human PTNR. PRC1.3 in its entirety is therefore predicted to play a pivotal role in formation of naïve PSCs by reprogramming. This project will analyse the importance of PRC1.3 in human PTNR in a comprehensive and detailed manner. By better understanding how critical factors facilitate the key molecular events that underpin PTNR, it will be easier to unlock and harness the full potential of naïve PSCs, producing more, higher quality naïve PSCs *in vitro* on a shorter timescale for desired primed PSC lines.

I will begin by creating and confirming knockout lines of PRC1.3 complex genes in human primed PSCs, which has never before been performed. Using cellular characterisation approaches such as immunofluorescence microscopy, flow cytometry and proliferation assays, the phenotype of PRC1.3-deficient primed PSCs can be determined. This will help to understand whether the PRC1.3 complex is functionally required for human primed pluripotency or only for PTNR, as a perturbed primed state of human pluripotency could help to explain defective PTNR. A primed pluripotency genome-wide screen performed alongside the reprogramming screen suggested that PRC1.3 was not required for human primed PSCs (Collier *et al.*, 2022), but this can be assessed more rigorously on an individual basis using the methods described above.

I will also confirm the importance of the PRC1.3 complex in PTNR using the PRC1.3 knockout (KO) lines. As the screen was performed under 5i/L/A reprogramming conditions (Theunissen *et al.*, 2014), I will also test the ability of PRC1.3-deficient primed PSCs to reprogramme under alternative reprogramming conditions, to establish whether the PRC1.3 complex is required only for 5i/L/A reprogramming or is more centrally important for the reprogramming cell state transition by any means. Several PRC1.3 complex specific genes were identified by the screen

and so creating knockouts of more than one PRC1.3 complex gene will be used to corroborate the findings of the screen and strengthen the evidence for the role of PRC1.3 in reprogramming.

Finally, I would also like to study the importance of the PRC1.3 complex in the established naïve state of human pluripotency. PRC1.3 could also play a role in the developmental establishment of the naïve pluripotency state in the human pre-implantation epiblast during this early window in embryo development. PRC1.3 deficiency in established naïve PSCs could act as a model system to test this hypothesis. An inducible knockdown system, which has the power to synchronously and reversibly remove the complex from naïve PSCs, would be ideal for this approach. A naïve-specific inducible PRC1.3 knockdown could determine whether the complex is essential only for reprogramming to naïve PSCs, or if it is more broadly essential for the human naïve state of pluripotency, which could explain the failure of naïve PSCs to form and expand during reprogramming.

3.1.2 – Aims

A better understanding of PTNR is required to fully harness the power of human naïve pluripotency. To this end, a genome-wide CRISPR-Cas9 screen performed by Amanda Collier identified the PRC1.3 complex as one of the most critical regulators of human PTNR (Collier *et al.*, 2022). This chapter will characterise the cellular role of the PRC1.3 complex in human pluripotency and reprogramming, confirming the central phenotype identified by the screen. Molecular and mechanistic characterisation of the function of the PRC1.3 complex will be explored in subsequent chapters.

1. Create knockouts (KO) of PRC1.3 complex genes to determine the phenotype of PRC1.3-deficient human primed PSCs
2. Demonstrate the essential role for PRC1.3 in human PTNR
3. Create an inducible knockdown for PRC1.3 complex genes to determine the phenotype of PRC1.3-deficient human naïve PSCs

3.2 – Investigating the role of PRC1.3 in human primed pluripotency

3.2.1 – Analysing the specific requirement for PRC1.3 from the genome-wide screen of human PTNR

PRC1.3 components were amongst the most significant factors identified by the screen (Fig. 1.5.4) and catalysis of H2AK119ub1 was a significant gene ontology term amongst all the ‘essential’ genes identified (Fig. 1.5.5). H2AK119ub1 however is catalysed by all PRC1 complexes, which are highly diverse in their composition, recruitment mechanisms and catalytic activities (Aranda *et al.*, 2015). I therefore began by assessing the statistical significance of all PRC1-associated genes from the genome-wide screen, to confirm whether PRC1.3 is specifically important in PTNR, or whether this was a more general characteristic of PRC1 complexes. This was done by analysing the calculated p-value for each of these genes with respect to a pre-determined significance threshold.

The exclusive, identity-determining component of PRC1.3 complexes, *PCGF3*, is the #2 factor identified in the screen when ranked by p-value. While both *PCGF1* (#430 factor) and *PCGF2* (#173 factor), of PRC1.1 and PRC1.2 complexes respectively, exceed the statistical significance threshold, none of the other genes encoding auxiliary components of the PRC1.1 or PRC1.2 complexes were identified as significant and no specific components of the PRC1.6 were identified as essential factors. However, auxiliary components of the PRC1.3 complex, *PCGF3*, *FBRS* (#10 factor) and *CSNK2A3* (#88 factor) of the CKII complex, along with the core factors present in all PRC1 complexes, *RNF2/RING1B* (#15 factor) and *RYBP* (#398 factor), were identified. Despite PRC1.3 and PRC1.5 complexes sharing most of the same auxiliary proteins, only *PCGF3*, and not *PCGF5* (#1,115 factor), was identified as essential for reprogramming (Fig. 3.2.1).

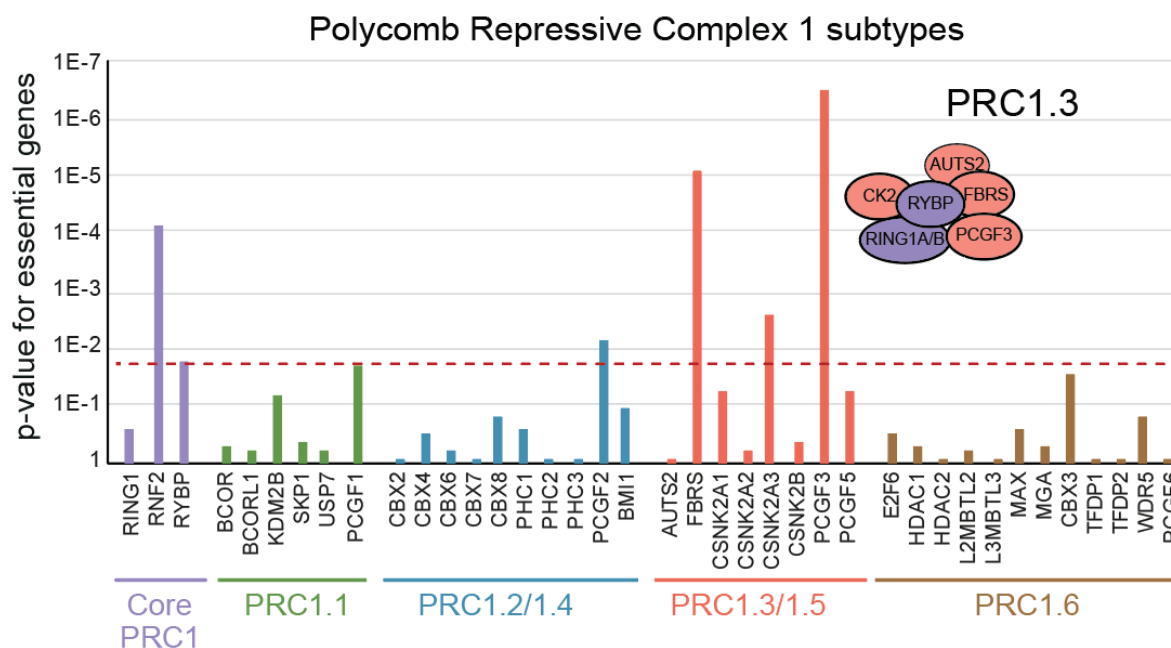


Figure 3.2.1 – Chart shows the significance of all PRC1 complex component genes in human PTNR

The dotted red line indicates a significance threshold of $p=0.02$ as determined by a permutation test. Schematic of the PRC1.3 complex is shown with essential genes outlined in black.

Thus, factors identified as statistically significant capture almost all of the components that specifically comprise the PRC1.3 complex, and not components of other PRC1 complexes, suggesting PRC1.3 is a key candidate for an essential epigenetic regulator of human PTNR.

I also analysed genes encoding components of Polycomb Repressive Complex 2 (PRC2) complexes due to the high degree of cooperativity between the PRC1 and PRC2 complexes (Blackledge *et al.*, 2014; Cooper *et al.*, 2016; Gao *et al.*, 2012; Kalb *et al.*, 2014; Tavares *et al.*, 2012). PRC2 complexes contain only two specific subtypes, named as PRC2.1 (containing EPOP, PHF1, MTF2, PHF19 and LCOR) and PRC2.2 (containing AEBP2 and JARID2) (Healy *et al.*, 2019; Holoch and Margueron, 2017; Mierlo *et al.*, 2019). Only *EZH2* of the core components of PRC2 complexes *EZH2*, *EED*, *SUZ12*, *RBBP4* and *RBBP7*, was identified as essential for PTNR, even though *EED* and *SUZ12* are essential for catalysis and propagation of H3K27me3 (Cao and Zhang, 2004; Montgomery *et al.*, 2005; Pasini *et al.*, 2004). None of the accessory proteins of PRC2.1 or PRC2.2 complexes exceeded the statistical threshold.

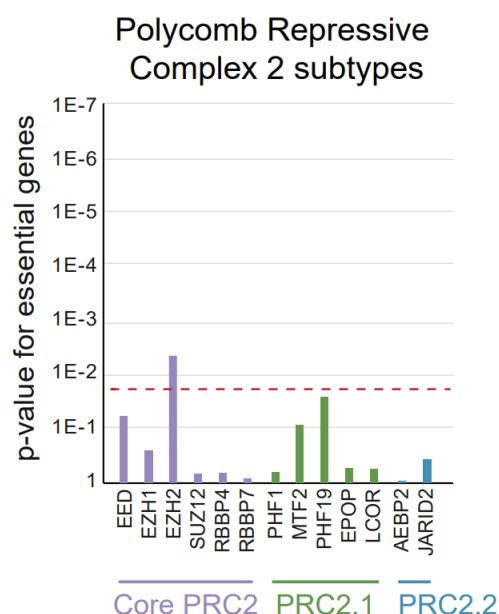


Figure 3.2.2 – Charts show the p-values of PRC2 complex components as essential regulators of human PTNR

The red line indicates $p=0.02$ as a cut-off (permutation test).

Taken together, these results show that the action of PRC1.3 is specifically required for human PTNR, with the action of other PRC1 subtypes and PRC2 complexes not playing a key role in this process.

3.2.2 – Assessing PRC1.3 component gene expression in human embryos and PSCs

I wanted to investigate whether PRC1.3 complex genes were expressed in human pre-implantation embryo development. If PRC1.3 complex components were not expressed in human epiblast cells, which human naïve and primed PSCs model *in vitro*, I would reason that this complex is unlikely to have functionality in human embryology during this early stage. PRC1.3 component gene expression was analysed in human pre-implantation embryo single-cell RNA-sequencing datasets (Petropoulos *et al.*, 2016; Stirparo *et al.*, 2018).

I confirmed that PRC1.3 component genes were expressed throughout human pre-implantation embryo development, including in pre-implantation epiblast cells that are the *in vivo* counterpart of naïve PSCs, meaning that understanding PRC1.3 function is human PSCs is also relevant to the study of human development (Fig. 3.2.3).

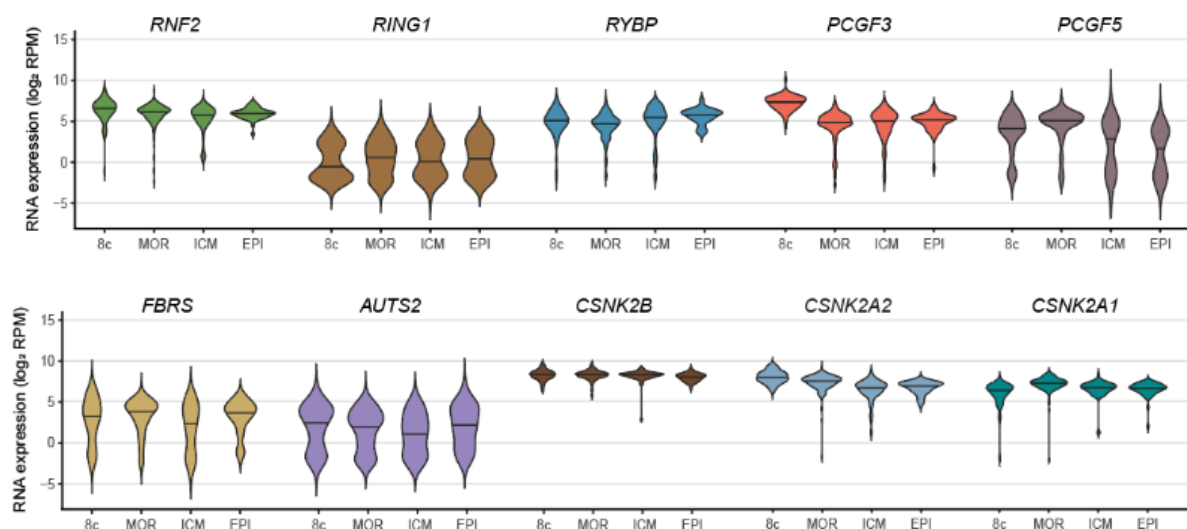


Figure 3.2.3 – Violin plots show the expression level distribution of PRC1.3/5 complex components in single cells throughout human preimplantation embryo development

8-cell stage (8c); morula (MOR); early inner cell mass (eICM); epiblast (EPI). Data obtained from (Petropoulos *et al.*, 2016) with annotation from (Stirparo *et al.*, 2018).

Additionally, I observed that in the majority of epiblast cells, *PCGF3* is expressed at higher levels than *PCGF5*. I analysed this pattern in published RNA-sequencing data from primed and naïve PSCs (Collier *et al.*, 2017) and saw the same pattern of increased expression of *PCGF3* compared to *PCGF5* (Fig. 3.2.4).

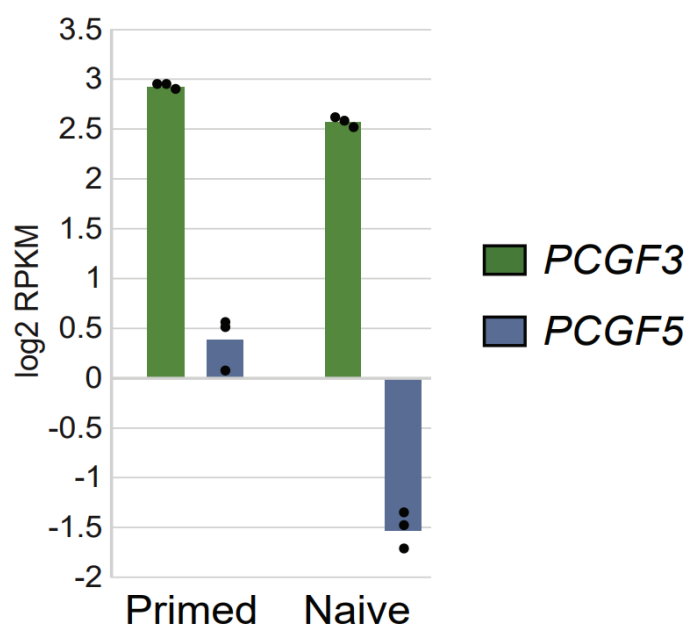


Figure 3.2.4 – Charts show *PCGF3* and *PCGF5* expression levels in log₂ RPKM (reads per kilobase transcript per million reads), in primed and naïve PSCs

Circles show each biological replicate (n=3 for both primed and naïve) for the *PCGF3* and *PCGF5* genes. Data obtained from (Collier *et al.*, 2017).

As PRC1.3 and PRC1.5 share auxiliary proteins that were identified as important for PTNR, the low level of *PCGF5* expression might explain why *PCGF5* was not also identified in the genome-wide screen.

In summary, PRC1.3 components are expressed in human pre-implantation embryos and in human naïve and primed PSCs. This finding prompted me to further explore the role of the PRC1.3 complex in human pluripotency and reprogramming.

3.2.3 – Generating PRC1.3-deficient human primed PSCs

As the role of PRC1.3 has not been studied in the context of human pluripotency, I sought to first investigate whether PRC1.3 is important for the maintenance of pluripotency in primed PSCs. Although the human primed pluripotency proliferation screen (Collier *et al.*, 2022) did not identify PRC1.3 complex genes as important for primed PSC proliferation, I sought to explore this possibility on an individual basis, using additional cellular assays. To do this, I generated individual genetic deletions of *PCGF3* in primed PSCs using CRISPR-Cas9. This approach could provide novel information about the regulation of human primed pluripotency and as primed PSCs are the starting population for reprogramming, this could identify whether the PRC1.3 dependence in reprogramming identified by the screen might reflect a more general pluripotency defect.

To disable *PCGF3*, I deleted exon 5 from the gene in H9 primed PSCs (Fig. 3.2.5), which was predicted to cause a frameshift mutation.

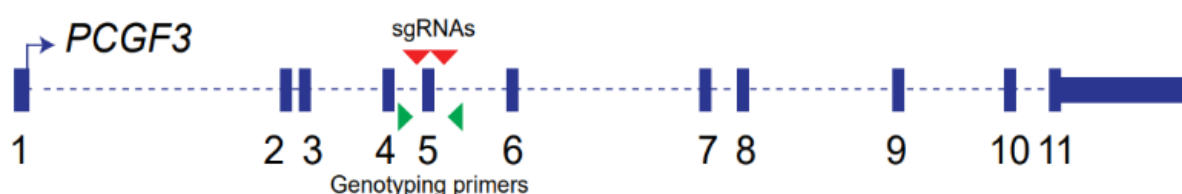
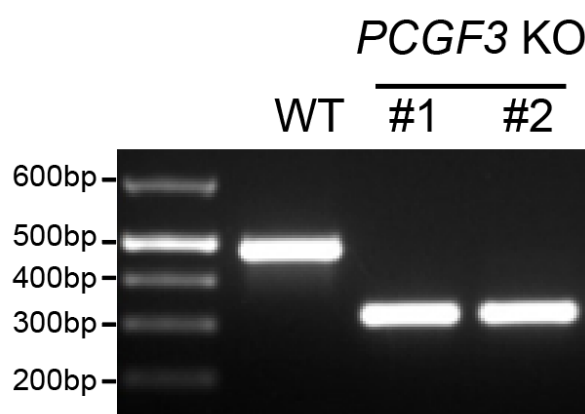


Figure 3.2.5 – Schematic of the CRISPR-Cas9 strategy used to create *PCGF3* KO primed PSCs

Schematic of *PCGF3* exon structure. Red triangles indicate the position of the gRNAs used and green triangles indicate the position of genotyping primers subsequently used to confirm the exon deletion.

CAS9 fused to a GFP reporter, and a pair of guide RNAs designed to target either side of exon 5 of the *PCGF3* gene, were co-transfected into cells on separate plasmids. 48 h after transfection, the GFP-positive cells were flow-sorted and expanded. Individual colonies were picked, and the clones were screened by PCR for successful *PCGF3* deletion in both alleles. Genotyping primers were designed to produce a shorter PCR product when exon 5 is deleted compared to their untargeted WT counterparts. In this genotyping screening, 25 clones were screened and identified 16 homozygous *PCGF3* KO clones, where exon 5 was deleted from both alleles, and 6 heterozygous clones, where the exon was deleted on one allele. Two *PCGF3* KO clones were selected for further experimentation (Fig. 3.2.6).

**Figure 3.2.6 – Electrophoresis image of PCR-based genotyping of the *PCGF3* gene**

WT parental cells show an expected band size of 471 bp. The reduced band in the two *PCGF3* KO primed PSC lines of 308 bp indicates homozygous deletion of exon 5. DNA fragment sizes in the ladder are shown in base pairs (bp).

PCR products from WT cells and the two *PCGF3* KO clones were sub-cloned into the pCR™2.1-TOPO® TA vector and Sanger sequencing of two bacterial clones per genotype confirmed the homozygous genetic deletion of exon 5 from both *PCGF3* KO primed PSCs (Fig. 3.2.7).

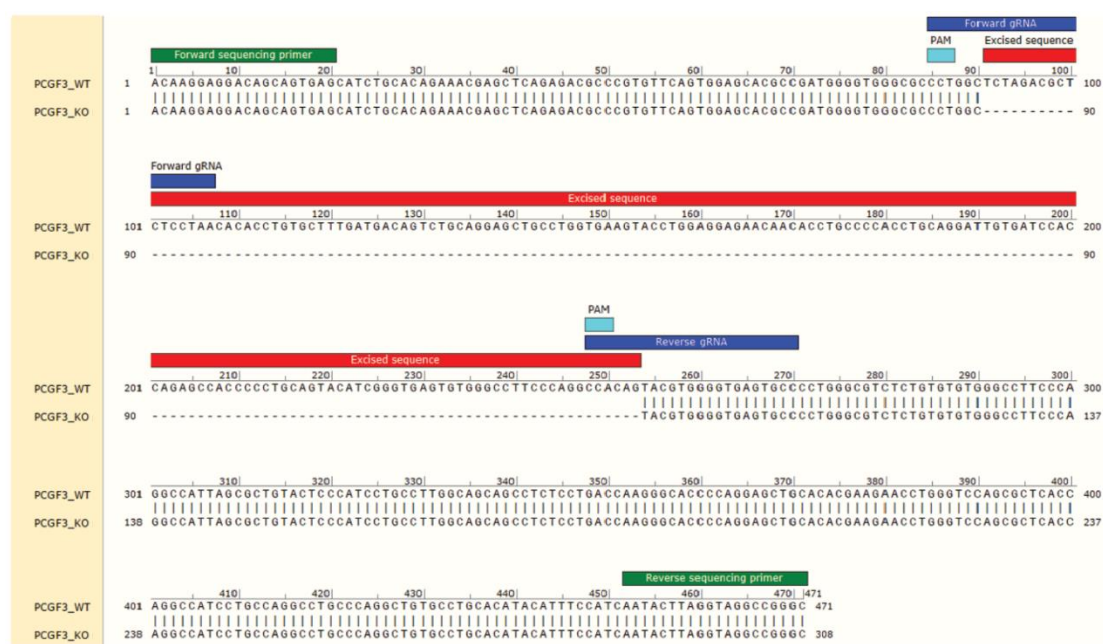


Figure 3.2.7 – Multiple sequence alignment of Sanger sequencing identifies successful exon deletion within *PCGF3*

WT parental primed PSCs show an uninterrupted sequence throughout exon 5 while the *PCGF3* KO (identical sequences for both KO clones) primed PSC lines shows 163bp of deleted sequence, indicating homozygous deletion of the targeted exon. Sequencing primers (green), guide RNA sequences (dark blue) protospacer adjacent motif (PAM) sequences (light blue) and the excised sequence within the *PCGF3* KO genotype (red) are highlighted.

I next used Western blotting to confirm the successful deletion of *PCGF3* protein from the two independent *PCGF3* KO clones. I used an antibody that detects both *PCGF3* and *PCGF5* proteins, as no *PCGF3*-specific antibody was of satisfactory quality when tested. *PCGF3* protein was detectable in WT, but not KO clones as expected, while *PCGF5* protein is undetectable in all cells, consistent with low levels of *PCGF5* expression in human epiblast cells and PSCs. This also shows that *PCGF5* is not upregulated to a detectable degree by the deletion of *PCGF3* (Fig. 3.2.8).

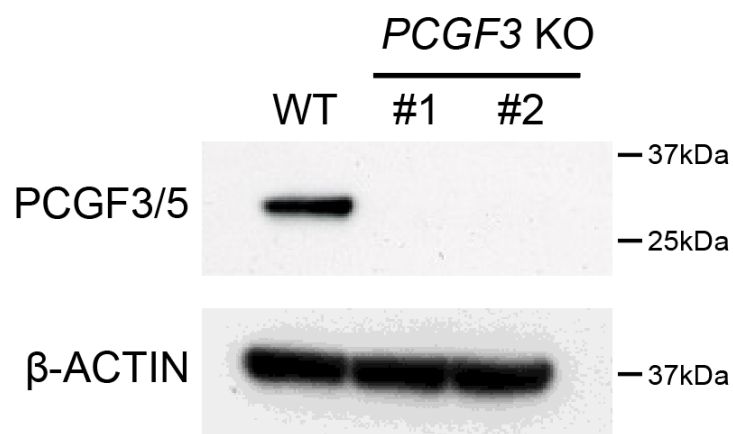


Figure 3.2.8 – Western blotting shows the absence of PCGF3 protein in two KO primed PSC lines

PCGF3 has a molecular weight of 28 kDa; β -ACTIN, loading control, has a molecular weight of 42 kDa. PCGF5 protein has a predicted molecular weight slightly lower than that of PCGF3.

In summary, I have successfully generated a number of putative *PCGF3* KO primed PSC lines, in which two have been validated, to test the impact of PRC1.3-deficiency in human primed pluripotency.

3.2.4 – Testing the impact of *PCGF3* deletion on the association of the PRC1.3 complex

In order to ensure that *PCGF3* deletion is equivalent to PRC1.3 complex deletion, I tested whether the PRC1.3 complex can remain intact in the absence of *PCGF3* using a co-immunoprecipitation Western blotting approach. By performing an immunoprecipitation for RING1B, common to all PRC1 complexes including PRC1.3, in WT primed PSCs, I was able to co-detect RING1B, RYBP, PCGF3 and AUTS2, indicating that variant PRC1.3 complexes are being pulled down (Gao *et al.*, 2012; Hauri *et al.*, 2016). As expected, in one of the two *PCGF3* KO primed PSC clones, while RING1B and RYBP were detected, PCGF3 was not. AUTS2, which is specifically associated with PRC1.3 and PRC1.5 complexes, was detected at lower levels in the inputs of *PCGF3* KO compared to WT PSCs, suggesting that the protein may be destabilised outside the PRC1.3 complex, which has been previously observed for PRC2 component proteins in *EZH2* KO human primed PSCs (Collinson *et al.*, 2016). Crucially, AUTS2 was not detected in the RING1B-immunoprecipitated material from *PCGF3* KO cells (Fig. 3.2.9).

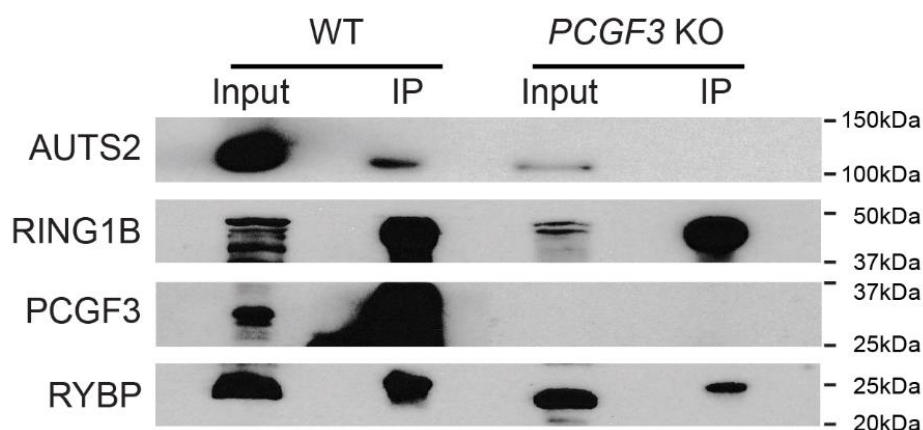


Figure 3.2.9 – Co-immunoprecipitation with RING1B antibody shows PRC1.3 complex dissociation in *PCGF3* KO primed PSC

RING1B acts as the immunoprecipitation target. PCGF3 confirms the genotype of cells. AUTS2 confirms the dissociation of PRC1.3 complexes in the absence of PCGF3. RYBP acts as a proxy for a loading control, present within all variant PRC1 complexes.

These findings show that in the absence of PCGF3, RING1B and AUTS2 do not interact with one another. The PRC1.3 complex therefore dissociates in the absence of PCGF3, and no PRC1.5 complexes that would also bridge RING1B to AUTS2 are present. Therefore, *PCGF3* deletion in human primed PSCs is equivalent to the removal of the entire PRC1.3 complex.

3.2.5 – Characterising the primed state of pluripotency in PRC1.3-deficient PSCs

With two PRC1.3 KO clones generated, I was able to extensively characterise these cells in order to ascertain the role, if any, for the complex in human primed pluripotency.

Firstly, the morphology of the cells was assessed in order to detect any signs of cellular differentiation or stress, which would indicate an important role for PRC1.3 in the human primed pluripotency network that cannot be compensated by other PRC1 complexes. The morphology of the two *PCGF3* KO primed PSC clones showed no detectable difference from WT cells, with all cells showing a typical flattened, dispersed colony structure with defined colony boundaries (Fig. 3.2.10), which is the classical morphology of pluripotent human primed PSCs.

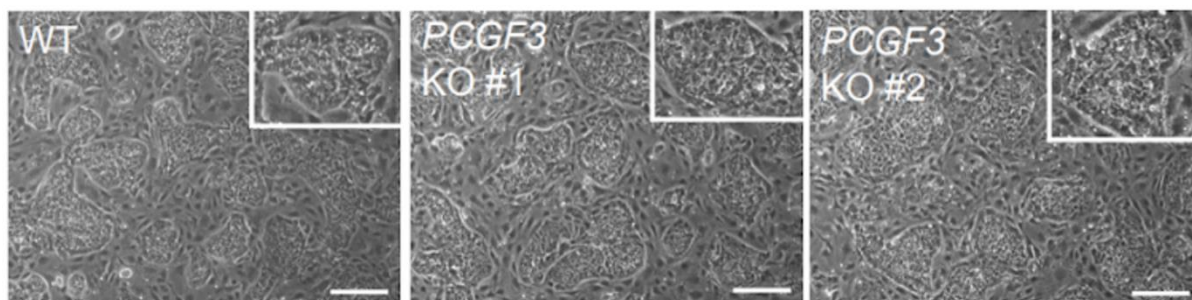


Figure 3.2.10 – Phase contrast images show the morphology of WT and *PCGF3* KO primed PSCs

The WT cells are untargeted parental H9 cells. Scale bars, 250 μm . Insets show a zoomed in view of colony morphology at high resolution.

Although morphologically similar, I also investigated the expression of the core pluripotency transcription factors *OCT4* and *NANOG* (Chambers *et al.*, 2003; Niwa, 2001) using immunofluorescence microscopy. Both the WT and the two *PCGF3* KO primed PSCs uniformly express both *OCT4* and *NANOG* in all cells, and so the cells are indeed pluripotent and *PCGF3* KO has not affected core pluripotency marker expression (Fig. 3.2.11).

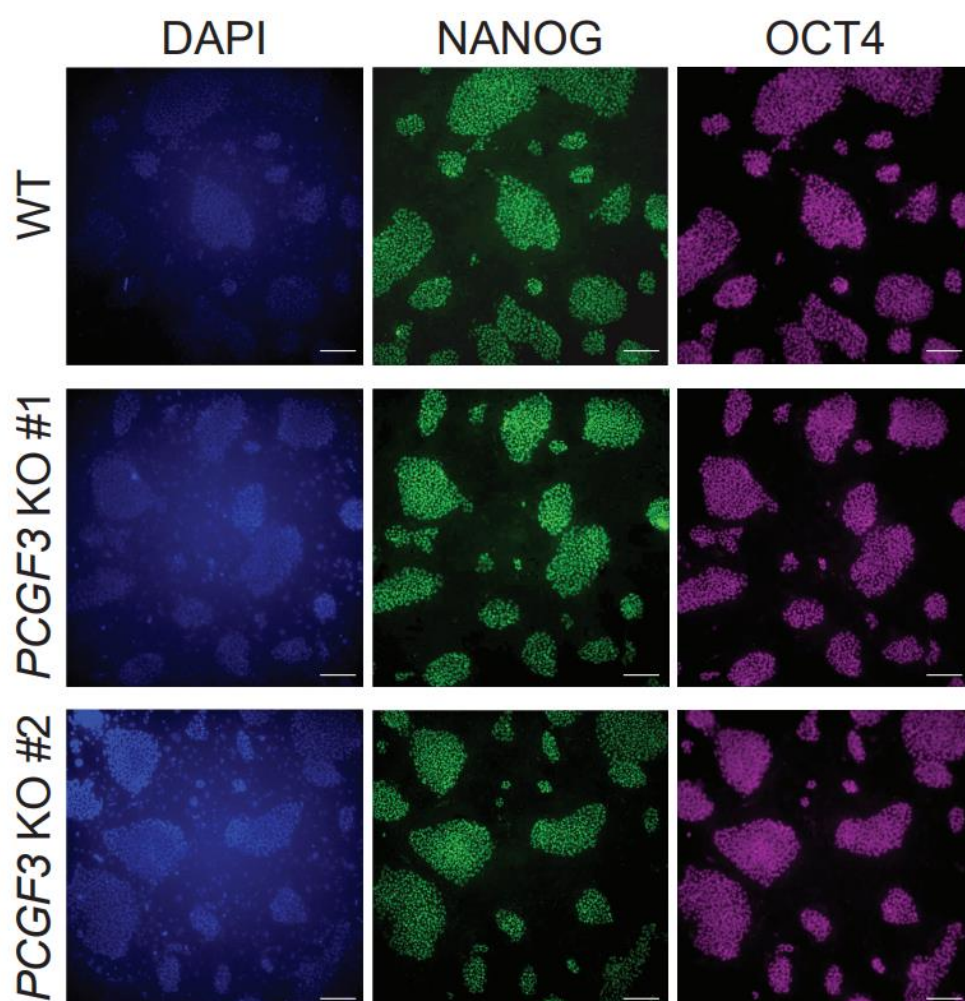


Figure 3.2.11 – Immunofluorescence microscopy shows the uniform expression of OCT4 and NANOG in WT and both *PCGF3* KO primed PSC lines

DAPI stains all nuclei. OCT4 and NANOG are both core pluripotency factors. Scale bars, 250 μ m.

The cell surface marker SSEA-4 also acts as a robust marker for undifferentiated cells and should be expressed on close to 100% of human primed PSCs (Henderson *et al.*, 2002; The International Stem Cell Initiative, 2007). Flow cytometry for SSEA-4 showed that almost all *PCGF3* KO primed PSCs express SSEA-4 (Fig. 3.2.12).

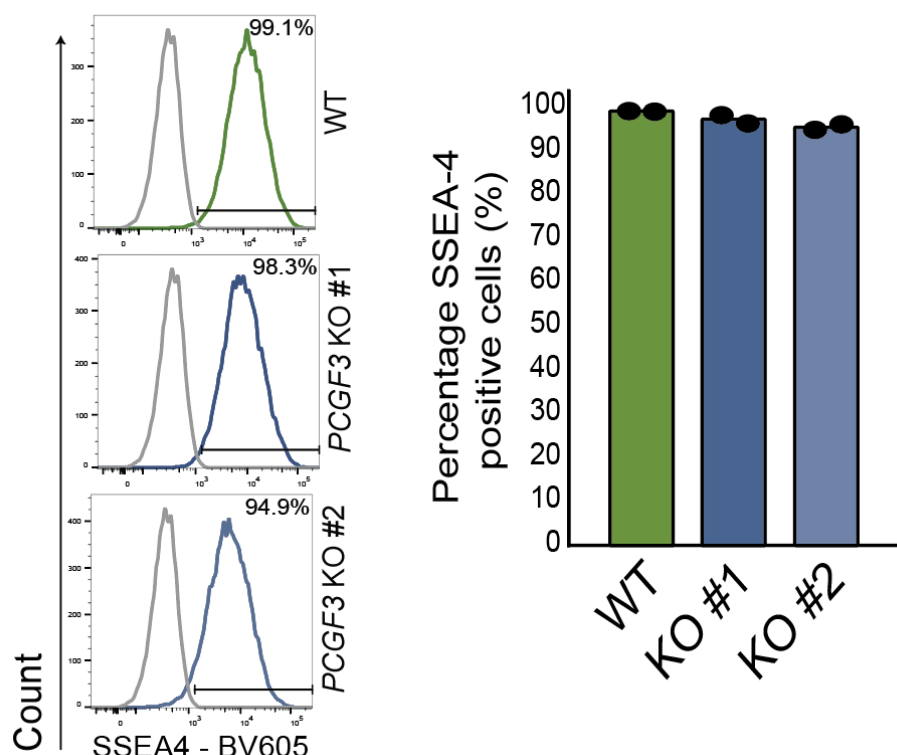


Figure 3.2.12 – SSEA-4 flow cytometry in WT and *PCGF3* KO primed PSC lines

Flow cytometry data represented as a histogram, representative of two independent biological replicates. Grey line shows omission of the SSEA-4 antibody. Bar chart quantifies the SSEA-4 positive cells as a percentage. Bar height shows the mean and circles show the individual data-points.

As *PCGF3* KO primed PSCs robustly express intracellular and cell-surface pluripotency markers, I analysed other properties that might perturb primed pluripotency, one of which was proliferation rate. As passaging takes place on defined days during the reprogramming protocol and human PSCs are very confluence sensitive, an accelerated or decreased proliferation rate could affect the efficiency of human PTNR. As PRC1.3 components were not detected in the primed proliferation CRISPR screen, it is unlikely that proliferation is severely impaired, but a more subtle increase or decrease in proliferation rates is possible. However, I confirmed that *PCGF3* KO clones show no significant difference in proliferation from WT primed PSCs (Fig. 3.2.13).

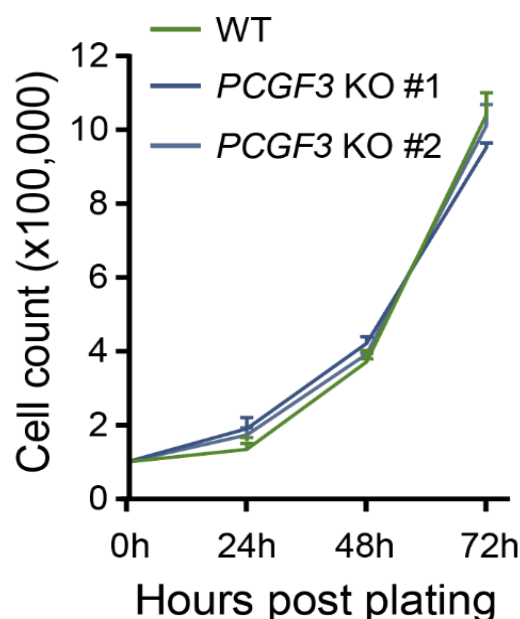


Figure 3.2.13 – Charts shows proliferation rates of WT and both *PCGF3* KO primed PSCs

Individual datapoints show the mean of three independent biological replicates and error bars represent the standard error of the mean.

As well as proliferation, I also checked the karyotype of the *PCGF3* KO primed PSCs, to detect any abnormalities such as major chromosome deletions, deletions or aneuploidy that could influence processes such as PTNR, for example by giving cells an increased or decreased propensity for passing through any molecular bottlenecks across the cell state transition. The karyotypes of the WT and *PCGF3* KO PSCs showed no major differences, with all cells containing a full complement of autosomal and sex chromosomes (46XX), with no major deletions or insertions detected (Fig. 3.2.14).

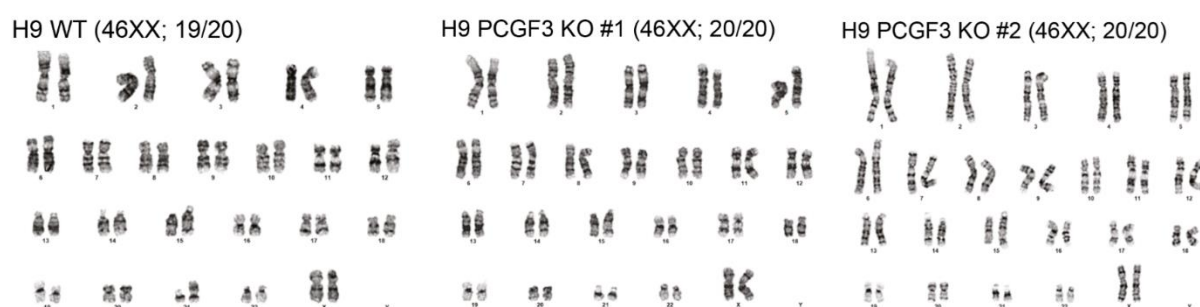


Figure 3.2.14 – Karyotyping of WT and two *PCGF3* KO primed PSC clones

Representative G-banded chromosome metaphase spreads are shown, as performed by a clinical cytogeneticist at Cell Guidance Systems. Twenty cells were examined for each cell line.

Altogether, by morphology, immunofluorescence microscopy, flow cytometry, proliferation and karyotyping analysis, I see no cellular differences between WT and PRC1.3 KO primed PSCs.

3.2.6 - *PCGF3* deletion causes very few transcriptional changes in primed PSCs

Cellular phenotyping shows that PRC1.3 is dispensable for the assessed characteristics of human primed pluripotency. However, the effect of PRC1.3 deletion on gene transcription has not been studied, which could inform on the gene regulatory role for PRC1.3 in human primed pluripotency. Strong perturbations to the underlying transcriptome in primed PSCs, albeit without characterised changes to the pluripotent state, might prejudice the starting population of cells, pre-disposing *PCGF3* KO primed PSCs to a failure to reprogramme. In order to determine whether there are transcriptional differences, RNA-sequencing was performed on WT and both *PCGF3* KO clones, on three independent biological replicates, which showed very similar global transcriptional profiles (Fig. 3.2.15).

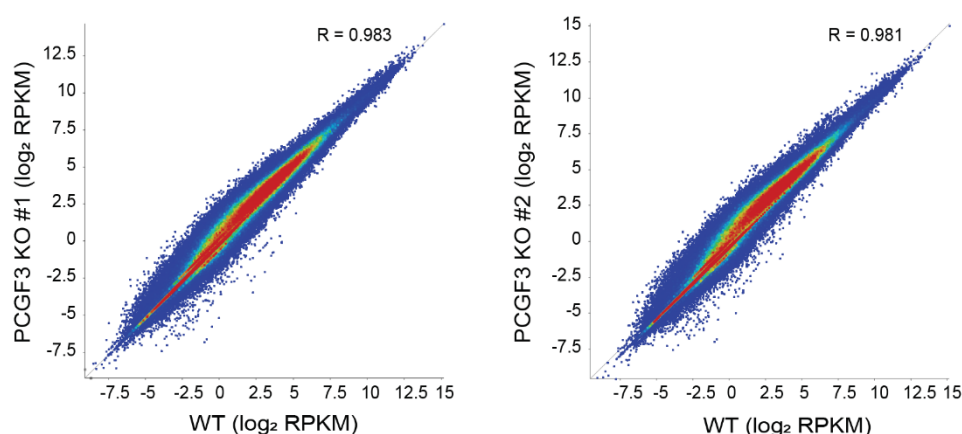


Figure 3.2.15 – Gene expression is nearly identical between WT and *PCGF3* KO primed PSCs

Pearson correlation coefficient of 0.983 for WT versus *PCGF3* KO clone 1 and 0.981 for WT versus *PCGF3* KO clone 2 indicate strong correlation in transcription. Three independent biological replicates were used for both WT primed and *PCGF3* KO primed PSCs and were grouped during analysis.

PCGF3 KO clone 1 possesses only nine differentially expressed genes from WT cells, and only fourteen are identified for *PCGF3* KO primed clone 2. This analysis was then focused towards a more directed subset of genes to confirm that differentially expressed genes did not include any pluripotency markers, differentiation markers or PRC1.3 components. All of these selected genes fail to show significant transcriptional differences between WT and *PCGF3* KO primed PSCs, with high expression levels of primed-specific and core pluripotency genes and

low levels of differentiation-specific gene expression, consistent with the expected gene expression signature of primed PSCs (Fig. 3.2.16).

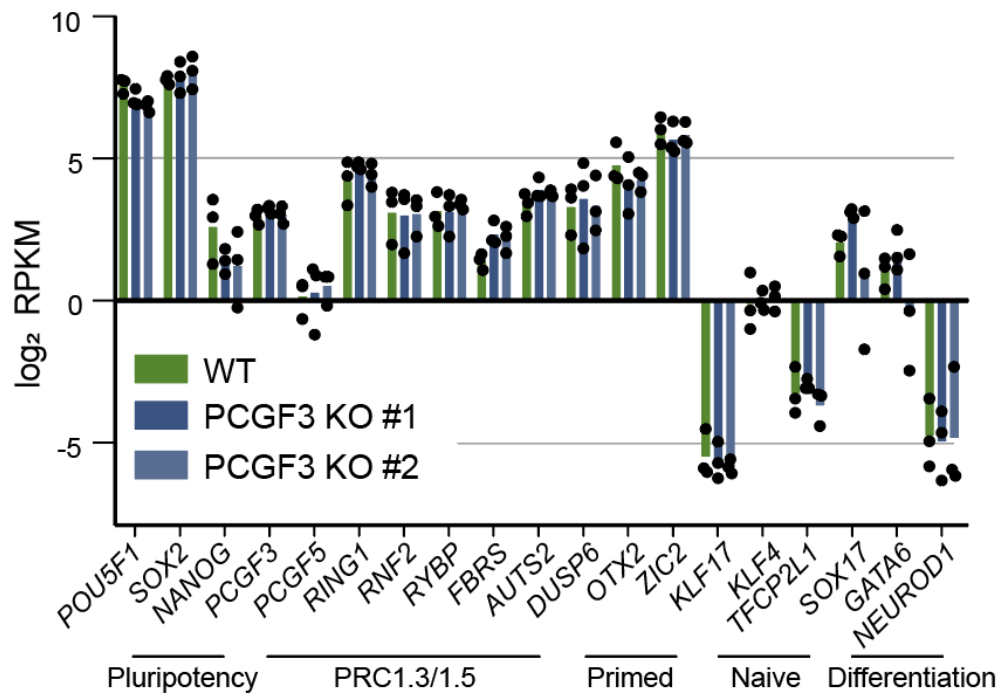


Figure 3.2.16 – Chart shows gene expression similarity between WT and *PCGF3* KO primed PSCs

Genes analysed include members of the core pluripotency network (*OCT4/POU5F1*, *SOX2* and *NANOG*) (Boyer *et al.*, 2005), components of PRC1.3/5 complexes (*PCGF3*, *PCGF5*, *RING1/RING1A*, *RNF2/RING1B*, *RYBP*, *FBRS* and *AUTS2*) (Gao *et al.*, 2012; Hauri *et al.*, 2016), primed pluripotency markers (*DUSP6*, *OTX2* and *ZIC2*) (Bernardo *et al.*, 2018; Buecker *et al.*, 2014) and naïve pluripotency markers (*KLF17*, *KLF4* and *TFCEP2L1*) (Lea *et al.*, 2021; Wang *et al.*, 2019; Weinberger *et al.*, 2016). Differentiation marker genes shown are *SOX17* (Wang *et al.*, 2011), *GATA6* (Heslop *et al.*, 2021) and *NEUROD1* (Choi *et al.*, 2020). Black dots represent individual datapoints. Three independent biological replicates were used for both WT and *PCGF3* KO primed PSCs.

RNA-sequencing data therefore indicates that there is no substantial change in gene expression programmes in the absence of PRC1.3, consistent with the normal growth and pluripotency characteristics of these cells.

3.3 – The PRC1.3 is complex is essential for successful human PTNR

3.3.1 – PRC1.3-deficient primed PSCs fail to produce KLF17-expressing colonies during 5i/L/A reprogramming

The genome-wide screen predicted that PRC1.3 is important for acquisition of naïve pluripotency by reprogramming. In order to test this and to analyse the phenotype of *PCGF3* KO during reprogramming from primed PSCs to the naïve state, the parental WT line and the two *PCGF3* KO lines were subjected to 5i/L/A reprogramming (Theunissen *et al.*, 2014), which was the same reprogramming condition used for the prior screen.

Early in reprogramming, there were no morphological differences between the WT and *PCGF3* KO lines. However, after 10 days of reprogramming in WT cells, human naïve PSC colonies can be seen, while the *PCGF3* KO cells fail to produce naïve PSC colonies and instead produce colonies that resemble neither naïve nor primed PSCs and instead appear differentiated (Fig. 3.3.1).

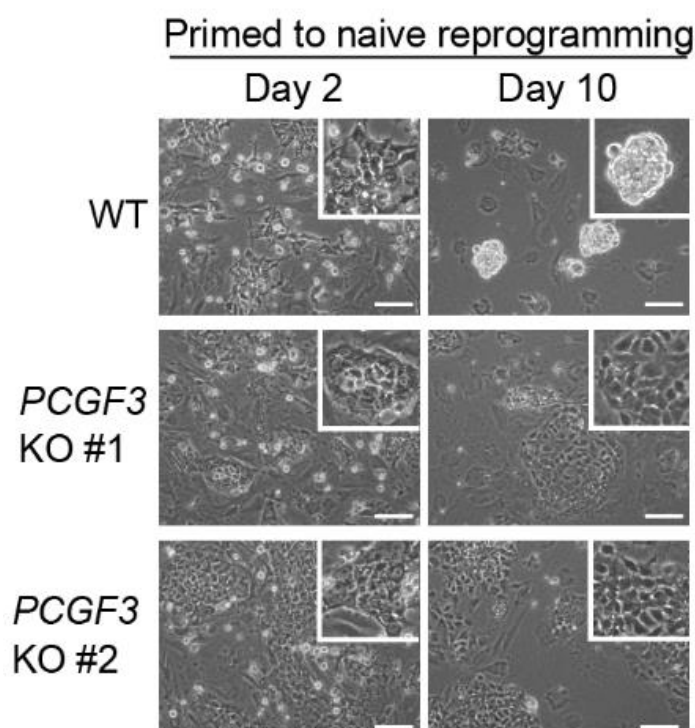


Figure 3.3.1 – Phase contrast images of parental WT and *PCGF3* KO PSCs at day 2 and day 10 of PTNR in 5i/L/A conditions

Images are representative of three independent experiments. On Day 10, the bright, domed and tightly compacted colonies are naïve PSCs, while flattened, more dispersed and spikier colonies instead

resemble cells that have exited pluripotency. Insets show a zoomed in view of colony morphology at high resolution. Scale bars, 100 μ m.

In order to rigorously assess the pluripotency dynamics of cells beyond their visible morphology, I performed immunofluorescence microscopy on both the WT and *PCGF3* KO cell populations after 10 days of 5i/L/A reprogramming, analysing the core pluripotency transcription factors OCT4 and NANOG, and the naïve-specific pluripotency marker KLF17 (Lea *et al.*, 2021; Wang *et al.*, 2022). KLF17-positive colonies were abundantly identified in the WT Day 10 populations, with these colonies also co-expressing OCT4 and NANOG, indicative of nascent naïve PSC colonies (KLF17+ OCT4+). Neither *PCGF3* KO clone day 10 culture contained detectable KLF17 positive colonies. Amongst the KLF17-negative cells, a number of these cells retain expression of OCT4 (KLF17– OCT4+), and I speculate that these cells retain some properties of primed pluripotency. Regardless of genotype, some cells also produced colonies that were negative for OCT4, NANOG and KLF17 (KLF17– OCT4–), indicating that these cells have exited pluripotency and undergone differentiation, as is known to occur in a population of non-reprogrammed cells during PTNR (Collier *et al.*, 2017) (Fig. 3.3.2).

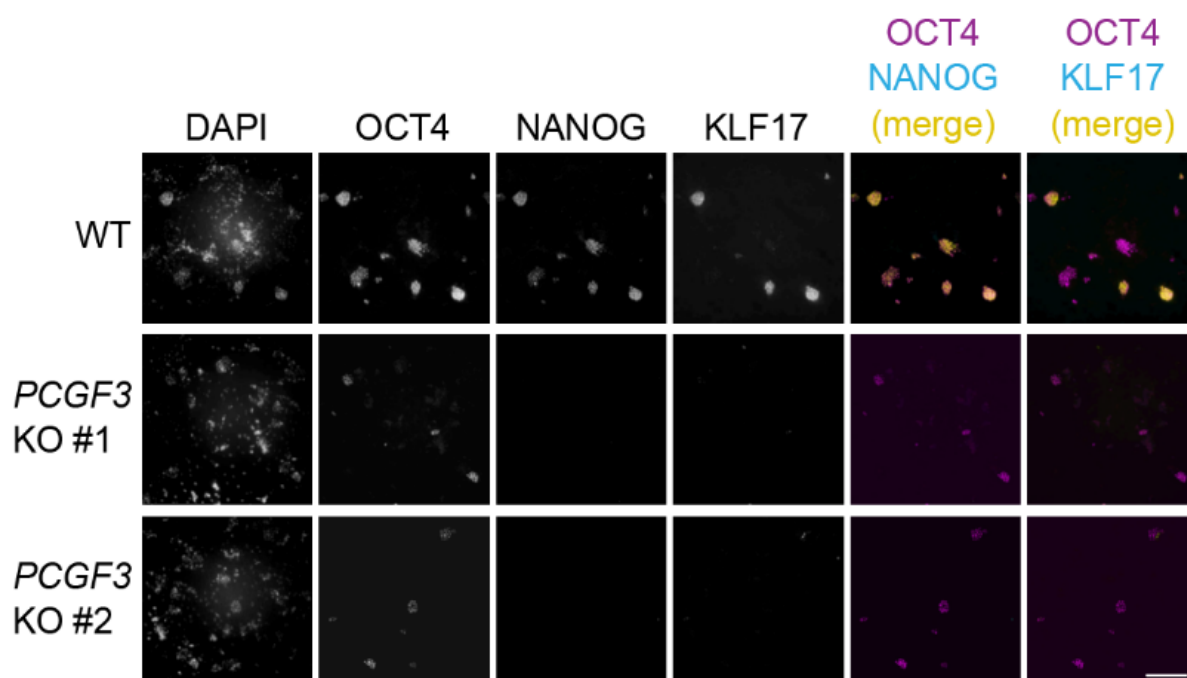


Figure 3.3.2 – Immunofluorescent microscopy of WT and *PCGF3* KO cells at day 10 of PTNR (5i/L/A conditions)

OCT4 and NANOG are pan-pluripotency markers, expressed in naïve and primed PSCs; KLF17 is a naïve-specific marker. Scale bar, 400 μ m.

I quantified the proportion of nascent naïve (KLF17+ OCT4+), primed (KLF17- OCT4+) and differentiated (KLF17- OCT4-) cells detected, with no KLF17+ OCT4- cells detected in any conditions. WT cells have an increased proportion of KLF17 and OCT4 positive colonies compared to *PCGF3* KO cells, which have more KLF17- OCT4- differentiated colonies. Notably in WT conditions, most OCT4+ KLF17- cells also co-express the pluripotency marker NANOG, while in *PCGF3* KO cells, almost no NANOG expression is detectable, even in OCT4 positive colonies. The identity of the OCT4+ NANOG- cell population is unclear but likely identifies a transient population, as NANOG downregulation results in cellular differentiation (Hyslop *et al.*, 2005) (Fig. 3.3.3).

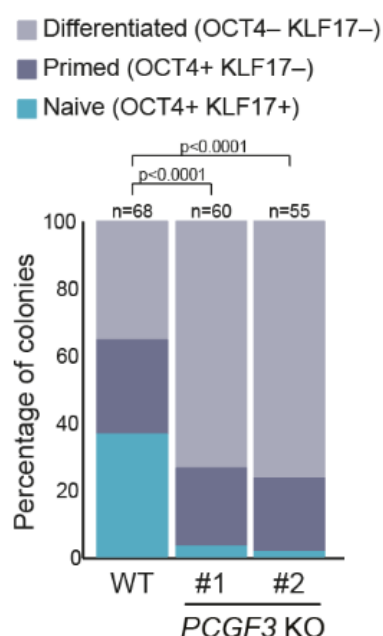


Figure 3.3.3 – Cell colony classification from Fig. 3.3.2 based on OCT4 and KLF17 expression

The number (n) of colonies counted in each category is shown above each bar. Colonies are graded as either naïve (OCT4+ KLF17+), primed (OCT4+ KLF17-) or differentiated (OCT4- KLF17-) in WT and both *PCGF3* KO clones. P-values were calculated using a Chi-Square Test (2 degrees of freedom).

Altogether, these results show that while WT primed PSCs are able to form nascent KLF17-expressing colonies with naïve PSC morphology, PRC1.3 KO primed PSCs fail to emulate

this, instead producing a mixed cell population, with a far greater proportion of cells not expressing pluripotency markers, that neither resemble naïve nor primed PSCs, and have likely exited pluripotency.

3.3.2 – PRC1.3-deficient primed PSCs fail to produce upregulate naïve-specific cell-surface markers during 5i/L/A reprogramming

To further assess the inability of *PCGF3* KO primed PSCs to reprogramme, I used flow cytometry to measure the expression of pluripotency state-specific cell-surface markers on live, single cells. I examined CD75 and SUSD2 as markers of human naïve pluripotency, and CD24 as a marker of primed PSCs (Brendenkamp et al., 2019a; Collier et al., 2017; Shakiba et al., 2015; Wojdyla et al., 2020), using a gating strategy that can distinguish naïve and primed PSCs, and stringently identify nascent naïve PSCs during reprogramming, using fewer cell-surface markers than the panel used in the genome-wide screen (Collier et al., 2022; Rugg-Gunn, 2022a) (Fig. 3.3.4).

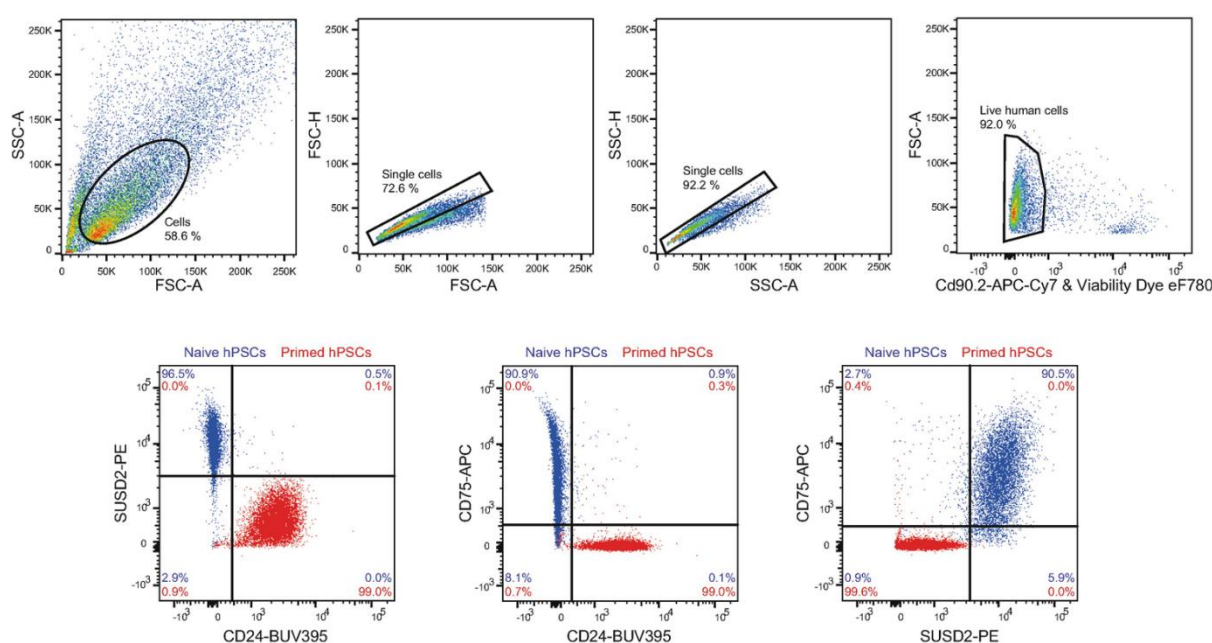


Figure 3.3.4 – Flow cytometry gating strategy to identify nascent naïve cell populations

The first three gates identify single cells. The fourth gate identifies live, human cells; as mouse ‘feeder’ embryonic fibroblasts, positive for Cd90.2 and dead cells, positive for the viability dye, are removed. The remaining gates apply a stringent strategy to identify nascent naïve PSCs, either SUSD2+ CD24-,

CD75+ CD24- or CD75+ SUSD2+. Naïve PSCs are labelled in blue and primed PSCs are labelled in red. Adapted from (Rugg-Gunn, 2022a).

Flow cytometry results using this panel in WT cells shows a population of cells considerably downregulates CD24, indicating a loss of primed PSC cell fate, and upregulates the naïve-specific markers CD75 and SUSD2 by day 14 of PTNR. In *PCGF3* KO cells, while a population also downregulates CD24, almost no cells upregulate CD75 and SUSD2. As CD24, CD75 and SUSD2 are assessed in parallel rather than hierarchically, two different gates are shown to comprehensively assess the cell-surface marker profiles of the cell populations. This approach confirms a severe defect in nascent naïve PSC formation in *PCGF3* KO cells (Fig. 3.3.5A, Fig. 3.3.5B and Fig. 3.3.5C).

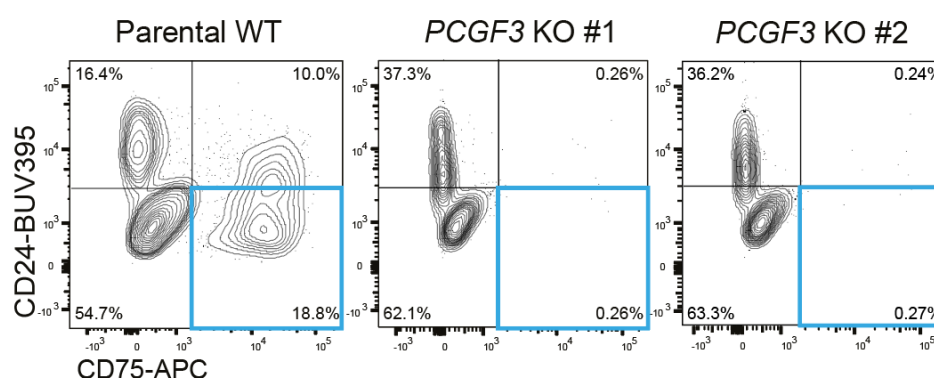


Figure 3.3.5A – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 14 of PTNR in 5i/L/A conditions

Cells were labelled with antibodies against the naïve marker CD75 and the primed marker CD24. Successfully reprogrammed naïve cells are shown in the lower right quadrant (CD75+ CD24-). Data are representative of three independent experiments.

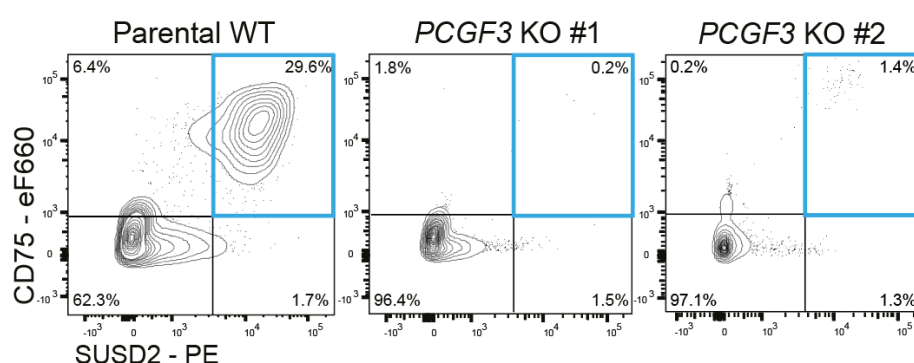


Figure 3.3.5B – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 14 of PTNR in 5i/L/A conditions

Cells were labelled with antibodies against the naïve markers, CD75 and SUSD2. Successfully reprogrammed naïve cells are shown in the upper right quadrant (CD75⁺ SUSD2⁺). Data are representative of three independent experiments.

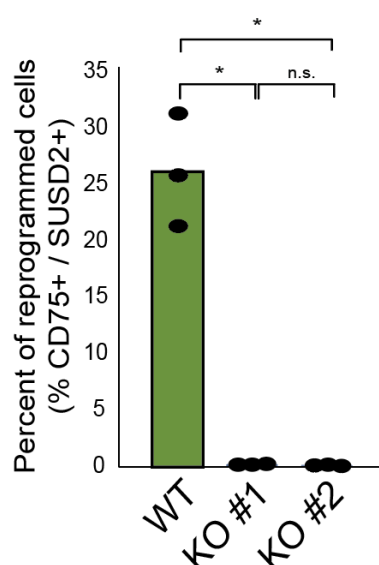


Figure 3.3.5C – Chart showing the quantification of flow cytometry data in Fig. 3.3.5B

Pairwise comparisons between samples analysed using a one-way ANOVA with Tukey correction. Three independent experiments, *, $p < 0.0001$.

Although nascent naïve PSCs are not detected in *PCGF3* KO cells after 14 days of reprogramming, I speculated that *PCGF3* KO cells may produce naïve PSCs at a later timepoint, as the phenotype identified in the day 10 genome-wide screen could represent a delay, rather than a fundamental defect, in human PTNR. To address this possibility, I also performed flow cytometry on WT and *PCGF3* KO cells after 24 days of continuous reprogramming.

This extended reprogramming time-course further increased the proportion of nascent naïve PSCs in WT cells. Although the *PCGF3* KO cells strongly downregulate CD24, they fail to upregulate CD75 and SUSD2, producing triple negative cells for pluripotency cell-surface markers and failing to reprogramme to the naïve state, even on a delayed timescale. The proportion of *PCGF3* KO cells that survive this late into reprogramming is also substantially lower compared to the parental WT cell line (Fig. 3.3.6A and Fig. 3.3.6B).

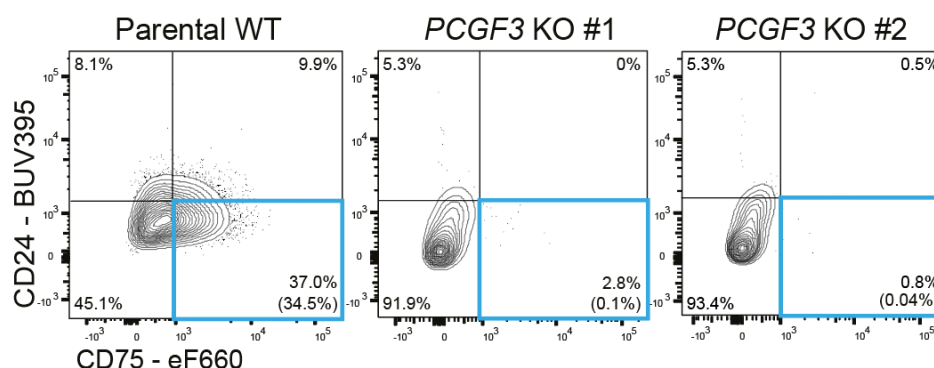


Figure 3.3.6A – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 24 of PTNR in 5i/L/A conditions

Cells were labelled with antibodies against the naïve marker CD75 and the primed marker CD24. Successfully reprogrammed naïve cells are shown in the lower right quadrant (CD75⁺ CD24⁻). The numbers in each quadrant show the percentage of cells out of the live human cells in the sample; the numbers in parentheses show the percentage of cells out of all single cells in the sample (including dead cells and MEFs). Data are representative of two independent experiments.

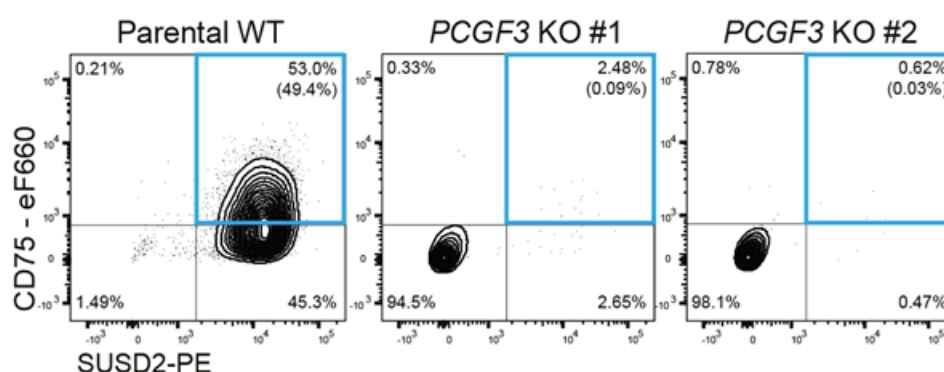


Figure 3.3.6B – Flow cytometry contour plots of parental WT and *PCGF3* KO PSCs on day 24 of PTNR in 5i/L/A conditions

Cells were labelled with antibodies against the naïve markers, CD75 and SUSD2. Successfully reprogrammed naïve cells are shown in the upper right quadrant (CD75⁺ SUSD2⁺). The numbers in each quadrant show the percentage of cells out of the live human cells in the sample; the numbers in parentheses show the percentage of cells out of all single cells in the sample (including dead cells and MEFs). Data are representative of two independent experiments.

The observed viability defect in *PCGF3* KO cells might represent the inability of any reprogrammed cells to proliferate and survive, while nascent naïve PSCs become enriched

under these conditions. Even the *PCGF3* KO cells that do survive until this time-point do not reprogramme.

In summary, these results show that by 5i/L/A based reprogramming, PRC1.3 is indeed essential for PTNR, as *PCGF3* KO primed PSCs show are unable to reprogramme, even across an expanded reprogramming time-course. *PCGF3* KO cells instead appear to undergo cellular differentiation, consistent with findings in section 3.2.1 and then ultimately die over an extended duration.

3.3.3 – PRC1.3 is also essential for PTNR by chemical resetting

I have confirmed that *PCGF3* is essential for human PTNR in 5i/L/A conditions, as predicted by the prior screen. I sought to corroborate this finding using an alternative reprogramming method known as chemical resetting (CR) (Guo et al., 2017; Rugg-Gunn, 2022b).

Mindful that the inability of the two *PCGF3* KO primed clones to reprogramme could be a clone-specific defect when compared to the parental WT population, I also decided to rescue *PCGF3* expression in one of the two KO clones. *PCGF3* cDNA was generated and cloned into an expression vector, which was transfected into *PCGF3* KO primed PSCs. Western blotting confirmed the expression of PCGF3 protein and a clone was selected with as close to WT levels of PCGF3 protein as possible, albeit it is still a PCGF3 overexpression (Fig. 3.3.11).

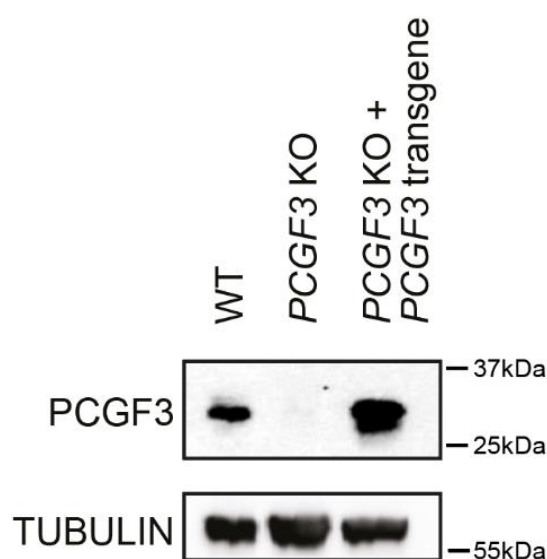
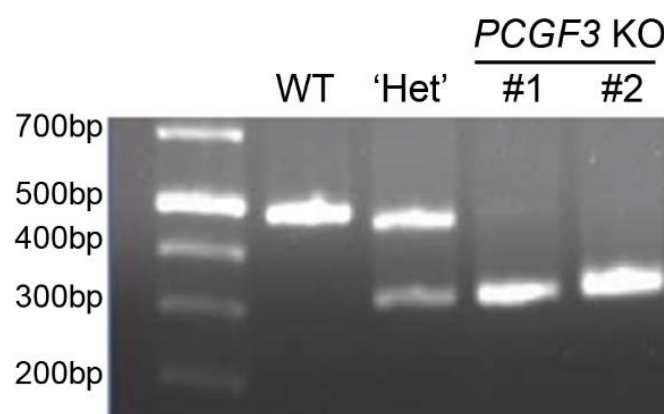


Figure 3.3.11 – Western blotting shows the production of a PCGF3 rescue clone

PCGF3 has molecular weight of 28 kDa; TUBULIN, loading control, has a molecular weight of 55 kDa.

As an additional control cell line, I took forward a clonal *PCGF3* heterozygous cell line, previously generated during the CRISPR-Cas9 targeting, in which *PCGF3* was deleted from a single allele (Fig. 3.3.12).

**Figure 3.3.12 – Electrophoresis image of PCR-based genotyping of the *PCGF3* gene**

WT and *PCGF3* KO primed PSCs show expected band sizes of 471 bp and 308 bp respectively. The presence of both bands indicates the presence of a heterozygous clone for *PCGF3*. The size of the DNA fragments in the ladder are shown in base pairs (bp).

The heterozygote and rescue clones, alongside both *PCGF3* KO lines and the parental WT cells were all subject to chemical resetting. Visualisation of the reprogramming cultures confirmed the presence of nascent naïve PSCs in WT, *PCGF3* heterozygote (Het) and *PCGF3* rescue clones, and not in the *PCGF3* KO cells (Fig. 3.3.13).

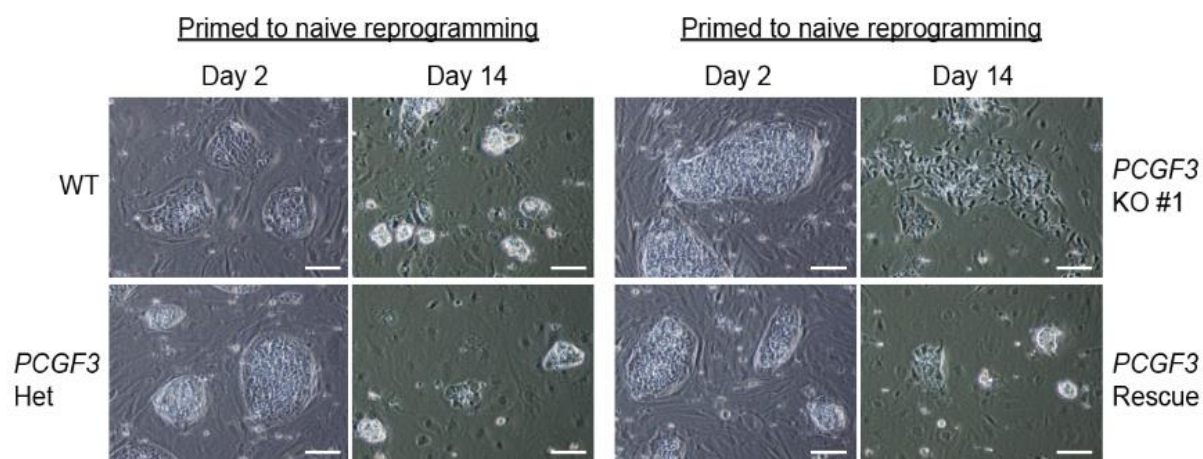


Figure 3.3.13 – Phase contrast images of parental WT, *PCGF3* KO, *PCGF3* Het and *PCGF3* rescue PSCs at day 2 and day 14 of PTNR in chemical resetting conditions

Images are representative of two independent experiments. On Day 14, the bright, domed shaped and tightly compacted colonies are naïve PSCs, while flattened, more dispersed and spikier colonies instead resemble cells that have exited pluripotency. Scale bars, 100 μ m.

This finding was corroborated by flow cytometry, again using the pluripotency state-specific cell-surface markers to interrogate the reprogramming cell populations. These results confirmed a statistically significant inability of *PCGF3* KO primed PSCs to reprogramme by chemical resetting, as well as by 5i/L/A reprogramming, compared to WT cells. The *PCGF3* heterozygous clone, which expresses *PCGF3* from a single allele and the *PCGF3* rescue clone, re-enables upregulation of CD75 and *SUSD2* and downregulation of CD24, confirming the formation of nascent naïve PSCs. This demonstrates that a failure to reprogramme is not an inherent property of the *PCGF3* KO clones but results from the absence of PCGF3 protein and its downstream effects, which are yet to be determined (Fig. 3.3.14A, Fig. 3.3.14B and Fig. 3.3.14C).

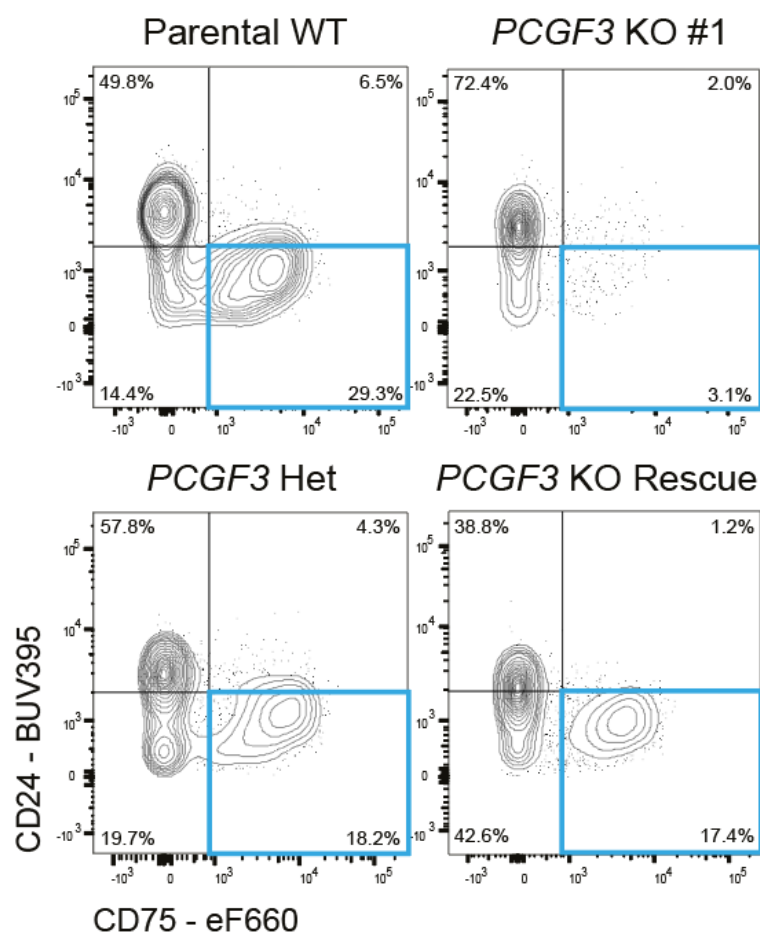


Figure 3.3.14A – Flow cytometry contour plots of parental WT, *PCGF3* heterozygous (Het), *PCGF3* KO and *PCGF3* rescue PSCs at day 10 of PTNR in chemical resetting (CR) conditions
Cells were labelled with antibodies against the naïve marker CD75 and the primed marker CD24. Successfully reprogrammed naïve cells are shown in the lower right quadrant (CD75⁺ CD24⁻). Data are representative of two KO lines and two independent experiments.

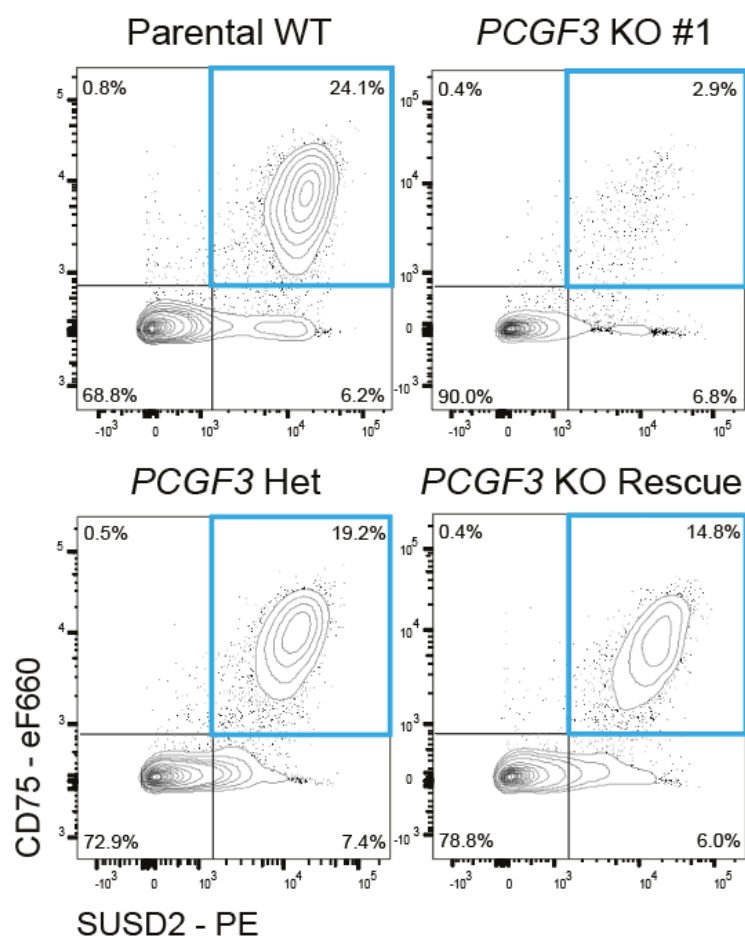


Figure 3.3.14B – Flow cytometry contour plots of parental WT, *PCGF3* heterozygous (Het), *PCGF3* KO and *PCGF3* rescue PSCs at day 10 of PTNR in CR conditions

Cells were labelled with antibodies against the naïve markers, CD75 and SUS2. Successfully reprogrammed naïve cells are shown in the upper right quadrant (CD75⁺ SUS2⁺). Data are representative of two KO lines and two independent experiments.

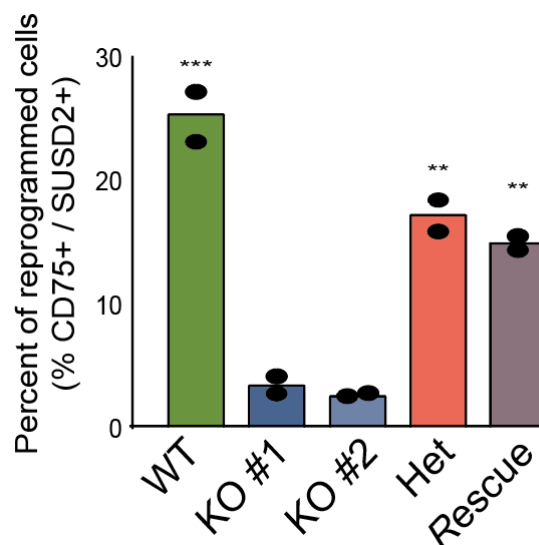


Figure 3.3.14C – Chart showing the quantification of flow cytometry data shown in Fig. 3.4.14B
 Samples are each compared to each of the *PCGF3* KO cell lines using a one-way ANOVA with Tukey correction. **, $p < 0.01$; ***, $p = 0.0005$. Two independent experiments.

This approach acts as a strong methodology to confirm the necessity of *PCGF3* in the PTNR process. Although the *PCGF3* rescue clone overexpresses PCGF3 protein compared to WT cells, the rescue clone does not reprogramme with a higher efficiency than WT primed PSCs. This preliminary finding suggests that there is no direct correlation between PCGF3 protein levels and the efficiency of naïve cell reprogramming, perhaps indicating a threshold level of PCGF3 protein is instead required for reprogramming, but more characterisation is needed.

These experiments establish that PRC1.3 is also essential for PTNR in chemical resetting conditions.

3.4 – Confirming that PRC1.3 is essential for PTNR using *FBRS* KO primed human PSCs

My experiments so far show that *PCGF3* is essential for human PTNR and that *PCGF3* deletion results in the complete dissociation of the PRC1.3 complex in human PSCs. To ensure that the essential role for *PCGF3* is acting within the PRC1.3 complex, and not simply *PCGF3* acting outside of PRC1.3 (although this has never been reported in the literature), I generated a second genetic deletion line of a PRC1.3 component gene. The screen predicted that *FBRS*, which is present within PRC1.3 and PRC1.5 complexes, is amongst the most highly-ranked genes for reprogramming (#10 factor). This is an interesting finding as *FBRS* function is woefully understudied. The PRC1.3 complex auxiliary component genes *AUTS2* and *FBRS* are paralogues, arising from a whole genome duplication event (Kondrychyn *et al.*, 2017; Sellers *et al.*, 2020; Singh *et al.*, 2015), but interestingly only *FBRS* is detected as a key regulator of human PTNR, while *AUTS2* is not (#14,751 factor).

I used CRISPR-Cas9 to delete *FBRS* in human primed PSCs, employing a similar CRISPR-Cas9 strategy as was used for the creation of *PCGF3* KO primed PSCs. I designed guide RNAs targeting the critical exons 14 and 15 of the *FBRS* gene for deletion, as these exons are common to all known *FBRS* isoforms and also their deletion would result in a frameshift (Fig. 3.4.1)

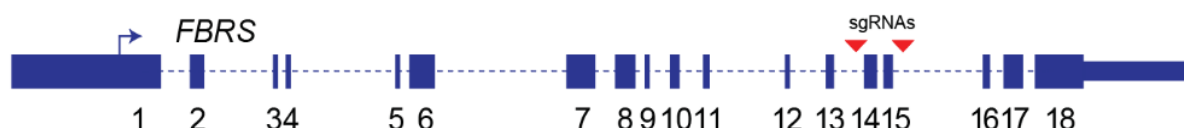


Figure 3.4.1 – Schematic of the CRISPR-Cas9 strategy for the creation of *FBRS* KO primed PSCs
Schematic of *FBRS* exon structure. The red triangles indicate the position of the gRNAs used.

Genotyping identified one single *FBRS* KO clone out of 96 clones screened. Western blotting using an anti-*FBRS* antibody detected a band of around 35 kDa in molecular weight in WT primed PSCs, which is absent from *FBRS* KO primed PSCs. Curiously, the antibody is predicted to detect a peptide of around 70 kDa, while human-specific isoforms for *FBRS* are reported by UniProt databases to produce proteins approximately 50 kDa in length. Given the understudied nature of *FBRS*, more isoforms might exist than have been currently reported, with human primed PSCs expressing a shorter isoform of *FBRS* compared to the antibody-

validating cell lines, as the antibody specifically detects residues 175-225 rather than the entire protein (Fig. 3.4.2).

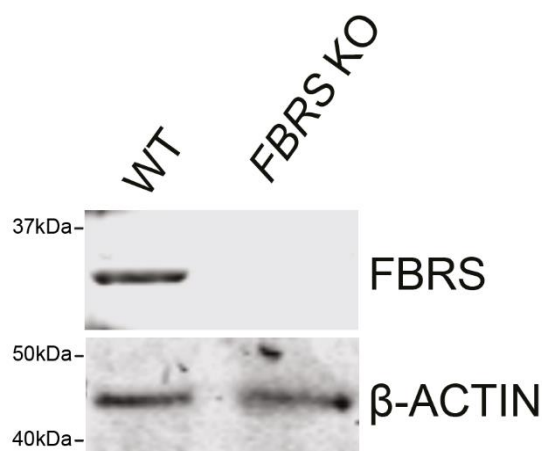


Figure 3.4.2 – Western blotting does not detect FBRS protein in the KO primed PSC line

FBRS has predicted molecular weight of 50 or 70 kDa but shows a band at approximately 35 kDa; β-ACTIN, loading control, has a molecular weight of 42 kDa.

Once established, I also characterised the *FBRS* KO primed PSCs using the cellular characterisation assays, including a visual assessment of cell morphology (Fig. 3.4.3), OCT4/NANOG immunofluorescence microscopy (Fig. 3.4.4), SSEA-4 flow cytometry (Fig. 3.4.5) and karyotyping (Fig. 3.4.6). These experiments showed that *FBRS* KO primed PSCs all possess a WT like morphology, uniform expression of OCT4, NANOG and SSEA-4 and a normal karyotype, identical to WT primed PSCs.

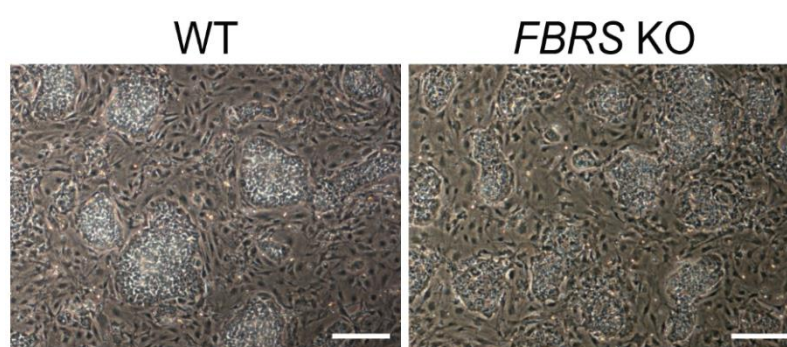


Figure 3.4.3 – Phase contrast images showing the morphology of WT and *FBRS* KO primed PSCs

The WT cells are untargeted parental H9 cells. Scale bars, 250 μm.

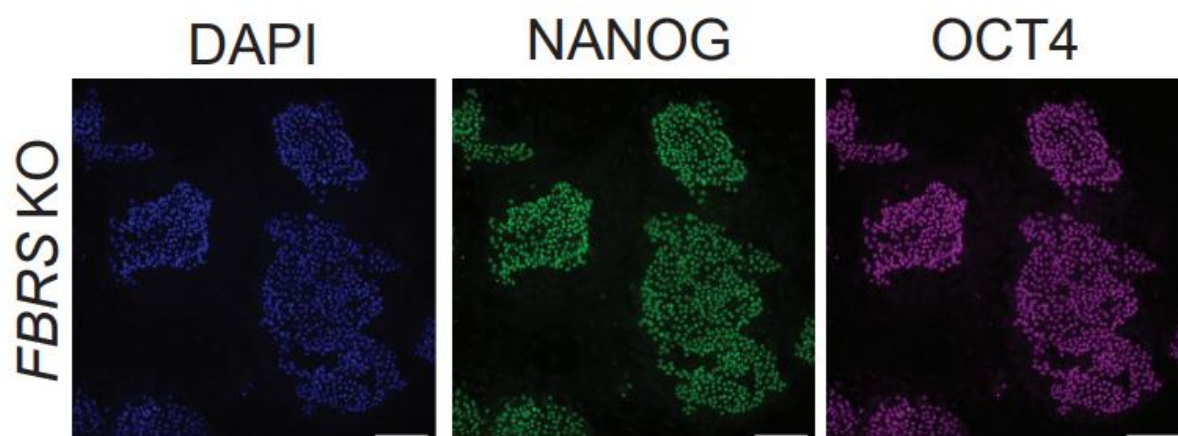


Figure 3.4.4 – Immunofluorescence microscopy shows the uniform expression of OCT4 and NANOG in WT and *FBRS* KO primed PSC lines

OCT4 and NANOG are both core pluripotency factors. Scale bars, 250 μ m.

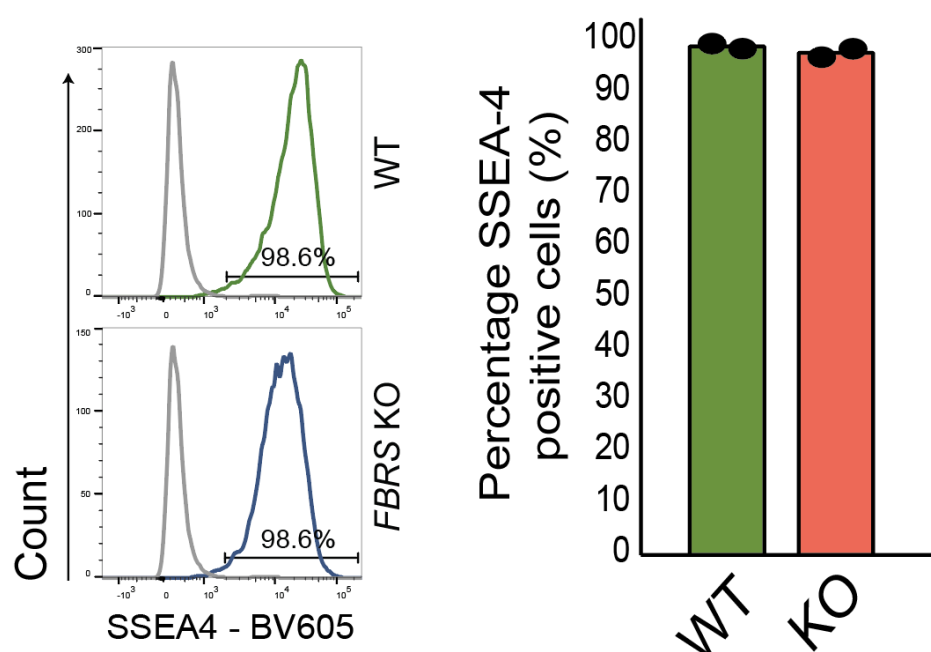


Figure 3.4.5 – SSEA-4 flow cytometry in WT and *FBRS* KO primed PSC lines

Flow cytometry data represented as a histogram, representative of two independent biological replicates. Grey line shows omission of the SSEA-4 antibody. Bar chart quantifies the SSEA-4 positive cells as a percentage. Bar height shows the mean and circles show the individual datapoints.

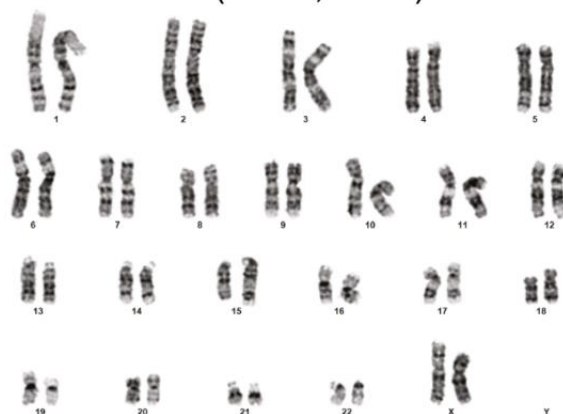
H9 *FBRS* KO (46XX;20/20)

Figure 3.4.6 – Karyotyping of the *FBRS* KO primed PSC line

Representative G-banded chromosome metaphase spreads are shown, as performed by a clinical cytogeneticist at Cell Guidance Systems. Twenty cells were examined.

Collectively, these results suggest that the PRC1.3 component gene *FBRS* is also dispensable for human primed pluripotency. *FBRS* is very poorly understood and has been primarily studied as an immune lymphokine (Prakash *and* Robbins, 1998), and could function outside of the PRC1.3/5 complex. The identification of *FBRS* as dispensable for human primed pluripotency is a novel finding itself and also allows this cell line to act as an appropriate control to test whether the PRC1.3 complex specifically, requiring both *PCGF3* and *FBRS*, is essential for PTNR. This was tested by chemical resetting and *FBRS* KO primed PSCs also showed a strong visual impairment in their ability to undergo PTNR (Fig. 3.4.7).

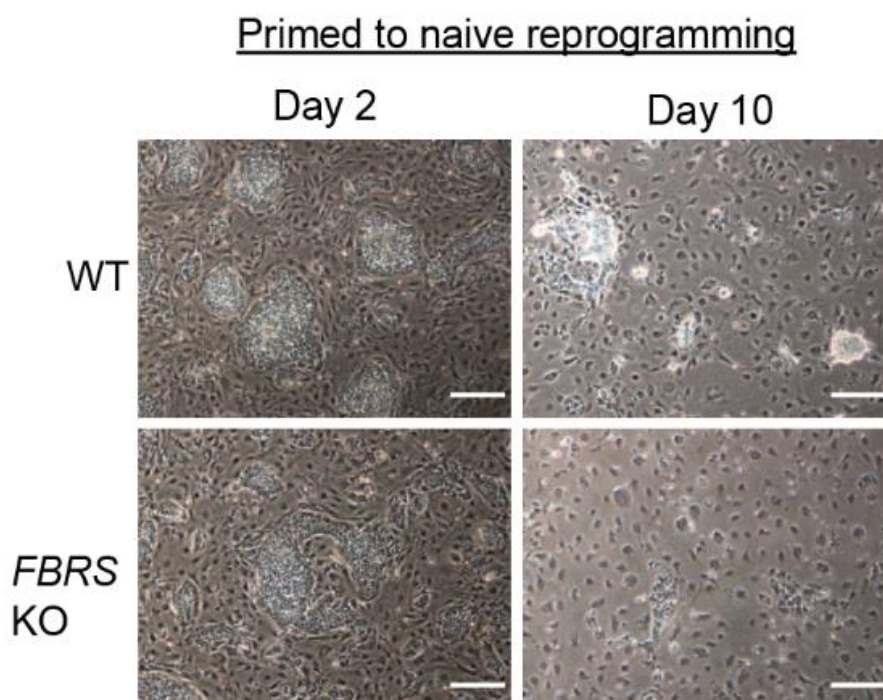


Figure 3.4.7 – Phase contrast images of parental WT and *FBRS* KO PSCs at day 2 and day 10 of PTNR in CR conditions

Images are representative of three independent experiments. On Day 10, the bright, domed shaped and tightly compacted colonies are naïve PSCs, while flattened, more dispersed and spikier colonies instead resemble cells that have exited pluripotency. Scale bars, 100 μ m.

I corroborated this visual finding, again using the naïve and primed PSC discriminating antibody panel to analyse the reprogramming cell populations using flow cytometry (Fig. 3.4.8A, Fig. 3.4.8B and Fig. 3.4.8C), which showed that *FBRS*, like *PCGF3*, is essential for human PTNR.

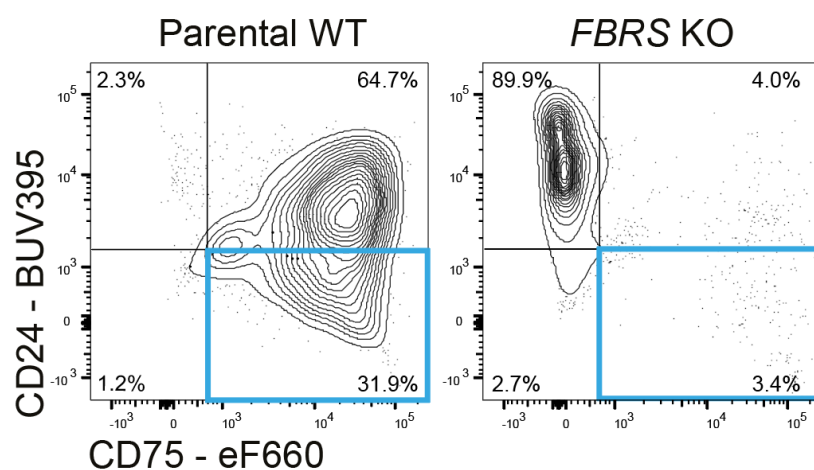


Figure 3.4.8A – Flow cytometry contour plots of parental WT and *FBRS* KO PSCs on day 12 of PTNR in CR conditions

Cells were labelled with antibodies against the naïve marker CD75 and the primed marker CD24. Successfully reprogrammed naïve cells are shown in the lower right quadrant (CD75+ CD24-). Data are representative of one KO line and two independent experiments.

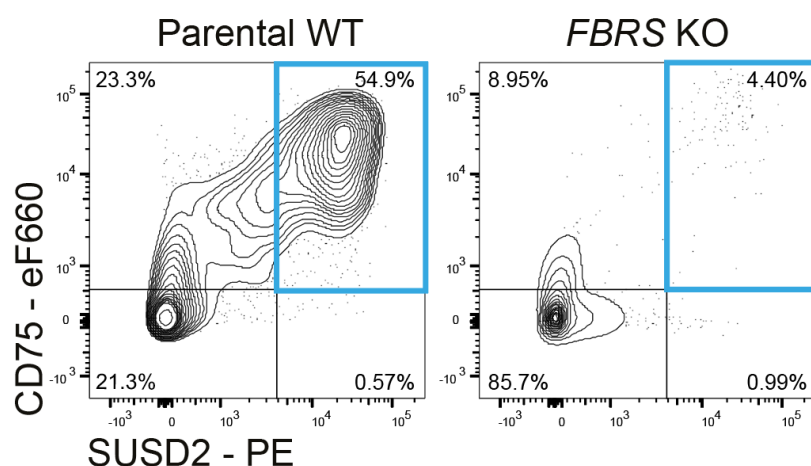


Figure 3.4.8B – Flow cytometry contour plots of parental WT and *FBRS* KO PSCs on day 12 of PTNR in CR conditions

Cells were labelled with antibodies against the naïve markers, CD75 and SUSP2. Successfully reprogrammed naïve cells are shown in the upper right quadrant (CD75+ SUSP2+). Data are representative of one KO line and two independent experiments.

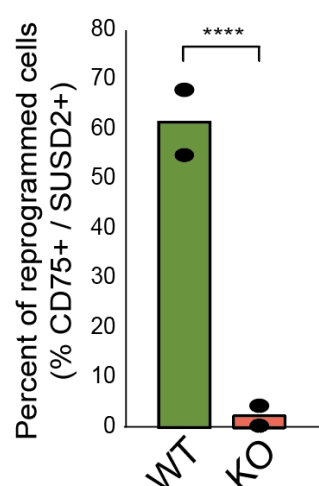


Figure 3.4.8C – Chart showing the quantification of flow cytometry data shown in Fig. 3.4.8B

Samples are each compared to each of the *PCGF3* KO cell lines using a one-way ANOVA. ****, $p < 0.05$. Two independent experiments.

Although the percentage of *FBRS* KO naïve marker positive cells is higher than in the previous CR experiments, this also reflects a higher efficiency of reprogramming in the WT control cells. Compared to WT cells, *FBRS* KO primed PSCs still demonstrate a severe reprogramming defect, and the differing efficiencies could simply reflect variation, which is dependent upon a large number of technical factors.

In summary, I have shown that both the *PCGF3* and *FBRS* genes, as components of the PRC1.3 complex, are essential for PTNR, despite being dispensable for the assessed metrics of human primed pluripotency.

3.5 – Generating PRC1.3 knockdown naïve PSCs

The PRC1.3 complex is critical for reprogramming primed PSCs to the naïve state. However, I do not know whether the complex is critical simply for acquisition of the naïve state of pluripotency through reprogramming or if it is also required for its maintenance. Findings in human PSCs could inform on the function of PRC1.3 in the developmental establishment of naïve pluripotency in the human pre-implantation epiblast, which might be important also to sustain this transient period of embryonic pluripotency. In order to specifically study the role of PRC1.3 in human naïve PSCs in the latter, which is not possible using my established knockout lines due to their inherent inability to reprogramme, I utilised an inducible CRISPR interference (CRISPRi) system to knockdown the expression of *PCGF3* in human naïve PSCs. CRISPRi represses transcription of a gene of interest by targeting a transcriptional repressor to the transcription start site (TSS). The CRISPRi system uses specific TSS-targeting guide RNAs to recruit a catalytically dead CAS9 (dCas9) to this region, which in turn is fused to the KRAB (Krüppel associated box) transcriptional repressor protein (Gilbert *et al.*, 2013).

A human primed iPSC cell line has been established (Mandegar *et al.*, 2016), in which the dCas9-KRAB construct, along with an mCherry reporter, is expressed in the presence of Doxycycline as a Tet-On system (T. Das *et al.*, 2016). *PCGF3* knockdown can therefore be produced by the introduction of constitutively expressed guide RNAs targeting the *PCGF3* gene TSS and induction of dCas9-KRAB by the addition of Doxycycline (Fig. 3.5.1), and these primed PSCs can be converted to naïve PSCs by reprogramming.

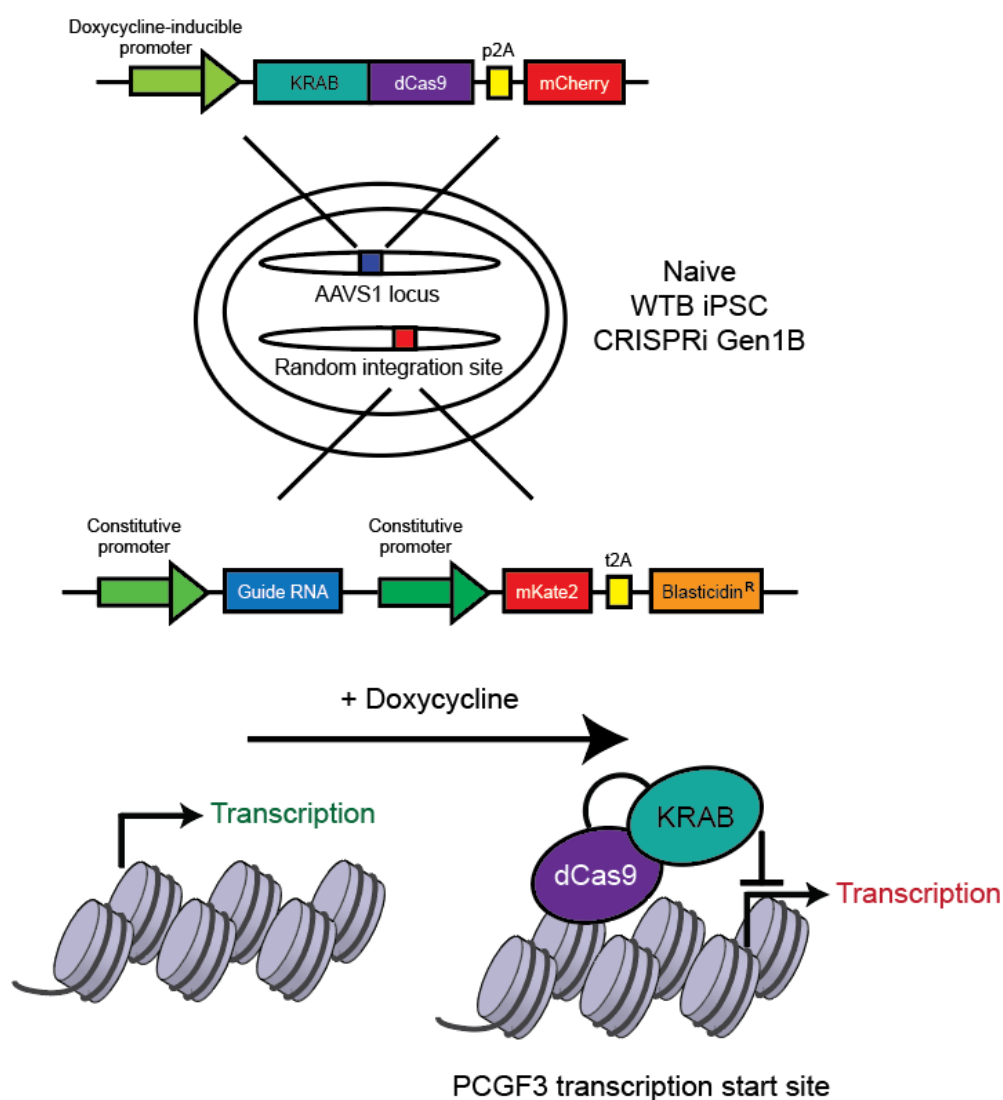


Figure 3.5.1 – Schematic of CRISPRi-mediated inducible *PCGF3* knockdown

Guide RNAs that can target dCas9-KRAB to the *PCGF3* TSS are cloned into a vector with a fluorescent mKate2 reporter and a blasticidin resistance gene for selection. This vector is then transfected into primed PSCs that inducibly express dCas9-KRAB, along with an mCherry reporter, from the human safe harbour site AAVS1 when Doxycycline is applied to cells.

Prior to transfection, I confirmed the fidelity of the CRISPRi primed PSC line by assessing the ability of the dCas9-KRAB-mCherry construct to be induced by Doxycycline, using a starting concentration of 1 μ M. When doxycycline is not present, no cells express any detectable mCherry reporter, showing that the dCas9-KRAB-mCherry construct is not induced in the absence of Doxycycline. 1 μ M Doxycycline is sufficient to activate the mCherry construct in

close to 100 % of cells, showing that induction is optimal at this Doxycycline concentration (Fig. 3.5.2).

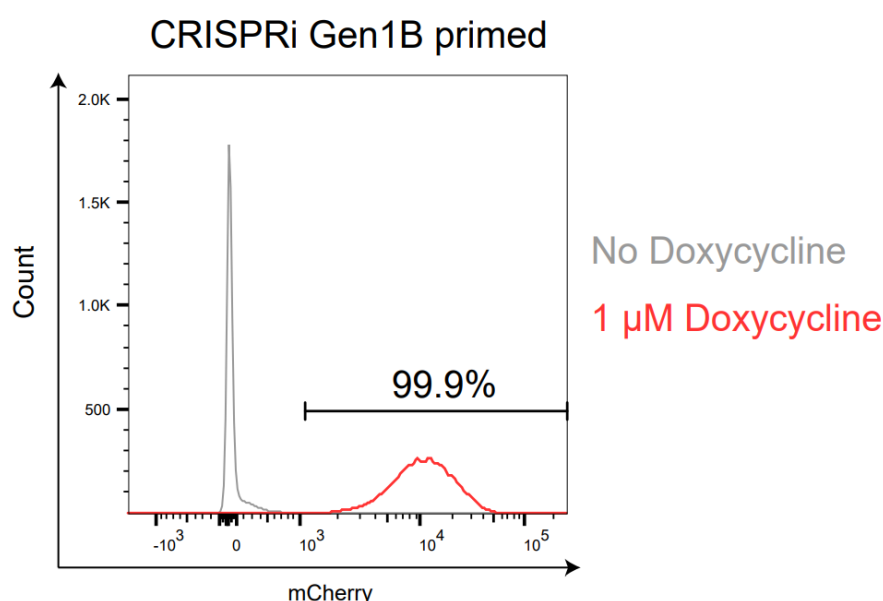


Figure 3.5.2 – mCherry flow cytometry in Doxycycline treated and untreated CRISPRi primed iPSCs

Flow cytometry data represented as a histogram from a single experiment. Grey line shows the omission of Doxycycline while the red line represents the addition of 1 μ M Doxycycline.

Two independent *PCGF3* guide RNAs were selected based on their proximity to the *PCGF3* transcription start site – named gRNA 16 and 17 (Gilbert *et al.*, 2014) (Fig. 3.5.3).

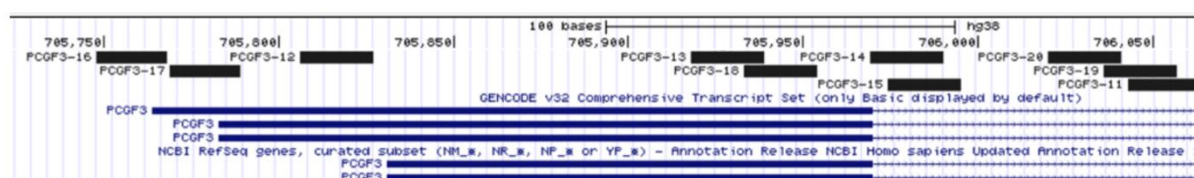


Figure 3.5.3 – Mapping of the guide RNAs predicted to successfully knockdown *PCGF3* expression

Black boxes indicate the location of predicted *PCGF3* guide RNAs for CRISPRi. Blue boxes indicate the position of transcribed *PCGF3* gene sequences. Guide RNAs 16 and 17 show the closest proximity to the reported *PCGF3* TSS in the human genome (Gilbert *et al.*, 2014)

Guide RNAs 16 and 17 were separately cloned into the pgRNA-CKB vector and confirmed by Sanger sequencing and then transfected into human primed PSCs, which are easier to transfect than their naïve counterparts. Transfected cells were visualised 48 h post-transfection, during Blasticidin antibiotic selection. The detection of the mKate2 reporter signal within cells showed that the cells contain the vector, which under antibiotic selection should randomly integrate into the human genome to produce a constitutive guide RNA expressing CRISPRi primed iPSC line. (Fig. 3.5.4).

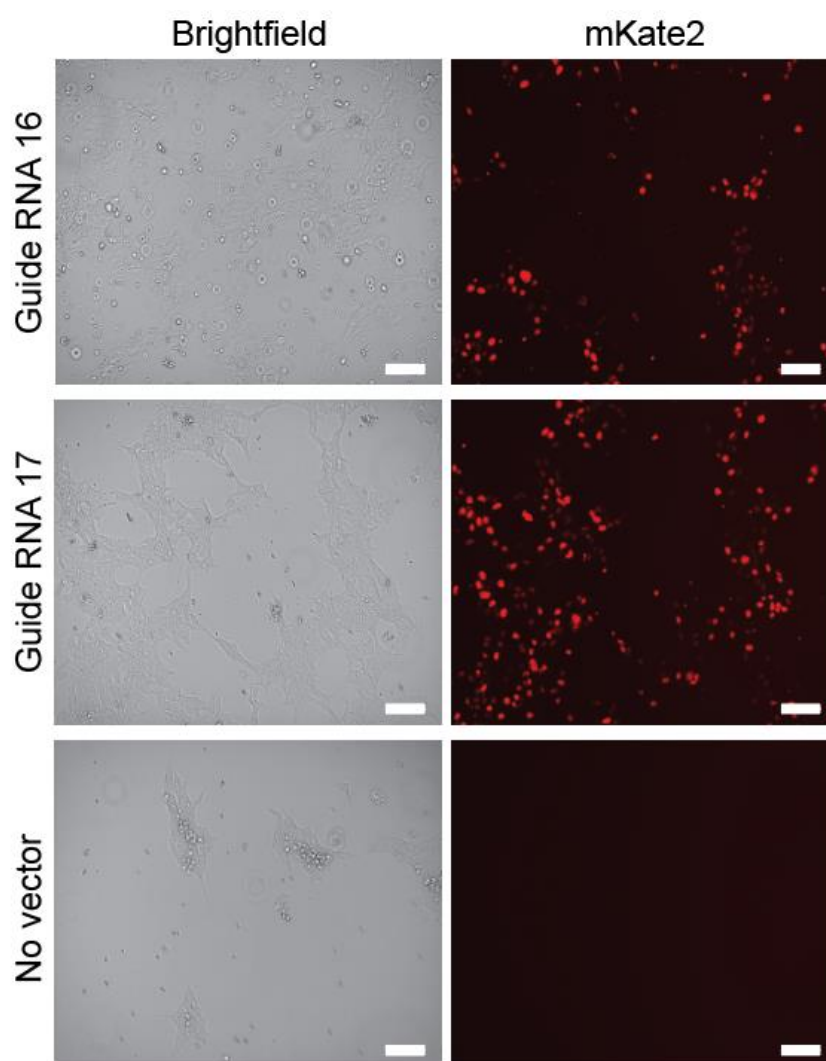


Figure 3.5.4 – Fluorescence microscopy confirms successful vector transfection

Brightfield and mKate2 fluorescence are shown. Scale bars, 100 μm .

As the transfection was performed in primed PSCs, I first tested the ability of primed PSCs to knockdown *PCGF3* using RT-qPCR. Once a primed knockdown line has been established

and confirmed by Western blot, I planned to reprogramme these lines into naïve PSCs. I would then once again test the *PCGF3* knockdown in cells, to ensure that the guide RNA construct has not become silenced, before phenotyping the effect of *PCGF3* knockdown in human naïve PSCs. After antibiotic selection, I added Doxycycline to both primed PSC lines, containing guide RNA 16 and 17 and assessed the *PCGF3* mRNA expression level by RT-qPCR. Unfortunately, the bulk cell populations for both guide RNAs show only minimal knockdowns, and so were not suitable candidates for reprogramming for naïve knockdown experiments, although the extent of knockdown does increase with the length of Doxycycline addition (Fig. 3.5.5).

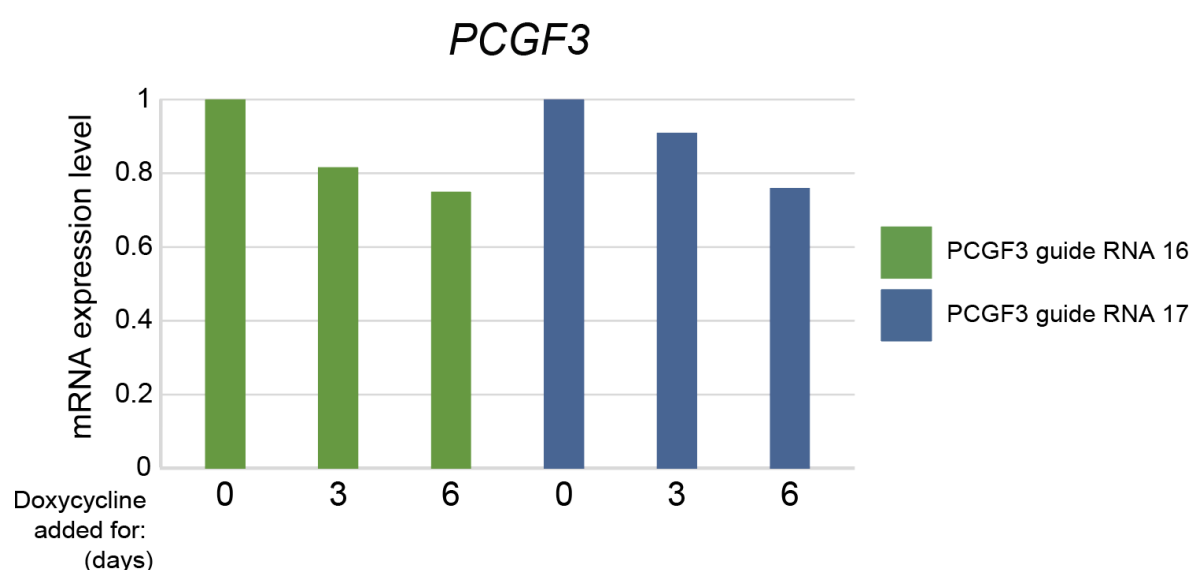


Figure 3.5.5 – RT-qPCR showing *PCGF3* expression across a Doxycycline addition time-course

Data from one experiment for two different *PCGF3* TSS-targeting guide RNAs. Samples were collected before (day 0), after 3 days and 6 days of Doxycycline treatment. Samples were normalised to the geometric mean of the housekeeper genes *HMBS* and *GAPDH*, and adjusted such that doxycycline untreated (Day 0) is equal to 1 for each gRNA.

I speculated two possibilities for this phenotype. Firstly, all of the cells in the population might be knocking down *PCGF3* expression to between 70-80 % of WT levels. Alternatively, between 20-30 % of cells might be strongly knocking down *PCGF3* expression while the remainder produced no knockdown, perhaps requiring a higher concentration or longer duration of antibiotic treatment to remove any non-transfected or guide RNA silenced cells.

In order to test whether a fraction of cells are producing a strong knockdown, I picked eight individual clones from *PCGF3* guide RNA 16 transfected primed CRISPRi cells and screened for *PCGF3* expression levels in these cells by RT-qPCR. If around 25 % of cells had a substantial *PCGF3* knockdown, I would expect that one or two of the eight clones would be successful knockdown lines. Disappointingly, this experiment showed that all of the clones failed to produce a strong *PCGF3* knockdown, with none reducing *PCGF3* to even half of WT expression levels (Fig. 3.5.6).

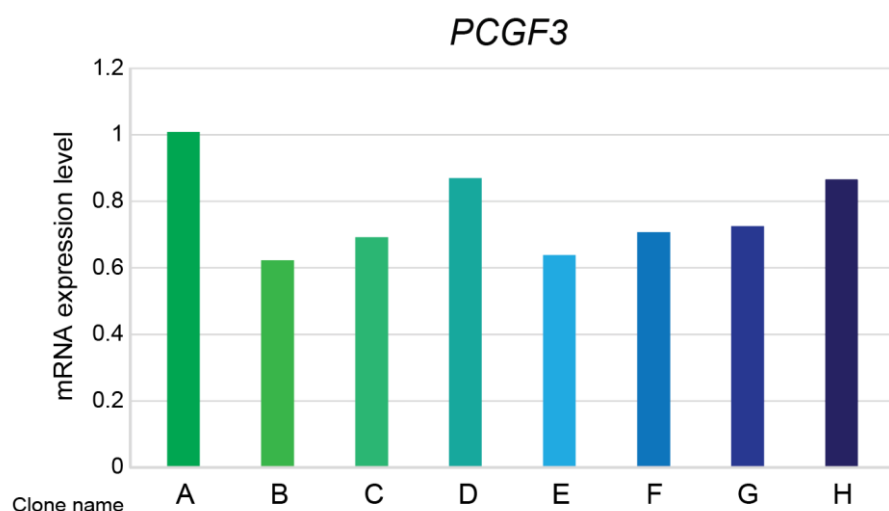


Figure 3.5.6 – RT-qPCR showing *PCGF3* expression across eight CRISPRi clones

Data from one experiment for eight clones transfected with guide RNA 16. Clones were treated with Doxycycline for six days and normalised to the geometric mean of the housekeeper genes *HMBS* and *GAPDH*, and were adjusted such that their corresponding Doxycycline untreated cells are equal to 1 for each clone.

As a result, none of these clones made attractive candidates for further experimentation and so I was unable to generate the ability to knock down or KO *PCGF3* specifically in human naïve PSCs. I have therefore determined that *PCGF3* is essential for human PTNR but it is not known whether *PCGF3* is needed only for the cell state transition, or more broadly for establishment, maintenance and propagation of the human naïve state of pluripotency, and by extension, this pluripotent state in human embryos.

3.6 – Chapter Discussion

3.6.1 – PRC1.3 is dispensable for the analysed characteristics of human primed PSCs

As the selected complex from the genome-wide screen for additional study, PRC1.3 had never been studied within the context of human pluripotency or reprogramming and PRC1.3 was identified surprisingly specifically, rather than PRC1 or PRC2 complexes more broadly.

Within this chapter, I generated individual gene knockouts of the PRC1.3 component genes *PCGF3* and *FBRS*. I confirmed that *PCGF3* deletion causes complete dissociation of the PRC1.3 complex and performed cellular characterisation assays, confirming that the PRC1.3 complex is dispensable for a number of metrics of human primed pluripotency – including morphology, pluripotency cell-surface and transcription factor expression, proliferation, karyotype and global transcriptome. However, the PRC1.3 complex might have roles in primed pluripotency for aspects that have not yet been examined, such as on the epigenome, and this will be studied in Chapter 4.

PCGF3 KO primed PSCs show few transcriptional differences from WT PSCs, suggesting either that PRC1.3 does not regulate gene expression or that any gene regulation exerted by the PRC1.3 complex can be compensated in primed PSCs, consistent with synergy between variant PRC1 complexes seen in mESCs (Fursova *et al.*, 2019; Scelfo *et al.*, 2019). *Pcgf3* and *Pcgf5* combined deletion caused loss of approximately 50 % of cellular H2AK119ub1 levels, but the combined deletion of *Pcgf1*, *Pcgf3*, *Pcgf5* and *Pcgf6* was required for widespread changes to gene expression (Fursova *et al.*, 2019). In another study, only 185 genes were reported to be specifically bound by PCGF3 and PCGF5 in mESCs (Scelfo *et al.*, 2019), and so the lack of characterised impact of *PCGF3* deletion in steady-state primed PSCs could result from only a small number of genes being bound and regulated by PRC1.3 complexes. The use of RNA-sequencing could have been improved by the use of calibration relative to an exogenous spike-in, for example mRNA detected from *Drosophila melanogaster*, in order to normalise expression levels in a quantitative manner between samples. It is possible that small, but important, PRC1.3-dependent transcriptional differences may exist between these samples in a genome-wide manner but have not been detected due to the lack of calibration.

One of the main hallmarks of pluripotent cells is their ability to undergo multi-lineage differentiation to produce cells contributing to the three germ lineages. As I demonstrated the differentiation capacity of PRC1.3-deficient primed PSCs, it is difficult to fully establish that the

primed state of pluripotency is entirely intact. Further investigation is required to ascertain the mechanism underpinning PRC1.3 function, but the lack of profound PRC1.3 KO phenotype under consistent growth conditions is not without precedent.

3.6.2 – The absence of PRC1.3 causes cells to fail to reprogramme to naïve PSCs

I have established that the PRC1.3 complex is required for the formation of nascent naïve PSCs during human PTNR. This has been confirmed using two different reprogramming methods, by analysing genetic deletions of two different PRC1.3 component genes, *PCGF3* and *FBRS*, and has been shown to represent a failure rather than a delay in reprogramming by performing extended duration experiments.

Cells that fail to reprogramme, which make up a substantial proportion of WT cells and all of the cells in *PCGF3* KO cells, have three possibilities for their cell fate. Refractory cells could represent the continued presence of primed PSCs under reprogramming conditions, the cells could exit pluripotency and undergo differentiation or cells could undergo apoptosis and die. Immunofluorescence microscopy on *PCGF3* KO cells during reprogramming indicate the presence of a large number of OCT4- KLF17- or OCT4+ NANOG- KLF17- cells, and live cells are visible that morphologically resemble neither primed nor naïve PSCs. Flow cytometry confirms that a large proportion of cells downregulate the primed-specific marker CD24, without upregulating naïve-specific markers, further indicating that *PCGF3* KO cells predominantly differentiate in early human PTNR. This finding is consistent with reports of differentiation during reprogramming even in WT cells (Collier *et al.*, 2017) alongside the formation of nascent naïve PSCs.

Differentiated refractory cells produced during reprogramming have been shown to resemble neural-like cells (Collier *et al.*, 2017), as well as trophectoderm and primitive endoderm like signatures (Bi *et al.*, 2022). I speculate that this may be due to the formation of these cell types being permitted during reprogramming, when the pluripotency network cannot be sustained. Over a longer duration, the refractory *PCGF3* KO cells largely undergo cell death, while WT cells are able to proliferate and expand the production of naïve PSCs. Cell death could arise due to the selective disadvantage of differentiated refractory cells to survive and proliferate under reprogramming conditions and during single-cell passaging, which is exploited to generate pure populations of naïve PSCs for research.

The molecular role of PRC1.3 in reprogramming is still not known. For example, PRC1.3 could transcriptionally upregulate factors that promote the naïve PSC transcriptional programme or silence genes involved in differentiation, such that a small fraction of cells can successfully reprogramme. Further studies to examine the molecular basis of this cellular phenotype are required, which will be the focus of Chapter 4. The question of why some, and not all, cells are able to reprogramme in any reprogramming system remains an enigma that is poorly understood.

3.6.3 – It is not known whether PRC1.3 specifically regulates PTNR or pluripotency cell state transitions more broadly

Due to the synergy between variant PRC1.3 complexes discussed in section 3.6.1, the specific essential requirement for the PRC1.3 complex in human PTNR is unexpected. In particular, the identification of *FBRS* as essential for reprogramming suggests that the *FBRS* and *AUTS2* paralogues may not be entirely redundant with one another, and this finding will be discussed in more detail in Chapter 5.

Reprogramming of primed PSCs to the naïve state is not a developmentally relevant process but the formation of a transient human pre-implantation epiblast state, which naïve PSCs mimic, is. These two processes do have some molecular parallels, such as genome-wide DNA demethylation to achieve a hypomethylated state. The study of PRC1.3 function in human naïve pluripotency is therefore worthy of study, and could provide a novel understanding of PRC1.3 function in human pre-implantation epiblast specification. As I identify that the PRC1.3 complex is essential for the transition from primed to naïve pluripotency, but not for self-renewal of human primed PSCs, it is possible that the essential role of the PRC1.3 complex is only highlighted when cells transition into an alternative cell state, either by differentiation or reprogramming. Many of the studies that have ascertained important specific functions for the PRC1.3 and PRC1.5 complex have focused on the context of cellular differentiation in mice or mESCs, which have identified PRC1.3/5 complexes as important for neuronal differentiation and central nervous system development (Gao *et al.*, 2014; Wang *et al.*, 2018; Yao *et al.*, 2018), angiogenesis (Kanki *et al.*, 2022) and transcription of mesoderm lineage marker genes during embryoid body differentiation (Zhao *et al.*, 2017).

A limitation of my current work is that I have not examined whether PRC1.3 is needed for differentiation of primed PSCs into specific or multiple lineages. Cell state transitions require a huge degree of transcriptional and epigenetic remodelling to promote lineage-specific

transcriptional pathways and to suppress transcriptional pathways of alternative lineages. Differentiation experiments could help to ascertain whether the PRC1.3 complex is important only for human PTNR, or is more generally required for pluripotency cell state transitions. By differentiating human *PCGF3* KO primed PSCs, I could also develop a novel understanding of the role, if any, of PRC1.3 in the specification of cell identity *in vitro* and during human post-implantation development. Future experiments to address this limitation could include using directed differentiation protocols or a more unbiased, broader analysis of lineage specification using systems such as embryoid body or gastruloid differentiation.

3.6.4 – A naïve PRC1.3 knockdown system could address a number of outstanding questions

Another limitation of my current work is that I have not been able to disentangle whether PRC1.3 is required only for the establishment of naïve pluripotency through reprogramming or is also important for maintenance of naïve PSCs. Naïve PSCs possess a large number of desirable properties, and from the perspective of human development are able to model the first lineage specification decisions in the embryo by the formation of blastoids (Kagawa *et al.*, 2022; Yanagida *et al.*, 2021; Yu *et al.*, 2021), and contain a sub-population of cells that mimic the 8-cell human embryo state (Taubenschmid-Stowers *et al.*, 2022). A role for PRC1.3 within established naïve PSCs could suggest a key role for the complex in human pre-implantation embryo development and the ability to perturb PRC1.3 levels in naïve PSCs could test this using blastoid and 8-cell-like cell models.

I attempted to establish a naïve PRC1.3 knockdown system using CRISPR interference, however the steps that I took were not fruitful in establishing a successful *PCGF3* knockdown. This experiment could have been repeated and optimised by testing a far greater number of *PCGF3*-targeting guide RNAs to maximise the likelihood of knockdown, as well as by combining different guide RNAs, with the hope of boosting dCas9-KRAB recruitment and repression across the *PCGF3* TSS. I could also have tried to deliver *PCGF3* guide RNAs into available WTB iPSC CRISPRi Gen1B naïve PSCs, simplifying the experiment by circumventing the time-consuming, inefficient reprogramming process for each cell line and avoiding potential guide RNA silencing. An inducible knockdown system in naïve PSCs has not been reported previously and this remains a technically challenging process, due to difficulties in transfecting naïve PSCs. The direct generation of naïve PRC1.3 knockdowns would also have minimised the potential power of the inducible system in primed PSCs to study the importance of the complex in reprogramming with temporal resolution. The PRC1.3

knockdown could have been applied, and also reversed, at specific timepoints in early PTNR in order to determine the exact timing of PRC1.3-dependent molecular events.

On the other hand, a naïve PRC1.3 knockdown system could have been used to determine the importance of the complex in the reverse direction, as naïve PSCs differentiate to the primed state, during which time human X chromosome inactivation occurs in female cells (Petropoulos *et al.*, 2016; Sahakyan *et al.*, 2017; Vallot *et al.*, 2017). As the genome-wide screen was performed in a male cell line, the essential role of PRC1.3 must be X chromosome independent. However, PRC1.3 and PRC1.5 complexes have been specifically identified in mESCs in *Xist* mediated gene silencing along the inactive X chromosome (Almeida *et al.*, 2017; Bousard *et al.*, 2019; Bowness *et al.*, 2022; Nesterova *et al.*, 2019; Pintacuda *et al.*, 2017) and a naïve knockdown system could have been used to establish conserved or distinct roles for the PRC1.3 complex and the wider Polycomb system in human X chromosome inactivation *in vitro*.

Using either approach, the PRC1.3 knockdown system could have been used to phenotype the importance of the complex in naïve PSCs. This could be assessed using naïve-specific cell-surface markers, naïve-specific and core pluripotency transcription factor expression and cell proliferation, as well as wider tests on the global transcriptional and epigenetic consequences of PRC1.3-deficiency in naïve PSCs. This system could also test the role of the complex in the transition into specific cell types including trophoblast stem cells and primitive endoderm, either individually in 2D systems or in blastoids as a 3D system, to determine the developmental significance of the complex for cell fate specification, which has never been considered or examined.

An alternative method to study the role of the PRC1.3 complex in human naïve PSCs would have been to directly create a genetic deletion of *PCGF3* in naïve PSCs, using CRISPR-Cas9, similarly to the creation of *PCGF3* KO primed PSCs. This would however have had two disadvantages, firstly if the complex was essential for maintenance of naïve PSCs, it is likely like all the successful homozygous deleted *PCGF3* KO naïve PSCs would have died or differentiated by the time clones were picked and were screened by genotyping. This would have limited our ability to characterise the molecular function of PRC1.3 in these cells. Secondly, this approach would not be inducible, which allows the immediate consequences of PRC1.3-deficiency to be analysed in a synchronised cell population with temporal resolution. The ability to restore the PRC1.3 levels within cells means that I could study whether any PRC1.3-dependent events are reversible and how the system recovers when PRC1.3 is restored.

On reflection, the inducible knockdown system selected could also have been improved. CRISPRi is a slow system, which can take up to 7 days to achieve a complete gene knockdown at the protein level (Mandegar *et al.*, 2016), while inducible knockdown systems with rapid kinetics now exist. For example, an auxin inducible degron (Nishimura *et al.*, 2009) has been demonstrated to cause complete protein knockdown of the PRC1 component RING1B in 2 h (Dobrinić *et al.*, 2021), and an FKBP12-dTag13 system (Nabet *et al.*, 2018) has even more rapid knockdown kinetics, allowing a complete protein knockdown of PCGF3 and PCGF5 in as little as 15 min (Bowness *et al.*, 2022). Both of these systems work by tagging the endogenous allele, meaning that the protein can be degraded in the presence of small molecular inhibitors. The rapid consequences of PRC1.3 knockdown could be studied with these systems, in order to analyse the role of PRC1.3 in reprogramming, as well as in human naïve PSCs more broadly.

It is difficult to speculate whether the PRC1.3 might be required to maintain naïve pluripotency or not, based on its critical requirement for reprogramming. There is an apparent disconnect between genes that characterise naïve PSCs and their requirement for reprogramming. Neither *KLF17* nor *SUSD2*, both of which I have used to profile naïve PSCs, as a transcription factor and a cell-surface protein, are required for human PTNR (Lea *et al.*, 2021; Wojdyla *et al.*, 2020). Factors that are required for the establishment of a particular cell state are not necessarily always required to maintain the cell state, and instead facilitate passage through key checkpoints during the reconfiguration of cell identity. For example, a number of the core factors for human induced pluripotent stem cell reprogramming of fibroblasts are required for the mesenchymal-to-epithelial transition (Collier *et al.*, 2022; Toh *et al.*, 2016), which is needed for a fibroblast to become a primed PSC, but may not be required in established human primed PSCs.

Additionally, a number of the common conditions that are used to grow and sustain naïve PSCs are not able to reprogramme primed PSCs to the naïve state with high efficiency. For example, transient broad spectrum histone deacetylase inhibition with valproic acid or transgene induction of *NANOG* and *KLF2*, or *KLF4*, are typically required to support high efficiency reprogramming in t2iL + PKCi and PXGL conditions (Bredenkamp *et al.*, 2019b; Guo *et al.*, 2017; Liu *et al.*, 2017; Takashima *et al.*, 2014). While 5i/L/A conditions are on the other hand able to reprogramme and sustain naïve PSCs, this condition is prone to genetic instability of human naïve PSCs, which can be alleviated by reduced concentrations of the MEK inhibitor PD0325901 (Di Stefano *et al.*, 2018) or transfer into another growth condition. This demonstrates that optimal maintenance versus efficient acquisition of naïve pluripotency often

have conflicting requirements, and thus the underlying molecular events may diverge strongly from one another also.

Taken together, I suggest that it should not be automatically assumed that the PRC1.3 complex is important for the maintenance of naïve pluripotency simply due to its importance for the acquisition of human naïve pluripotency, however the possibility still exists and generation of a knockdown or knockout system could still be used to comprehensively test this.

4 – PRC1.3-dependent transcriptional repression is associated with successful human PTNR

4.1 – Introduction

4.1.1 – Studying the molecular function of the PRC1.3 complex in human PTNR

Chapter 3 identified that PRC1.3-deficient primed PSCs are unable to form naïve PSCs during human PTNR. In Chapter 4, I aim to uncover how PRC1.3 might be regulating this process. PRC1.3 is a multi-protein chromatin modifying enzyme that catalyses H2AK119ub1 and this newly identified context of PRC1.3-specific function offers an opportunity to better understand the function of this complex in human pluripotency and in reprogramming, or in cell state transitions more broadly.

Previously reported functions of the PRC1.3 complex have been linked to regulation of transcription, by both activation and repression mechanisms (Gao *et al.*, 2014; Kanki *et al.*, 2022; Wang *et al.*, 2018; Yakushiji-Kaminatsui *et al.*, 2018; Yao *et al.*, 2018; Zhao *et al.*, 2017), in chromosome-wide blanketing of the inactive X chromosome with H2AK119ub1 to promote gene silencing (Almeida *et al.*, 2017; Bousard *et al.*, 2019; Bowness *et al.*, 2022; Nesterova *et al.*, 2019; Pintacuda *et al.*, 2017) and in the genome-wide blanketing of H2AK119ub1, with minimal transcriptional effect (Fursova *et al.*, 2019). As the X chromosome reactivates during PTNR, and the genome-wide screen was performed in a male cell line, I excluded a role in the regulation of X chromosome activity as the critical function of PRC1.3 in reprogramming.

As well as the potential for direct gene regulation, low-level genome-wide blanketing of H2AK119ub1 coverage by PRC1.3/5 complexes does not affect gene expression in steady state conditions, where cell fate is unchanged (Fursova *et al.*, 2019). During cell state transitions where epigenome remodelling is required, the PRC1.3-deposited epigenetic landscape could play a key role in modulating gene transcription. I therefore hypothesised that the function of PRC1.3 in PTNR involves regulation of gene expression to facilitate the reconfiguration of cell identity. This regulation could occur either by direct activation or repression of target genes or by indirect means such as the influence of the PRC1.3 complex on the global epigenome.

In order to address a potential gene regulatory function for PRC1.3 during reprogramming, I will first identify the transcriptional differences between WT and *PCGF3* KO cells during PTNR. Global transcription differences must be detectable as WT PSCs are able to produce nascent naïve PSCs, while *PCGF3* KO primed PSCs do not. PRC1.3 function must therefore play an early role in reprogramming, prior to the formation of the first nascent naïve PSCs, resulting

in the severe reprogramming defect, which persists even over an extended duration of reprogramming. However, the differences in global expression are not known and whether they are directly or indirectly dependent upon PRC1.3 or instead reflect underlying changes to the cell state.

PRC1.3/5 targeting to gene promoters has been identified in specific contexts, including neuronal differentiation genes (Gao *et al.*, 2014; Wang *et al.*, 2018), angiogenesis genes (Kanki *et al.*, 2022), TGF- β target genes (Yao *et al.*, 2018), mesoderm lineage genes (Zhao *et al.*, 2017), genes that are occupied by the USF1 transcription factor (Scelfo *et al.*, 2019) and the *Meis2* promoter in the mouse distal forelimb bud (Yakushiji-Kaminatsui *et al.*, 2018). The diversity of target genes suggests that PRC1.3 possesses a large degree of gene regulatory versatility depending on the circumstances. It is difficult to predict or infer a possible gene regulatory role for PRC1.3 during human PTNR based upon these reports.

The genes that are bound by the PRC1.3 complex in primed and naïve PSCs, and how they might be regulated during reprogramming to naïve PSCs, are not currently known. I will therefore use ChIP-sequencing to determine which genes are directly bound by PRC1.3 and thus are the strongest candidates for direct transcription regulation by the complex. Following the identification of these genes in human primed and naïve PSCs. I will integrate this information with their transcriptional changes across the cell state transition. This has scope for an enhanced understanding of Polycomb function in human pluripotency. It appears that at specific genes, PRC1.3 is able to act as both a transcriptional activator and a repressor, which appears to be context-dependent. This experimental approach could resolve these possibilities within the context of human PTNR, by determining if a gain in PRC1.3 complex occupancy is associated with transcriptional upregulation or downregulation.

This work will also expose gene regulatory mechanisms and their dependence on PRC1.3 in cells during reprogramming, which generates an opportunity to better understanding the transcriptional regulation of human pluripotent cell state transitions. A better understanding of the molecular changes underpinning human PTNR, including those orchestrated by the PRC1.3 complex, could enhance our understanding of developmental epiblast formation and could also facilitate optimisation of the reprogramming process. Pluripotent reprogramming is typically a low efficiency process (Ebrahimi, 2015), including human PTNR. Molecular interrogation of PRC1.3-dependent processes underlying reprogramming could identify bottlenecks during reprogramming which restricts the formation of naïve PSCs and interventions could be subsequently developed to manipulate cells to reprogramme with a higher efficiency. This in turn could help to unlock the full potential of human naïve PSCs,

which could be harnessed in biomedical applications and the study of human development and epigenetics.

4.1.2 – Studying the epigenetic function of PRC1.3 complexes in human primed PSCs

PRC1.3 complexes have been reported to activate transcription, independent of or counteracting H2AK119ub1 deposition (Gao *et al.*, 2014; Scelfo *et al.*, 2019). PRC1.3 complexes can however deposit H2AK119ub1 on chromatin, often with a high rate of catalytic activity (Almeida *et al.*, 2017; Fursova *et al.*, 2019; Huseyin and Klose, 2020). PRC1.3-dependent H2AK119ub1 activity, either at specific gene promoters or via genome-wide coverage, or both, could play an important functional role in human PTNR.

Variant PRC1 complex activity has been shown to act upstream of PRC2 activity in early mammalian embryo development (Chen *et al.*, 2021; Mei *et al.*, 2021) and X chromosome inactivation (Almeida *et al.*, 2017; Pintacuda *et al.*, 2017). These findings indicate that variant PRC1 complexes are often the drivers of developmentally-regulated Polycomb-mediated epigenetic changes. The abundance and distribution of both H2AK119ub1 and H3K27me3 differ strongly between human naïve and primed PSCs (Kumar *et al.*, 2022; Zijlmans *et al.*, 2022) and the Polycomb-dependent epigenetic landscape is extensively remodelled across this cell-state transition. As PRC1.3, but no other PRC1 complexes or PRC2 complexes, are crucial for PTNR, I speculate that PRC1.3 could be a key epigenetic driver across this cell state transition.

The contribution of an individual variant PRC1 complexes to the epigenetic regulation of human primed pluripotency has never been studied. Although the transcriptome of WT and *PCGF3* KO PSCs do not differ, this does not mean that the epigenetic landscape is also unchanged, and any epigenetic changes could help to explain why the complex is crucial for the cell state transition to naïve PSCs. Using the PRC1.3 KO system, I can determine the epigenetic regulation exerted by the PRC1.3 complex in primed PSCs, which unfortunately is not possible in naïve PSCs due to the absence of a knockout or knockdown system. Due to extensive interplay between epigenetic modifications (Murr, 2010), any epigenetic perturbations could have far-reaching consequences on chromatin regulation beyond simply H2AK119ub1, which could in turn influence reprogramming of these cells. The epigenetic consequences of PRC1.3 complex deletion can be determined in an agnostic manner to which modifications might be changing using a proteomics approach, as well as focusing specifically

on H2AK119ub1 and H3K27me3, due to the intense co-regulation of PRC1 and PRC2 complexes.

4.1.3 – Aims

By investigating the molecular and epigenetic role of PRC1.3 in human pluripotency and reprogramming, I aim to better understand the function of PRC1.3 in these contexts, as well as the events that occur during early human PTNR that are dependent on PRC1.3. In this chapter, I will use a combination of next-generation sequencing approaches and epigenetics assays to approach these aims:

1. Identify the transcriptional differences between WT and *PCGF3* KO cells during PTNR
2. Identify the genes bound by PRC1.3 and how these genes are regulated during PTNR
3. Use KO cells to analyse the epigenetic function of PRC1.3 in human primed PSC

4.2 – *PCGF3* KO primed PSCs fail to upregulate naïve pluripotency genes

4.2.1 - WT and *PCGF3* KO cells transcriptionally diverge during reprogramming

In order to identify the transcriptional differences between WT and *PCGF3* KO PSCs that underpin the failure of PRC1.3-deficient primed PSCs to reprogramme, as shown in Chapter 3, RNA-sequencing was performed on three independent biological replicates of parental WT and the two *PCGF3* KO clones on days 0 and day 10 of 5i/L/A (Theunissen *et al.*, 2014) reprogramming. Day 10 was selected as an early time-point in reprogramming when the PRC1.3-dependent phenotype has manifested due to the emergence of nascent naïve PSCs only in WT cells and aligns with the genome-wide screen, which was also performed on day 10 of 5i/L/A reprogramming. Although day 10 is an early time-point in reprogramming, by virtue of the heterogeneous population of cells arising from the reprogramming phenotype, it may be difficult to deconvolute the direct and indirect PRC1.3-dependent effects on transcription. Taking this finding a step further, RNA-sequencing can examine the phenotype of the *PCGF3* KO cells that fail to reprogramme, which I previously hypothesised were differentiated cells (as shown in Chapter 3).

RNA-sequencing was performed on bulk samples rather than using flow-sorting with cell-surface markers to collect nascent naïve PSCs or refractory cells. This approach was necessary as there were no suitable markers that could have been used, as there are no naïve PSCs in the KO cultures at day 10 of reprogramming (Chapter 3) and the markers of these *PCGF3* KO refractory cells are not known. In addition, sorting cells using markers might bias the cells that are studied, by obtaining or excluding specific populations.

I firstly assessed the similarity of the replicates within my dataset using a correlation matrix. This identified a particularly high degree of similarity between the homogeneous day 0 samples, as expected (Chapter 3) and reduced levels of similarity between day 0 and day 10 cells. Notably, within conditions, the two WT Day 10 cells types showed a relatively low level of similarity with one another (Fig. 4.2.1).

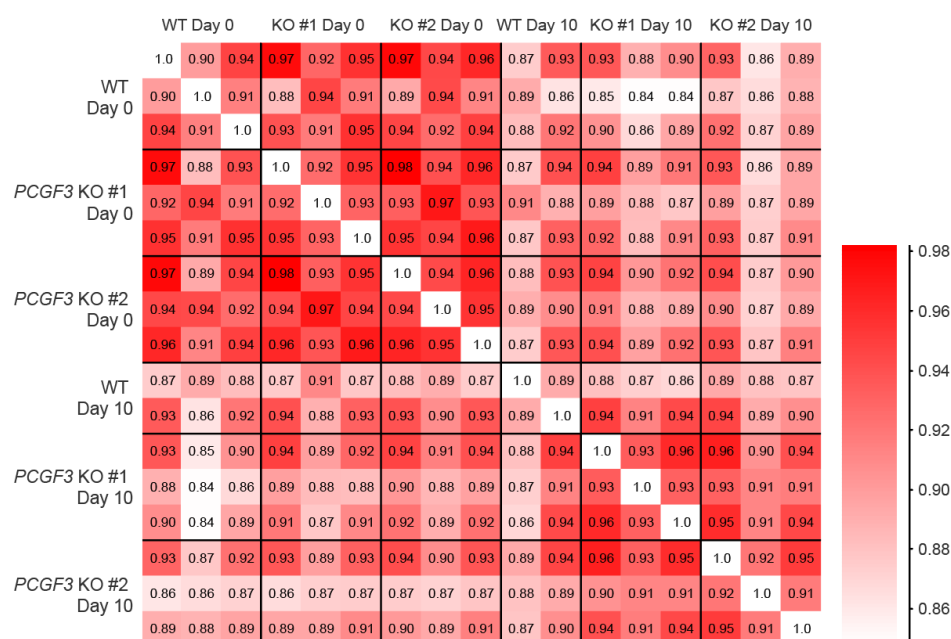


Figure 4.2.1 – Correlation matrix of RNA-sequencing data at days 0 and 10 of PTNR in 5i/L/A conditions

All replicates are shown for parental WT and the two *PCGF3* KO clone populations, with each column and row representing a single RNA-sequencing condition. Three biological replicates for all samples except WT Day 10, with two biological replicates.

The WT Day 10 cells are likely the most heterogeneous samples within the dataset, due to the presence of nascent naïve PSCs, primed-like cells and refractory cells (Chapter 3), and their relative proportions depend upon the reprogramming efficiency. In order to identify global patterns between samples, I performed principal components analysis on grouped averages of each condition within the dataset. This provided a useful biological insight into the transcriptional differences during reprogramming. Primed (Day 0) parental WT and *PCGF3* KO PSCs cluster closely due to their strong transcriptional similarity (Chapter 3). After 10 days of 5i/L/A reprogramming, both *PCGF3* KO clones cluster closely together, indicating that both clones fail to reprogramme in a similar manner and produce similar transcriptional profiles as a result. WT and *PCGF3* KO cells strongly diverge transcriptionally after 10 days of reprogramming (Fig. 4.2.2).

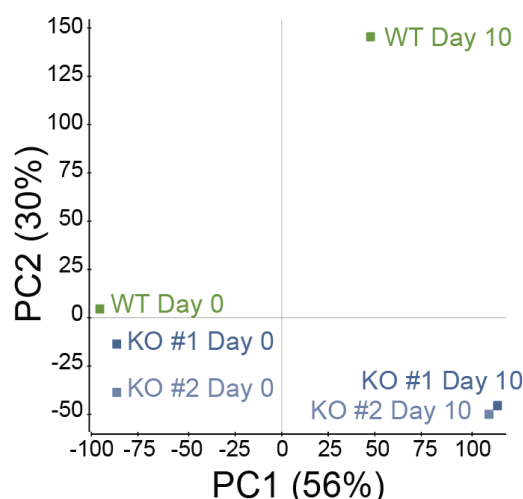


Figure 4.2.2 – Principal component analysis (PCA) of RNA-sequencing data at days 0 and 10 of PTNR in 5i/L/A conditions

Samples from bulk cell populations for the parental WT and *PCGF3* KO cell populations. Each data point represents the average of three biological replicates, except WT Day 10, which is the average of two replicates.

Principal component 1 (PC1) separates day 0 and 10 of reprogramming, regardless of the presence of nascent naïve PSCs. This shows that there are large transcriptional differences between cells before and after reprogramming. Principal component 2 (PC2) instead segregates WT Day 10 cells from all other cell types, with WT and *PCGF3* KO primed PSCs undergoing distinct transcriptional trajectories over ten days of 5i/L/A reprogramming.

4.2.2 - *PCGF3* KO cells upregulate neural marker genes during reprogramming

I further interrogated the transcriptional differences between WT and *PCGF3* KO PSCs on day 10 of 5i/L/A reprogramming. DESeq2 (Love *et al.*, 2014) was used to identify statistically significant differentially expressed genes between WT and *PCGF3* KO conditions. Amongst the genes expressed at significantly elevated levels in WT cells on day 10 of 5i/L/A reprogramming compared to the *PCGF3* KO population are the naïve pluripotency marker genes *DPPA5* (Kilens *et al.*, 2018; Yan *et al.*, 2013), *DNMT3L* (Ware, 2017; Yan *et al.*, 2013), *FGF4* (Petropoulos *et al.*, 2016) and *ALPP* (Bi *et al.*, 2020), and the pan-pluripotency marker *OCT4*, consistent with previous findings (Chapter 3), and *NODAL*, which is expressed more highly in human pre-implantation epiblast and whose signalling is proposed to inhibit the

neuroectodermal differentiation of human primed PSCs (Molè *et al.*, 2021; Vallier *et al.*, 2004) (Fig. 4.2.3).

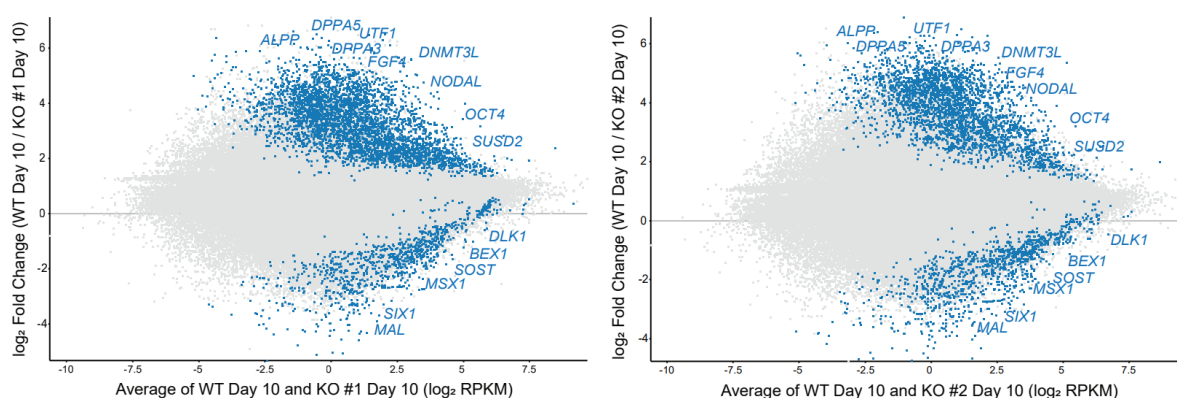


Figure 4.2.3 – MA plot comparing WT and *PCGF3* KO cells on day 10 of PTNR in 5i/L/A conditions Differentially expressed genes highlighted in blue (FDR < 0.05; Wald test with Benjamini-Hochberg correction). Data show average of two (WT) or three (KO) biological replicates.

These findings confirm the failure of *PCGF3* KO cells to upregulate naïve pluripotency markers. Conversely, genes significantly upregulated in the *PCGF3* KO cells compared to WT include the marker genes of neuronal differentiation including *DLK1* (Surmacz *et al.*, 2012) and *BEX1* (Vilar *et al.*, 2006), consistent with reports that show that refractory cells that fail to form nascent naïve PSCs have a neural-like transcriptional profile (Collier *et al.*, 2017), as well as trophectoderm and primitive endoderm marker upregulation (Bi *et al.*, 2022).

Looking specifically at pluripotency transcription factors, RNA-sequencing, corroborated by RT-qPCR, was used to look specifically at pluripotency transcription factor gene expression. I see that expression of core, pan-pluripotency markers of *OCT4*, *NANOG* and *PRDM14* (Boyer *et al.*, 2005; Chia *et al.*, 2010; Tsuneyoshi *et al.*, 2008) are downregulated in *PCGF3* KO cells during reprogramming and naïve-specific pluripotency markers fail to become upregulated, with *DPPA3* (Bowles *et al.*, 2003; Theunissen *et al.*, 2014), *TFCP2L1* (Wang *et al.*, 2019) and *KLF4* (Weinberger *et al.*, 2016) induced after 10 days of 5i/L/A reprogramming only in WT cells (Fig. 4.2.4 and Fig. 4.2.5).

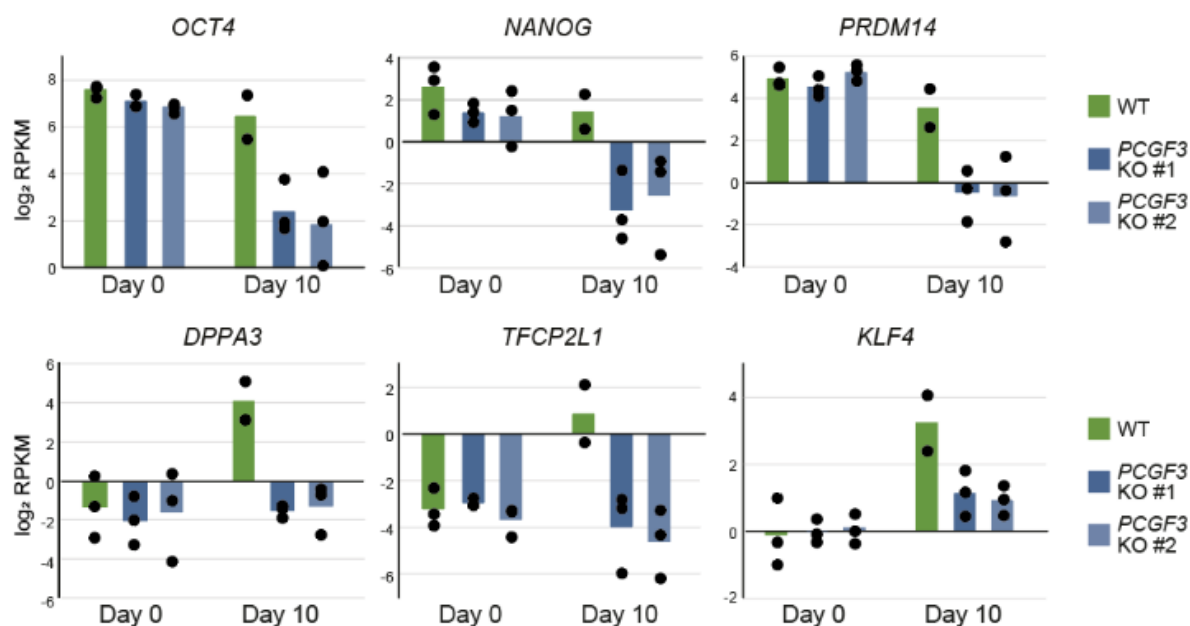


Figure 4.2.4 – Charts show gene expression levels in log₂ RPKM (reads per kilobase transcript per million reads), in WT and PCGF3 KO cells at days 0 and 10 of PTNR in 5i/L/A conditions

Circles show each biological replicate (n=2 for WT Day 10; n=3 for other samples) for the pluripotency genes *OCT4*, *NANOG* and *PRDM14*, and the naïve PSC markers *DPPA3*, *TFCEP2L1* and *KLF4*.

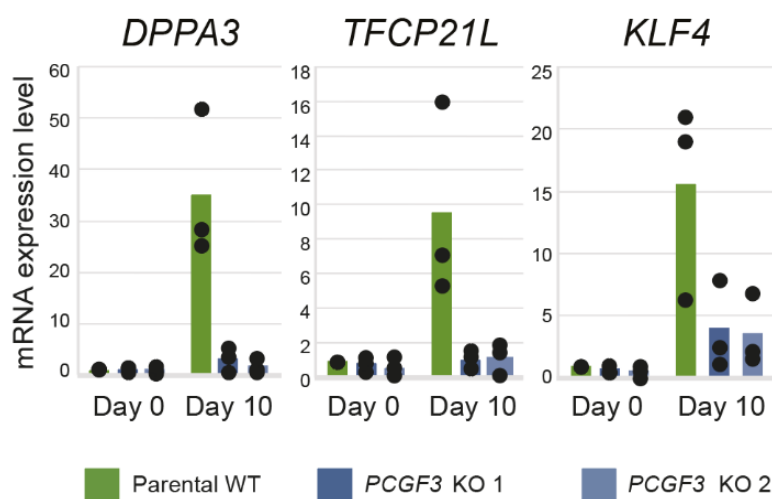


Figure 4.2.5 – RT-qPCR showing gene expression levels in WT and PCGF3 KO cells at days 0 and 10 of PTNR in 5i/L/A conditions

Individual data points are shown for three independent experiments for the naïve PSC markers *DPPA3*, *TFCEP2L1* and *KLF4*. Samples were collected on day 0 and day 10 of 5i/L/A reprogramming. Samples

were normalised to the geometric mean of the housekeeper genes *HMBS* and *GAPDH*, and adjusted such that WT Day 0 is equal to 1 for each gene.

Clustering of this gene set is able to group these differentially expressed genes, which allows further investigation of the identity of refractory cells on day 10 of reprogramming. The most enriched gene ontology biological processes term for the genes upregulated specifically in *PCGF3* KO cells at day 10 of reprogramming is 'axonogenesis', further supporting a neural-like differentiation pattern of gene expression (Fig. 4.2.6). These data confirm a failure of pluripotency self-renewal under reprogramming conditions when *PCGF3* is removed.

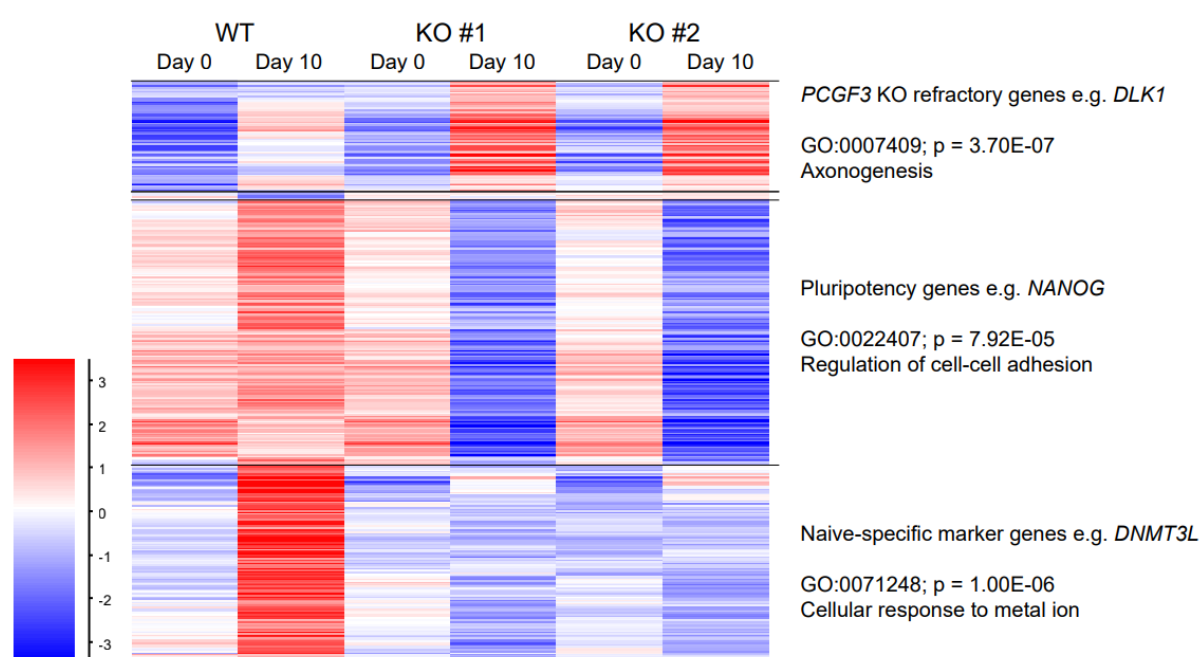


Figure 4.2.6 – Clustering of RNA-sequencing within differentially expressed genes

Hierarchical clustering performed on differentially expressed genes between WT and *PCGF3* KO #1 cells on day 10 of reprogramming. Three clusters are identified with > 100 genes and their highest statistically significant gene ontology biological processes term and an example gene are provided.

In summary, RNA-sequencing has identified that after 10 days of 5i/L/A reprogramming, WT PSCs upregulate naïve marker genes and maintain expression of core pluripotency marker genes. On the other hand, *PCGF3* KO cells do not upregulate naïve marker genes and fail to maintain expression of core pluripotency marker genes, instead upregulating neural differentiation genes, consistent with findings from WT refractory cells (Collier *et al.*, 2017).

This shows the transcriptional consequences of *PCGF3* deletion that renders these cells unable to reprogramme to naïve PSCs. Whether the same transcriptional signatures of *PCGF3* KO cells in alternative reprogramming conditions is not known, as I did not test this.

However, the identification of a large number of differentially expressed genes does not inform on which of these genes are directly regulated by PRC1.3, as a large number will change indirectly as a result of alterations to cell identity under reprogramming conditions. As it is not known where the PRC1.3 complex binds to chromatin in primed and naïve PSCs, this data alone is also not able to indicate exactly how the PRC1.3 complex regulates transcription.

During reprogramming, the PRC1.3 complex could be important for gene repression or gene activation, in a manner that facilitates the establishment of the naïve pluripotency transcriptional network. The PRC1.3 complex could for example play a role in suppression of differentiation genes, or could activate expression of naïve pluripotency promoting genes, such that a small proportion of cells evade differentiation into a permissive cell type. Equally, the scope of PRC1.3 transcriptional regulation could be limited to a very small number of, or even a single, master regulator - with all identifiable phenotypic changes being an indirect by-product of dysregulation of this gene set.

4.3 – PRC1.3 target genes were identified by ChIP-sequencing

To help understand the function of PRC1.3 in human reprogramming, as well as an understanding of the genes regulated by the PRC1.3 in primed and naïve PSCs, a ChIP-sequencing approach was required to determine where the PRC1.3 complex binds to chromatin in the two pluripotent cell types.

4.3.1 - Generating PRC1.3 and PRC1.5-deficient (DKO) human primed PSCs

For PRC1.3-specific ChIP-sequencing, I needed to use the anti-PCGF3+5 antibody to detect the complex, due to the shortcomings of PCGF3-specific antibodies. Although PCGF5 protein was not detectable in human primed PSCs, I decided to generate *PCGF3* and *PCGF5* double knockout (DKO) primed PSCs as a robust negative control for the ChIP-sequencing experiment. For this, CRISPR-Cas9 was used to delete the *PCGF5* gene in primed PSCs where *PCGF3* had already been deleted, using the same strategy as used previously, with a pair of guide RNAs to target the critical sixth exon of the *PCGF5* gene (Fig. 4.3.1).

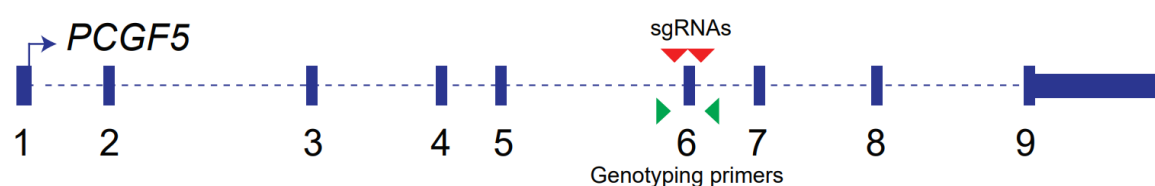


Figure 4.3.1 – Schematic for the creation of *PCGF3/5* DKO primed PSCs by *PCGF5* gene deletion

Schematic of *PCGF5* exon structure. Red triangles indicate the position of the gRNAs used and green triangles indicate the position of genotyping primers subsequently used to confirm the exon deletion.

After transfection and CAS9-GFP sorting, individual colonies were picked and screened using specific primers to genotype whether exon 6 of the *PCGF5* gene had been successfully excised. 71 clones were screened and produced 3 homozygous *PCGF5* deletions, one of which was selected for further study (Fig. 4.3.2). As PCGF5 is not detectable in human primed PSCs, Western blotting was not used to confirm the deletion of *PCGF5*.

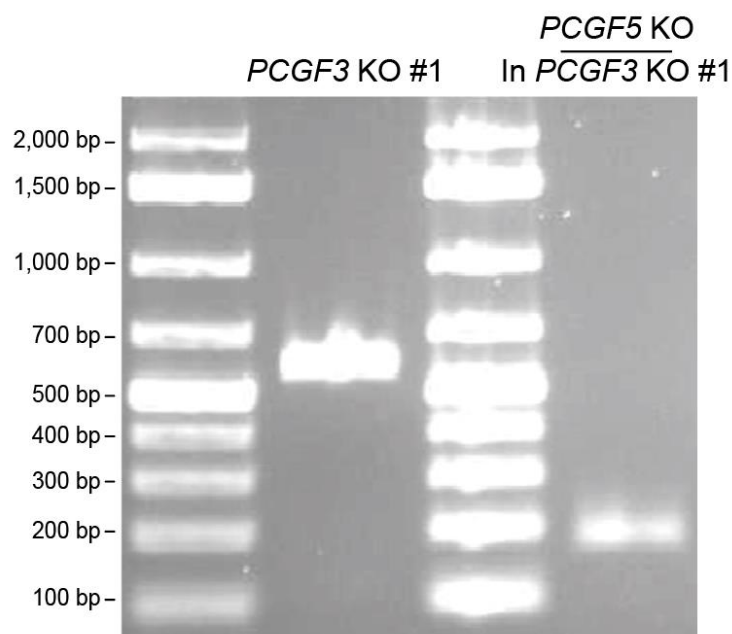


Figure 4.3.2 – Electrophoresis image of PCR-based genotyping of the *PCGF5* gene

The reduced PCR product size in the *PCGF5* targeted cells compared to a *PCGF3* single KO control, indicating a homozygous deletion of exon 6 of *PCGF5*. The size of the DNA fragments in the ladder are shown in base pairs (bp).

4.3.2 – ChIP-seq identifies PRC1.3 target genes

Following the generation of *PCGF3* and *PCGF5* DKO primed PSCs, I performed ChIP-seq in human primed and naïve PSCs, with two biologically independent replicates per condition and using *PCGF3/5* DKO primed PSCs as a negative control to confirm the specificity of the identified sites of PRC1.3 binding on chromatin. To stringently identify the PRC1.3 complex, ChIP-seq was performed separately against two different proteins within the complex – RING1B and PCGF3/5. RING1B is present within all PRC1 complexes, including PRC1.3, while PCGF3 (and PCGF5) are identity-determining components of PRC1.3 (and PRC1.5) complexes (Gao *et al.*, 2012; Hauri *et al.*, 2016). The co-localisation of both RING1B and PCGF3 bound at the same sites on chromatin was used to determine the presence of PRC1.3-specific binding sites and PRC1.3 target genes.

In analysing the ChIP-seq data, I first used model-based analysis of ChIP-seq (MACS) (Zhang *et al.*, 2008) to call RING1B peaks, using an algorithm to identify predicted chromatin binding sites with statistical significance, based on the enrichment of mapped ChIP-seq reads against an input sample. A large number of RING1B peaks were detected

in both primed (11,046) and naïve (5,771) PSCs, while far fewer peaks were detected for PCGF3 on naïve (405) and primed (338) chromatin. Any PCGF3 peaks also detected in *PCGF3/5* DKO primed PSCs were excluded from the primed (20) and naïve (4) peak list and this confirmed the ability to specifically detect PCGF3 binding to chromatin, as the anti-PCGF3/5 antibody has not been previously validated for ChIP-sequencing. PRC1.3 binding sites were established by overlapping sites of RING1B and PCGF3 binding in naïve (378) and primed (138) PSCs. While previous findings in mESCs showed only 20 % of PCGF3 peaks were identified at RING1A/B binding sites on chromatin (Scelfo *et al.*, 2019), this was around 93 % in naïve PSCs and 40 % in primed PSCs.

I analysed the enrichment of naïve-specific and primed-specific PCGF3 peaks from ChIP-sequencing across a 20 kb window over TSS, where much of the future analysis will be directed in order to study the gene regulation by the PRC1.3 complex and where 60 % of naïve PRC1.3 peaks are identified. Enrichment of PCGF3 signal is seen in WT, but not in the *PCGF3/5* DKO primed PSCs, showing the specificity of identified PCGF3 binding to chromatin. The magnitude of PCGF3 enrichment at naïve target peaks is much higher than at primed target peaks, which are somewhat less convincing, albeit still specific compared to *PCGF3/5* DKO primed PSCs (Fig. 4.3.3).

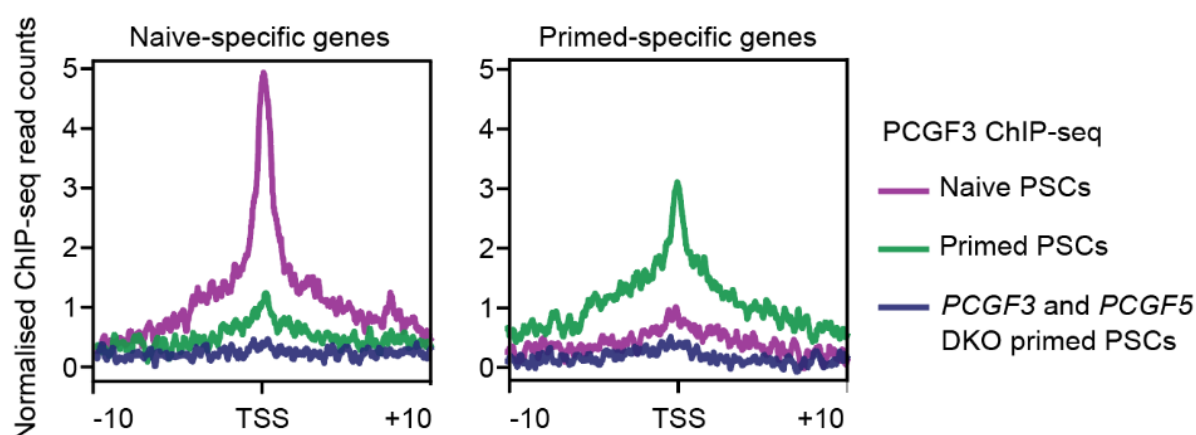


Figure 4.3.3 – Metaplot of PCGF3 ChIP-seq. signal over naïve-specific and primed-specific target genes in naïve PSCs, WT and DKO primed PSCs

Data is centred on a 20 kb window surrounding transcriptional start sites. Plot produced by Andrew Malcolm.

Notably, the vast majority of genes bound by the PCGF3 and RING1B are bound either in naïve PSCs or in primed PSCs, with fewer than 5 % of PRC1.3-bound TSS regions sharing PRC1.3 occupancy in both naïve and primed PSCs (Fig. 4.3.4).

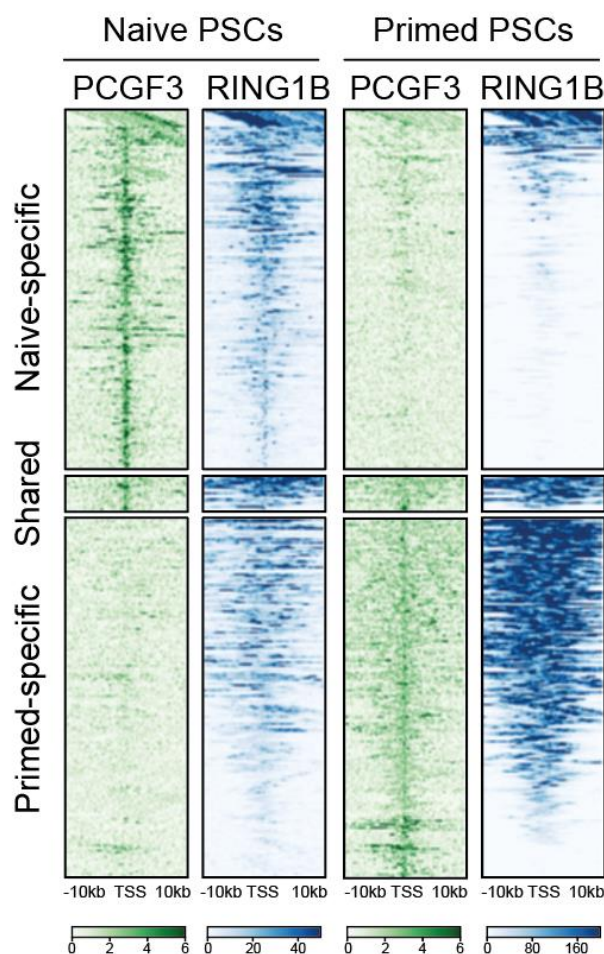


Figure 4.3.4 – Heatmaps of normalised PCGF3 and RING1B ChIP-seencing read counts within a 20 kb transcription start site (TSS)-centred window in naïve and primed PSCs

Regions were separated into naïve-specific, primed-specific and shared PRC1.3 binding sites. Heatmap produced by Andrew Malcolm.

These data suggest that during reprogramming, PRC1.3 undergoes a substantial redistribution on chromatin and the genes that could be directly regulated by the PRC1.3 complex are largely distinct between the two cell states.

To examine how PRC1.3 modulates PTNR, I focused specifically on genes that gain PRC1.3 binding during reprogramming – henceforth named naïve PRC1.3 target genes. As PRC1.3 is crucial for reprogramming, it is possible that these naïve-specific PRC1.3 target genes might be crucial for enacting PRC1.3 function during reprogramming, whether they be transcriptionally upregulated or downregulated by the action of PRC1.3.

I also performed Cleavage Under Targets and Tagmentation (CUT&Tag), as a low input chromatin approach which works well for histone post-translational modifications, for the

PRC1-deposited H2AK119ub1 modification in naïve, and WT and DKO primed PSCs, along with an IgG negative control alongside each sample for peak calling. I also developed a calibrated approach to normalise the data using a spike-in with exogenous *Drosophila* cells. Unfortunately this dataset did not achieve an optimal number of sequencing reads and so should be treated with some caution. Using this approach, I identified a specific enrichment for H2AK119ub1 over naïve PRC1.3 target genes in naïve PSCs (Fig. 4.3.5 and Fig. 4.3.6).

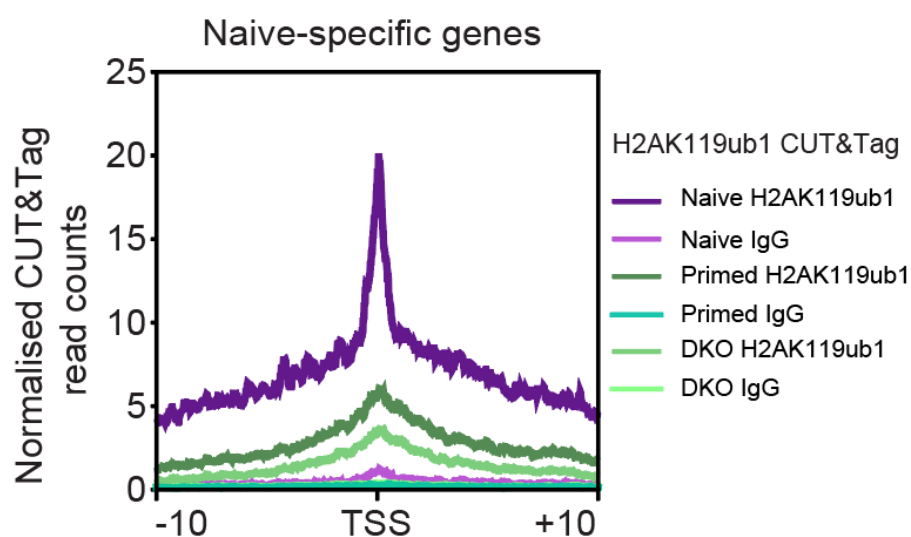


Figure 4.3.5 – Metaplot of H2AK119ub1 CUT&Tag signal over naïve-specific PRC1.3 target genes in naïve PSCs, WT and DKO primed PSCs

Data is centred on a 20 kb window surrounding transcriptional start sites. Plot produced by Andrew Malcolm.

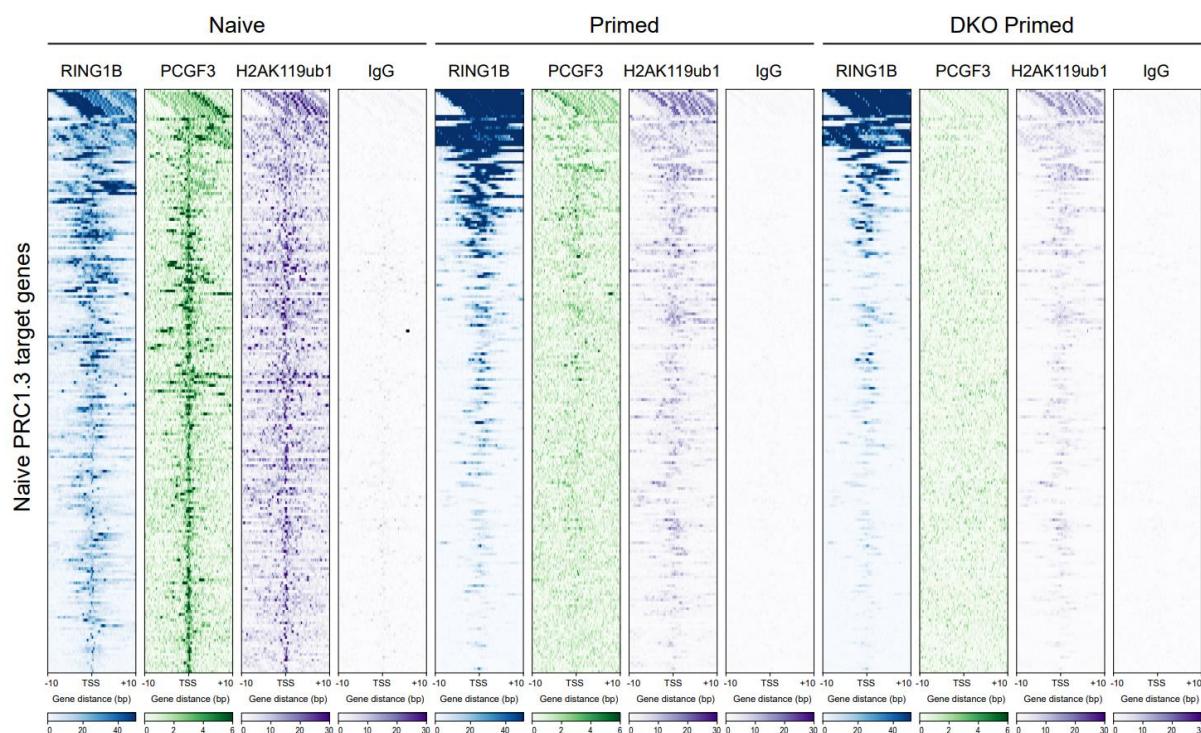


Figure 4.3.6 – Heatmaps of normalised PCGF3 and RING1B ChIP-seq, and H2AK119ub1 + IgG CUT&Tag, read counts over a 20 kb TSS-centred window in naïve, WT and DKO primed PSCs

Analysed over naïve PRC1.3 target genes. Heatmap produced by Andrew Malcolm.

The presence of H2AK119ub1 over this gene set acts as the first indicator that these genes might be undergoing PRC1.3-dependent transcriptional repression as cells reprogramme from primed to naïve. An example naïve PRC1.3 target gene is *LETM1*, which is not bound by either RING1B or PCGF3 in primed PSCs, but is co-bound by both RING1B and PCGF3 in naïve PSCs, and marked with H2AK119ub1 specifically in naïve PSCs (Fig. 4.3.7).

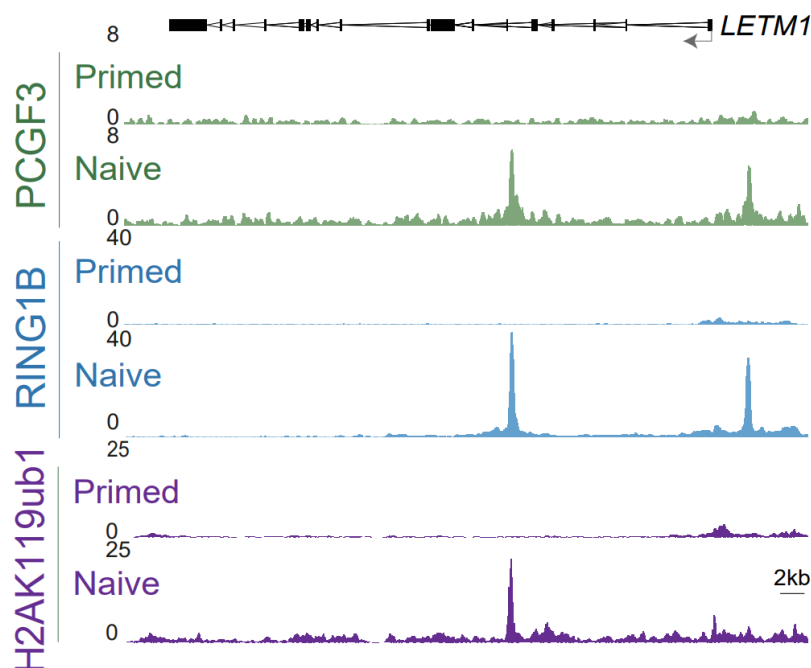


Figure 4.3.7 – Chromatin profiling shows *LETM1* is a naïve-specific PRC1.3 target gene

Genome browser tracks shown for PCGF3 and RING1B (ChIP-seq) and H2AK119ub1 (CUT&Tag) in naïve and primed PSCs at the *LETM1* gene locus.

Amongst the naïve PRC1.3 target gene subset, with PRC1.3 bound at the TSS only in naïve PSCs, are developmental mediators; *PAX7*, *OTX1* and *TCF7L1*, cell signalling facilitators; *DUSP6*, *BRAF* and *RARA*; and chromatin-based regulators such as *BCORL1*, *AUTS2* and *MED12*. These candidates could play a central role in translating PRC1.3 function into successful reprogramming, which would allow new hypotheses to be developed and tested, to study regulation of PTNR downstream of PRC1.3.

However, as I am specifically interested in PRC1.3 function in the molecular events that underpin human PTNR, I focused my analysis on how PRC1.3 regulates this gene subset. The exact role that these PRC1.3-regulated genes might play in this process are not currently clear and would require further analysis to determine this.

4.4 – PRC1.3 target genes are downregulated during reprogramming

4.4.1 - PRC1.3 occupancy is associated with gene downregulation during PTNR

The PRC1.3 target genes identified by ChIP-sequencing are the strongest candidates for how the function of the PRC1.3 complex is translated into the essential PRC1.3-dependent processes that occur during human PTNR. Using the knowledge of which genes are naïve-specific PRC1.3 target genes, I used both my own and published (Collier *et al.*, 2017) RNA-sequencing data in order to identify the effect of gaining PRC1.3 occupancy on the expression of these genes during reprogramming. This is instrumental to understand how the PRC1.3 complex modulates reprogramming to the naïve state, for example via either transcriptionally activating or transcriptionally repressive modes.

Initially, I analysed the fold change in the expression levels of naïve PRC1.3 target genes between the naïve and primed PSCs. This analysis showed that as an ensemble, naïve PRC1.3 target genes are expressed more highly in primed PSCs compared to naïve PSCs and become transcriptionally downregulated when cells reprogramme (Fig. 4.4.1).

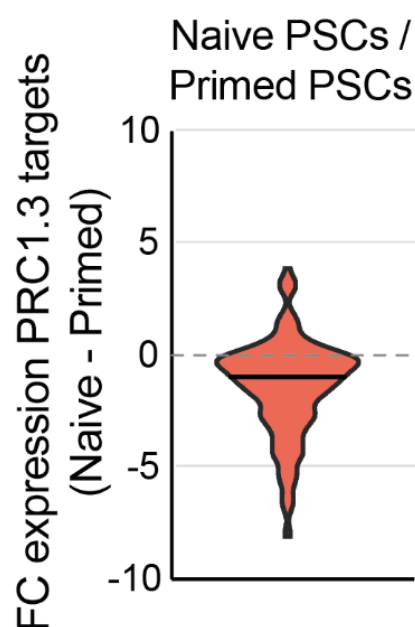


Figure 4.4.1 – Violin plot showing the fold-change in the expression of PRC1.3 target genes in naïve compared to primed PSCs

Data shown as an average expression value of three biological replicates for both primed and naïve PSCs for naïve PRC1.3 target genes (n=191), where the black line represents the median value. RNA-sequencing data obtained from (Collier *et al.*, 2017).

There is a correlation between a gain of PRC1.3 occupancy and a decrease in gene expression from the primed to naïve PSCs. When represented on a gene-by-gene basis, the proportion of PRC1.3 target genes that are downregulated can be seen. Over 75% (n=147) of the naïve PRC1.3 target genes (n=191) are expressed to a higher degree in primed PSCs and are transcriptionally downregulated, albeit to different degrees, when primed PSCs reprogramme to the naïve state. (Fig. 4.4.2).

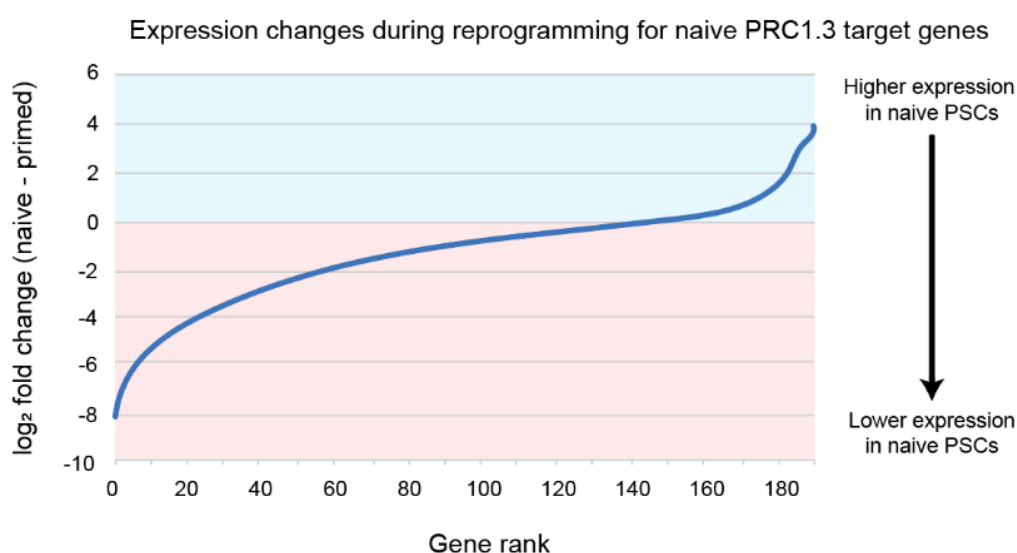


Figure 4.4.2 – Chart shows the gene expression change of PRC1.3 target genes between naïve and primed PSCs

Genes are ranked according to their fold change in expression. RNA-sequencing data obtained from (Collier *et al.*, 2017).

A very small number of naïve PRC1.3 target genes either minimal expression differences between the two states or are transcriptionally upregulated from primed to naïve, but these are far from representative of the trend within the gene set. The correlation between gaining PRC1.3 complex occupancy and becoming downregulated does not however directly associate the action of the PRC1.3 complex with transcriptional repression of these genes.

In order to more directly analyse the dependence of the PRC1.3 complex on transcriptional repression of these genes, I assessed their expression levels on day 0 and day 10 of 5i/L/A reprogramming in both WT and *PCGF3* KO cells. On day 10 in WT PSCs, where nascent naïve PSCs are present, the expression of PRC1.3 naïve target genes are show a statistically significant decrease compared to primed (day 0) expression levels. In contrast, these genes are not significantly downregulated in the two clones where PRC1.3 is deleted, with an almost identical median expression level to their starting primed state (Fig. 4.4.3).

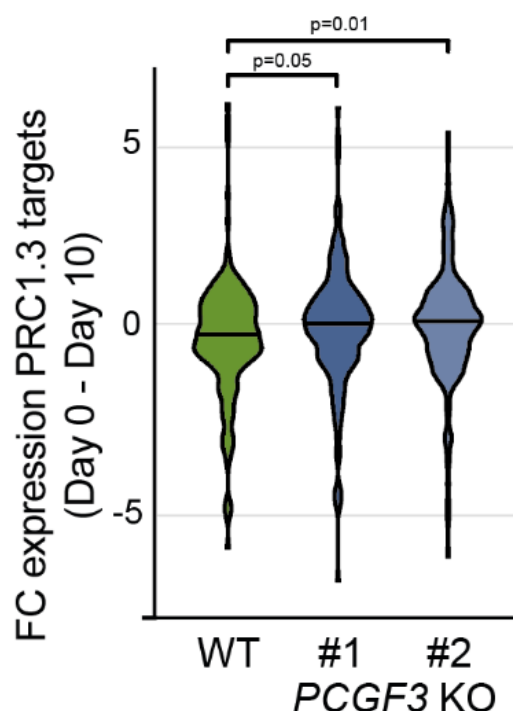


Figure 4.4.3 – Violin plots showing the fold-change in expression of PRC1.3 target genes in WT and *PCGF3* KO cells at day 0 compared to day 10 of PTNR (5i/L/A conditions)

Data shown is the mean expression for WT Day 10 (n=2) and for all other samples (n=3) for naïve PRC1.3 target genes (n=191). Horizontal black line represents the median value. P-values were calculated using a Kruskal-Wallis test with Dunn's multiple comparison test.

The magnitude of repression of these genes is somewhat lower in day 10 reprogramming cells compared to established naïve PSCs (Fig. 4.4.1). I hypothesise that this could be due to repression occurring only in nascent naïve PSCs and as reprogramming cultures are heterogeneous, only a proportion of the cells are nascent naïve PSCs, compared to pure naïve cultures and so fewer of the cells in the population may be downregulating these genes. These

findings collectively indicate that PRC1.3 complex is unlikely to be activating transcription, as has been seen in other contexts.

4.4.2 - Failure to repress PRC1.3 targets is associated with a failure to reprogramme

Looking again on a gene-by-gene basis at the expression of a subset of naïve PRC1.3 target genes, the PRC1.3-dependent transcriptional repression signature can be seen. After 10 days of 5i/L/A reprogramming, these genes are either transcriptionally downregulated in WT cells or are expressed to much lower levels in than KO cells, where the genes are aberrantly expressed (Fig. 4.4.4).

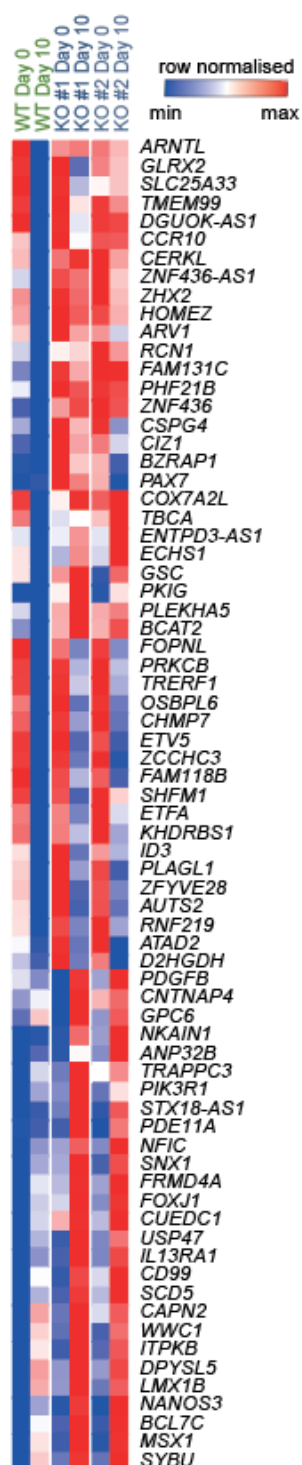


Figure 4.4.4 – Heatmap showing the row normalised expression levels for a subset of naïve PRC1.3 target genes in WT and *PCGF3* KO cells at days 0 and 10 of PTNR (5i/L/A conditions)

Blue represents lowest expression levels and red represents highest expression levels.

In order to determine whether the transcriptional repression of naïve PRC1.3 targets is instructive for successful reprogramming, I analysed the expression dynamics of these genes in flow-sorted cell populations linked to human PTNR in WT cells (Collier *et al.*, 2017). I compared the expression of naïve PRC1.3 target genes in primed PSCs, sorted nascent naïve PSCs (positive for naïve-specific and negative for primed-specific cell-surface markers) and refractory cells (negative for naïve-specific cell surface markers), after ten days of 5i/L/A reprogramming. Using this approach, the expression levels of these PRC1.3 target genes can be linked to the cellular phenotype seen during successful or failed reprogramming. PRC1.3 target genes are only significantly transcriptionally repressed in nascent naïve PSCs, where early reprogramming has been successful (Fig 4.4.5).

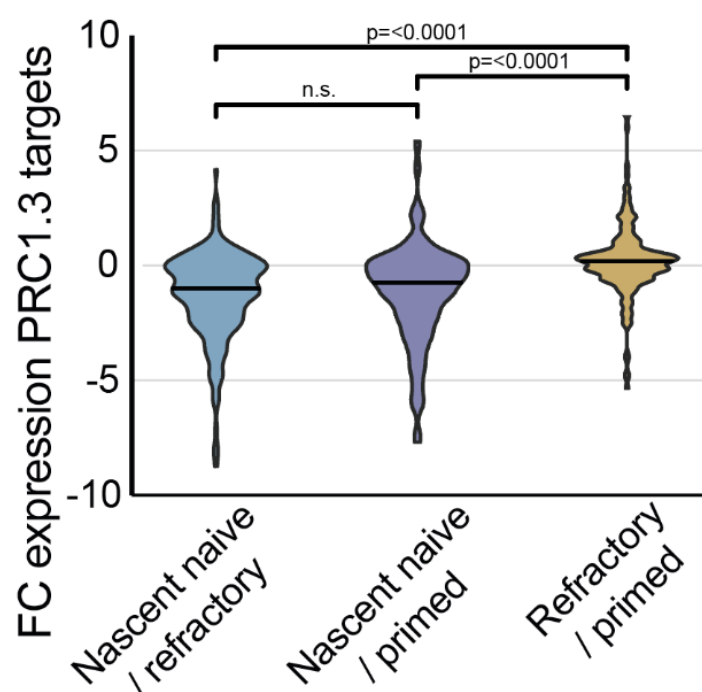


Figure 4.4.5 – Violin plots showing the fold-change in the expression of PRC1.3 target genes in cell-sorted populations at day 0 (primed) and day 10 (nascent naïve and refractory) of PTNR (5i/L/A conditions)

Data shown for three independent biological replicates for each sample (nascent naïve, refractory and primed) for naïve PRC1.3 target genes (n=191). Horizontal black line represents the median value. P-values were calculated using a Kruskal-Wallis test with Dunn's multiple comparison test. Data obtained from (Collier *et al.*, 2017).

The expression level of these genes shows no significant differences between primed PSCs and refractory cells, despite primed and refractory cells having distinct cellular identities. This shows that the changes in the expression of these genes in WT cells is likely not due to an indirect, underlying change in cell identity but are instead directly dependent upon the formation of nascent naïve PSCs, which in turn relies on PRC1.3.

In summary, PRC1.3 binds a subset of genes specifically in naïve PSCs and the repression of these genes appears to be a pre-requisite for successful reprogramming to naïve pluripotency. When the PRC1.3 complex is deleted, repression of these genes fails to occur, and cells instead form a refractory cell population under conditions that otherwise induce naïve PSC formation by reprogramming.

4.5 – PRC1.3 influences the epigenetic landscape of primed PSCs

4.5.1 - H2AK119ub1 and H3K27me3 are more abundant in human naïve PSCs

PRC1 complexes traditionally repress transcription by the deposition of the epigenetic modification H2AK119ub1. As the PRC1.3 complex is associated with transcriptional repression in this context, I hypothesised that this PRC1.3-dependent repression is mediated by catalysis of H2AK119ub1. However, the activity of PRC1 complexes, including the dynamics of H2AK119ub1 in human naïve and primed pluripotency has not been studied until very recently (Kumar *et al.*, 2022).

To study the epigenetic influence of PRC1 complexes in human PSCs, I firstly assessed the global abundance of the H2AK119ub1 modification, as a proxy for PRC1 complex activity, in human naïve and primed PSCs. Unfortunately, it is very difficult to study H2AK119ub1 levels using proteomics, as sample preparation using propionylation and trypsin digestion gives poor coverage of this region of the H2A histone protein (Daled *et al.*, 2021). Instead, I used histone Western blotting, which shows that global H2AK119ub1 levels are elevated in naïve PSCs compared to primed PSCs (Fig. 4.5.1).

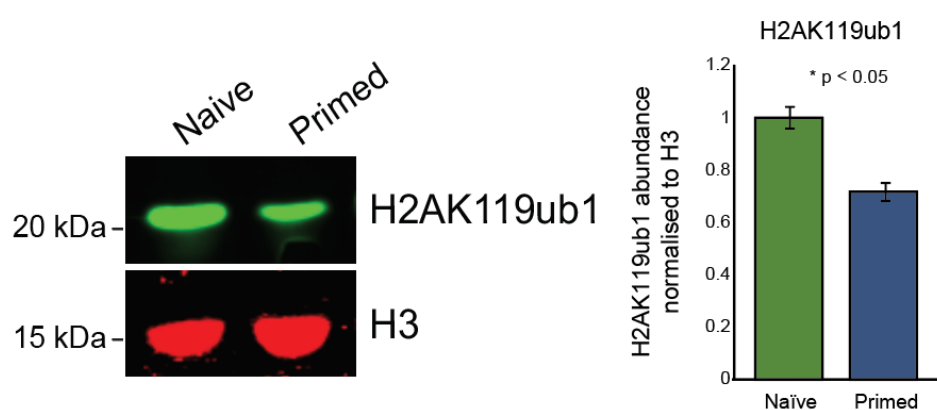


Figure 4.5.1 – Western blotting shows the abundance of H2AK119ub1 in naïve and primed PSCs

H2AK119ub1 has a molecular weight of 23 kDa. Histone H3, loading control, has a molecular weight of 15 kDa. Data are representative of three independent experiments. Quantification of H2AK119ub1 signal is shown in the bar chart, with bar height representing sample mean and error bars showing the standard error of the mean. Statistical difference determined by paired t-test.

This finding is consistent with results from quantitative ChIP-sequencing approach for H2AK119ub1 in human naïve and primed PSCs (Kumar *et al.*, 2022), which suggests that PRC1 complex activity might be higher in naïve than primed PSCs. Due to the extensive interplay between PRC1 and PRC2 complexes, I also analysed the levels of PRC2-deposited modification H3K27me3 using a quantitative histone proteomics approach. Indeed, like H2AK119ub1, H3K27me3 has a higher global abundance in naïve over primed PSCs (Fig. 4.5.2), as has been previously reported in human PSCs (De Clerck *et al.*, 2019; Kumar *et al.*, 2022; Zijlmans *et al.*, 2022).

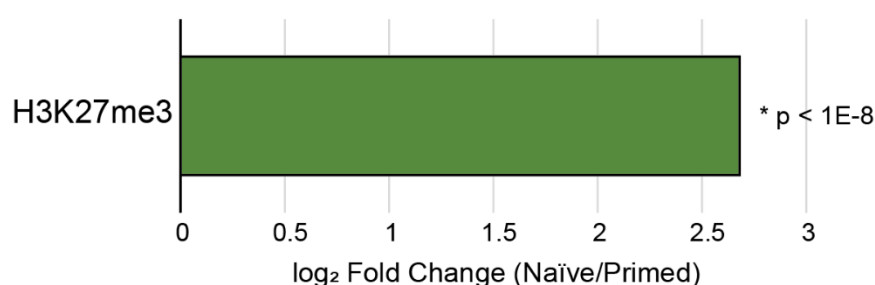


Figure 4.5.2 – Histone proteomics shows the relative abundance of the H3K27me3 modification in naïve and primed PSCs

Bar length represents log₂ fold change between naïve and primed PSCs. P-value calculated by student's t-test. Performed on five (primed) and six (naïve) independent biological replicates. Sample processing, mass spectrometry and analysis performed by Sigrid Verhelst and Maarten Dhaenens.

4.5.2 - DKO primed PSCs show reduced levels of H2AK119ub1 and H3K27me3

It would appear that PRC1 and PRC2 complexes are more active in the naïve state of pluripotency due to increased H2AK119ub1 and H3K27me3 levels, or the mechanisms antagonistic to Polycomb enzymatic activity, such as H2AK119ub1 deubiquitylation and H3K27me3 demethylation, are more active in the primed state of pluripotency, or both. Following on from this, I wanted to understand the contribution of the variant PRC1.3 complex to the overall levels of H2AK119ub1 and H3K27me3, as well as broader changes to the epigenome. This is not possible in naïve PSCs due to the inherent failure of *PCGF3* KO cells to reprogramme and my inability to generate a naïve knockdown or knockout line, and so this study was limited to primed PSCs. Prior to examining the specific impact of PRC1.3, I studied

the relative contributions of variant and canonical PRC1 complexes in the catalysis of H2AK119ub1 in primed PSCs. To achieve this, I used an EZH1/2 inhibitor, that inhibits the catalytic activity of PRC2 complexes, named UNC1999 (Konze *et al.*, 2013). UNC1999 was applied to primed PSCs for 4 days, due to the slow turnover of H3K27me3 on chromatin, which dilutes out with each DNA replication and cell division (Jadhav *et al.*, 2020). UNC1999 treatment was able to ablate H3K27me3 from chromatin and thus remove the substrate for the recruitment of canonical, but not variant, PRC1 complexes. In mESCs, H2AK119ub1 was shown to have a very high rate of turnover on chromatin (Dobrinić *et al.*, 2021; Huseyin and Klose, 2020), and if this was conserved in human PSCs, any canonically-deposited H2AK119ub1 would be removed and not re-established, leaving only variant PRC1-deposited H2AK119ub1 on chromatin. When H3K27me3 is removed, some, but critically not all, H2AK119ub1 is lost from chromatin (Fig. 4.5.3).

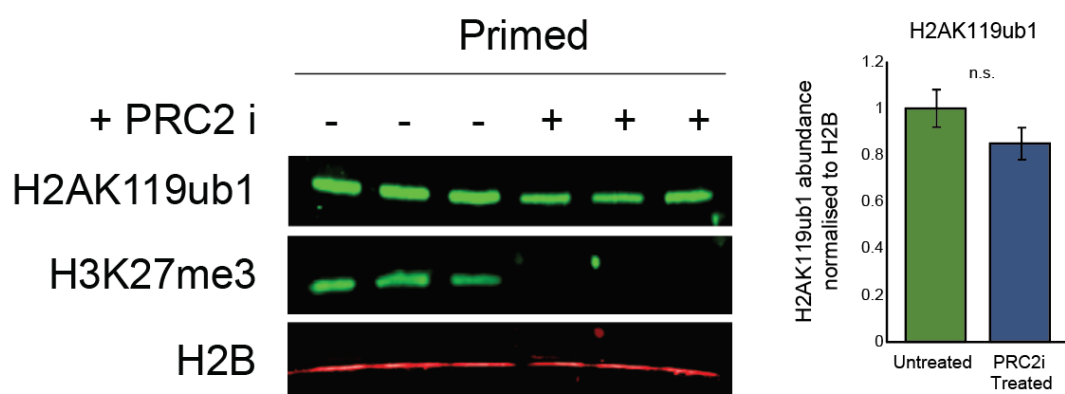


Figure 4.5.3 – Western blotting shows the abundance of H2AK119ub1, with and without PRC2 inhibition, in primed PSCs

H3K27me3 confirms the action of the PRC2 inhibitor. Histone H2B, loading control. Three independent biological replicates are shown for PRC2 inhibitor untreated and treated cells. Quantification of H2AK119ub1 signal is shown in the bar chart, with bar height representing sample mean and error bars showing the standard error of the mean. No statistical difference determined by paired t-test.

This finding suggests that both canonical and variant PRC1 complexes contribute to H2AK119ub1 levels in human primed PSCs, however there is no statistically significant difference in H2AK119ub1 levels between the two conditions. As variant PRC1 complexes can contribute to H2AK119ub1 catalysis in human primed PSCs, PRC1.3 complex could play an important epigenetic role in this cell type, despite the absence of strong transcriptional

differences between WT and *PCGF3* KO primed PSCs. For example, in mESCs, the combined deletion of *Pcgf3* and *Pcgf5* resulted in a loss of approximately half of total H2AK119ub1 levels and PRC1.3/5 complexes have been demonstrated to have a high level of catalytic activity (Fursova *et al.*, 2019; Huseyin *and* Klose, 2020). I tested the global abundance of H2AK119ub1 in DKO human primed PSCs to determine the contribution of PRC1.3 to overall PRC1 catalytic activity. I saw that PRC1.3 deletion causes a strong reduction in the abundance of the H2AK119ub1 modification by Western blot (Fig. 4.5.4).

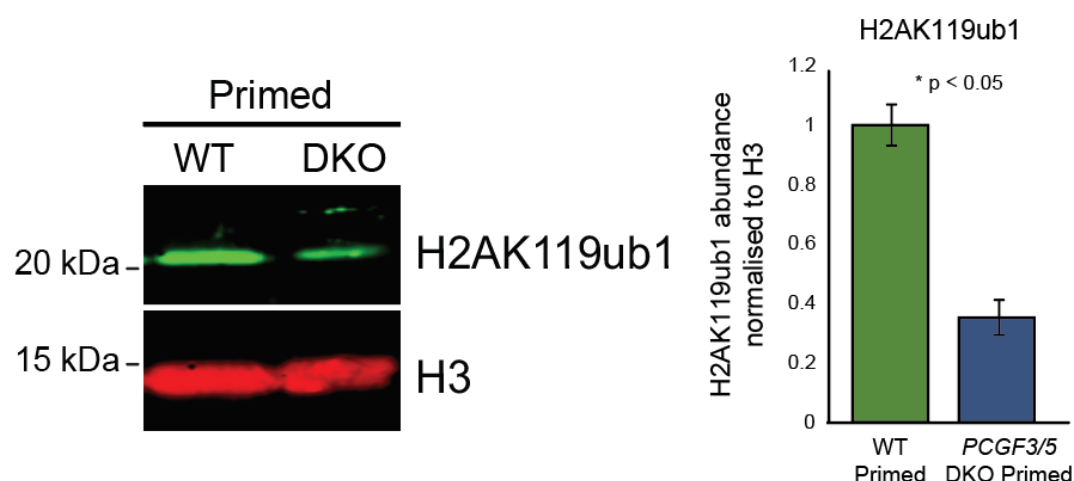


Figure 4.5.4 – Western blotting shows the abundance of H2AK119ub1 in WT and *PCGF3/5* DKO primed PSCs

Histone H3, loading control. Data are representative of three independent experiments. Quantification of H2AK119ub1 signal is shown in the bar chart, with bar height representing sample mean and error bars showing the standard error of the mean. Statistical difference determined by paired t-test.

I also explored my calibrated CUT&Tag data for H2AK119ub1 in naïve, and WT and DKO primed PSCs, and corroborated both that H2AK119ub1 is more abundant in naïve than primed PSCs, and that H2AK119ub1 levels are reduced in DKO primed PSCs (Fig. 4.5.5).

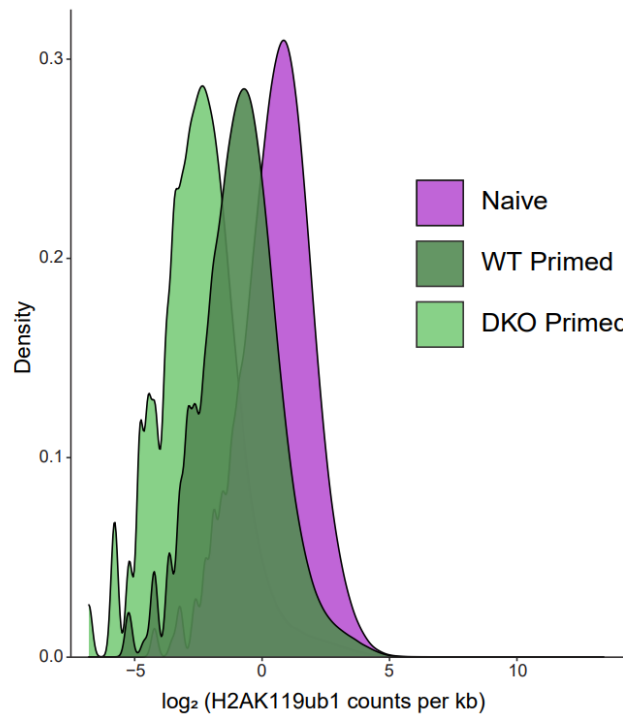


Figure 4.5.5 – Kernel density estimate of normalised H2AK119ub1 CUT&Tag reads in naïve, WT and DKO primed PSCs

The genome was divided into 1 kb bins and the $\log_2$ -transformed number of H2AK119ub1 reads in each bin was quantitated. Two independent biological replicates of naïve, WT and DKO primed PSCs. Plot produced by Andrew Malcolm.

I decided to investigate the distribution of PRC1.3-dependent H2AK119ub1 in primed PSCs, in particular due to the lack of transcription effect of PRC1.3 deletion. Across an autosomal chromosome, I saw that H2AK119ub1 is more abundant across the entire chromosome in naïve PSCs compared to primed PSCs. PRC1.3 contributes to H2AK119ub1 across the entire length of the chromosome in primed PSCs, and suggests that PRC1.3 might contribute to in particular to global H2AK119ub1 deposition (Fig. 4.5.6).

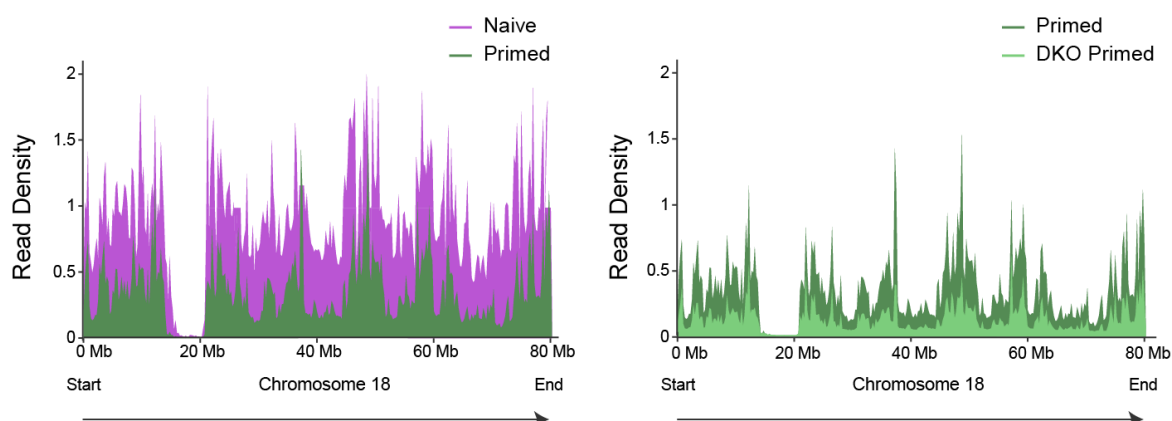


Fig. 4.5.6 – A plot showing coverage of H2AK119ub1 CUT&Tag reads across chromosome 18

The genome was divided into 250 kb bins and the averaged read density is shown for each bin across the length of chromosome 18. Plot produced by Andrew Malcolm.

Due to the strong impact of PRC1.3 deletion on global H2AK119ub1 abundance, I profiled the impact of PRC1.3 deletion on the wider epigenome using histone proteomics. As JARID2-containing PRC2.2 complexes are recruited to chromatin by H2AK119ub1 and catalyse H3K27me₃, PRC1.3/5 deletion might be expected to strongly impact the trimethylation of lysine 27 on histone H3 in particular. Deletion of PRC1.3 from primed PSCs shows dysregulation of the abundance of post-translational modifications on lysine 27 of histone H3 (Fig. 4.5.7).

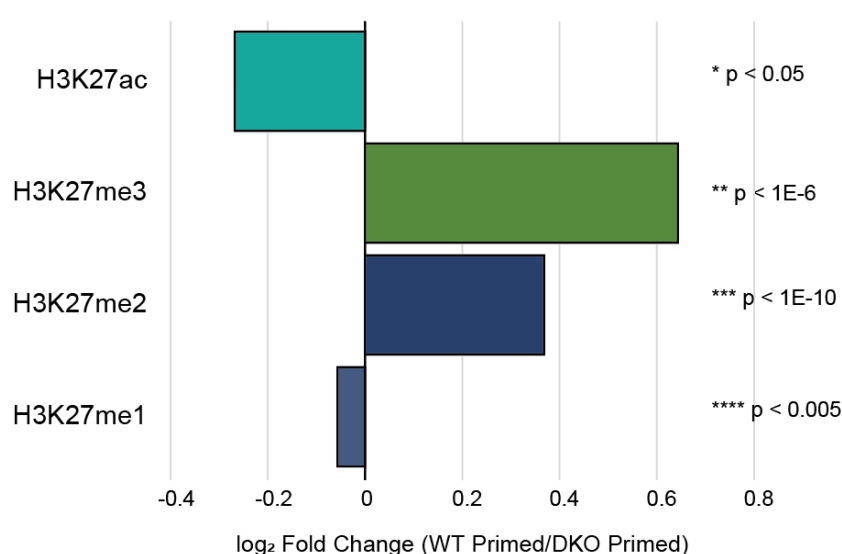


Figure 4.5.7 – Histone proteomics show the relative abundance of the H3K27me1, H3K27me2, H3K27me3 and H3K27ac modifications in WT and DKO primed PSCs

Bar length represents log2 fold change between WT and DKO primed PSCs. P-value calculated by student's t-test. Performed on five (WT Primed) or six (DKO Primed) independent biological replicates. Sample processing, mass spectrometry and analysis performed by Sigrid Verhelst and Maarten Dhaenens.

In PRC1.3-deficient primed PSCs, the levels of both H3K27me2 and H3K27me3 are significantly reduced. H3K27me2 has been reported in mESCs to be a diffuse modification covering the majority of histone H3 on chromatin to prevent aberrant acetylation events, which are mutually exclusive with histone methylation at this residue (Ferrari *et al.*, 2014). Consistent with this, levels of H3K27ac increase in the absence of PRC1.3, albeit remaining much lower in abundance than H3K27 methylation. Loss of H3K27me2 and H3K27me3 is accompanied by an increase in levels of H3K27me1. The significance of the increase in abundance of H3K27me1 under these conditions is unknown (Ferrari *et al.*, 2014).

Beyond simply lysine 27 of histone H3, I also compared the fold change in abundance of all detected histone H3 and H4 post-translational modifications between WT and DKO primed PSCs. While H3K27me2 and H3K27me3 are the most differentially enriched histone modifications in WT primed PSCs, other modifications show small but significant differences between WT and DKO primed. Monomethylated lysine 36 on histone H3 (H3K36me1) is elevated in WT PSCs and is associated with the introns of actively expressed gene bodies (Dhami *et al.*, 2010). In DKO primed PSCs, there are higher levels of the poorly understood crotonylation of lysine 12 on histone H4 (H4K12Cr) (Li and Wang, 2021) and the acetylation of lysine 16 on histone H4 (H4K16Ac), associated with active genes and enhancer elements (Taylor *et al.*, 2013) (Fig. 4.5.8).

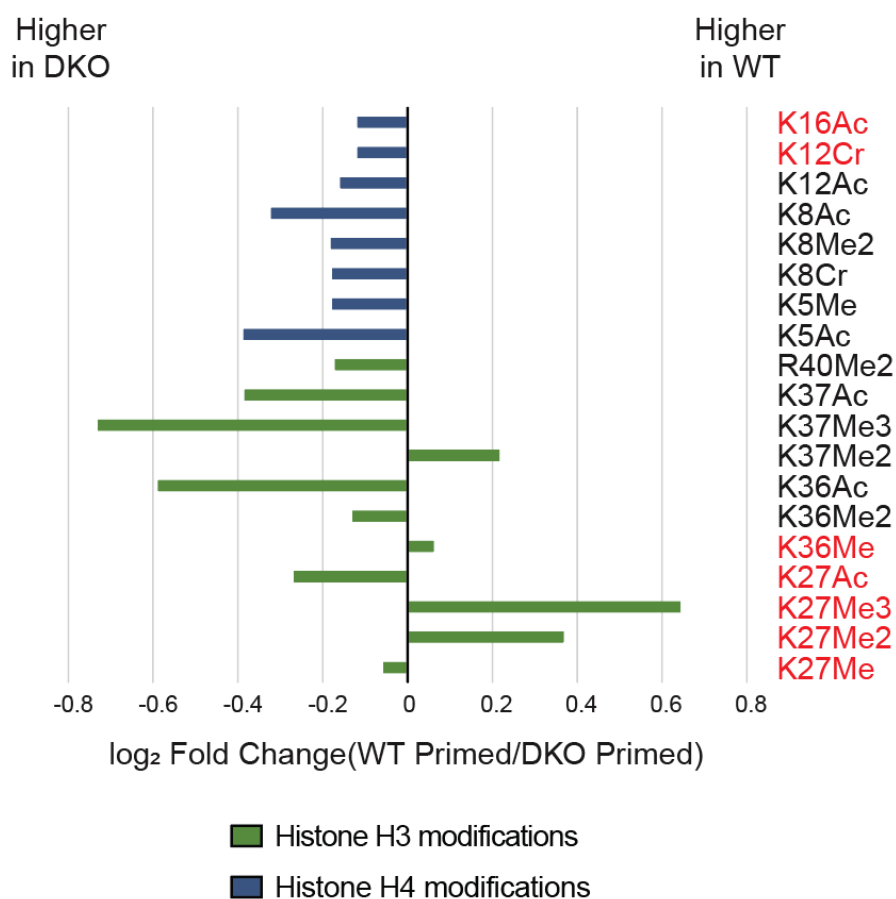


Figure 4.5.8 – Histone proteomics shows the relative abundance of all detected histone H3 and H4 modifications detected in WT and DKO primed PSCs

Bar length represents log₂ fold change between WT and DKO primed PSCs. P-value calculated by student's t-test and modifications with a $p < 0.05$ are highlighted in red. Performed on five (WT Primed) and six (DKO Primed) independent biological replicates. Sample processing, mass spectrometry and analysis performed by Sigrid Verhelst and Maarten Dhaenens.

In summary, a large proportion of H2AK119ub1 in primed PSCs is catalysed by PRC1.3 complexes, and a large proportion of H3K27me3 is catalysed by PRC2.2 complexes that depend upon PRC1.3-deposited H2AK119ub1 for their recruitment to chromatin. PRC1.3-deficiency also affects the abundance of other epigenetic modifications in human primed PSCs, to a lesser degree.

4.5.3 - PRC1.3-deficiency influences the histone H2A variants present within nucleosomes

Along with histone post-translational modification proteomics, I analysed the proteomics of acid extracted proteins that bind to DNA with a high affinity, including histone proteins themselves. In particular, I saw that the composition of variants of histone H2A, which is the enzymatic substrate for PRC1 complexes, were shifted between WT and DKO primed PSCs. The histone H2A variants H2AW and H2AY are more abundant in DKO primed PSCs while H2AZ and H2AV variants are more abundant in WT primed PSCs (Fig. 4.5.9).

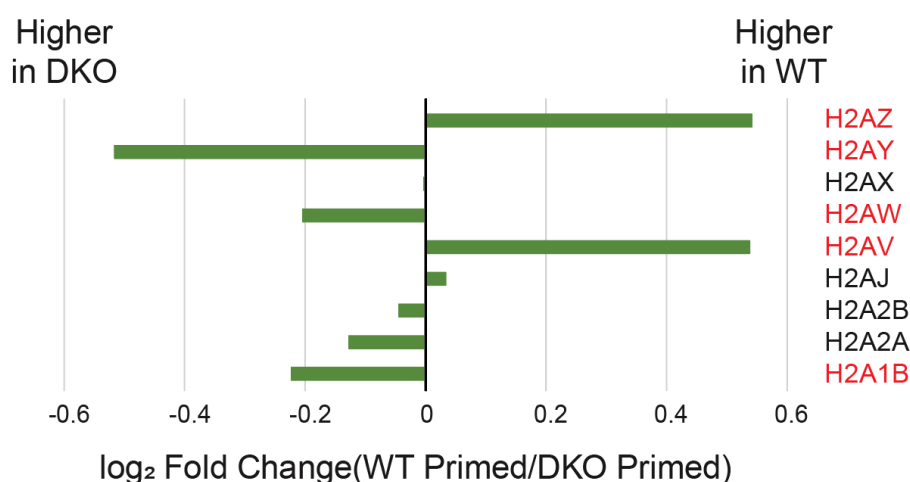


Figure 4.5.9 – Histone proteomics shows the relative abundance of histone H2A variants in WT and DKO primed PSCs

Bar length represents log₂ fold change between WT and DKO primed PSCs. P-value calculated by student's t-test and histone variants with a $p < 0.05$ are highlighted in red. Performed on five (WT Primed) and six (DKO Primed) independent biological replicates. Sample processing, mass spectrometry and analysis performed by Sigrid Verhelst and Maarten Dhaenens.

This provides a novel insight into how PRC1.3 complex function influences the overall chromatin dynamics, either directly or indirectly. Further experimentation is required to better understand this phenomenon and the link between epigenetic modifiers and the abundance of different histone variants within nucleosomes.

4.6 – Discussion

4.6.1 – WT and *PCGF3* KO cells transcriptionally diverge during PTNR

WT and *PCGF3* KO primed PSCs are remarkably similar transcriptionally (Chapter 3), but after 10 days of reprogramming, their gene expression profiles diverge. This could have been assessed further using an RNA-sequencing time-course with finer sampling through the first ten days of reprogramming. This would have increased the level of dynamic gene expression information and helped to establish the earliest transcriptional differences between WT and *PCGF3* KO cells during reprogramming, prior to detectable phenotypic differences.

A limitation of the reprogramming RNA-sequencing dataset I produced was that it contained a heterogeneous population of cells that could not be sorted, and were sequenced in bulk. This dataset could be improved using a single-cell RNA-sequencing approach (Hwang *et al.*, 2018), which would transcend the heterogeneity of the populations by identifying different subpopulations, both abundant and rare, rather than performing ensemble averaging. Assessing more time-points and using single-cell sequencing across the first ten days of reprogramming could identify the timing of the earliest PRC1.3-dependent transcriptional changes and determine the impact of these divergent transcriptional changes on the cell types and within transcriptional subpopulations.

It is unclear whether the PRC1.3 complex is essential specifically for human PTNR, as identified by our prior genetic screen, or whether the complex is a central transcriptional regulator that also affects other pluripotency cell state transitions. This would be consistent with literature in which mESCs that are lacking PRC1.3/5 are deficient in differentiation to certain differentiated cell types (Gao *et al.*, 2014; Yao *et al.*, 2018; Zhao *et al.*, 2017). However, it is unlikely that all cell state transitions are impacted, as the *PCGF3* KO cells appear adept at differentiation into neural-like cells. Differentiation of *PCGF3* KO primed PSCs, either towards a specific directed lineage, or in a more indirect manner using embryoid bodies or gastruloids, could determine whether PRC1.3 plays a broader role in cell state transitions and reconfiguration of cellular identity. Nevertheless, the PRC1.3 complex is indeed essential for the generation nascent naïve PSCs.

4.6.2 – PRC1.3 naïve target genes are repressed during reprogramming

ChIP-sequencing, calibrated CUT&Tag and RNA-sequencing identified genes that gain PRC1.3 binding throughout reprogramming, gain the H2AK119ub1 modification and become transcriptionally repressed during successful reprogramming in a PRC1.3-dependent manner. This is important for understanding PRC1.3 function in coordinating the acquisition of naïve pluripotency. Fewer than 200 genes were identified as being naïve PRC1.3 target genes, and this may be in part due to the very fast chromatin binding kinetics of PRC1.3 identified in mESCs (Huseyin *and* Klose, 2020), meaning that transient interactions with chromatin are not detected within the fixation time of ChIP-sequencing. The gene targets detected may represent the longest residency interactions but knowledge of the identity of these genes is nevertheless highly informative.

Naïve-specific PRC1.3 target genes include *TCF7L1*, *DUSP6*, *B-RAF* and *AUTS2*. *TCF7L1* regulates Wnt signalling-dependent transcriptional regulation, with Wnt signalling activated by many reprogramming conditions including 5i/L/A (Sierra *et al.*, 2018; Takashima *et al.*, 2014; Theunissen *et al.*, 2014). *DUSP6* is a primed-specific marker gene and pluripotency factor whose expression levels closely correlate with ERK activity, which is specifically inhibited in naïve PSCs (Theunissen *et al.*, 2014; Yoo *et al.*, 2020). *B-RAF* is a component of the MEK/ERK signalling pathway that is specifically inhibited in 5i/L/A reprogramming conditions (Theunissen *et al.*, 2014) and *AUTS2* is a component of the PRC1.3 complex (Gao *et al.*, 2012; Hauri *et al.*, 2016), suggesting a degree of autoregulation of the complex, which will be discussed further in chapter 5.

A number of the naïve PRC1.3 target genes have functions that could feasibly make their repression important for the acquisition of naïve pluripotency and whose continued expression may impede reprogramming. Further analysis is required to develop a comprehensive understanding of the key PRC1.3-dependent transcriptional effects during reprogramming. This in turn could determine PRC1.3-dependent remodelling of the reprogramming transcriptional networks which could then establish hierarchy, and elucidate a sequence of events that must be followed for nascent naïve PSCs to form. It is important to note that a limitation of my experimental approach is that I am using established naïve PSCs to predict where the PRC1.3 complex binds to chromatin in nascent naïve PSCs, and this assumption may not be correct. However, due to the inefficiency of reprogramming and the large amount of cell death during reprogramming, obtaining sufficient cell numbers to perform ChIP-sequencing on sorted nascent naïve PSCs on day 10 of reprogramming simply would not be possible. The best way to approach this type of experiment with low cell numbers would have been to instead test the compatibility of the PCGF3/5 and RING1B antibodies with CUT&RUN (Cleavage Under Targets and Release Using Nuclease) (Meers *et al.*, 2019; Skene *and* Henikoff, 2017) or CUT&Tag approaches (Kaya-Okur *et al.*, 2019).

To comprehensively determine the critical downstream effects of PRC1.3 in reprogramming, a screening approach could be used. As the PRC1.3 complex is associated with transcriptional repression, this screen would utilise either a CRISPR knockout or CRISPR-interference method. Silencing of a key reprogramming regulatory gene, ordinarily silenced by PRC1.3, would enable PRC1.3-independent reprogramming and would allow an understanding of which target genes are important for PRC1.3 to repress during reprogramming.

4.6.3 – The PRC1.3 complex influences the primed PSC epigenome

The PRC1.3 complex contributes to global H2AK119ub1 levels, and subsequently H3K27me3 levels in human primed PSCs. This provides a further example in which PRC2 complexes and H3K27me3 act downstream of H2AK119ub1 deposited by variant PRC1 complexes. However, despite these differences to the epigenome, PRC1.3 deletion has little effect on gene transcription in steady-state primed PSCs.

A possibility for the disconnect between the global H2AK119ub1 levels and the transcriptome is that as only a small number of genes were identified as being PRC1.3 bound, PRC1.3-dependent H2AK119ub1 and H3K27me3 may rarely be associated with gene promoters, particularly as loss of H2AK119ub1 from gene promoters has been reported to cause rapid gene activation in mESCs (Dobrinić *et al.*, 2021; Stock *et al.*, 2007). This would be consistent with the identification of genome-wide blanketing of PRC1.3/5 dependent H2AK119ub1 and the lack of transcription impact of PRC1.3/5 deletion observed in mESCs (Fursova *et al.*, 2019). Instead, PRC1.3-dependent H2AK119ub1 and H3K27me3 could blanket the genome at low levels in human primed PSCs. This is supported by the unconvincing appearance of the PCGF3 ChIP-sequencing data in primed PSCs, suggesting a lack of gene-specific targeting of PRC1.3 in human primed PSCs, which was more apparent in naïve PSCs. H2AK119ub1 and H3K27me3 show diffuse genome-wide distributions across chromatin in naïve PSCs (Kumar *et al.*, 2022; Zijlmans *et al.*, 2022), and my data shows that in primed PSCs, this can be seen for H2AK119ub1 also, albeit to a lesser extent, in a manner that is partially PRC1.3-dependent.

As well as H2AK119ub1 and H3K27me3, PRC1.3 deletion showed effects on other histone post-translational modifications and chromatin-associated proteins. Quantitative chromatin profiling could therefore be extended to study the effects of PRC1.3 deletion on H3K27me1, H3K27me2 and H3K27ac, and H2A specific histone variants also. Histone H2A variants are known to play key roles in chromatin regulation (Shaytan *et al.*, 2015), such as X chromosome inactivation (Costanzi and Pehrson, 1998). Performing these experiments in PRC1.3 KO

primed PSCs and in a naïve PRC1.3 knockdown would improve our understanding of the epigenetic regulation of human pluripotency and reprogramming, which could provide insight into the epigenetic regulation of human epiblast cells during embryonic development also.

As primed PSCs reprogramme to naïve PSCs, H2AK119ub1 and H3K27me3 must be accumulated across the genome at high abundance, but the timing of this epigenetic remodelling throughout reprogramming is not known. Although I have identified a potential role for the PRC1.3 complex in repressing target gene expression during reprogramming, the crucial role of the complex could instead be linked to the establishment of the naïve PSC epigenome during reprogramming. As the only Polycomb complex identified in the genome-wide screen, which is also associated with genome-wide H2AK119ub1 deposition in mESCs, PRC1.3 could act as a strong candidate for epigenetic remodelling across this cell state transition. I have established that DKO primed PSCs are already deficient in genome-wide H2AK119ub1 and a failure to sufficiently accumulate these modifications genome-wide across chromatin during reprogramming could result in a failure in acquisition of naïve pluripotency.

Genome-wide H2AK119ub1 and H3K27me3 are hypothesised to repress repetitive and transposable elements in DNA hypomethylated naïve PSCs (Cooper *et al.*, 2014; Pastor *et al.*, 2016; van Mierlo *et al.*, 2019a; Zijlmans *et al.*, 2022). PRC1.3 KO primed PSCs might aberrantly express repetitive or non-coding elements when DNA methylation is removed during reprogramming, and the timing of DNA methylation dynamics in PTNR is also not known. This could prove fatal for cells and specifically select against the formation of a hypomethylated cell state during reprogramming. This possibility would not have been identified in my RNA-sequencing data, which specifically enriches for polyadenylated messenger RNAs. My hypothesis could be tested by assessing the ability of *PCGF3* KO primed PSCs to reprogramme in alternative reprogramming conditions which have reported, at least partially, to retained a DNA hypermethylated state within naïve PSCs (Lynch *et al.*, 2020). I hypothesise that if genome-wide, rather than gene specific transcriptional repression by PRC1.3 is important for human PTNR, PRC1.3 KO cells would reprogramme at much higher efficiency under conditions where DNA methylation remains intact for repression of constitutive heterochromatin regions.

5 – PRC1.3 undergoes pluripotency state-specific regulation between human naïve and primed PSCs

5.1 – Background

5.1.1 – Investigating the PRC1.3 protein interaction network in human PSCs

PRC1.3 is dispensable for human primed pluripotency (as shown in the chapter 3), despite considerable epigenetic differences between WT and PRC1.3 KO primed PSCs (as shown in chapter 4). However, PRC1.3 is essential for human PTNR (as shown in chapter 3) and undergoes considerable redistribution on chromatin between primed and naïve PSCs (as shown in chapter 4), binding a unique subset of genes in naïve PSCs which are transcriptionally downregulating during PTNR. These findings indicate that the PRC1.3 complex has acquired novel functionality during reprogramming and that the recruitment and targeting mechanisms that govern PRC1.3 binding to chromatin differ between the two states.

None of the components of the PRC1.3 complex have characterised DNA-binding ability and previous reports have shown the existence of context-specific protein-protein interactions with the PRC1.3 complex, which can regulate chromatin recruitment and the function of the complex. Upon induction of *Xist* expression from the inactive X chromosome (Xi) in mESCs, the PRC1.3 complex interacts with the RNA binding protein hnRNPK (Pintacuda *et al.*, 2017). This interaction mediates PRC1.3/5 recruitment to *Xist*-coated regions along Xi in female cells, where it catalyses the H2AK119ub1 to facilitate early gene silencing (Almeida *et al.*, 2017; Bousard *et al.*, 2019; Bowness *et al.*, 2022; Nesterova *et al.*, 2019; Pintacuda *et al.*, 2017). An additional recruitment factor identified in mESCs was USF1, which interacts with PRC1.3 and whose knockdown causes a loss of PRC1.3 binding to target genes (Scelfo *et al.*, 2019). However, the recruitment factor or mechanism for PRC1.3 binding to chromatin in human PSCs has never been identified.

Wider interactions with PRC1.3 are also able to modulate the function of the complex. AUTS2-PRC1.3 complexes can activate transcription of key neurodevelopmental genes in mice, promoted by an interaction between PRC1.3 and the histone acetyltransferase complex P300 (Gao *et al.*, 2014). The interaction between PRC1.3 and P300 is proposed to be bridged by the protein TEX10, and also enables PRC1.3-dependent activation of mesoderm lineage genes in mESC-derived embryoid bodies (Zhao *et al.*, 2017). Wider functional influence of protein interactions with PRC1.3 have been largely reported in the transcriptional activation by PRC1.3 as the core complex does not innately possess the ability to activate transcription. However, I have not seen any evidence to support PRC1.3-dependent transcriptional activation in human pluripotency and reprogramming. PRC1.3 may not require wider

interactions for the H2AK119ub1 catalysing enzymatic function of the complex, but they may instead be required for targeting of the complex to chromatin.

I hypothesise that the distinct function and distribution of PRC1.3 on chromatin between human primed and naïve PSCs may be due to differences in the protein interaction network of the complex. Pluripotency state-specific interactions could allocate distinct chromatin-targeting, functional or catalytic properties to PRC1.3 during reprogramming. In this chapter, I will identify PRC1.3 complex interactors in primed, reprogramming and naïve human PSCs using Quantitative, Multiplexed Rapid Immunoprecipitation Mass Spectrometry of Endogenous protein (qPLEX-RIME) (Mohammed *et al.*, 2016; Papachristou *et al.*, 2018). This technique allows profiling of multiple samples, using tandem mass tag labelling, to provide dynamic, quantitative information about the interactors of a protein of interest between pluripotency states (Papachristou *et al.*, 2018).

As I am interested by the manner in which the PRC1.3 complex facilitates human PTNR, I will filter interacting proteins that are identified by cross-referencing with the genes that were identified in the genome-wide screen as being essential for PTNR. This should identify a small number of candidates for state-specific interactions with the PRC1.3 complex, which can then be mechanistically interrogated to gain insight into whether the interaction influences the function or recruitment of PRC1.3 during reprogramming.

5.1.2 – Investigating the PRC1.3 complex catalytic activity in primed and naïve PSCs

In Chapter 4, I identified that the H2AK119ub1 and H3K27me3 modifications are more abundant in naïve than primed PSCs, and reports have demonstrated the differing distribution of these modifications between the two states (Kumar *et al.*, 2022; Zijlmans *et al.*, 2022). I also identified that PRC1.3 is responsible for a large proportion of global H2AK119ub1, and consequently H3K27me3, levels in primed PSCs, and PRC1.3 could be equally, if not more, important in the Polycomb regulation of human naïve pluripotency, where PRC1.3 binds different targets on chromatin. Chromatin proteomics has previously identified that the abundance of most PRC1.3 complex components, PRC2 components and the H2AK119 deubiquitinase BAP1 do not significantly differ between human naïve and primed PSCs (Zijlmans *et al.*, 2022). This therefore leads me to hypothesise that the increased H2AK119ub1 and H3K27me3 on chromatin is due to an increased catalytic activity of H2AK119ub1-catalysing enzymes, including potentially PRC1.3 on chromatin. However, the catalytic activity

of PRC1.3, and the influence of any pluripotency state-specific interactors, has not been characterised to date.

5.1.3 – Aims

A better understanding of the regulatory interaction network surrounding the PRC1.3 complex is needed to reconcile the epigenetic and transcriptional changes associated with the PRC1.3 complex in human pluripotency and cell state transitions. To this end, qPLEX-RIME was performed to identify state-specific interactors with the PRC1.3 complex component PCGF3 between primed, early reprogramming and naïve PSCs. Candidates of interest will be individually interrogated to determine the mechanistic role that they might play in regulating PRC1.3 function or recruitment to chromatin.

1. Perform qPLEX-RIME for the PRC1.3 complex across the reprogramming cell state transition.
2. Interrogate state-specific PRC1.3 qPLEX-RIME interactors with possible regulatory functions, such as acquisition of novel properties or recruitment to chromatin.
3. Test the catalytic activity of the PRC1.3 complex associated with human primed and naïve pluripotency.

5.2 – Using qPLEX-RIME to identify PRC1.3 complex interactors

5.2.1 – Optimisation of the qPLEX-RIME methodology in human PSCs

Rapid Immunoprecipitation Mass Spectrometry of Endogenous Protein (RIME) allows the study of protein-protein interactions on chromatin for a specific protein or complex of interest. This technique works by fixing cells, which are then lysed and the nuclei are extracted. Chromatin is then sonicated and an immunoprecipitation is performed using an antibody against the protein of interest. Following stringent washing, the proteins that interact with the complex of interest are identified by mass spectrometry (Mohammed *et al.*, 2016). Multiple samples from a number of different conditions can be combined by multiplexed tandem mass tag labelling, which allows quantitative comparison between samples – known as Quantitative Multiplexed RIME (qPLEX-RIME) (Papachristou *et al.*, 2018). This approach is ideal for understanding the regulation of the PRC1.3 complex in multiple conditions, such as across reprogramming. qPLEX-RIME could identify state-specific co-interactors, although not necessarily direct interactors, of the PRC1.3 complex with statistical significance, such as a novel protein that may recruit the PRC1.3 complex to distinct sites on chromatin between primed and naïve PSCs.

A qPLEX-RIME experiment was planned using 4 replicates of primed PSCs, 4 replicates of early reprogramming cells, after 48 h in 5i/L/A reprogramming conditions, 4 replicates of naïve PSCs and 4 replicates of *PCGF3/5* DKO primed PSCs as a negative control (16 multiplexed samples in total). This study was performed in collaboration with the Cancer Research UK Cambridge Proteomics Facility. I first needed to optimise the RIME experimental protocol, which has never before been conducted in human PSCs, using the anti-*PCGF3/5* antibody that I have used previously. It was necessary to grow human PSCs on mouse embryonic fibroblasts (MEFs) for this experiment to produce data that best recapitulated primed and early reprogramming cells that were used in 5i/L/A reprogramming experiments. I was concerned that by growing cells on MEFs, I may enrich for mouse derived *PCGF3/5* proteins if MEFs express these proteins, as well as also detecting mouse-derived contaminating proteins.

I therefore sought to determine whether mass spectrometry could distinguish human and mouse *PCGF3* protein. Amino acid sequence alignment of human and mouse *PCGF3* protein shows a 97 % sequence similarity and therefore strong conservation between the two species (Fig. 5.2.1).

```

CLUSTAL O(1.2.4) multiple sequence alignment

SP|Q3KNV8|PCGF3_HUMAN  MLTRKIKLWDINAHITCRLCSGYLIDATTVTECLHTFCRSLVKYLEENNTCPTCRIVIH  60
SP|Q8BTQ0|PCGF3_MOUSE  MLTRKIKLWDINAHITCRLCSGYLIDATTVTECLHTFCRSLVKYLEENNTCPTCRIVIH  60
                        *****

SP|Q3KNV8|PCGF3_HUMAN  QSHPLQYIGHDRTMQDIVYKLVPGQEAEMRKQREFYHKLGMVPGDIKGETCSAKQHLD  120
SP|Q8BTQ0|PCGF3_MOUSE  QSHPLQYIGHDRTMQDIVYKLVPGQEAEMRKQREFYHKLGMVPGDIKGEACSAKQHLD  120
                        *****

SP|Q3KNV8|PCGF3_HUMAN  SHRNGETKADDSSNKEAAEEKPEEDNDYHRSDEQVSICLECNSSKLRGLKRWIRCSAQA  180
SP|Q8BTQ0|PCGF3_MOUSE  P-RNGETKADDNSNKETAEEKQEEDNDYHRSDEQVSICLECNSSKLRGLKRWIRCSAQA  179
                        *****

SP|Q3KNV8|PCGF3_HUMAN  TVLHLKKFIAKKLNLSSFNELDILCNEEILGKDHTLKFVVVTRWRFKKAPLLLHYRPFKMD  240
SP|Q8BTQ0|PCGF3_MOUSE  TVLHLKKFIAKKLNLSSFNELDILCNEEILGKDHTLKFVVVTRWRFKKAPLLLHYRPFKMD  239
                        *****

SP|Q3KNV8|PCGF3_HUMAN  LL  242
SP|Q8BTQ0|PCGF3_MOUSE  LL  241
                        **

```

Figure 5.2.1 – Amino acid multiple sequence alignment shows similarity between human and mouse PCGF3

Human PCGF3 amino acid sequence, top; mouse PCGF3 amino acid sequence, below. Identical amino acids denoted by asterisk (*) seen across the vast majority of the residues in the proteins.

Due to the similarity in amino acid sequence, I predicted that the anti-PCGF3/5 antibody might detect both mouse and human PCGF3 and PCGF5. I therefore conducted a pilot experiment in which the PRC1.3 RIME protocol was tested on WT and *PCGF3/5* DKO primed PSCs, both grown on MEFs. The aim was to determine whether the protocol, adapted for use with human naïve and primed PSCs, was working well, and to also identify whether the protocol was only detecting human-derived PCGF3. This experiment identified PCGF3-specific peptides in both WT and DKO primed PSCs, with detection of a peptide sequence that is identical between human and mouse. Detection of this peptide in DKO primed PSCs suggests that MEFs might express PCGF3, and that this is detected by the anti-PCGF3/5 antibody (Fig. 5.2.2).

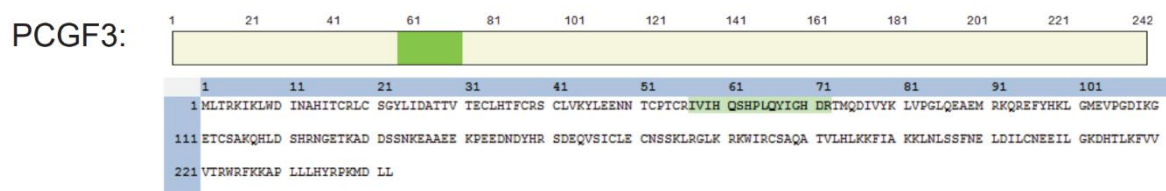


Figure 5.2.2 – RIME shows PCGF3 peptides can be detected in *PCGF3/5* DKO primed PSCs grown on MEFs

Green highlighted region shows peptide detected by RIME. Comparison with Fig. 5.2.1 shows that this peptide sequence is identical in human and mouse PCGF3.

Using immunofluorescence microscopy of human primed PSCs, grown on MEFs, with the anti-PCGF3/5 antibody, I confirmed that MEFs express PCGF3 or PCGF5 to a level that can be detected by the anti-PCGF3/5 antibody (Fig. 5.2.3).

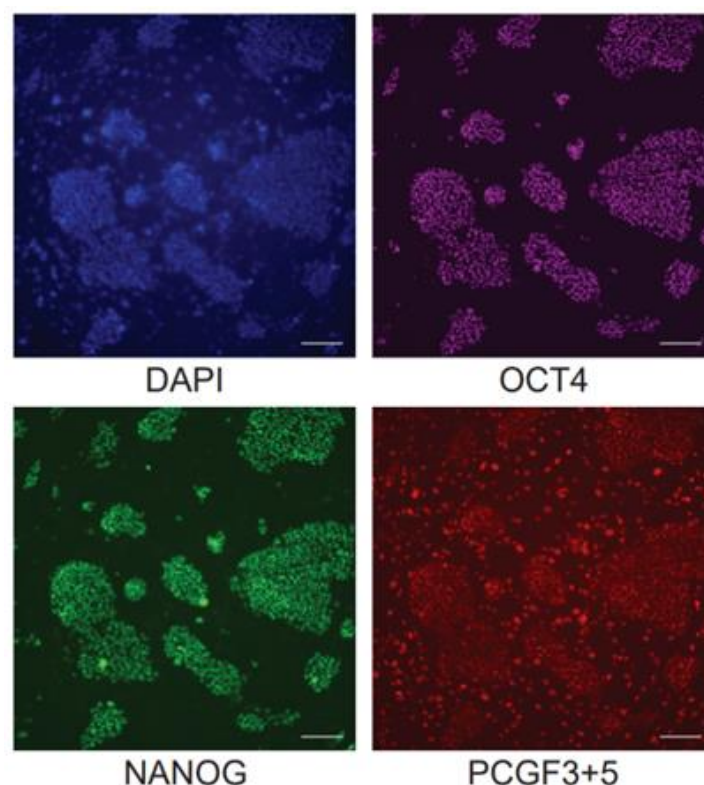


Figure 5.2.3 – Immunofluorescence detects PCGF3/5 in MEFs and human primed PSCs

DAPI stains all nuclei, OCT4 and NANOG stain pluripotent primed PSCs. PCGF3/5 stains primed PSCs and MEFs. Scale bar = 250 μ m.

To overcome this experimental issue, I used magnetic activated cell sorting (MACS) to selectively deplete MEFs from samples. Using a biotinylated antibody against the MEF-specific cell-surface marker Cd90.2 (Ishihara *et al.*, 1987), which is not expressed on human PSCs, allows specific labelling of MEFs. Biotin labelled MEFs are then bound by Streptavidin microbeads and retained on a magnetic column, while human PSCs that do not express Cd90.2 pass through the column into the flow-through. To assess the fidelity of MACS in MEF

depletion, flow cytometry was performed. Unpurified cells possess a mixed population of CD147 expressing (human) and Cd90.2 (mouse) expressing cells, as well as double negative cells which is likely an antibody staining defect. In the flow-through cells, MEFs are depleted from 20 % to less than 1 % of the cell population, while cells retained on the column are enriched for MEFs, but some human cells are lost also (Fig. 5.2.4).

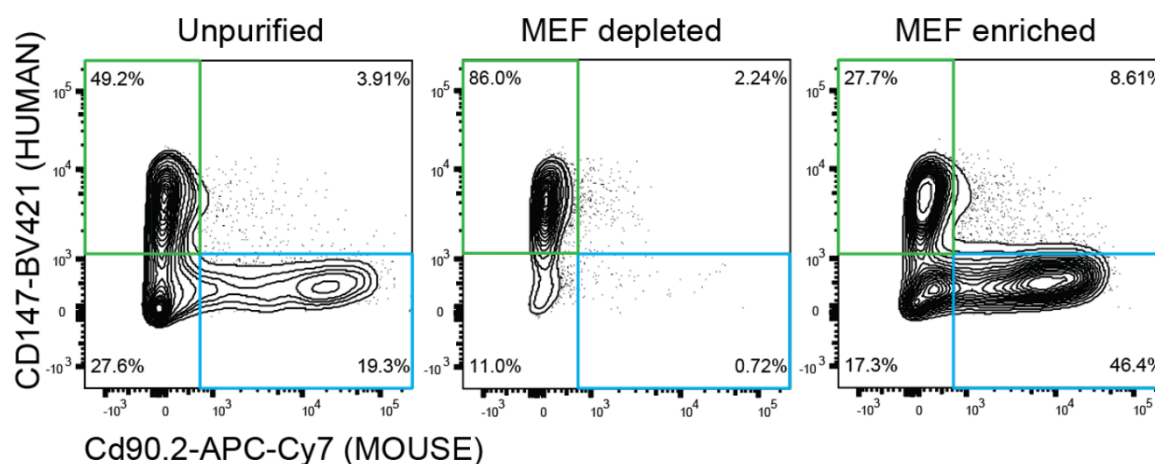


Figure 5.2.4 – Flow cytometry confirms the MACS-dependent depletion of Cd90.2 positive MEFs and enrichment of CD147 positive human cells

MEFs are positive for Cd90.2 and negative for CD147 (lower right quadrant, blue), while human primed PSCs are positive for CD147 and negative for Cd90.2 (top left quadrant, green).

Although some human cells are lost using MEF depletion, I decided to prioritise the purity of the human cells in the samples over cell yield, and this methodology was integrated into the qPLEX-RIME protocol.

5.2.2 – qPLEX-RIME can identify the PRC1.3 complex

The full 16-plex qPLEX-RIME protocol was subsequently performed, in primed and DKO primed PSCs, naïve PSCs and 48 h 5i/L/A reprogramming cells. The 48 h 5i/L/A reprogramming time-point was selected to determine whether any changes in the PRC1.3 complex interactome are detected, as these cells remain homogeneous and primed-like but have been treated with reprogramming signalling factors for 48 h. This can help to identify

whether state-specific interactors detected between primed and naïve PSCs are due to changes in the external signalling conditions or are the result of the underlying changes to cell identity.

Following data acquisition, quality control and normalisation of the qPLEX-RIME dataset was performed. Assessing the correlation between unnormalized replicates identified a weak correlation between two of the naïve PSC replicates (naïve 3 and 4) and the other samples within the dataset, including the other two naïve replicates (naïve 1 and 2) (Fig. 5.2.5).

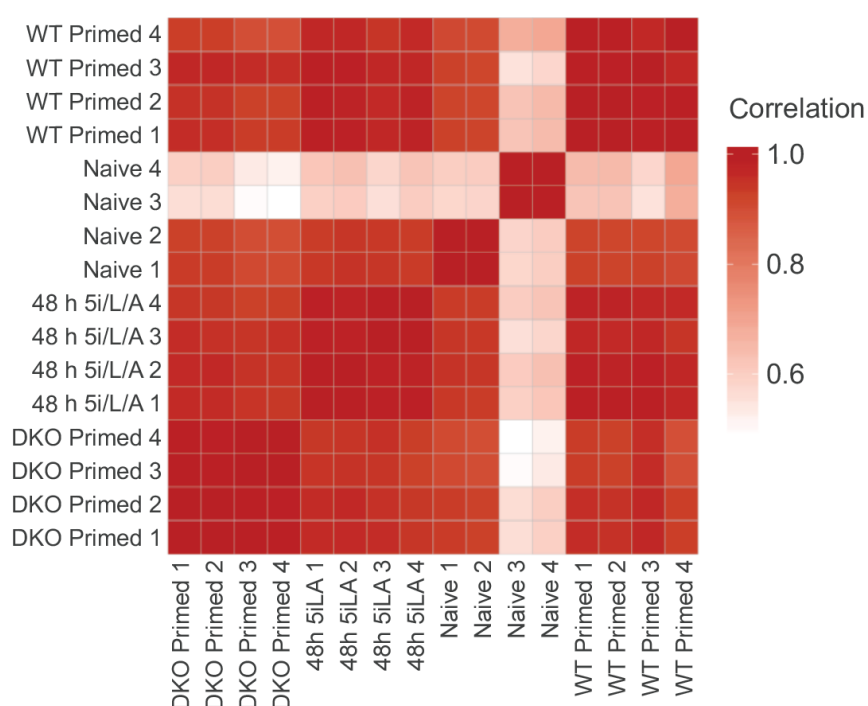


Figure 5.2.5 – Correlation matrix of qPLEX-RIME dataset

Each column and row represent a single qPLEX-RIME condition, plotted based on correlation of normalised protein intensity. All conditions except DKO primed represent the WT state. 48 h 5iLA; primed PSCs cultured in 5i/L/A reprogramming conditions for 48 hours. Four biological replicates for all samples.

The intensity of peptide abundances were plotted for all 16 samples within the dataset. Notably, the naïve samples 3 and 4 show lower protein abundance intensities than the other WT samples in all conditions, and were instead comparable to the four DKO primed PSC samples, where the immunoprecipitation target is genetically deleted from cells (Fig. 5.2.6).

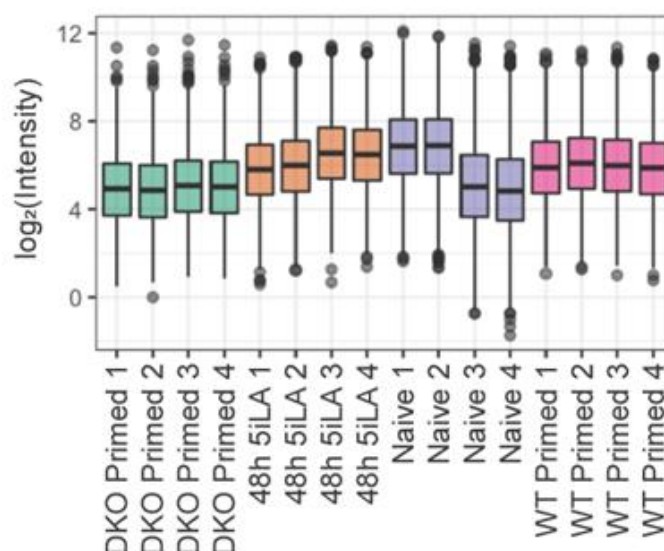


Figure 5.2.6 – Box and whisker plot shows the peptide abundance intensity in each biological replicate

$\log_2$ (Intensity) represents the raw unnormalised abundance of the tandem mass tag label of each of the 16 samples. Box represents interquartile range, with central horizontal line showing the median value and outliers are shown.

I speculated that the immunoprecipitation failed in the naïve 3 and 4 samples, and therefore these samples were omitted from the final dataset. On the 14 remaining samples, principal components analysis was used to analyse the clustering of the samples. Individual replicates in the same condition closely together, showing the reproducibility of the experiments. A stark difference in clustering can be seen between WT and DKO primed samples, consistent with the absence of PCGF3/5 acting as the immunoprecipitation target protein in DKO primed PSCs. 48 h 5i/L/A samples cluster much more closely to primed PSCs than naïve PSCs (Fig. 5.2.7).

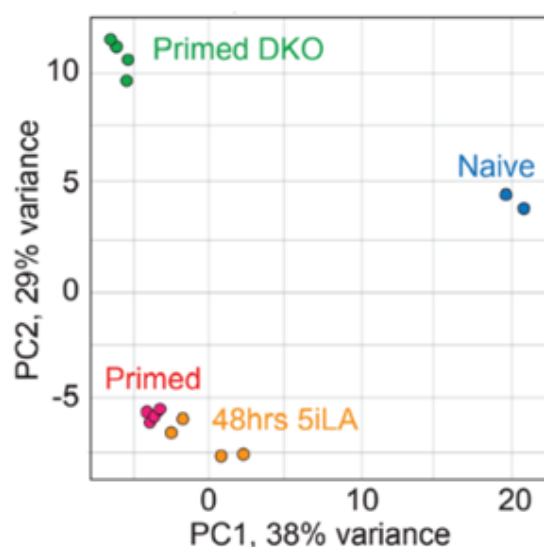


Figure 5.2.7 – Principal Components Analysis of the 14-plex qPLEX-RIME data

Circles represent independent biological replicates.

The similarity between the primed and 48 h 5i/L/A conditions shows that extensive changes in the PRC1.3 complex interactome between primed and naïve PSCs do not occur immediately at the onset of reprogramming but resembles a longer-term remodelling of the PRC1.3 interaction network. An important barometer of the PRC1.3 qPLEX-RIME experiment is the ability to detect known components of the PRC1.3 complex itself, indicating that the immunoprecipitation was successful. Quantitative comparison of WT and DKO primed PSCs confirmed a significant enrichment of the known PRC1.3 components PCGF3, RNF2 (RING1B), RING1 (RING1A), CSNK2B (Casein Kinase II complex), AUTS2 and FBRS in WT cells (Gao *et al.*, 2012; Hauri *et al.*, 2016) (Fig. 5.2.8).

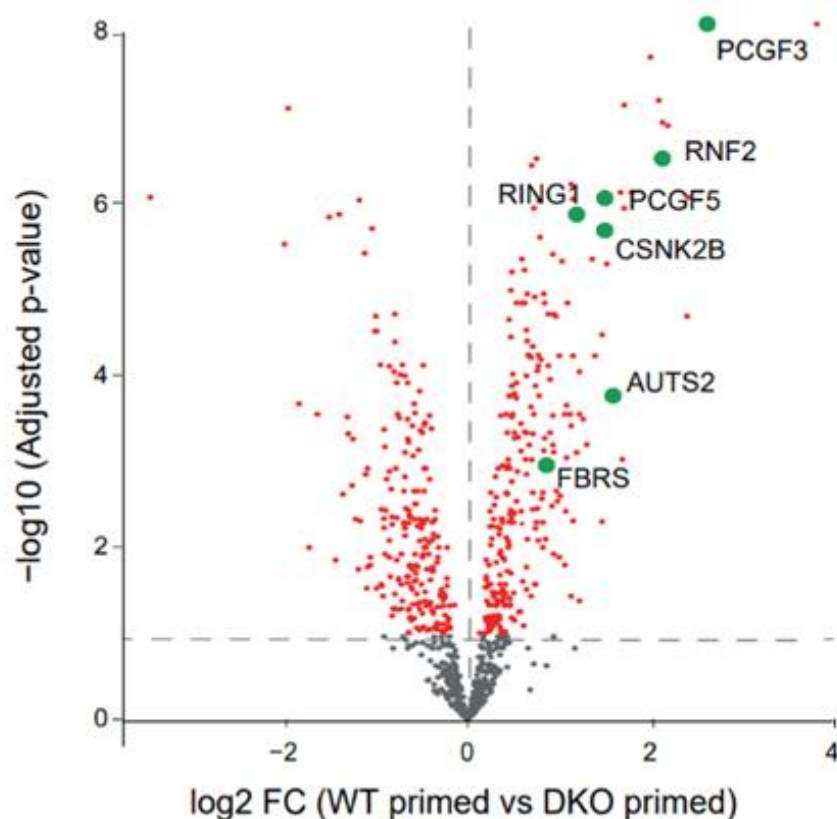


Figure 5.2.8 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change (FC) in abundance and significance of PCGF3-interacting proteins between WT and DKO in primed PSCs

Significant proteins (393) possess an $-\log_{10}(\text{adjusted p-value}) > 1$, red dots, and non-significant proteins (394) with an adjusted p-value < 1 , grey dots. PRC1.3 complex components, green circles. Four independent biological replicates for primed and DKO primed.

Interestingly, neither RYBP nor YAF2 were detected in any of the 14 samples in the complete dataset, despite being reported to be central to all variant PRC1 complexes (Gao *et al.*, 2012; Hauri *et al.*, 2016). In summary, the 14-plex qPLEX-RIME experiment is able to detect the PRC1.3 complex and its interacting proteins in different states of human pluripotency.

5.2.3 - qPLEX-RIME identifies PRDM14 as a naïve-specific PRC1.3 co-interactor

In total, the qPLEX-RIME experiment detected peptides from 787 distinct proteins across the 14 samples. Previously reported PRC1.3 interacting proteins in other cellular contexts; P300 (Gao *et al.*, 2014), TEX10 (Zhao *et al.*, 2017) and USF1 (Scelfo *et al.*, 2019), were not detected in any of the conditions of the qPLEX-RIME experiment. The PRC1.3 interactor hnRNPK (Pintacuda *et al.*, 2017) was identified, but interacted with no significant difference in PRC1.3 enrichment between pluripotency states, making it an unlikely candidate for state-specific regulation of the complex.

Of the 787 detected proteins, 26 were also identified in the genome-wide genetic screen as key regulators of human PTNR. In identifying potential candidate genes for state-specific regulation of the PRC1.3 complex, I focused specifically on the essential genes, which presented the likeliest regulators that link the PRC1.3 complex to reprogramming. The proteins detected in the qPLEX-RIME screen were plotted, comparing primed and 48 h 5i/L/A reprogramming cells (Fig. 5.2.9), and primed and naïve PSCs (Fig. 5.2.10).

As predicted by principal components analysis, a smaller number of proteins show a significant difference between primed and 48 h 5i/L/A cells ($n=290$) than between primed and naïve PSCs ($n=466$). The magnitude of the fold change in enrichment of these proteins is also much smaller between primed and 48 h 5i/L/A cells. Amongst the significant state-specific proteins also identified as essential for reprogramming, are PHF6 (#70 factor), which is enriched with PRC1.3 in 48 h 5i/L/A conditions compared to primed, and DDX5 (#96 factor), which is instead enriched with PRC1.3 in primed PSCs compared to 48 h 5i/L/A. *PHF6* has been reported as a chromatin regulator in haematopoietic stem cells (Todd *et al.*, 2015) but its role in embryonic pluripotency is unclear. *DDX5*, on the other hand, has been reported to suppress iPSC reprogramming by upregulation of variant PRC1 component *RYBP* and subsequent H2AK119ub1 deposition at lineage-specific genes (Li *et al.*, 2017). Intriguingly, *DDX5* is critically important for human PTNR and interacts with PRC1.3 in primed PSCs, but the possible interplay between PRC1.3 and *DDX5* in this context is unclear.

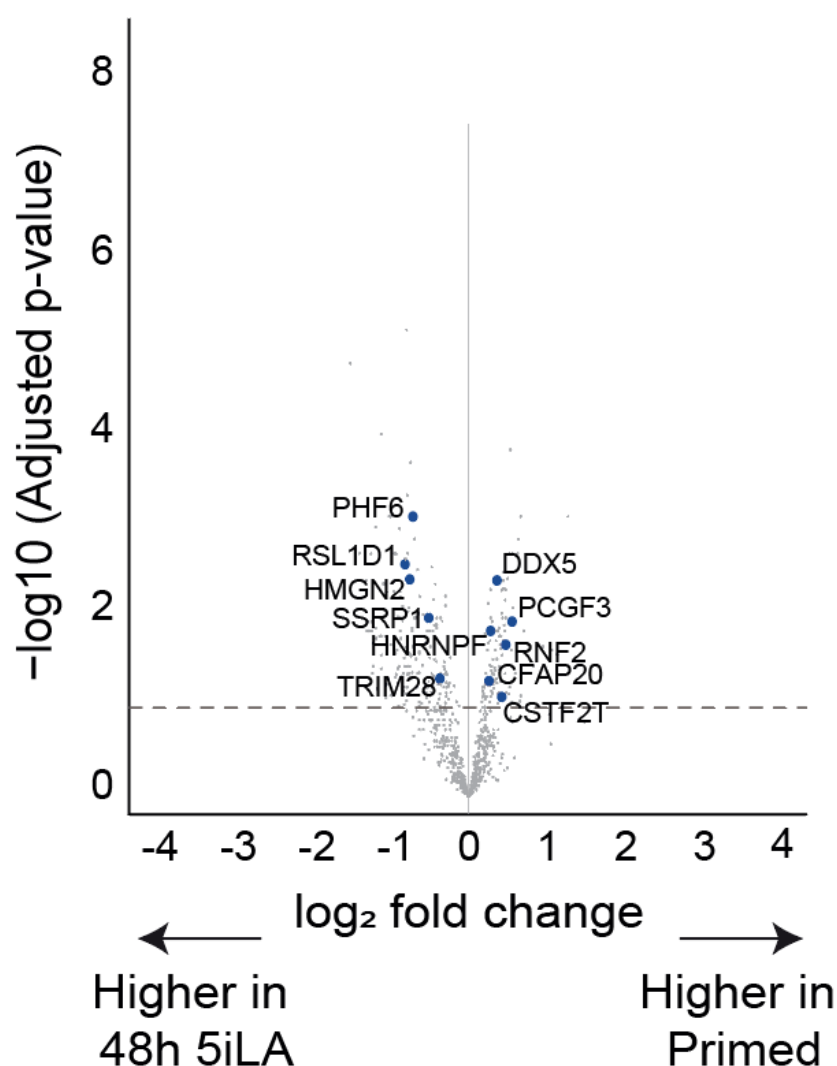


Figure 5.2.9 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change and the significance of proteins identified as key for human PTNR when comparing primed and 48 h 5i/LA PSCs

Proteins significantly differentially enriched and essential for human PTNR, blue circles; proteins not significant or not essential for human PTNR, grey circles. Four biological replicates for both conditions.

When comparing primed and naïve PSCs, the only key reprogramming gene enriched in primed PRC1.3 complexes is CSTF2T (#170 factor), which has a characterised function in mammalian mRNA processing (Yao *et al.*, 2013). The most significant naïve enriched PRC1.3 co-interactor and reprogramming essential gene (#74 factor) is the pluripotency transcription factor *PRDM14* (Fig. 5.2.10). Despite this association being specifically identified in human

naïve PSCs, PRDM14 binds to chromatin in both human naïve and primed PSCs, with PRDM14 binding somewhat enriched on naïve PSC chromatin (Zijlmans *et al.*, 2022).

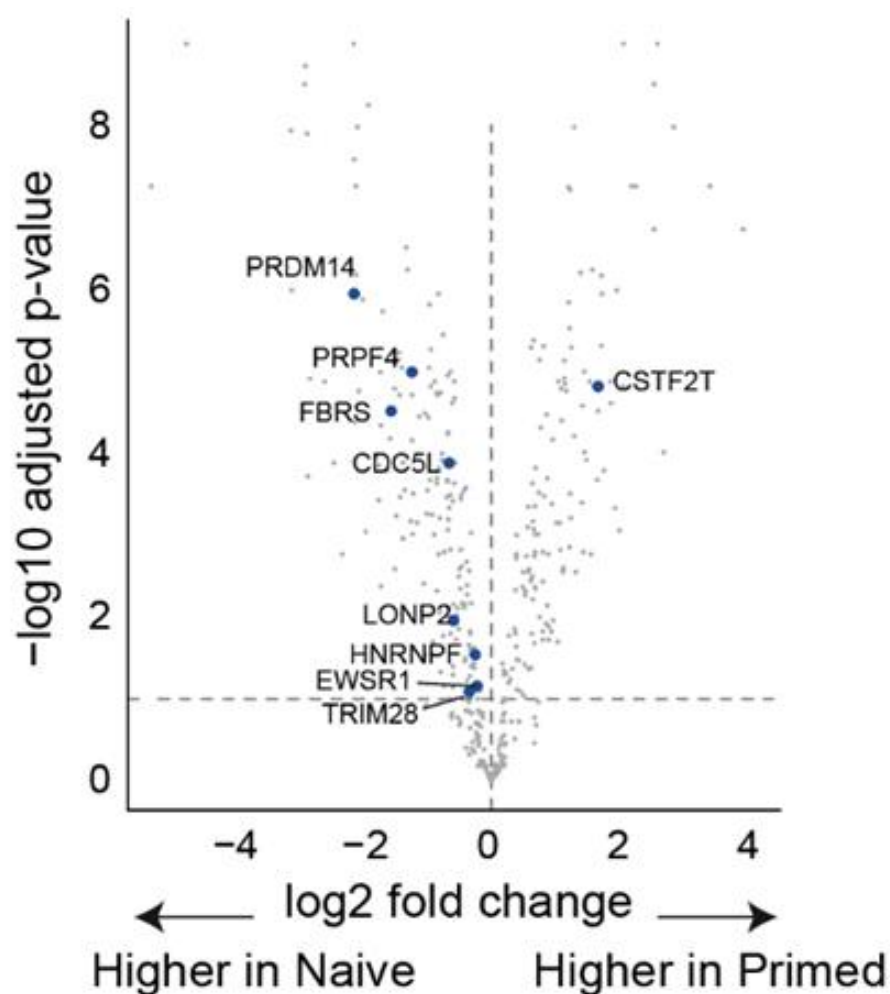


Figure 5.2.10 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change and the significance of proteins identified as key for human PTNR when comparing primed and naïve PSCs

Proteins significantly differentially enriched and essential for human PTNR, blue circles; proteins not significant or not essential for human PTNR, grey circles. Four biological replicates for primed PSCs and two biological replicates for naïve PSCs.

PRDM14 was a promising candidate, as a key transcription factor in PSCs, and so I focused further on how *PRDM14* might regulate the function of the PRC1.3 complex during PTNR.

5.3 – PRC1.3 and PRDM14 cooperate to repress transcription during human PTNR

5.3.1 – PRC1.3 and PRDM14 interact in naïve PSCs

PRDM14 is specifically expressed in pluripotent cells, where it acts as a transcriptional repressor of early differentiation genes in order to maintain self-renewal and maintenance of the primed state of pluripotency (Tsuneyoshi *et al.*, 2008) and also activates transcription of the core pluripotency transcription factor *OCT4* (Chia *et al.*, 2010). A genome-wide RNA interference screen identified *PRDM14* as an important gene for the human primed pluripotent state (Chia *et al.*, 2010) and is integrated within the core *OCT4*, *NANOG* and *SOX2* pluripotent transcription factor circuitry (Chia *et al.*, 2010). An interaction with PRDM14 could therefore link PRC1.3 centrally to the transcriptional regulation of human naïve pluripotency.

Studies in human primed PSCs have previously identified a connection between PRC2 and PRDM14, in the repression of developmental genes, with *PRDM14* depletion causing loss of H3K27me3 binding and gene derepression (Chan *et al.*, 2013a). However, a co-interaction between PRDM14 and the PRC1 complexes, particularly PRC1.3, has never been identified before now. This may be in part due to the understudied nature of the PRC1.3 complex in human pluripotency, but also because PRC1.3 and PRDM14 are associated to a much greater extent in human naïve PSCs, where *PRDM14* also has not been studied in depth.

PRDM14 has been shown to have an important role in the differentiation of human PSCs, with the protein being required for the specification of human germ cell fate from primed PSCs (Sybirna *et al.*, 2020), setting a precedent for *PRDM14* mediating cell fate specification and transitions such as PTNR. I decided to investigate the interaction between PRDM14 and PRC1.3, including whether PRDM14 influences the function or recruitment of the PRC1.3 complex to chromatin.

A native co-immunoprecipitation was used to confirm the co-interaction between PCGF3 and PRDM14 in naïve PSCs. This approach used an immunoprecipitation with the anti-PCGF3/5 antibody in primed and naïve PSCs, and then blotting to detect the presence of PRDM14. Input samples were also assessed, which were not subject to immunoprecipitation. Both PCGF3 and PRDM14 were detected in the inputs of primed and naïve PSCs, with PRDM14 equally abundant in both cell types. However, in the PRC1.3 co-immunoprecipitation samples,

PRDM14 was more strongly enriched in naïve PSCs, while the level of PCGF3 detected was broadly similar (Fig. 5.3.1).

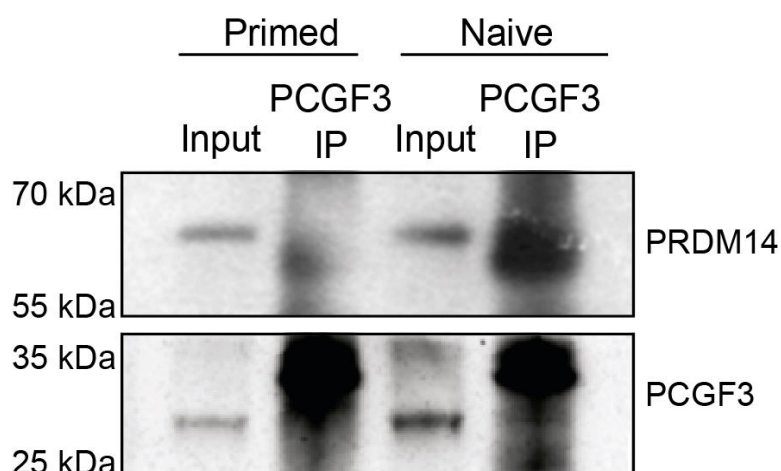


Figure 5.3.1 – Co-immunoprecipitation-Western blotting confirms a co-interaction between PRDM14 and PCGF3 in naïve PSCs

PRDM14 has a predicted molecular weight of 65 kDa.

The enriched co-interaction between PRDM14 and PRC1.3 in naïve PSCs supports the qPLEX-RIME data. The immunoprecipitation experiment included benzonase in the buffers, and therefore PCGF3 and PRDM14 must interact, directly or indirectly, via protein-protein interactions, and are not simply bridged by DNA as is possible within the qPLEX-RIME protocol. I attempted a reciprocal co-immunoprecipitation experiment, by pulling down PRDM14 and blotting for PCGF3. However, this approach yielded no specific signal even for PRDM14 and suggested that the PRDM14 antibody was not compatible with immunoprecipitation.

Analysis of the PRC1.3 ChIP-sequencing data from Chapter 4, identified another link between PRDM14 and PRC1.3 in naïve PSCs. Motif enrichment analysis of naïve PRC1.3 target sites identified the motif most significantly enriched at these sites is the PRDM14 motif, with a high level of statistical significance (Fig. 5.3.2).

Motif enrichment analysis for PRC1.3 sites in naïve PSCs

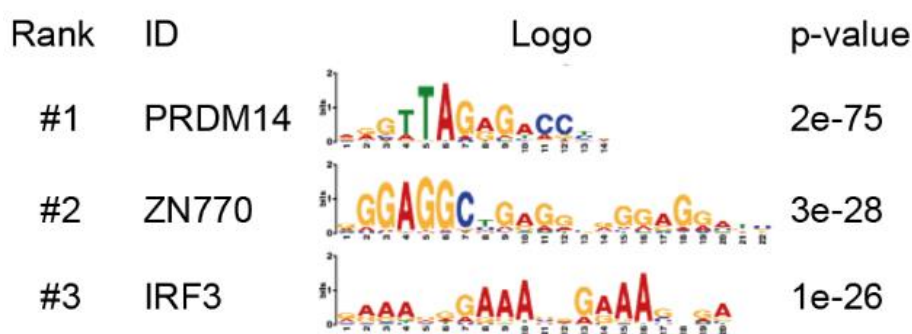


Figure 5.3.2 – Motif enrichment analysis for PRC1.3 target sites in naïve PSCs

The top three motifs are shown, ranked by adjusted p-value, which was calculated using a one-tailed Fisher's exact test with Bonferroni correction.

When the same motif enrichment analysis was performed on primed PRC1.3 target genes, other motifs were significantly enriched but the PRDM14 motif was not. This further establishes that the interaction between PRC1.3 and PRDM14 occurs specifically in naïve PSCs and suggests that the PRC1.3 complex could be recruited to sites of PRDM14 binding on chromatin in naïve PSCs.

5.3.2 – PRDM14 binds with PRC1.3 at naïve target genes

To confirm whether PRC1.3 and PRDM14 are co-localised on chromatin in naïve PSCs, consistent with a naïve-specific interaction, I needed to know the sites of PRDM14 binding in naïve PSCs.

A previous study (Sybirna *et al.*, 2020) generated a cell line in which PRDM14 is epitope tagged with a VENUS fluorescence reporter and an auxin inducible degron tag (Nishimura *et al.*, 2009). ChIP-sequencing was performed on these PRDM14-AID-VENUS cells in 4i naïve media conditions, using an anti-GFP antibody to detect the binding of VENUS-fused PRDM14

to chromatin (Sybirna *et al.*, 2020). I analysed the PRDM14 ChIP-sequencing signal at naïve PRC1.3 target genes and verified the specificity of this signal by also analysing the signal when PRDM14 was depleted by the presence of indole-3-acetic acid (IAA). This analysis confirms that PRDM14 binding can be specifically detected at the vast majority of naïve PRC1.3 target genes in 4i naïve media conditions (Fig. 5.3.3).

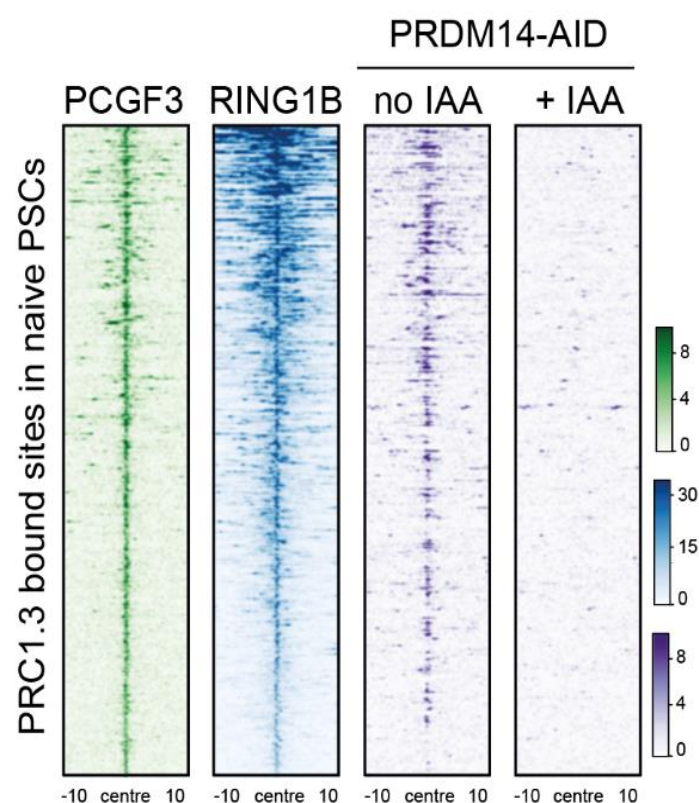


Figure 5.3.3 – Normalised PCGF3, RING1B and PRDM14-AID-VENUS ChIP-sequencing read counts, displayed in a heatmap, within a 20 kb peak-centered window in naïve PSCs

PCGF3 and RING1B signal shown as PRC1.3-bound naïve sites (n=378) and PRDM14 signal at these sites in 4i naïve PSCs, both with IAA (Indole-3-acetic acid) and without (no IAA) PRDM14 depletion. PRDM14 data obtained from (Sybirna *et al.*, 2020). Heatmap generated by Andrew Malcolm.

A limitation of this analysis is that the PRC1.3 and PRDM14 ChIP-sequencing experiments were conducted in different naïve PSC media conditions – t2iL + PKCi for PCGF3 and RING1B, and 4i for PRDM14-VENUS. Standardisation of these conditions would have been preferable for optimal comparison between the datasets. Nevertheless, this finding suggests that naïve PRC1.3 target genes are enriched for PRDM14 binding.

I expanded my analysis, in order to determine whether PRDM14 binding is common to all PRC1 binding sites in naïve PSCs, or whether its detection was specific to PRC1.3 target genes. PRDM14 ChIP-sequencing signal was quantified for all naïve PRC1.3 target sites and across a similar number of RING1B only sites, matched in terms of RING1B signal. The specific PRDM14 binding, eliminated by PRDM14 degradation, was seen specifically at PRC1.3 binding sites on chromatin and not more generally at PRC1 binding sites (Fig. 5.3.4).

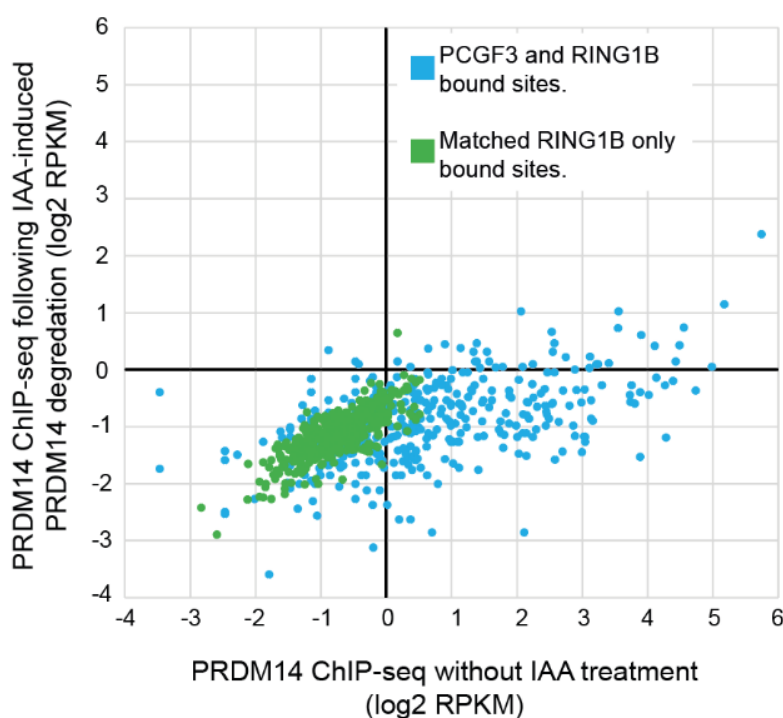


Figure 5.3.4 – The distribution of the PRDM14-AID-VENUS ChIP-sequencing normalised read counts are shown in a scatter plot, with and without IAA-induced PRDM14 degradation

PRC1.3 naïve target sites (PCGF3 & RING1B, n=378); blue circles, matched signal intensity RING1B-only sites (n=397); green circles. PRDM14 ChIP-seq data are obtained from (Sybirna *et al.*, 2020).

These data identify a novel link between PRDM14 and PRC1.3 in human naïve pluripotency. These findings are consistent with PRDM14 possibly recruiting PRC1.3 to its naïve target sites, which possess PRDM14 binding motifs, via a direct protein-protein interaction.

5.3.3 – *PRDM14* is also essential for human PTNR

During human PTNR, an association is established between *PRDM14* and *PRC1.3* and *PRDM14* was identified as a key factor in reprogramming (#74 factor) by the genome-wide screen. I wanted to establish whether the key roles of *PRC1.3* and *PRDM14* are linked and whether these factors cooperate with one another during reprogramming.

I decided to first establish that *PRDM14* is in fact critical for reprogramming, as I did for *PCGF3*. To this end, I obtained the *PRDM14*-AID-VENUS inducible loss-of-function primed PSC cell line (Sybirna *et al.*, 2020) and I characterised that the *PRDM14*-AID-VENUS primed PSCs express SSEA-4 on almost 100 % of cells (Fig. 5.3.5), uniformly express both OCT4 and NANOG (Fig. 5.3.6) and have an unperturbed 46 XX karyotype (Fig. 5.3.7).

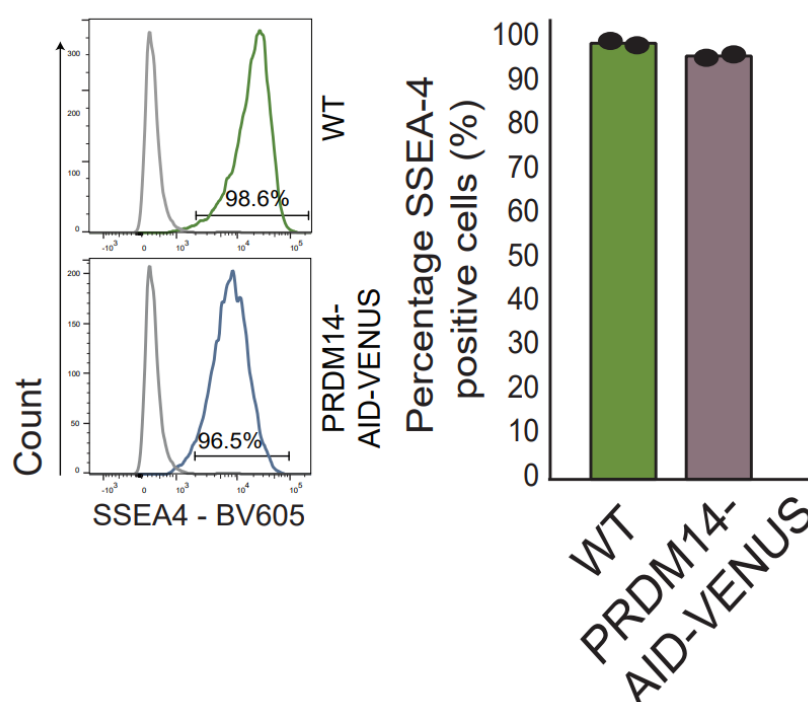


Figure 5.3.5 – Flow cytometry histogram of SSEA4 expression in *PRDM14*-AID-VENUS primed PSCs

Flow cytometry data represented as a histogram, representative of two independent biological replicates. Grey line shows omission of the SSEA-4 antibody. Bar chart quantifies the SSEA-4 positive cells as a percentage. Bar height shows the mean and circles show the individual datapoints.

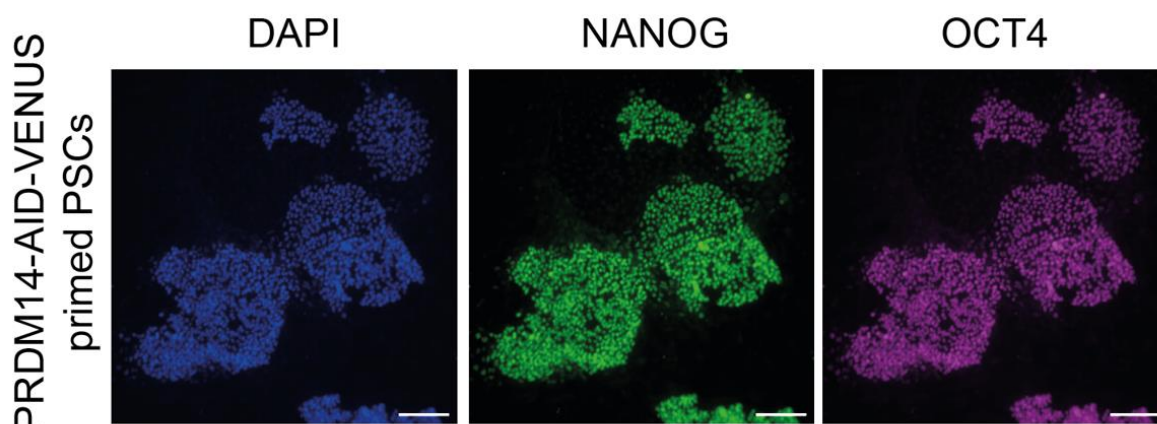


Figure 5.3.6 – Immunofluorescence microscopy images show the uniform expression of OCT4 and NANOG in PRDM14-AID-VENUS primed PSCs

DAPI stains all nuclei. OCT4 and NANOG are core pluripotency markers. Scale bars, 200 μ m.

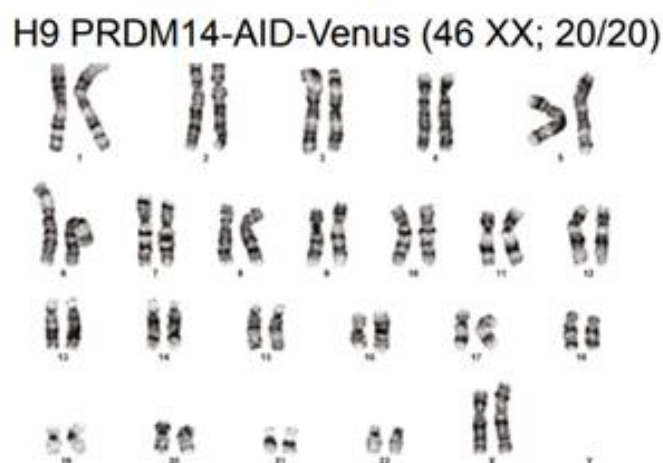


Figure 5.3.7 – Karyotyping of the PRDM14-AID-VENUS primed PSC line

Representative spreads of G-banded chromosomes for the indicated cell lines. Twenty cells were examined for each cell line.

I also confirmed the ability of PRDM14-AID-VENUS primed PSCs to inducibly deplete PRDM14-VENUS in the presence of IAA, with immunofluorescence microscopy confirming that the majority of detectable nuclei do not express PRDM14-VENUS under knockdown conditions (Fig. 5.3.8).

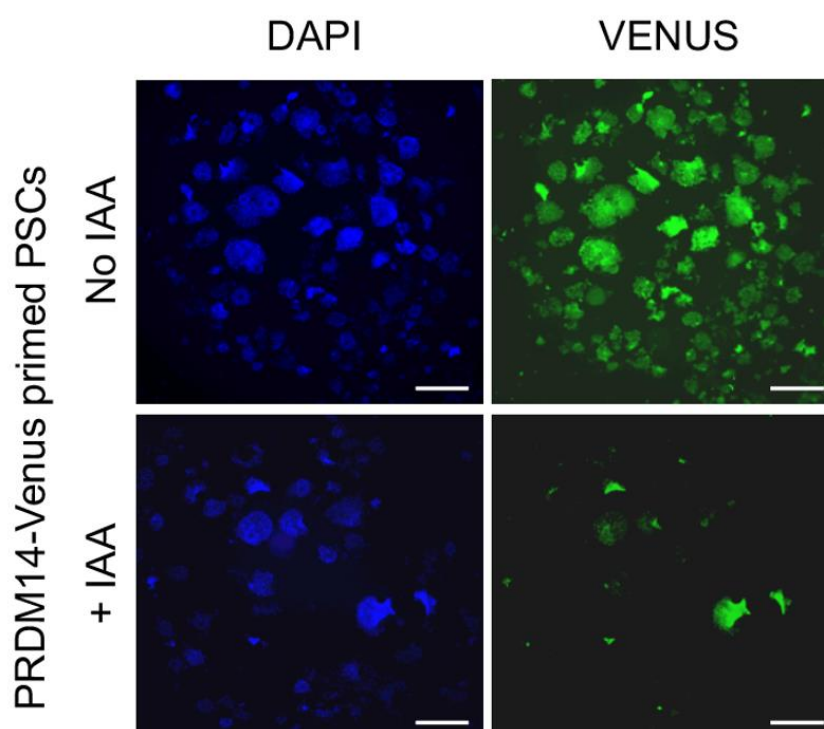


Figure 5.3.8 – Immunofluorescence in PRDM14-AID-VENUS primed PSCs in the presence and absence of IAA treatment, confirming PRDM14 depletion

DAPI stains all nuclei. VENUS stains for PRDM14 expression. IAA was applied at 100 μ M for 24 h. Scale bars, 500 μ m.

I then exploited the inducible knockdown system to confirm that *PRDM14* is indeed essential for PTNR. I performed chemical resetting with PRDM14-AID-VENUS primed PSCs both with and without inducible PRDM14 depletion. Notably, the efficiency of PRDM14-AID-VENUS primed PSC reprogramming without IAA treatment was very low, compared to a positive control H9 line that was reprogrammed concurrently. Previous reports in different reprogramming contexts have documented cell line-to-line variability in reprogramming ability (Vitale *et al.*, 2012). In this instance, poor reprogramming efficiency could be a consequence of the multiple targeting and selection steps required for successful knock-in of the AID-VENUS tag to both endogenous alleles (Sybirna *et al.*, 2020) or due to leaky PRDM14 degradation in the absence of IAA (Natsume *et al.*, 2016; Yesbolatova *et al.*, 2019). IAA was added from 48 h of reprogramming onwards, as this is the time-point at which the first reprogramming media cocktail is added to chemical resetting cells and the qPLEX-RIME experiment using 5i/L/A reprogramming showed that PCGF3 and PRDM14 did not

significantly interact after 48 h under reprogramming conditions. PRDM14 depletion led to a significant failure to produce nascent naïve PSCs under reprogramming conditions (Fig. 5.3.9).

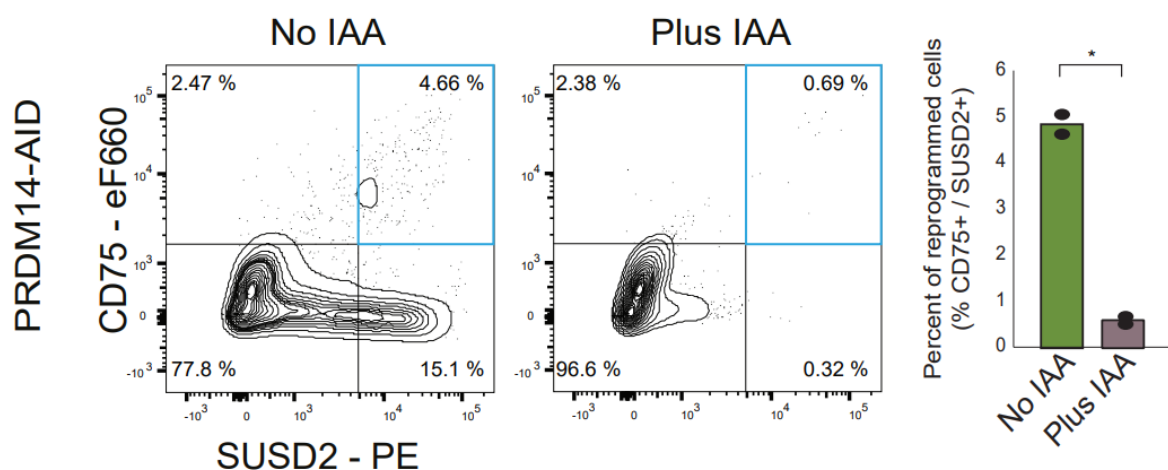


Figure 5.3.9 – Flow cytometry contour plots of PRDM14-AID-VENUS PSCs after 10 days of chemical resetting, in the absence (left) or presence (right) of IAA-induced PRDM14 degradation IAA was added from 48 h onwards, for 8 days in total. Cells were labelled with antibodies against the naïve markers, CD75 and SUSD2. Successfully reprogrammed naïve cells are shown in the upper right quadrant (CD75⁺ SUSD2⁺). Data are representative of one KO line and two independent experiments. P-value calculated by one-way ANOVA; *, p<0.05.

Phase contrast imaging also showed a low efficiency, but detectable, formation of nascent naïve PSCs only in the no IAA-treated cells (Fig. 5.3.10).

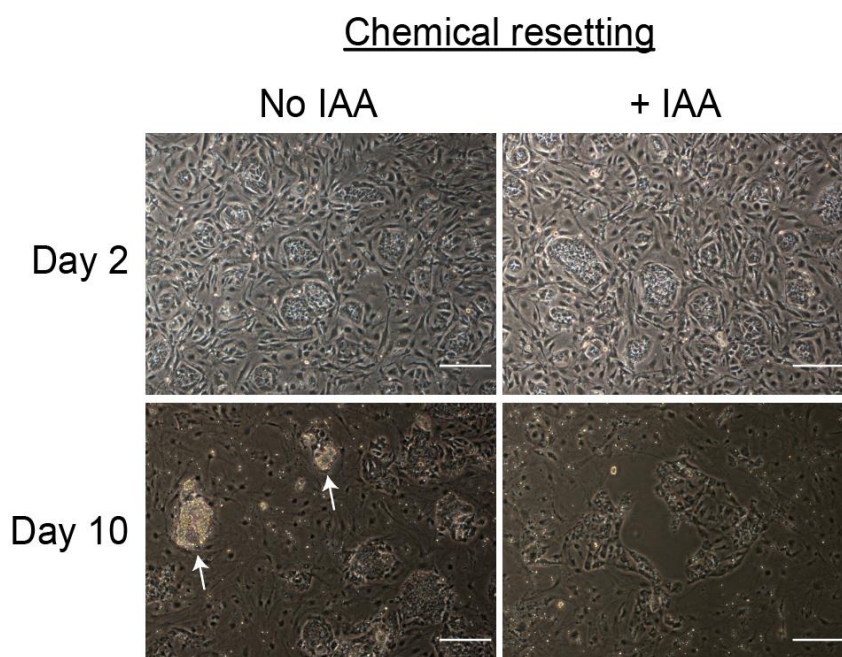


Figure 5.3.10 – Phase contrast images of PRDM14-AID-VENUS PSCs at day 2 and day 10 of naïve PSC reprogramming, with and without IAA (CR conditions)

Images are representative of three independent experiments (n=3). Scale bars, 100 μ m. Day 2 cells, prior to PRDM14 depletion, are very morphologically similar in both conditions. Arrows indicate the naïve PSC colonies in the day 10 no IAA cultures and were not seen in the day 10 plus IAA cultures.

I therefore concluded that like *PCGF3* and *FBRS*, *PRDM14* is also essential for human PTNR, as predicted by the prior genome-wide screen.

5.3.4 – PRDM14 and PRC1.3 cooperate to repress PRC1.3 target genes

To better understand the potential functional cooperation between PRC1.3 and PRDM14 during reprogramming and in naïve PSCs, I needed to stably produce PRDM14-AID-VENUS naïve PSCs. To do this, chemical resetting of an IAA untreated population was extended to propagate and stabilise a pure population of naïve PSCs in culture, which required FACS sorting for cells that express the naïve markers CD75 and SUSD2, and do not express the primed marker CD24. I was only able to produce one PRDM14-AID-VENUS naïve PSC line

as the cells reprogrammed with a very low efficiency and were difficult to stabilise under naïve growth conditions.

To confirm the naïve state of pluripotency in PRDM14-AID-VENUS cells, I used immunofluorescence microscopy and identified that the cells co-express the pan-pluripotency transcription factors OCT4 and NANOG, and the naïve-specific transcription factor KLF17 (Fig. 5.3.11).

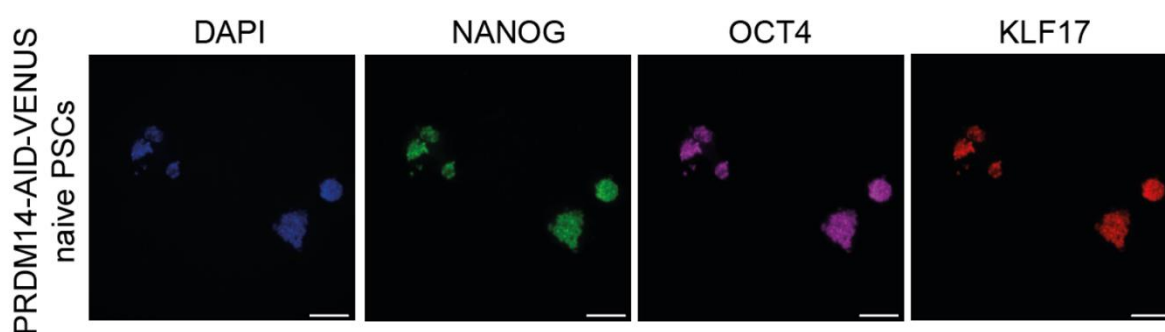


Figure 5.3.11 – Immunofluorescence microscopy shows the uniform expression of OCT4, NANOG and KLF17 in PRDM14-AID-VENUS naïve PSCs

DAPI stains all nuclei. OCT4 and NANOG are pan-pluripotency markers. KLF17 is a naïve-specific pluripotency marker. Scale bars, 200 μ m.

Immunofluorescence microscopy also confirmed the ability of PRDM14-AID-VENUS naïve PSCs to deplete PRDM14-VENUS in the presence of IAA. (Fig. 5.3.12).

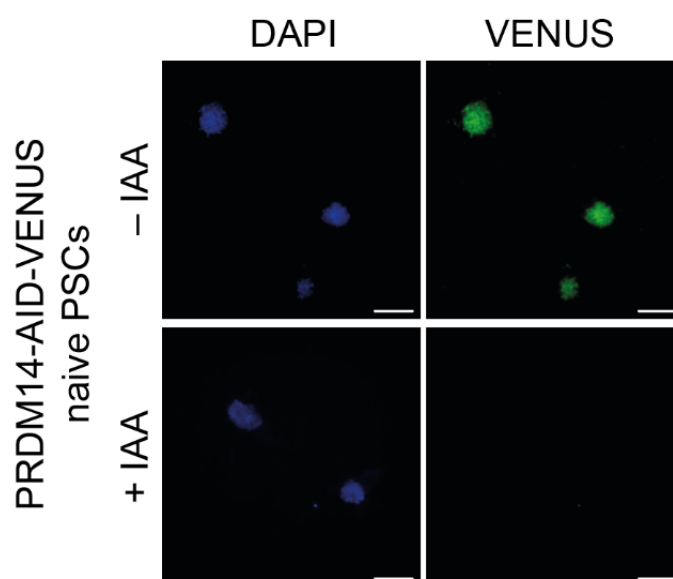


Figure 5.3.12 – Immunofluorescence microscopy in PRDM14-AID-VENUS naïve PSCs in the presence and absence of IAA treatment, confirming PRDM14 depletion

DAPI stains all nuclei. VENUS stains for PRDM14 expression. IAA was applied at 100 μ M for 24 h. Scale bars, 200 μ m.

I next performed RNA-sequencing on PRDM14-AID-VENUS naïve PSCs, using three independent biological replicates both with and without IAA. IAA was added for 24 h to deplete PRDM14 protein and this time-point was selected both to identify early consequences of PRDM14 depletion and because longer exposures to IAA (48 h) strongly impacted cell morphology and viability. This suggested that *PRDM14* might be crucial for naïve PSC self-renewal and maintenance, although this was not followed up experimentally.

I was particularly interested in whether PRDM14 cooperates with PRC1.3 to repress the expression of naïve PRC1.3 target genes. To study this, a gene set enrichment analysis (GSEA) (Subramanian *et al.*, 2005) of the dataset was performed. GSEA works by determining whether a specific gene set shows a statistically significant difference between two biological conditions. In this case, I wanted to determine whether naïve PRC1.3 target genes are enriched in the genes whose expression increases or decreases when PRDM14 is depleted. GSEA confirmed that naïve PRC1.3 target genes show a statistically significant increase in expression when PRDM14 is depleted, as confirmed by the positive enrichment score (Fig. 5.3.13).

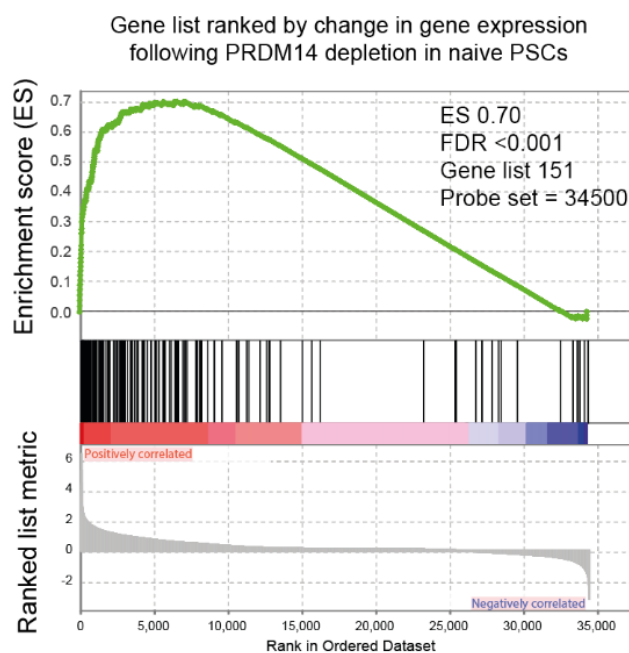


Figure 5.3.13 – Gene set enrichment analysis (GSEA) of naïve PRC1.3 target genes, ranked according fold change in transcription, following IAA-induced PRDM14 depletion

A positive enrichment score (ES) reveals that PRC1.3 naïve target genes are significantly derepressed in the absence of PRDM14 (FDR < 0.001; Kolmogorov-Smirnov test).

Two example naïve PRC1.3 target genes, *MED12* and *CIZ1*, are shown below with an increase in gene expression evident when PRDM14 is degraded (Fig. 5.3.14).

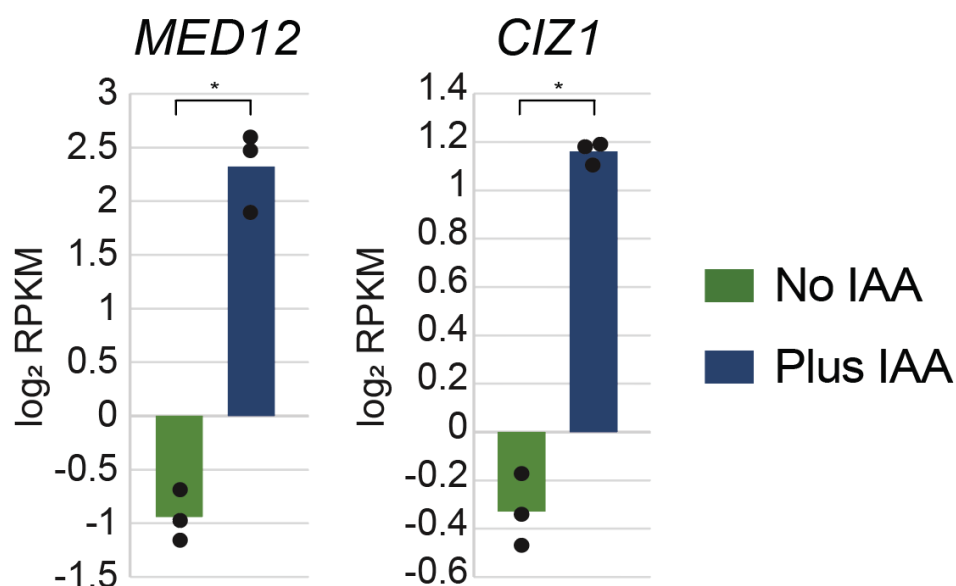


Figure 5.3.14 – Charts show $\log_2$ RPKM of gene expression in PRDM14-AID-VENUS naïve PSCs, in the presence (no IAA) and absence (plus IAA) of PRDM14

Circles show each biological replicate (n=3 for both conditions) for naïve PRC1.3 target genes *MED12* and *CIZ1*. P-value determined by paired t-test; *, p<0.01.

Derepression of naïve PRC1.3 target genes when PRDM14 is depleted is analogous to the failure to repress this gene set during reprogramming in *PCGF3* KO cells (Chapter 4), indicating that PRC1.3 and PRDM14 cooperate to repress this subset of genes, with both fundamentally required. Failure to repress these genes in WT cells is a signature of cells that fail to reprogramme, and as *PRDM14* is also essential for reprogramming, it is possible that this cooperative gene repressive mechanism is critical for acquisition of human naïve pluripotency.

5.4 – The PRC1.3 complex undergoes a developmentally-regulated composition switch

5.4.1 – *FBRS* and *AUTS2* are enriched in human naïve and primed PSCs respectively

As well as identification of an interaction between PRC1.3 and PRDM14, the qPLEX-RIME experiment also provided a fascinating insight into regulation within the PRC1.3 complex itself and its composition. When comparing naïve and primed PSCs, while core factors PCGF3 and RNF2/RING1B show no significant difference, the two auxiliary protein paralogues *AUTS2* and *FBRS* show an unexpected, statistically significant differences in abundance across the two pluripotency states. *FBRS* is enriched in naïve PRC1.3 complexes, while *AUTS2* is instead enriched in primed PRC1.3 complexes (Fig. 5.4.1).

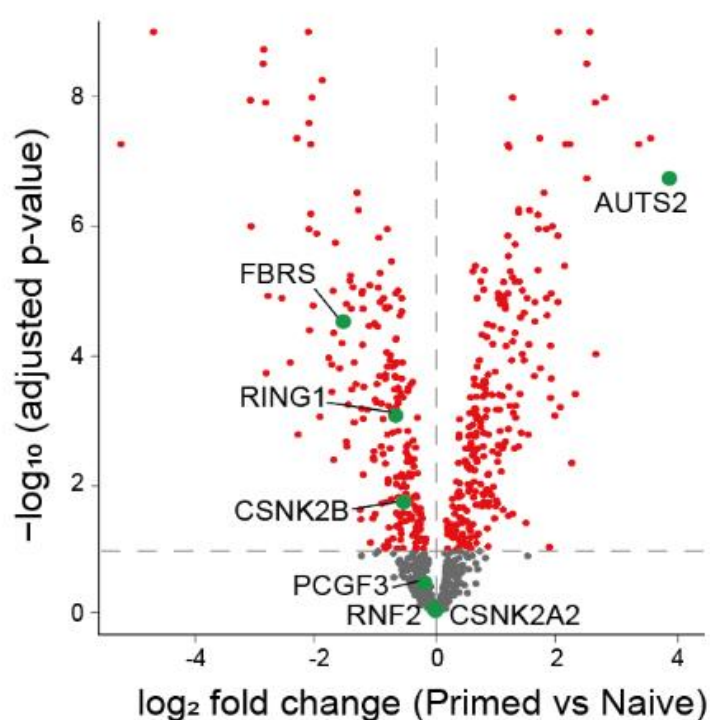


Figure 5.4.1 – Volcano plot of PRC1.3 qPLEX-RIME data shows the fold change in abundance and statistical significance of PCGF3-interacting proteins in naïve and primed PSCs

Proteins with an adjusted p-value > 1, red dots; proteins with an adjusted p-value < 1, grey dots; PRC1.3 complex components are labelled, green dots. Limma test with Benjamini-Hochberg correction; biological replicates: n=4 primed; n=2 naïve.

AUTS2 and FBRS are paralogues and it was previously assumed that both proteins were simultaneously present within PRC1.3 complexes (Gao *et al.*, 2014, 2012). It is not known how the inclusion of these paralogues within PRC1.3 complexes is regulated, but my work suggests for the first time that PRC1.3 composition could vary, whereby PRC1.3 contains either FBRS or AUTS2. Interestingly, this mutual exclusivity seems to be pluripotency state specific.

Extending this finding, I analysed published chromatin proteomics data to determine the abundance of PRC1.3 component proteins globally on chromatin (Zijlmans *et al.*, 2022). Consistent with the qPLEX-RIME findings, the majority of PRC1.3 complex components do not alter on naïve and primed PSC chromatin, although both showed an increased abundance of RING1/RING1A on naïve chromatin. Chromatin proteomics showed a significantly increased FBRS abundance on naïve chromatin and an increased AUTS2 abundance on primed chromatin (Fig. 5.4.2).

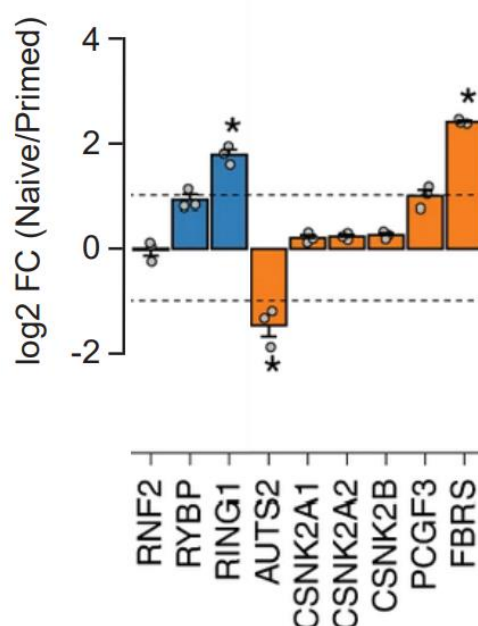


Figure 5.4.2 – Chart shows the fold change in abundance on PRC1.3 complex components using chromatin enrichment for proteomics (ChEP) between naïve and primed PSCs

Primed PSCs were grown in E8 media and naïve PSCs were grown in PXGL media. Error bars represent standard error of the mean. Asterisks indicate p-value < 0.05 and dashed line shows a fold change > 2 (student's t-test). Data obtained from (Zijlmans *et al.*, 2022).

To better understand this pluripotency state specific regulation of the *AUTS2* and *FBRS* paralogues, I analysed the expression dynamics of these paralogues from published RNA-sequencing data (Collier *et al.*, 2017; Rostovskaya *et al.*, 2019), specifically the cell state transition between human primed and naïve pluripotency, both in the reprogramming of primed to naïve PSCs (Fig. 5.4.3) and the capacitation of naïve to primed PSCs (Fig. 5.4.4).

As cells reprogramme to naïve PSCs, *FBRS* is significantly upregulated and *AUTS2* downregulated (Fig. 5.4.3). Conversely, when naïve PSCs capacitate to the primed state, the reverse expression pattern is instead seen (Fig. 5.4.4).

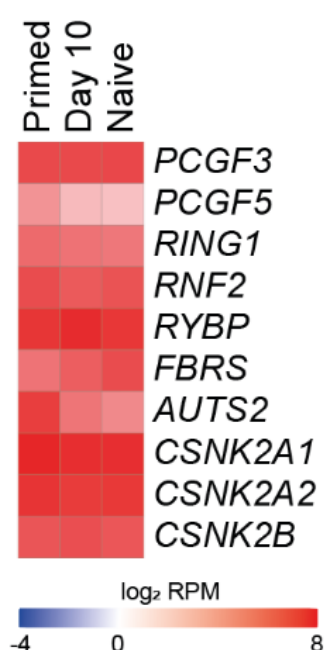


Figure 5.4.3 – Heatmap shows PRC1.3/5 component expression levels during PTNR, on day 0 (primed PSCs), day 10 (nascent naïve PSCs) and day 34 (established naïve PSC lines)

Average of three independent biological replicates. Data obtained from (Collier *et al.*, 2017).

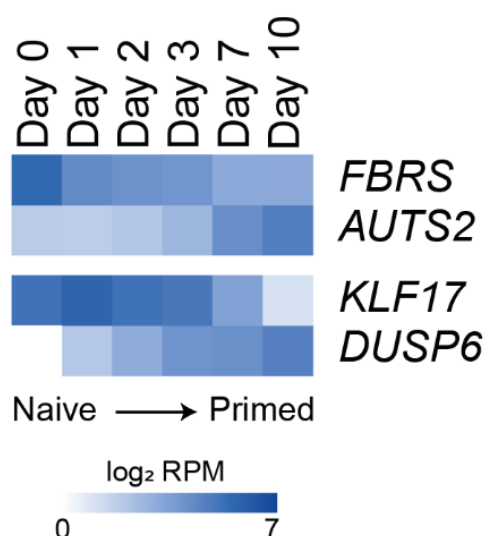


Figure 5.4.4 – Heatmap shows *FBRS* and *AUTS2* expression levels over a naive (day 0) to primed (day 10) capacitation time course

Average of three independent biological replicates. Data obtained from (Rostovskaya *et al.*, 2019). Naïve-specific marker gene *KLF17* and primed-specific marker gene *DUSP6*, are also shown.

These findings provide further evidence that the expression of the *AUTS2* and *FBRS* paralogues is developmentally regulated across human pluripotency. I also performed Western blotting for PRC1.3 complex proteins with specific antibodies on whole cell lysates. Consistent with my previous findings, *FBRS* protein is more abundant in naïve PSCs and *AUTS2* protein was detectable only in primed PSCs (Fig. 5.4.5).

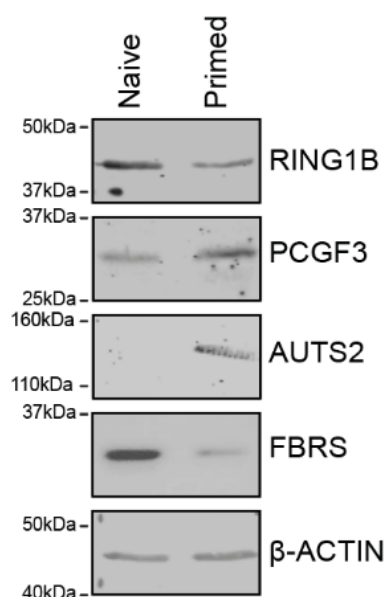


Figure 5.4.5 – Western blotting shows protein levels of PRC1.3 components in naive and primed PSCs

Primed PSCs were grown in KSR+FGF2 media and naive PSCs were grown in t2iL + PKCi media. Molecular weight in kDa. B-ACTIN; loading control.

Two main isoforms have been reported for the *AUTS2* gene (Monderer-Rothkoff *et al.*, 2019), but it is not known which isoforms are expressed in human PSCs. I was only able to detect long isoform specific peptides of *AUTS2* by qPLEX-RIME and Western blotting. However, this is likely indicative of only the long isoform of *AUTS2* acting as a component of PRC1.3 and PRC1.5 complexes (Monderer-Rothkoff *et al.*, 2019). Additionally, the anti-*AUTS2* antibody I used detects residues 522-770 in the long isoform, with the first residue of the reported human short isoform corresponding to residue 556 of the long isoform, and so I have likely not been able to detect any short isoform *AUTS2* protein. The predominant *AUTS2* protein isoform in human PSCs is therefore not known.

I speculated that the paralogue present within the PRC1.3 complex switches during reprogramming, with *AUTS2*-PRC1.3 complexes becoming *FBRS*-PRC1.3 complexes. The genome-wide screen identified *FBRS* (#10 factor), but not its paralogue *AUTS2* (#14,751 factor), as a key regulator of human PTNR, and so *FBRS*-PRC1.3 and *AUTS2*-PRC1.3 could have functional differences, with only *FBRS*-PRC1.3 complexes able to promote PTNR. This could however reflect that *FBRS*-PRC1.3 complexes are the dominant form of the complex during reprogramming, as *AUTS2* is downregulated.

I analysed the expression dynamics of the *AUTS2* and *FBRS* paralogues during reprogramming to determine whether they can be in any way linked to the success of reprogramming. For this, I used published RNA-sequencing data in primed and naïve PSCs and in flow-sorted nascent naïve PSCs and refractory cells after 10 days of 5i/L/A reprogramming (Collier *et al.*, 2017). *FBRS* is upregulated after 10 days of reprogramming, in both nascent naïve and refractory cells, while *AUTS2* is downregulated only in nascent naïve PSCs but not in refractory cells (Fig. 5.4.6).

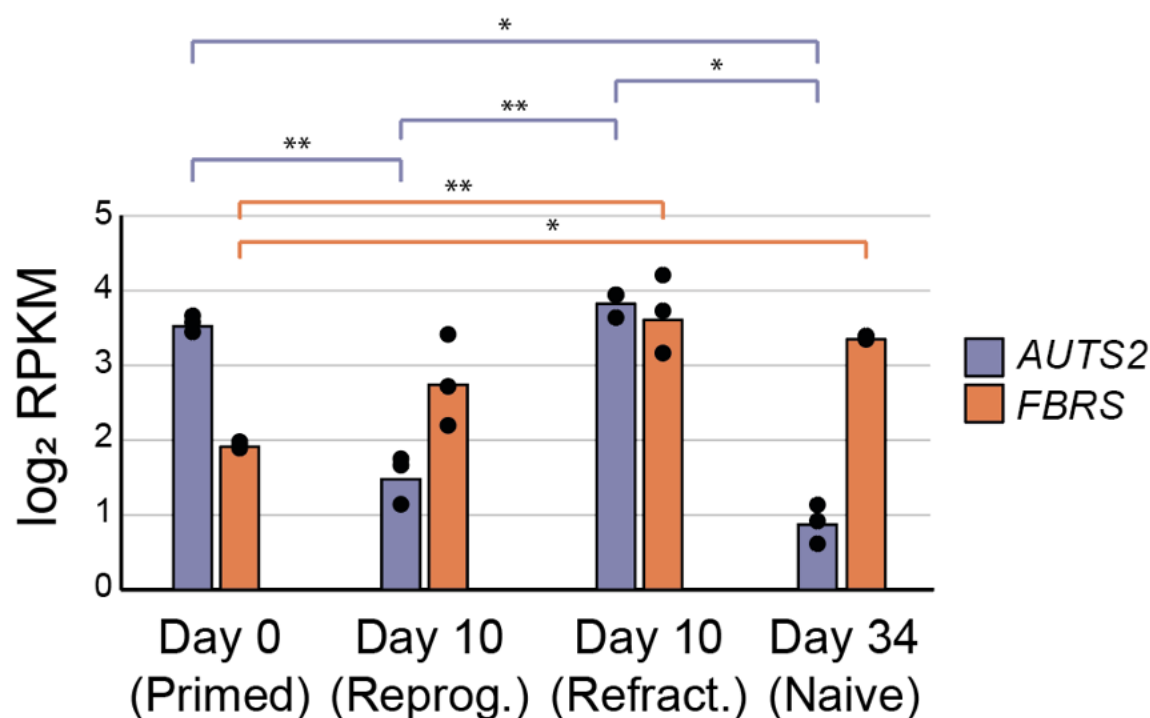


Figure 5.4.6 – Chart shows the *FBRS* and *AUTS2* expression levels in primed PSCs, nascent naïve PSCs (day 10, reprog.), refractory cells (day 10, refract.) and established naïve PSC lines

Individual replicates are shown (n=3). P-values were calculated using a Kruskal-Wallis test with Dunn's multiple comparison test; *, p<0.01, **, p<0.001. Data obtained from (Collier *et al.*, 2017).

Notably, *AUTS2* is a naïve PRC1.3 target gene (Chapter 4) and its failure to undergo transcriptional repression is seen in refractory cells, characteristic of other naïve PRC1.3 target genes. This indicates some degree of autorepression by PRC1.3 complexes during reprogramming that may reinforce the paralogue switch from *AUTS2*-PRC1.3 to *FBRS*-PRC1.3 complexes, which are able to successfully facilitate PTNR.

Although *FBRS* is enriched in naïve PRC1.3 complexes and *AUTS2* is enriched in primed PRC1.3 complexes, refractory cells express both *FBRS* and *AUTS2* at high levels. It is unclear whether refractory cells express both *FBRS* and *AUTS2*, or if this is a mixed population of cells, some of which express high levels of *FBRS* and others which express high levels of *AUTS2*, which would require single-cell RNA-sequencing to distinguish. Although upregulated at the mRNA level, a limitation of my work is that I have not confirmed whether *FBRS* protein is upregulated, or if *AUTS2* is downregulated, during reprogramming by Western blotting, looking in both nascent naïve or refractory cells – having only looked in primed and naïve PSCs.

These findings indicate that upregulation of *FBRS* and the formation of *FBRS*-PRC1.3 complexes alone is insufficient for successful reprogramming, and in fact *AUTS2* downregulation is a stronger signature of successful PTNR, although *AUTS2* was not identified in the genome-wide screen as an impediment to reprogramming, which could be the result of its transcriptional downregulation during reprogramming. Taken together, I show that *FBRS* is the dominant paralogue expressed in naïve PSCs and within naïve PRC1.3 complexes, while *AUTS2* is the dominant paralogue expressed in primed PSCs and within primed PRC1.3 complexes.

5.4.2 – *FBRS* and *AUTS2* expression is mutually exclusive in human PSCs and embryos

It was not previously known if both *AUTS2* and *FBRS* are present within PRC1.3 complexes at the same time. My findings suggest that the two paralogues are mutually exclusive within PRC1.3 complexes, which echos their overall expression patterns.

To determine whether *FBRS* and *AUTS2* are expressed mutually exclusively in primed and naïve PSCs, I analysed a published single-cell RNA-sequencing dataset (Messmer *et al.*, 2019). Single-cell RNA-sequencing has the advantage over bulk RNA-sequencing that it can provide an expression value for both genes in the same cell, rather than relying on averaging from a large number of cells. Analysis of *AUTS2* and *FBRS* expression identified that very few primed or naïve PSCs express both *FBRS* and *AUTS2*, with the majority of naïve PSCs exclusively expressing *FBRS* and the majority of primed PSCs exclusively expressing *AUTS2*. Fewer than 2 % of cells possess an RPKM (reads per kilobase million) greater than 1 for both genes, as a stringent threshold (Fig. 5.4.7).

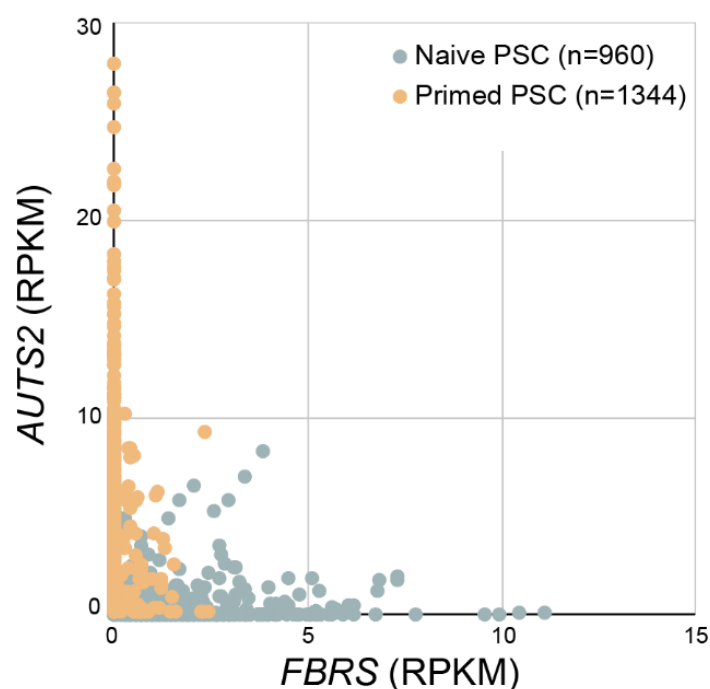


Figure 5.4.7 – Plot shows the expression levels of *FBRS* and *AUTS2* in single cells in primed and naïve PSCs

Naïve PSCs, blue; primed PSCs, orange. Data obtained from (Messmer *et al.*, 2019).

Within primed and naïve PSC cultures, there is a degree of heterogeneity in the expression of *AUTS2* and *FBRS*, but also strong mutual exclusivity in expression of *AUTS2* and *FBRS* are evident at single-cell resolution.

In recent years, single-cell RNA-sequencing has been performed in human embryos, profiling different stages of early development (Blakeley *et al.*, 2015; Huang *et al.*, 2014; Molè *et al.*, 2021; Nakamura *et al.*, 2016; Petropoulos *et al.*, 2016; Stirparo *et al.*, 2018; Tyser *et al.*, 2021; Xiang *et al.*, 2020), including the epiblast cells that are recapitulated *in vitro* by naïve and primed PSCs. I therefore extended my single-cell analysis of *AUTS2* and *FBRS* expression in PSCs into human embryos, to identify whether the paralogue shift only occurs in pluripotent stem cell lines, or if there is potential for developmental relevance in embryos also.

Three distinct cell annotations from human embryos were included in this analysis; inner cell mass, pre-implantation epiblast and post-implantation epiblast (Xiang *et al.*, 2020). In the inner cell mass, a larger proportion of cells express *FBRS* and at higher levels than *AUTS2*. In pre-implantation epiblast cells, *FBRS* and *AUTS2* are expressed at approximately equal levels in

a similar number of cells, but the majority of cells still express either *FBRS* or *AUTS2*. In post-implantation epiblast cells, more cells express *AUTS2* and at higher levels than *FBRS*. This confirms the developmental presence of a paralogue switch, with *FBRS* expression predominant in the inner cell mass and concurrently with embryo implantation, *AUTS2* becomes the dominantly expressed paralogue. Re-applying the stringent threshold shows that fewer than 8 % of cells express both *FBRS* and *AUTS2* with an RPKM > 1 (Fig. 5.4.8 and Fig. 5.4.9).

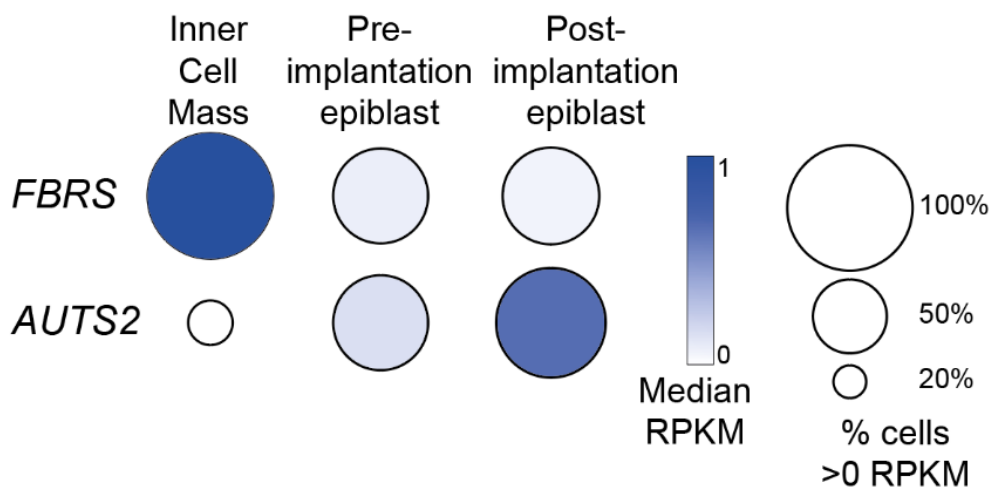


Figure 5.4.8 – Heatmap shows *FBRS* and *AUTS2* expression levels in early human embryos

Circle colour corresponds to the median gene expression level of single cells. Circle size indicates the percentage of single cells with an RPKM value > 0. Inner cell mass (n=10); pre-implantation epiblast cells (n=61); post-implantation epiblast cells (n=64). Data and cell annotations obtained from (Xiang *et al.*, 2020).

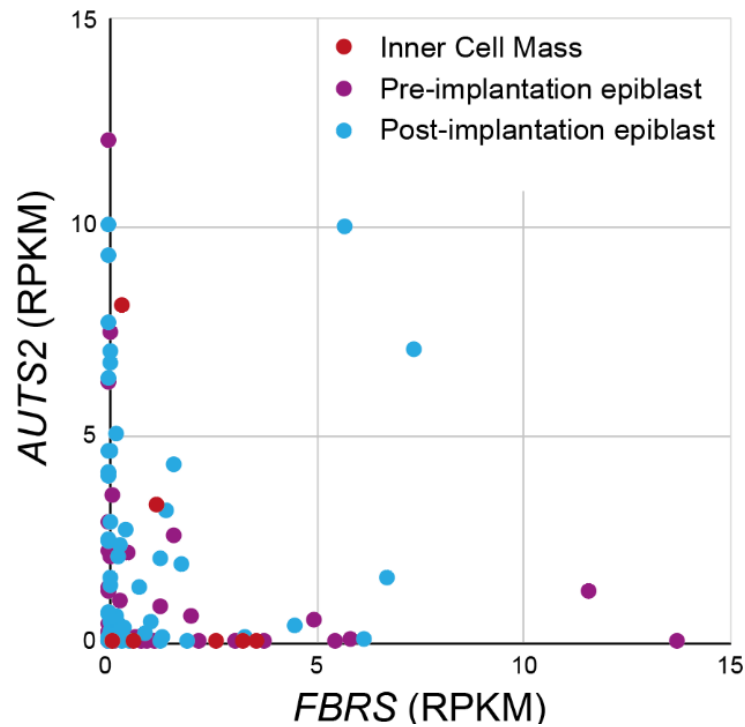


Figure 5.4.9 – Plot shows the *FBRS* and *AUTS2* expression levels in single cells in early human embryos

Data and cell annotations obtained from (Xiang *et al.*, 2020).

A developmentally-regulated shift in paralogue expression and mutually exclusive expression of the paralogues is a feature of human embryo development. Single-cell RNA-sequencing has been performed on cynomolgus monkey embryos as a non-human primate model organism. I also analysed the expression dynamics of the *FBRS* and *AUTS2* paralogues in primate embryos. The general expression trend of the paralogues is conserved between humans and cynomolgus monkeys, with developmental progression from the inner cell mass to the late post-implantation epiblast state causing decreased *FBRS* expression and increased *AUTS2* expression (Fig. 5.4.10).

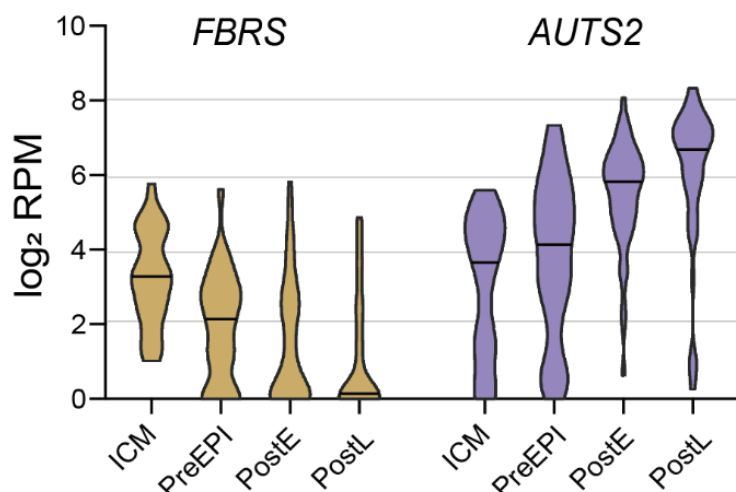


Figure 5.4.10 – Violin plots show the expression levels of *FBRS* and *AUTS2* in single cells during cynomolgus monkey embryo development

ICM, inner cell mass; PreEPI, pre-implantation epiblast; PostE, early post-implantation epiblast; PostL, late post-implantation epiblast. Data obtained from (Nakamura *et al.*, 2016).

The *FBRS* and *AUTS2* paralogue shift is therefore seems to be conserved across primates. Although not confirmatory, this evolutionary conservation of this expression pattern could indicate its importance across primates. The model below (Fig. 5.4.11) summarises my findings with respect to the regulation of the PRC1.3 complex during human PTNR.

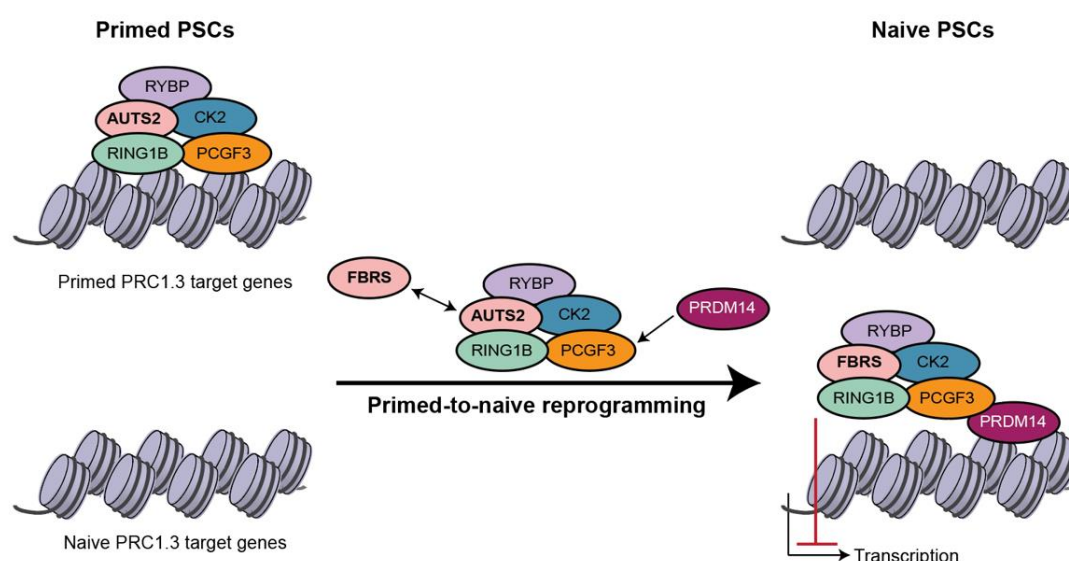


Figure 5.4.11 – Schematic showing the developmentally-regulated paralogue switch of FBRS and AUTS2 within the PRC1.3 complex

During human PTNR, AUTS2-containing PRC1.3 complexes switch to FBRS-containing PRC1.3 complexes and interact with PRDM14, become bound at a distinct set of target genes along with PRDM14. The expression of this gene subset is repressed during reprogramming. These genes fail to become repressed when PRDM14 is depleted or when PRC1.3 is deleted, both of which cause a failure to successfully reprogramme to naïve PSCs.

This is a novel finding in the biology of the PRC1.3 complex and provides an intriguing insight into developmental regulation of paralogue expression in human pluripotency, which could be explored further, as well as a further understanding of a key molecular event during PTNR and human embryo development.

5.5 – Interrogating PRC1.3 enzymatic activity in primed and naïve PSCs

In Chapter 4, H2AK119ub1 is more abundant in naïve PSCs than primed PSCs. The essential requirement for PRC1.3 in PTNR and the change in the composition and interactions of the complex across the cell state transition makes PRC1.3 a strong candidate for the remodelling of the Polycomb-associated epigenome between primed and naïve PSCs.

I wanted to investigate whether PRC1.3 and its potential for distinct primed-specific and naïve-specific regulation contributed to this phenotype, as PRC1.3 is responsible for a large proportion of H2AK119ub1 in human primed PSCs. The total chromatin-associated abundance of most core PRC1.3 components did not significantly change between primed and naïve PSCs and neither did the PRC1.3 deubiquitylase BAP1 (Zijlmans *et al.*, 2022). A potential explanation for this phenotype could be that PRC1.3 complexes in naïve PSCs possess a higher level of enzymatic activity than in primed PSCs.

Due to the absence of a naïve PRC1.3 knockout or knockdown system, I instead devised an *in vitro* histone post-translational modification assay to assess the catalytic activity of PRC1.3 complexes from naïve and primed PSCs. *In vitro* H2AK119ub1 assays have been conducted previously (Elderkin *et al.*, 2007; Gao *et al.*, 2014, 2012; Groh *et al.*, 2016; Rose *et al.*, 2016; Shukla *et al.*, 2021; H. Wang *et al.*, 2004) and these studies utilised recombinant PRC1 proteins produced *in vitro*. However, this approach would mean that the pluripotency state specific regulation of PRC1.3 that I have identified would be lost. As, I wished to study the potential influence of these factors on enzymatic activity, I instead immunoprecipitated PRC1.3 from human naïve and primed PSCs and used cell-derived PRC1.3 in the place of a recombinant E3 ubiquitin ligase.

The ubiquitylation machinery in cells requires a number of components (Pickart, 2001); an E1 ubiquitin-activating enzyme, an E2 ubiquitin-conjugating enzyme, an E3 ubiquitin protein ligase, free ubiquitin, an E3 ligase substrate and ATP. By mixing these components *in vitro*, but using a cell-derived E3 ubiquitin protein ligase, an *in vitro* assay should be possible to establish. This experiment can compare the enzymatic activity of the PRC1.3 complex derived from primed and naïve PSCs (Fig. 5.5.1).

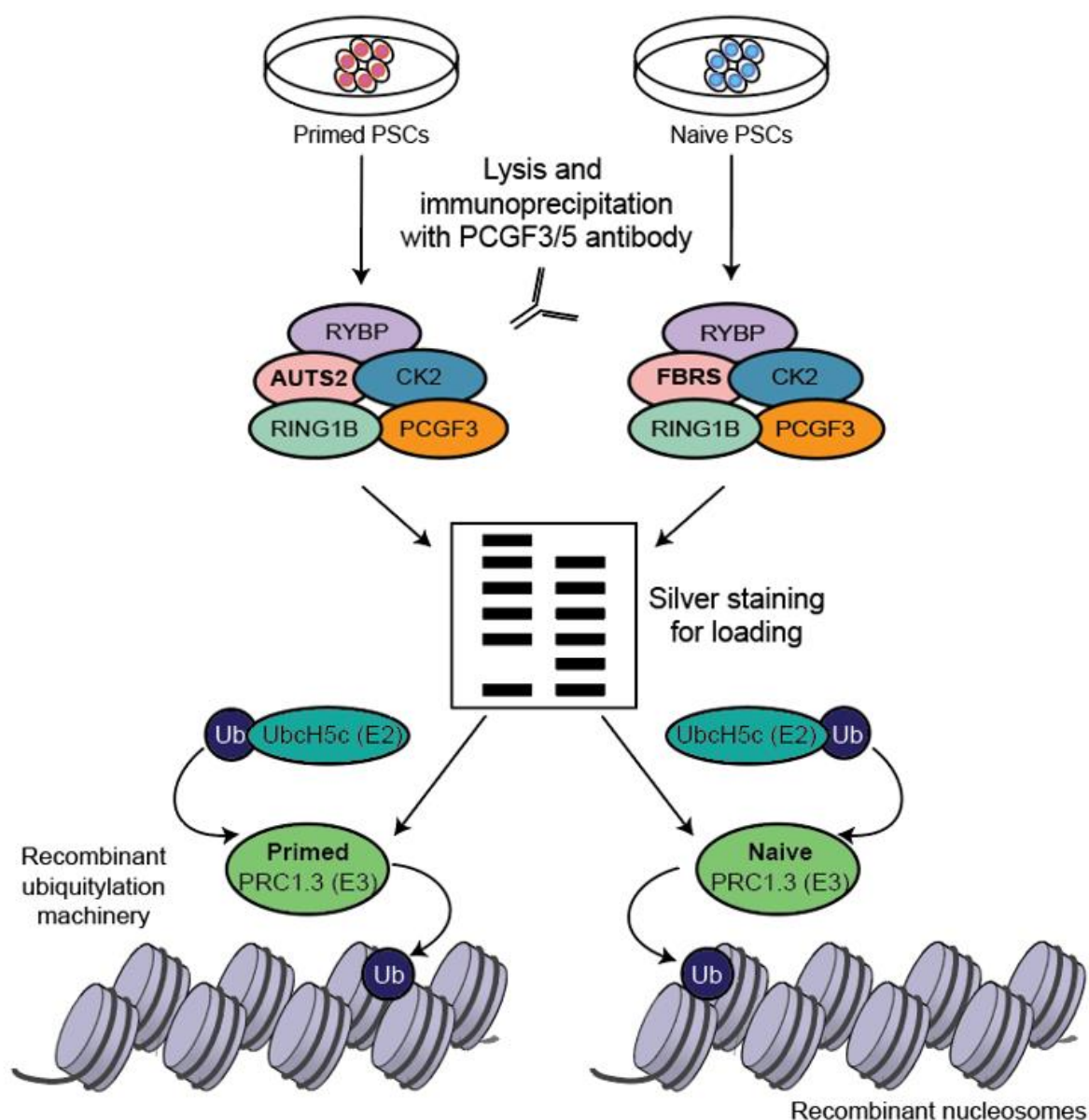
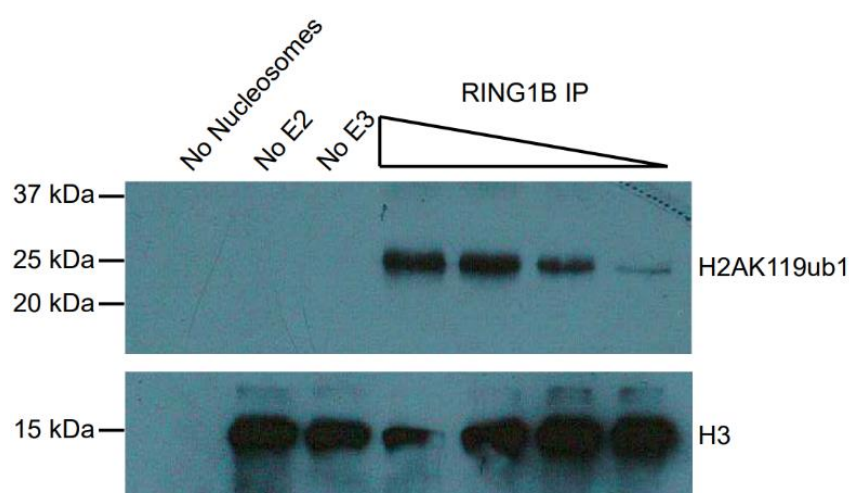


Figure 5.5.1 – Schematic showing the method used for the PRC1.3 *in vitro* ubiquitylation assay

Primed and naïve PSCs were lysed and nuclei extracted. An anti-PCGF3/5 antibody was used to specifically isolate the PRC1.3 complex and stringent washing was performed. Silver staining was used to quantify the concentration of PRC1.3 complex detected, and equal concentrations of PRC1.3 from naïve and primed PSCs were added to separate *in vitro* H2AK119ub1 assays, with the amount of H2AK119ub1 generated on recombinant polynucleosomes detected by Western blotting.

The *in vitro* assay used a ubiquitin-charged E2 ubiquitin-conjugating enzyme, negating the need for the addition of free ubiquitin, an E1 enzyme and ATP. This E2 was combined with commercial recombinant polynucleosomes and the added PRC1.3 E3 enzyme. The quantities of the components in the *in vitro* assay were kept the same to ensure that the only factor affecting the level of H2AK119ub1 detected by Western blot was the catalytic activity of PRC1.3 itself.

I conducted a pilot experiment to confirm that it was possible to produce H2AK119ub1 *in vitro* using this approach. This was performed by immunoprecipitating RING1B from primed PSCs in order to pull down all PRC1 complexes, to maximise the amount of H2AK119ub1 produced. This result confirmed the ability of this approach to generate H2AK119ub1 *in vitro* using intracellular PRC1 complexes. (Fig. 5.5.2).



This experimental approach was extended to the specific pull down of PRC1.3 using the anti-PCGF3/5 antibody and comparisons were performed, using silver staining to adjust loading. Due to technical issues with silver staining, it was difficult to ascertain the bands corresponding to known PRC1.3 complex components. It was therefore not possible to accurately quantify the levels of PRC1.3 complex components post-immunoprecipitation, making it difficult to add equal amounts of the PRC1.3 complex from both primed and naïve PSCs to the *in vitro* assay with confidence. To assess the amount of PRC1.3 present in the *in vitro* assay, I therefore had to retrospectively probe the Western blot for the catalytic component of the PRC1.3 complex, RING1B. Immunoprecipitated PRC1.3 is able to generate H2AK119ub1 while the modification was not detected in immunoprecipitated material from *PCGF3/5* DKO primed PSCs. As the quantity of PRC1.3 could only be determined retrospectively, more PRC1.3 from primed PSCs was added to the *in vitro* assay than from naïve PSCs, and meant that more H2AK119ub1 was detected from primed cells (Fig. 5.5.3).

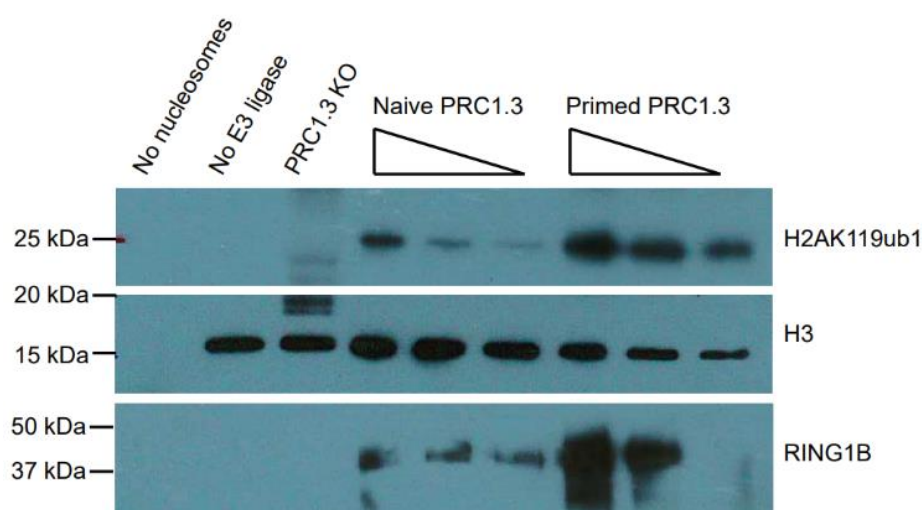


Figure 5.5.3 – H2AK119ub1 deposited by PRC1.3 complexes can be detected *in vitro*

PRC1.3 KO lane represents an anti-PCGF3/5 pulldown in *PCGF3/5* DKO primed PSCs and was used as a negative control to confirm the specificity of PRC1.3-deposited H2AK119ub1 *in vitro*. Molecular weight in kDa. Triangles indicate a titration range of the amount of PRC1.3 complex added.

The experimental set-up was subsequently repeated to improve the loading, but this remained difficult to control. When assessing RING1B levels, the lower concentrations from primed and naïve PSCs were not comparable. However, the two upper concentrations of PRC1.3 were

approximately equal in RING1B levels and the intensity of H2AK119ub1 *in vitro* was similar between these two conditions (Fig. 5.5.4).

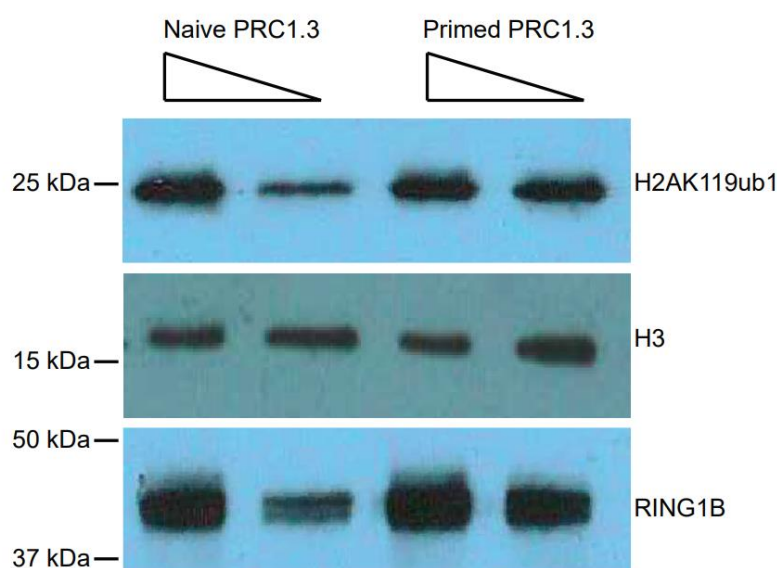


Figure 5.5.4 – H2AK119ub1 deposited by PRC1.3 complexes can be detected *in vitro*

Molecular weight in kDa.

Comprehensive repetition with improved loading and optimisation of the protocol is required to produce conclusive results. As a preliminary result, the increased abundance of H2AK119ub1 on naïve PSC chromatin is therefore unlikely due to pluripotency state-specific regulation of PRC1.3 enzymatic activity, such as the detectable shift between AUTS2 and FBRS paralogues. Another mechanism instead must govern the epigenetic remodelling of Polycomb associated modifications during reprogramming.

5.6 – Discussion

5.6.1 – PRC1.3 and PRDM14 cooperate in transcriptional repression during reprogramming

qPLEX-RIME was used to identify a number of state-specific PRC1.3 interacting proteins, with potential roles in modulating the function of PRC1.3 or its recruitment to chromatin. As PRC1.3 does not contain any proteins with characterised DNA binding ability, it is unclear how exactly PRC1.3 is targeted to chromatin. Previously identified PRC1.3 interactors P300 (Gao *et al.*, 2014), TEX10 (Zhao *et al.*, 2017), USF1 (Scelfo *et al.*, 2019) and hnRNP (Pintacuda *et al.*, 2017), did not provide promising candidates for state-specific regulators of PRC1.3.

This work did however identify a novel co-interaction between PRDM14 and PRC1.3 in naïve PSCs. The presence of a targeting motif within chromatin at naïve PRC1.3 target genes, the binding of PRDM14 at these regions and the naïve-specific interaction between PRDM14 and PRC1.3 all led me to hypothesise that PRDM14 may recruit PRC1.3 to its naïve target genes, which are repressed as reprogramming progresses. Transcriptional repression of these genes relies on both PRC1.3 and PRDM14, suggesting that the two factors cooperate, with loss of either PRC1.3 or PRDM14 causing a failure of cells to reprogramme. I am however unable to conclusively determine whether PRDM14 does in fact recruit PRC1.3 to these genes.

Whether PRDM14 directly recruits PRC1.3 to chromatin could have been tested in a planned experiment using the PRDM14-AID-VENUS naïve knockdown cells that I established. ChIP-sequencing for RING1B and PCGF3 in these cells when PRDM14 is depleted, as well as in undepleted control cells, could have established whether PRC1.3 dissociates from its naïve target genes in the absence of PRDM14. I tried to conduct this experiment but ultimately the instability of the PRDM14-AID-VENUS naïve PSCs was prohibitive for sufficient expansion of this cell line, perhaps due to a leaky PRDM14 degron, which could have been improved by culture with auxinole (Hayashi *et al.*, 2012; Yesbolatova *et al.*, 2019), or clonal instability. CUT&RUN or CUT&Tag (Bendall *and* Semprich, 2022; Kaya-Okur *et al.*, 2019; Meers *et al.*, 2019; Skene *and* Henikoff, 2017) experiments that require a lower cell input could instead have been attempted, although there is no guarantee that these experiments would have been successful with the RING1B or PCGF3/5 antibody. Future research could explore further the molecular mechanism by which PRDM14 and PRC1.3 target and regulate specific genes, and the functions played by these genes in reprogramming, enhancing our overall comprehension of this cell state transition.

Within the scope of this study, I only had the opportunity to interrogate PRDM14 as a state-specific PRC1.3 co-interacting protein but a number of the other identified factors could plausibly play an important, and as yet undefined, role in the regulation of PRC1.3 in human pluripotency and reprogramming. Essential factors that were detected as co-interacting specifically with PRC1.3 in primed PSCs included DDX5 and CSTF2T, both of which are reported as RNA binding proteins that function in nascent mRNA splicing and post-transcriptional processing (Legrand *et al.*, 2019; Yao *et al.*, 2013), and co-interactions between PRC1.3 and RNA binding proteins have been previously reported (Pintacuda *et al.*, 2017). In naïve PSCs, two of the highest affinity co-interactors of the PRC1.3 complex were DNTM3L and HIST1H1A, albeit neither were identified as essential for reprogramming in the genome-wide screen (#6,117 and 5,522 factors respectively). Both proteins are highly abundant on chromatin in naïve PSCs, suggesting that they could instead be contaminating proteins within the dataset (De Clerck *et al.*, 2019; Zijlmans *et al.*, 2022), with further study needed to confirm an interaction.

5.6.2 – *AUTS2* and *FBRS* undergo a paralogue shift in human PSCs and development

For the first time, by using qPLEX-RIME, my work has identified a developmentally-regulated shift in the composition of the PRC1.3 complex itself, which has previously been seen only in limited studies of PRC2 complexes in mammals and plants (De Lucia *et al.*, 2008; Xu *et al.*, 2015).

The *AUTS2* and *FBRS* paralogues undergo a developmentally-regulated shift within the complex, with naïve PRC1.3 complexes being enriched for *FBRS* and primed PRC1.3 complexes enriched instead for *AUTS2*. Whether there are two distinct forms of the complex has not been completely established however. A co-immunoprecipitation or another qPLEX-RIME study, using either *AUTS2* or *FBRS* as the immunoprecipitation target, could definitively establish whether *AUTS2* and *FBRS* are ever present within the same PRC1.3 complex.

Further analysis could be used to examine the composition of the PRC1.3 complex. Structural studies, such as X-ray crystallography or cryo-electron microscopy, could reveal the conformation and composition of *AUTS2*- and *FBRS*-containing PRC1.3 complexes. This could establish where the binding sites of *AUTS2* and *FBRS* are within PRC1.3 complexes and whether they overlap, meaning that binding of the paralogues must be mutually exclusive, which has been structurally determined for RYBP and CBX7 in PRC1 complexes previously

(Chittock *et al.*, 2017). If *AUTS2* and *FBRS* binding within PRC1.3 complexes is structurally mutually exclusive, this finding could be extrapolated to all other cell types.

The developmental regulation of *FBRS* and *AUTS2* expression and how these paralogues are incorporated into PRC1.3 complexes is also unclear, and these processes could be linked. CHIP-sequencing identified *AUTS2* as a naïve PRC1.3 target gene, that becomes transcriptionally repressed as cells reprogramme (Chapter 4). It is therefore possible that as reprogramming begins, PRC1.3 autorepresses *AUTS2* transcription, and by alternative means, *FBRS* becomes transcriptionally upregulated. *FBRS*-PRC1.3 complexes then regulate the crucial molecular events for reprogramming, consistent with *AUTS2* downregulation acting as a hallmark for successful reprogramming. The developmental regulation of the expression of the two paralogues requires further investigation to build a comprehensive molecular understanding.

It is not known whether *FBRS* and *AUTS2* are functionally redundant with one another within PRC1.3 complexes. *AUTS2*, *FBRS* and *FBRS1* share a region with a high degree of homology, called the 'AUTS2 family domain' (Kondrychyn *et al.*, 2017), as the paralogues arose from a whole genome duplication event (Singh *et al.*, 2015). Notably, *FBRS1* was not detected in the PRC1.3 qPLEX-RIME data, which suggests it is not abundant within PRC1.3 complexes in human pluripotency. The schematic below shows the amino acid sequence similarity of the *AUTS2* and *FBRS* paralogues (Fig. 5.6.1).

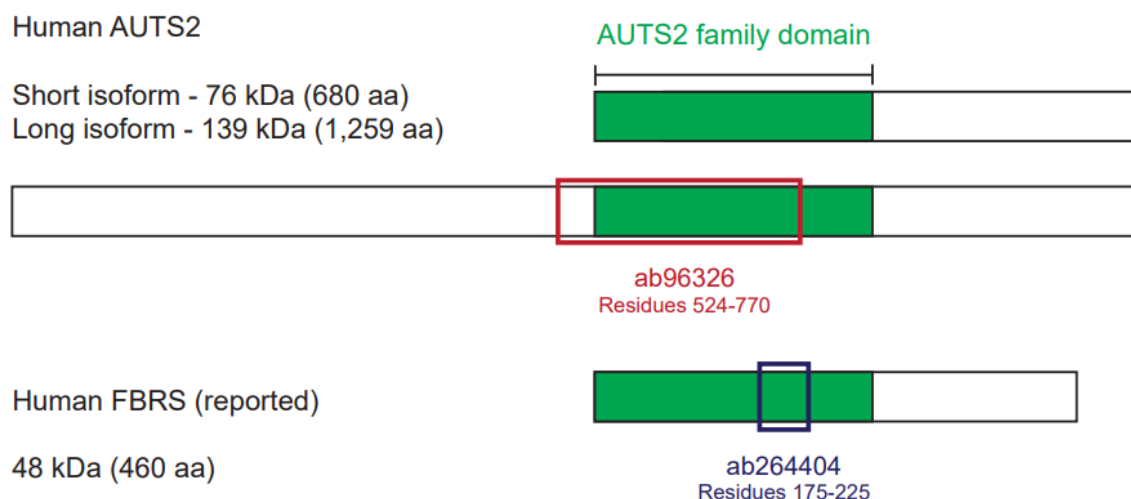


Figure 5.6.1 – Schematic showing the homology between AUTS2 and FBRS proteins

The AUTS2 family domain is a region of homology between *AUTS2* and *FBRS* (Kondrychyn *et al.*, 2017). Both a long and a short isoform have been reported for the *AUTS2* gene (Monderer-Rothkoff *et al.*, 2019). Red and blue boxes show the regions detected by the AUTS2 and FBRS antibodies used in this study. The antibodies are specific for AUTS2 and FBRS, despite the homology between the genes.

The short isoform of *AUTS2* more closely resembles the amino acid sequence of *FBRS*, but only the longer *AUTS2* isoform however interacts with PCGF3 and PCGF5 (Monderer-Rothkoff *et al.*, 2019). The much larger size of the long isoform of *AUTS2* than *FBRS* means that *AUTS2*-PRC1.3 complexes could possess a different structural conformation to *FBRS*-PRC1.3 complexes, but the potential effect of this is unknown.

Although only *FBRS*, and not *AUTS2*, is critical for human PTNR, this does not mean that the two paralogues are not functionally redundant. For example, when it would ordinarily be downregulated, overexpression of *AUTS2* could re-facilitate *FBRS* KO primed PSCs to reprogramme. This could be tested experimentally, and if these cells still could not reprogramme, this would suggest that *AUTS2* and *FBRS* are not functionally redundant, and only the *FBRS* can facilitate reprogramming.

One of the ways that *AUTS2* and *FBRS* could be functionally distinct and affect PRC1.3 complex function across reprogramming could be in regulating the interaction between PRC1.3 and PRDM14. I established that PRC1.3 and PRDM14 interact with one another specifically in naïve PSCs, but how this interaction exactly occurs and is pluripotency state specific is not known. The interaction between PRC1.3 and PRDM14 could be mediated by *FBRS*, and *AUTS2*-PRC1.3 complexes may be unable to interact with PRDM14. This hypothesis could be tested by inducibly overexpressing *FBRS* in human primed PSCs, where *FBRS* is ordinarily expressed at very low levels and testing for an interaction between *FBRS* and PRDM14 using co-immunoprecipitation Western blotting. As the PRDM14 antibody I used was not suitable for immunoprecipitation, this overexpression could be performed in PRDM14-AID-VENUS primed PSCs, and the anti-GFP antibody could be used to pull down PRDM14-VENUS and assess whether *FBRS* binding can be detected.

The paralogue shift was seen both in PSCs and in human embryos, indicating there could be developmental relevance underpinning this finding. The developmental importance of this finding is however unclear, including whether two distinct forms of PRC1.3 complex possess different functions. The function of *AUTS2* has been largely described in transcriptional

activation of mammalian neurodevelopmental genes (Gao *et al.*, 2014). It is possible that the upregulation of *AUTS2* and enrichment of *AUTS2* within PRC1.3 complexes in post-implantation epiblast cells is important for subsequent lineage specification during gastrulation and beyond, where PRC1.3/5 functionality has been reported (Gao *et al.*, 2014; Monderer-Rothkoff *et al.*, 2019; Wang *et al.*, 2018; Yao *et al.*, 2018; Zhao *et al.*, 2017). This could be assessed *in vitro* by the generation and differentiation of *AUTS2* KO primed PSC lines. Differentiation could use a directed methodology towards a neural or mesoderm lineage, where PRC1.3/5 has been previously demonstrated as important in mouse (Gao *et al.*, 2014; Yao *et al.*, 2018; Zhao *et al.*, 2017), or using an unbiased analysis such as the generation of gastruloids (Moris *et al.*, 2020). If *AUTS2* KO causes defective differentiation or transcriptional perturbations, this could provide some evidence to support a post-implantation specific role for *AUTS2*-PRC1.3 complexes, which could then be mechanistically investigated. On the other hand, *FBRS* function in PRC1.3 complexes is not at all understood but the generation of an *FBRS* knockdown or knockout in naïve PSCs could shed some light on this.

The study of the cellular, molecular and epigenetic impact of *AUTS2* and *FBRS* paralogues within PRC1.3 complexes is important to enhance our understanding of the regulation of human pluripotency and reprogramming and by extension this could be hugely informative about events occurring at some of the earliest stages of human embryo development.

5.6.3 – The PRC1.3 complex does not appear to differ in catalytic activity between the primed and naïve state of pluripotency

Although the focus of my experimental approach was on PRC1.3 targeting to gene loci, I am aware that PRC1.3, and PRC1.3-deposited H2AK119ub1, could have additional genome-wide epigenetic functions in human PSCs that could also be critical for reprogramming (Conway *et al.*, 2021; Fursova *et al.*, 2021, 2019; Huseyin and Klose, 2020). I have discovered that the PRC1.3 complex undergoes distinct exogenous regulation between the naïve and primed pluripotent states, such as the naïve-specific interaction with PRDM14, as well as intrinsic regulation of the composition of the complex itself, with a shift in the *AUTS2* and *FBRS* paralogues. This regulation could bestow primed-specific and naïve-specific properties to the complex. I have already determined that in primed PSCs, *AUTS2*-PRC1.3 complexes must contribute to a large proportion of global H2AK119ub1, while levels of both H2AK119ub1 and H3K27me3 are elevated in naïve over primed PSCs (Chapter 4). As the levels of the majority of PRC1.3 components on chromatin do not significantly change (Zijlmans *et al.*, 2022), I

speculated that regulation of PRC1.3 could increase the catalytic activity of the complex in naïve PSCs, and that this could be important for reprogramming.

Without a PRC1.3 knockout or knockdown system in naïve PSCs, which I attempted to generate without success (Chapter 3), the specific epigenetic contribution of FBRS-PRC1.3 complexes in naïve PSCs cannot be determined, but this remains important to establish. Instead, I performed an *in vitro* H2AK119ub1 assay using PRC1.3 from naïve and primed PSCs to test the catalytic activity of the complex in naïve and primed PSCs. These reactions were difficult to control due to technical difficulties, which required retrospective assessments of protein loading and iterative improvements. Despite this, my preliminary experiments suggest that, contrary to my predictions, the activity of AUTS2-PRC1.3 complexes from primed PSCs and FBRS-PRC1.3 complexes from naïve PSCs appear to be similar.

This experimental approach requires more optimisation, particularly in the silver staining step to ascertain the PRC1.3 component protein concentration. The presence of strong bands detected at 50 kDa and 25 kDa, at the molecular weight of the antibody heavy and light chain, as the protocol used an immunoprecipitation step. These bands obscured the detection of specific proteins on the blot, and are not an issue when recombinant proteins are purified *in vitro*. However, I specifically wanted to study the effect of intracellular regulation on the PRC1.3 complex enzymatic activity, rather than using recombinant protein, meaning an immunoprecipitation step was always necessary. One possible solution to this issue would have been to knock-in an epitope tag to the *PCGF3* gene, for example a FLAG tag, for which an antibody has been proven to work extremely well for immunoprecipitation. This approach could have reduced background in the protein pulldown and meant that less antibody would be required, improving the detection of specific PRC1.3 proteins, and optimisation of the post-immunoprecipitation wash steps could also improve the signal to noise of protein detection.

However, the lack of detection of catalytic differences for PRC1.3 between primed and naïve PSCs suggests that the differences in global H2AK119ub1 levels between the two cell states may be the result of other canonical or variant PRC1 complexes, or as a result of developmental regulation of the H2AK119 deubiquitylase *BAP1* (Campagne *et al.*, 2019; Fursova *et al.*, 2021). Published chromatin proteomics however shows that amongst the other PCGF proteins detected, PCGF2 is significantly enriched on primed PSC chromatin while PCGF6 and the H2AK119ub1 deubiquitinase *BAP1* shows no difference between the two cell types (Zijlmans *et al.*, 2022). Therefore, increased H2AK119ub1 levels resulting from other PRC1 complexes or from reduced *BAP1* function would also likely be the result of regulation of their enzymatic activity, rather than mirroring their abundance on chromatin. In all, it is not known how the genome-wide remodelling of Polycomb-associated epigenetic modifications is

regulated between human primed and naïve PSCs and whether it is in fact PRC1.3-dependent or not.

6 – Summary and Conclusions

6.1 – Reprogramming remains an enigmatic process

A little over a decade after the first study generated a human naïve state of pluripotency *in vitro* (Hanna *et al.*, 2010), our understanding of naïve PSCs and their remarkable properties continues to expand, alongside our understanding of human pre-implantation embryogenesis. However, the ways in which naïve PSCs can be generated remain restricted to either direct derivation from the inner cell mass of human pre-implantation blastocysts (Guo *et al.*, 2016b), induced reprogramming of somatic cells (Giulitti *et al.*, 2019; Kilens *et al.*, 2018) or primed-to-naïve reprogramming (PTNR) (Bayerl *et al.*, 2021; Chan *et al.*, 2013b; Chen *et al.*, 2015; Duggal *et al.*, 2015; Gafni *et al.*, 2013; Guo *et al.*, 2017; Hanna *et al.*, 2010; Khan *et al.*, 2021; Lynch *et al.*, 2020; Qin *et al.*, 2016; Szczerbinska *et al.*, 2019; Takashima *et al.*, 2014; Theunissen *et al.*, 2014; Ware *et al.*, 2014; Zimmerlin *et al.*, 2016).

The genome-wide CRISPR-Cas9 screen performed by Amanda Collier identified novel regulators of human PTNR in an unbiased manner. This screen produced a rich dataset where 446 genes were identified as important for human PTNR and 540 genes were identified as impediments of human PTNR (Collier *et al.*, 2022). PTNR is a poorly understood process, that shares broad characteristics with iPSC reprogramming, as both processes are slow and inefficient, with reported human and mouse iPSC reprogramming efficiencies between 0.00001 and > 1 %, meaning that the vast majority of cells do not reprogramme (Cieślak-Pobuda *et al.*, 2017). However, this screen identified that there is minimal overlap in the genes that are important for iPSC reprogramming and PTNR (Collier *et al.*, 2022).

Reprogramming inefficiency generates cellular heterogeneity even when all cells are exposed to the same external conditions. Whether an individual cell will reprogramme or not is likely stochastic, but successfully reprogrammed cells must undergo a number of defined molecular events. For example, during iPSC reprogramming, a 'labyrinth to pluripotency' has been postulated, whereby cells must initially counteract apoptosis and continue to proliferate, progress through the mesenchymal-to-epithelial transition (MET), upregulate endogenous pluripotency markers and lose the epigenetic memory of the starting cell type (Papp and Plath, 2011). I would predict that a similar 'labyrinth of PTNR' must involve counteracting apoptosis, continued proliferation, transition of primed-specific to naïve-specific pluripotency transcription networks and epigenetic remodelling. However, the mechanisms that govern these processes and the factors involved are not known. Mechanistic exploration of the candidates identified by the genome-wide screen posed an opportunity to derive novel information about the poorly understood PTNR process.

This project broadly aimed to provide insight into the key factors and specific molecular events that underpin PTNR. Continued exploration of this process could in the long-term build a ‘molecular roadmap’ to naïve pluripotency and identify any molecular bottlenecks that might restrict reprogramming efficiency. With this knowledge, interventions could be developed to increase the ability of cells to reprogramme to high quality naïve PSCs, from a range of primed cell lines or backgrounds, with high efficiency and on an expedited timescale. Increasing the accessibility of naïve PSCs as an *in vitro* tool could further unlock their unique beneficial properties in the study of human pre-implantation development and generation of desirable cell types. These findings could also more broadly inform on the mechanisms that cells use to change their identity during cell state transitions, which could be more generally applicable to developmental cell fate remodelling events in embryos.

Surprisingly, the genome-wide screen identified almost all of the components of the human Polycomb Repressive Complex PRC1.3 subtype as key regulators of PTNR, and this was unexpectedly specific to PRC1.3 itself, rather than PRC1 or PRC2 complexes more broadly. As PRC1.3 has never before been studied in human pluripotency and reprogramming, the study of PRC1.3 in this context provides novel insight into how the complex regulates human pluripotent cells and cell state reconfiguration. The central question of this project was: how does the PRC1.3 complex modulate the reprogramming of primed PSCs to the naïve state?

6.2 – PRC1.3 is essential for human PTNR

The first outcome of my work was the generation of genetic deletions of the PRC1.3 component genes *PCGF3* and *FBR3* in primed PSCs using CRISPR-Cas9. As the role of PRC1.3 in human primed pluripotency has never been explored and the starting cell state could exert strong influence on the outcome of reprogramming, I characterised the primed state of pluripotency in these PRC1.3-deficient cells. I was not able to detect any cellular or transcriptional differences between WT and PRC1.3 KO primed PSCs. Interestingly however, I subsequently determined that PRC1.3 KO strongly decreases the abundance of Polycomb-associated H2AK119ub1 and H3K27me3 modifications in primed PSCs. The discrepancy between the epigenome and transcriptome is yet to be fully explored, as I do not know where PRC1.3-dependent H2AK119ub1 and H3K27me3 is distributed on chromatin in primed PSCs, but this finding is not without precedent (Fursova *et al.*, 2019). The PRC1.3-dependent decrease in H3K27me3 levels is consistent with PRC2 complexes being recruited to chromatin in a H2AK119ub1-dependent manner (Blackledge *et al.*, 2014; Cooper *et al.*, 2016, 2014; Kalb

et al., 2014), which has also been observed in early mouse embryogenesis (Chen *et al.*, 2021; Mei *et al.*, 2021).

A caveat of my work is that I have not established whether the PRC1.3 complex is important for the differentiation of human primed PSCs. Although this seemed tangential to my central question regarding reprogramming, it is plausible that the identification of PRC1.3 from the genome-wide screen could reflect PRC1.3 being more broadly critical for pluripotency cell state transitions, rather than specifically in reprogramming. However, two pieces of evidence suggest that this is not the case. Firstly, PRC1.3 KO cells form neural-like cells during PTNR and so cells without PRC1.3 can at least differentiate into these cell types. Secondly, a CRISPR screen to identify genes important for differentiation of primed PSCs into the ectoderm, mesoderm and endoderm germ layers did not identify any PRC1.3 components as important for the formation of any of these three cell types (Yilmaz *et al.*, 2020).

The epigenome is extensively remodelled during both reprogramming and differentiation of primed PSCs, and ablation of H3K27me3 has been previously demonstrated to impair the ability of primed PSCs to differentiate competently into mature cell types (Collinson *et al.*, 2016; Shan *et al.*, 2017). The role of PRC1 complexes or specific PRC1 subtypes in human differentiation has not been specifically studied, but the dependence of a large proportion of global H3K27me3 on PRC1.3 indicates the potential for a defective differentiation phenotype, perhaps in terminal cell fate specification from multipotent cells. Loss of H2AK119ub1 and H3K27me3 from bivalent promoters could severely disrupt the finely coordinated expression of developmental genes (Azuara *et al.*, 2006; Bernstein *et al.*, 2006; Brookes *et al.*, 2012; Stock *et al.*, 2007), but the loss of PRC1.3-dependent non-promoter localised H2AK119ub1 and H3K27me3 (Conway *et al.*, 2021; Fursova *et al.*, 2019) could have different, or as of yet unknown, effects in cell state transitions. As well as changes to H2AK119ub1 and H3K27me3 abundance, I discovered that PRC1.3 deletion also showed wider effects on detectable epigenetic modifications and in particular, changes in the abundance of the histone H2A variants detected on chromatin. This is a novel finding, however its significance and the role that the PRC1.3 complex might play in pluripotency chromatin dynamics is yet to be understood.

My second notable outcome of this work is that I was able to comprehensively establish that the PRC1.3 complex is in fact essential for successful human PTNR, as predicted by the genome-wide screen. This finding was confirmed using two different reprogramming protocols. Different reprogramming methodologies use different combinations of intracellular signalling modulators, likely possess different underlying transcriptional dynamics that influence how naïve pluripotency is acquired and produce naïve PSCs with different molecular

properties (Huang *et al.*, 2014). Regardless of this, both 5i/L/A (Theunissen *et al.*, 2014) and chemical resetting (Guo *et al.*, 2017) based reprogramming are PRC1.3-dependent. PRC1.3 KO cells under reprogramming conditions downregulate *NANOG* expression, and this preceded *OCT4* downregulation. *NANOG* and *KLF2* overexpression has been previously shown to be sufficient to reprogramme human primed PSCs to the naïve state in minimal 2i/LIF conditions (Takashima *et al.*, 2014), indicating that *NANOG* is master regulator of PTNR. This finding highlighted that the failure of PRC1.3-deficient cells to reprogramme manifests in the collapse of the core pluripotency network, which dooms the ability of these cells to reprogramme and causes exit from pluripotency. Nevertheless, identifying PRC1.3-dependent phenomena, albeit indirect, that are required for maintenance of the core pluripotency network can contribute to the molecular roadmap of successful PTNR.

An essential role for PRC1.3 in human PTNR could essentially implicate the complex in other types of reprogramming, such as iPSC reprogramming. However, a genome-wide screen of fibroblast to primed iPSC reprogramming did not identify any components of the PRC1.3 complex as essential regulators of this process. This implicates PRC1.3 specifically in reprogramming to naïve pluripotency. As no comparable screen of PTNR has been performed in other species, it is not currently known whether the requirement for PRC1.3 is human-specific, or is conserved across mammals. Many of the properties of human and mouse PSCs are conserved, with mESCs providing a framework for the properties of human naïve pluripotency prior to the ability to use human pre-implantation embryos as a benchmark. However, the signalling requirements for mouse and human naïve PSCs reprogramming and maintenance are not identical, and neither is the timing, with mouse primed cells able to much more quickly reprogramme to the naïve state, and so species-specific mechanisms are also operative. Using a *Pcgf3* knockout or knockdown system in mouse pluripotency, and testing the ability of this cells to convert from mEpiSCs to mESCs could identify whether this critical function of PRC1.3 is conserved across mammalian pluripotency. Collectively, the findings from Chapter 3 advance our understanding of PTNR by demonstrating the critical nature of PRC1.3 in the formation of naïve PSCs, thus affirming the cellular phenotype predicted by the genome-wide screen.

Despite confirming that PRC1.3 is essential for human PTNR, an important limitation of my work is the inability to generate a naïve knockdown system to remove the PRC1.3 complex from established human naïve PSCs. This experimental approach could have proven remarkably informative on various aspects of this investigation, into how PRC1.3 regulates human naïve pluripotency and reprogramming. For example, I have not been able to establish whether PRC1.3 is essential for only reprogramming to naïve PSCs or if the complex is instead more broadly important for the survival or maintenance of the human naïve pluripotent state.

During reprogramming, these two possibilities could possess distinct molecular mechanisms and timing, but the cellular phenotype would ultimately present in the same way as an absence of naïve PSCs during reprogramming. A naïve knockdown system could also be used to compare the cellular, transcriptional and epigenetic function of PRC1.3 between primed and naïve pluripotency. A PRC1.3 knockdown system could also be used to test hypotheses regarding the function of PRC1.3 in human pre-implantation embryo development. For example, it is not known whether PRC1.3 is important for human X chromosome inactivation, as it is in mESCs (Almeida *et al.*, 2017), and this could be tested *in vitro* during the capacitation of naïve PSCs. This would advance our understanding of the mechanisms used by human cells to epigenetically mediate dosage compensation during development.

6.3 – PRC1.3 is associated with transcriptional repression during reprogramming

The next outcome of my project was to use RNA-sequencing to identify the genome-wide transcriptional divergence between WT and PRC1.3 KO cells after 10 days of reprogramming. This experiment further upholds the central finding that PRC1.3 is essential for PTNR and as PRC1.3 KO cells exit pluripotency, neural differentiation marker genes are instead upregulated. Neural-like cells can also form alongside nascent naïve PSCs during WT reprogramming (Collier *et al.*, 2017). This suggests that the formation of neural cell types is not specifically regulated by the absence of PRC1.3, but instead reflects a default exit from pluripotency and that the formation of neural-like cells is permitted under reprogramming conditions. A caveat of this finding is that a number of these gene expression changes will reflect an alteration in the underlying cell identity, as an indirect effect of PRC1.3 deletion. However, in order to understand how PRC1.3 directly modulates reprogramming, I wanted to identify genes that might be specifically regulated by the PRC1.3 complex. This led to the next outcome of my work, which was the identification of genome-wide PRC1.3 binding sites on chromatin in naïve and primed PSCs using ChIP-sequencing. This approach showed that very few genes share PRC1.3 occupancy in both states and the vast majority of chromatin-bound PRC1.3 redistributes during reprogramming. Identification of a small, naïve-specific PRC1.3 target gene subset provided a promising list of candidates to mechanistically determine how PRC1.3 influences successful naïve PSC reprogramming in a proportion of cells. The genes within this subset include modulators of intracellular signalling, such as *TCF7L1* and *BRAF*, and developmental transcription factors including *DUSP6* and *PAX7*. Regulation of these genes, and subsequent downstream effects, could plausibly translate PRC1.3 function into cell fate reconfiguration.

As so little is known about the events that underpin human PTNR, the ability to identify PRC1.3 as a factor that is essential for reprogramming and link this complex to the regulation of specific genes already represents a substantial advance on our current understanding. However, the manner in which these genes are regulated by PRC1.3 binding was not known. I therefore integrated the knowledge of this gene subset with my RNA-sequencing data and published data (Collier *et al.*, 2017), in order to understand what happens to naïve PRC1.3 target gene expression during reprogramming. The majority of these genes are transcriptionally downregulated during reprogramming and this downregulation is lost when PRC1.3 is deleted, or when WT cells fail to reprogramme. This finding suggests that PRC1.3-dependent transcriptional downregulation of this gene set is important for the successful generation of nascent naïve PSCs, and that the gene-regulatory function of PRC1.3 in human reprogramming is more likely to be transcriptional repression than transcriptional activation, which has been associated with PRC1.3 function in other contexts (Gao *et al.*, 2014; Kanki *et al.*, 2022; Scelfo *et al.*, 2019; Wang *et al.*, 2018; Yao *et al.*, 2018; Zhao *et al.*, 2017).

The association between PRC1.3 occupancy, H2AK119ub1 deposition and transcriptional downregulation does not however definitively confirm that PRC1.3 is directly repressing the transcription of these genes. It is not clear whether PRC1.3 alone is catalysing this H2AK119ub1 and if this H2AK119ub1 is directly inhibiting gene expression on a molecular level. Notably, none of the PRC1.3 target genes were identified as impediments of human PTNR by the genome-wide screen, likely as they are transcriptionally repressed during reprogramming, and so the importance of these genes individually for reprogramming would have to be tested in PRC1.3 KO cells. Collectively however, a failure to repress this gene set from their expression levels in primed PSCs is a signature of cells that fail to reprogramme. This finding presents the first identification of secondary targets that occur downstream of a gene that was identified as essential by the genome-wide screen. By understanding the genes that are co-regulated during reprogramming, a more comprehensive picture can be determined of the molecular roadmap to reprogramming. The strict PRC1.3-dependency on transcriptional repression requires further evaluation, for example using a screening approach, rather than focusing on individual genes, to establish how exactly PRC1.3 is necessary for reprogramming. This can help to ascertain how exactly PRC1.3 regulates gene expression and ascertain a hierarchy of PRC1.3-dependent events that translate exogenous signalling cues into transformation of the underlying state of pluripotency.

6.4 – PRC1.3 undergoes pluripotency state-specific regulation

As PRC1.3 does not have any accessory proteins with known DNA-binding ability, it is not clear how PRC1.3 becomes targeted to specific genes and how this targeting changes during reprogramming, which could allow PRC1.3 to fulfil its essential function in reprogramming. To this end, another key outcome of my work was the use of qPLEX-RIME to quantitatively identify proteins that co-interact with the PRC1.3 complex with pluripotency state-specific enrichment. qPLEX-RIME identified a number of potential pluripotency state-specific PRC1.3 co-interactors, many of which I have not had the opportunity to characterise, including genes that are also identified by the genome-wide screen as critical regulators of reprogramming.

Previously reported proteins that can recruit the PRC1.3 complex to chromatin were either not detected by PRC1.3 qPLEX-RIME or co-interacted with the complex with an equal affinity across all pluripotent states. I did however identify a novel co-interaction between PRC1.3 and the pluripotency transcription factor PRDM14 in human naïve PSCs. As *PRDM14* has been linked centrally to human pluripotency (Chia *et al.*, 2010), this interaction could link PRC1.3 deletion to the collapse of the core pluripotency network during reprogramming. To better understand the molecular events that occur during reprogramming, it will be important to determine exactly how the naïve-specific interaction between PRC1.3 and PRDM14 is mediated, and why these proteins do not co-interact with a similar affinity in human primed PSCs. This co-interaction could be regulated by factors such as pluripotency state-specific post-translational modification of either PRDM14 or a component of PRC1.3. The molecular basis of the PRC1.3-PRDM14 interaction could in turn inform on whether PRDM14 and PRC1.3 might interact in pre-implantation embryonic epiblast, and could therefore influence human development.

I also confirmed that *PRDM14*, like PRC1.3, is essential for human PTNR. Beyond the identification of a naïve-specific co-interaction between PRDM14 and PRC1.3, integrating my findings and published data (Sybirna *et al.*, 2020), suggest that PRDM14 binds at naïve PRC1.3 target genes and cooperates with PRC1.3 in repression of this subset of genes. PRDM14 is also a compelling candidate for the redistribution of PRC1.3 on chromatin during reprogramming, as PRDM14 is enriched only at naïve, but not primed PRC1.3 target genes. I was not however able to definitively establish that the co-interaction does in fact target PRC1.3 to chromatin and ironically, in order to determine this, I need to be able to improve the reprogramming of PRDM14-AID-VENUS naïve PSCs, to produce a more stable inducible PRDM14 knockdown system. These findings nevertheless advance our knowledge of the molecular events that occur during reprogramming. If the recruitment of PRC1.3 by PRDM14 to chromatin was confirmed, questions would still remain about this recruitment mechanism. PRDM14 binds to a large number of motif sequences on naïve chromatin at sites where

PRC1.3 binding is not detected (Sybirna *et al.*, 2020), and this indicates that additional layers of regulation may be involved in PRC1.3 recruitment to its target genes to coordinate successful reprogramming, potentially including other PRC1.3 interactors identified by qPLEX-RIME.

Another outcome from the PRC1.3 qPLEX-RIME experiment was the identification of a change in the composition of the PRC1.3 complex itself, with *AUTS2* enriched within primed PRC1.3 complexes and *FBRS* instead enriched within naïve PRC1.3 complexes. This was a novel finding, and regulation of the composition of PRC1 complexes has never been previously identified. This work also provides the first strong indicator that the PRC1.3 complex paralogues *AUTS2* and *FBRS* may be mutually exclusive within PRC1.3, as single cells developmentally express either *AUTS2* or *FBRS*. I therefore speculate that two different classes of PRC1.3 complex might exist within cells – *AUTS2*-PRC1.3 and *FBRS*-PRC1.3 complexes. The fact that *FBRS* is essential for human PTNR and *AUTS2* is not, implies that these different subtypes of PRC1.3 complex could in fact possess different functionalities, with an intriguing possibility being whether the paralogue switch from *AUTS2* to *FBRS* might influence the ability of PRDM14 to interact with PRC1.3 in naïve PSCs. However, further investigation is required to conclusively determine whether this is the case.

Importantly, the developmentally-regulated paralogue shift between *AUTS2* and *FBRS* expression was corroborated by the paralogue expression patterns in the early human embryo (Xiang *et al.*, 2020). This finding vindicates the study of processes such as human PTNR, even though they are not directly applicable to developmental processes, to advance our understanding of human embryonic development. The study of PRC1.3 in human PSCs expands our understanding of the epigenetic or transcriptional regulation of human development, and how this differs considerably between pre-implantation and post-implantation epiblast cells, when this paralogue shift takes place. Due to the restrictions on the genetic editing of human embryos, further understanding of the developmental implications of this paralogue shift could be made with the continued use of PSCs. The creation of knockdowns for the *AUTS2* and *FBRS* paralogues in the state of pluripotency in which they are highly expressed, or an overexpression when cells where they are not ordinarily expressed, could develop our understanding of the importance of the expression switch of the *AUTS2* and *FBRS* paralogues in human embryos.

As previously described, without a naïve PRC1.3 knockdown system, it is not possible to comprehensively ascertain the function of the PRC1.3 complex in human naïve PSCs. H2AK119ub1, and subsequently H3K27me3, have been shown to shape the epigenetic

landscape of the earliest stages of mouse embryogenesis (Chen *et al.*, 2021; Mei *et al.*, 2021). It is possible that PRC1.3 might play a role in shaping the epigenome of naïve PSCs and human pre-implantation embryos. A study that has profiled epigenetic modifications in early human embryos (Xia *et al.*, 2019) did not profile the H2AK119ub1 modification. To try to enhance my understanding of H2AK119ub1 catalysis by PRC1.3 without a knockdown system, the final outcome of my work is the development of an experimental system to test the *in vitro* activity of PRC1.3 complexes from naïve and primed PSCs. The intention of this approach was to assess whether PRC1.3 has a higher catalytic activity in naïve than primed PSCs, in order to test the potential functional impact of the pluripotency state-specific PRC1.3 regulation that I have identified, which would align with the increased abundance of H2AK119ub1 and H3K27me3 in naïve PSCs. The preliminary outcome of this activity assay suggested that the AUTS2-PRC1.3 and FBRS-PRC1.3 complexes do not appear to have a different enzymatic activity. However, the combined use of an E3 ubiquitin ligase enzyme derived from cells in an otherwise *in vitro* system presented technical challenges, which made execution of a more comprehensive, better controlled experiment difficult. This blended system may also limit the ability to model wider the intracellular regulation of enzymatic activity. For example, naïve and primed PSCs may have differences in the activity of the E1 and E2 enzymes upstream of PRC1.3, in the homeostasis of free ubiquitin itself as an enzyme substrate or differences in the activity of the H2AK119 deubiquitinase BAP1. In addition, I identified differences in the abundances of H2A variants when PRC1.3 is deleted from primed PSCs, alongside a reduced global abundance of the H2AK119ub1 modification. It is possible that different H2A variant proteins have different propensities for their lysine 119 to become monoubiquitinated, and H2A variants enriched in naïve PSCs are simply more susceptible to catalytic modification. As the *in vitro* system uses recombinant supplementation of E1, ubiquitin-conjugated E2 and polynucleosomes, this system would not be able to model any of these regulatory differences that could influence PRC1.3 activity. However, with added optimisation, the principle of this methodology could be adapted to test the activity of other chromatin-modifying epigenetic enzymes.

6.5 – Final conclusions

In this thesis, I sought to understand how PRC1.3 critically modulates human PTNR. CRISPR-Cas9 mediated deletion of PRC1.3 complex component genes in primed PSCs was used as a starting point to assess this complex for the first time in human pluripotency. I identified that PRC1.3 was dispensable for all of the characteristics of primed pluripotency that I assessed.

As PRC1.3-deficient primed PSCs were unable to undergo PTNR, I conclude that PRC1.3 is essential for cells to enter the naïve state of pluripotency, which was a novel finding.

To understand how the complex might mechanistically influence the cell-state transition, I performed ChIP-sequencing for PRC1.3 in primed and naïve PSCs and for the first time identified a small number of genes that might be directly regulated by PRC1.3. I identified that these genes are transcriptionally repressed, and not activated, as cells reprogramme. These findings help me to build a molecular roadmap of how previously reported transcriptional changes during reprogramming might be executed.

I observed that deletion of PRC1.3 from human primed PSCs did however decrease the global abundance of the PRC1-deposited H2AK119ub1, and PRC2-deposited H3K27me3, modifications, as well as causing wider changes to the epigenome and chromatin associated histone H2A variants. This finding leads me to conclude that PRC1.3 is important for the epigenetic regulation of human pluripotency, which has never before been known.

Using a proteomics approach, I identified a novel, naïve-specific interaction between PRC1.3 and PRDM14, and I confirmed that PRC1.3 and PRDM14 cooperate in transcriptional repression of PRC1.3 target genes. I also discovered a developmentally-regulated shift in the composition of PRC1.3 complexes, with the *AUTS2* and *FBRs* paralogues, which I was able to corroborate in published data from early human embryos. These findings allow me to build a picture of the pluripotency-state specific regulation of PRC1.3 across the reprogramming and illuminates further the specific events that take place, which I suggest could be developmentally relevant in the early progression of human embryogenesis.

These findings advance our knowledge of molecular events that underpin successful naïve cell reprogramming. By expanding on my findings further, steps could be taken to influence the ability of cells to reprogramme more effectively, which in turn unlocks the exciting potential of human naïve PSCs as an *in vitro* developmental model and a cell type with considerable therapeutic potential.

Bibliography

- Almeida, M., Pintacuda, G., Masui, O., Koseki, Y., Gdula, M., Cerase, A., Brown, D., Mould, A., Innocent, C., Nakayama, M., Schermelleh, L., Nesterova, T.B., Koseki, H., Brockdorff, N., 2017. PCGF3/5–PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* 356, 1081–1084.
<https://doi.org/10.1126/science.aal2512>
- Aranda, S., Mas, G., Di Croce, L., 2015. Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* 1, e1500737. <https://doi.org/10.1126/sciadv.1500737>
- Arents, G., Moudrianakis, E.N., 1993. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc. Natl. Acad. Sci.* 90, 10489–10493. <https://doi.org/10.1073/pnas.90.22.10489>
- Arrigoni, R., Alam, S.L., Wamstad, J.A., Bardwell, V.J., Sundquist, W.I., Schreiber-Agus, N., 2006. The Polycomb-associated protein Rybp is a ubiquitin binding protein. *FEBS Lett.* 580, 6233–6241. <https://doi.org/10.1016/j.febslet.2006.10.027>
- Atsuta, T., Fujimura, S., Moriya, H., Vidal, M., Akasaka, T., Koseki, H., 2001. Production of Monoclonal Antibodies Against Mammalian Ring1B Proteins. *Hybridoma* 20, 43–46.
<https://doi.org/10.1089/027245701300060427>
- Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., Lovell-Badge, R., 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* 17, 126–140. <https://doi.org/10.1101/gad.224503>
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., Fisher, A.G., 2006. Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* 8, 532–538.
<https://doi.org/10.1038/ncb1403>
- Bajusz, I., Kovács, G., Pirity, M., 2018. From Flies to Mice: The Emerging Role of Non-Canonical PRC1 Members in Mammalian Development. *Epigenomes* 2, 4.
<https://doi.org/10.3390/epigenomes2010004>
- Bannister, A.J., Kouzarides, T., 2011. Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. <https://doi.org/10.1038/cr.2011.22>
- Bar, S., Seaton, L.R., Weissbein, U., Eldar-Geva, T., Benvenisty, N., 2019. Global Characterization of X Chromosome Inactivation in Human Pluripotent Stem Cells. *Cell Rep.* 27, 20–29.e3. <https://doi.org/10.1016/j.celrep.2019.03.019>
- Bayerl, J., Ayyash, M., Shani, T., Manor, Y.S., Gafni, O., Massarwa, R., Kalma, Y., Aguilera-Castrejon, A., Zerbib, M., Amir, H., Sheban, D., Geula, S., Mor, N., Weinberger, L., Naveh Tassa, S., Krupalnik, V., Oldak, B., Livnat, N., Tarazi, S., Tawil, S., Wildschutz, E., Ashoukhi, S., Lasman, L., Rotter, V., Hanna, S., Ben-Yosef, D., Novershtern, N., Viukov, S., Hanna, J.H., 2021. Principles of signaling pathway modulation for enhancing human naive pluripotency induction. *Cell Stem Cell* 28, 1549–1565.e12. <https://doi.org/10.1016/j.stem.2021.04.001>
- Beddington, R.S., Robertson, E.J., 1989. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development* 105, 733–737. <https://doi.org/10.1242/dev.105.4.733>
- Bedogni, F., Hodge, R.D., Nelson, B.R., Frederick, E.A., Shiba, N., Daza, R.A., Hevner, R.F., 2010. Autism susceptibility candidate 2 (Aut2) encodes a nuclear protein expressed in developing brain regions implicated in autism neuropathology. *Gene Expr. Patterns* 10, 9–15. <https://doi.org/10.1016/j.gep.2009.11.005>
- Bendall, A., Semprich, C.I., 2022. Chromatin Profiling of Human Naïve Pluripotent Stem Cells, in: Rugg-Gunn, P. (Ed.), *Human Naïve Pluripotent Stem Cells, Methods in Molecular Biology*. Springer US, New York, NY, pp. 181–200.
https://doi.org/10.1007/978-1-0716-1908-7_12

- Bergmann, S., Penfold, C.A., Slatery, E., Siriwardena, D., Drummer, C., Clark, S., Strawbridge, S.E., Kishimoto, K., Vickers, A., Tewary, M., Kohler, T.N., Hollfelder, F., Reik, W., Sasaki, E., Behr, R., Boroviak, T.E., 2022. Spatial profiling of early primate gastrulation in utero. *Nature*. <https://doi.org/10.1038/s41586-022-04953-1>
- Beringer, M., Pisano, P., Di Carlo, V., Blanco, E., Chammass, P., Vizán, P., Gutiérrez, A., Aranda, S., Payer, B., Wierer, M., Di Croce, L., 2016. EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. *Mol. Cell* 64, 645–658. <https://doi.org/10.1016/j.molcel.2016.10.018>
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L., Lander, E.S., 2006. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* 125, 315–326. <https://doi.org/10.1016/j.cell.2006.02.041>
- Betto, R.M., Diamante, L., Perrera, V., Audano, M., Rapelli, S., Lauria, A., Incarnato, D., Arboit, M., Pedretti, S., Rigoni, G., Guerineau, V., Touboul, D., Stirparo, G.G., Lohoff, T., Boroviak, T., Grumati, P., Soriano, M.E., Nichols, J., Mitro, N., Oliviero, S., Martello, G., 2021. Metabolic control of DNA methylation in naive pluripotent cells. *Nat. Genet.* 53, 215–229. <https://doi.org/10.1038/s41588-020-00770-2>
- Beunders, G., Voorhoeve, E., Golzio, C., Pardo, L.M., Rosenfeld, J.A., Talkowski, M.E., Simoncic, I., Lionel, A.C., Vergult, S., Pyatt, R.E., van de Kamp, J., Nieuwint, A., Weiss, M.M., Rizzu, P., Verwer, L.E.N.I., van Spaendonk, R.M.L., Shen, Y., Wu, B., Yu, T., Yu, Y., Chiang, C., Gusella, J.F., Lindgren, A.M., Morton, C.C., van Binsbergen, E., Bulk, S., van Rossem, E., Vanakker, O., Armstrong, R., Park, S.-M., Greenhalgh, L., Maye, U., Neill, N.J., Abbott, K.M., Sell, S., Ladda, R., Farber, D.M., Bader, P.I., Cushing, T., Drautz, J.M., Konczal, L., Nash, P., de Los Reyes, E., Carter, M.T., Hopkins, E., Marshall, C.R., Osborne, L.R., Gripp, K.W., Thrush, D.L., Hashimoto, S., Gastier-Foster, J.M., Astbury, C., Ylstra, B., Meijers-Heijboer, H., Posthuma, D., Menten, B., Mortier, G., Scherer, S.W., Eichler, E.E., Girirajan, S., Katsanis, N., Groffen, A.J., Sistermans, E.A., 2013. Exonic Deletions in AUTS2 Cause a Syndromic Form of Intellectual Disability and Suggest a Critical Role for the C Terminus. *Am. J. Hum. Genet.* 92, 210–220. <https://doi.org/10.1016/j.ajhg.2012.12.011>
- Bi, Y., Tu, Z., Zhang, Y., Yang, P., Guo, M., Zhu, X., Zhao, C., Zhou, J., Wang, H., Wang, Y., Gao, S., 2020. Identification of ALPPL2 as a Naive Pluripotent State-Specific Surface Protein Essential for Human Naive Pluripotency Regulation. *Cell Rep.* 30, 3917–3931.e5. <https://doi.org/10.1016/j.celrep.2020.02.090>
- Bi, Y., Tu, Z., Zhou, J., Zhu, X., Wang, H., Gao, S., Wang, Y., 2022. Cell fate roadmap of human primed-to-naive transition reveals preimplantation cell lineage signatures. *Nat. Commun.* 13, 3147. <https://doi.org/10.1038/s41467-022-30924-1>
- Biswas, S., Rao, C.M., 2018. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur. J. Pharmacol.* 837, 8–24. <https://doi.org/10.1016/j.ejphar.2018.08.021>
- Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., Ishikura, T., Long, H.K., Sheahan, T.W., Brockdorff, N., Kessler, B.M., Koseki, H., Klose, R.J., 2014. Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. *Cell* 157, 1445–1459. <https://doi.org/10.1016/j.cell.2014.05.004>
- Blackledge, N.P., Fursova, N.A., Kelley, J.R., Huseyin, M.K., Feldmann, A., Klose, R.J., 2020. PRC1 Catalytic Activity Is Central to Polycomb System Function. *Mol. Cell* 77, 857–874.e9. <https://doi.org/10.1016/j.molcel.2019.12.001>
- Blackledge, N.P., Rose, N.R., Klose, R.J., 2015. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. *Nat. Rev. Mol. Cell Biol.* 16, 643–649. <https://doi.org/10.1038/nrm4067>
- Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaita, S.E., Hu, T.X., Elder, K., Snell, P., Christie, L., Robson, P., Niakan, K.K., 2015. Defining the three cell lineages of the

- human blastocyst by single-cell RNA-seq. *Development* 142, 3151–3165. <https://doi.org/10.1242/dev.123547>
- Boroviak, T., Loos, R., Bertone, P., Smith, A., Nichols, J., 2014. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* 16, 513–525. <https://doi.org/10.1038/ncb2965>
- Boroviak, T., Nichols, J., 2017. Primate embryogenesis predicts the hallmarks of human naïve pluripotency. *Development* 144, 175–186. <https://doi.org/10.1242/dev.145177>
- Bousard, A., Raposo, A.C., Żylicz, J.J., Picard, C., Pires, V.B., Qi, Y., Gil, C., Syx, L., Chang, H.Y., Heard, E., da Rocha, S.T., 2019. The role of *Xist* -mediated Polycomb recruitment in the initiation of X-chromosome inactivation. *EMBO Rep.* 20. <https://doi.org/10.15252/embr.201948019>
- Bowles, J., Teasdale, R.P., James, K., Koopman, P., 2003. *Dppa3* is a marker of pluripotency and has a human homologue that is expressed in germ cell tumours. *Cytogenet. Genome Res.* 101, 261–265. <https://doi.org/10.1159/000074346>
- Bowman, G.D., Poirier, M.G., 2015. Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. *Chem. Rev.* 115, 2274–2295. <https://doi.org/10.1021/cr500350x>
- Bowness, J.S., Nesterova, T.B., Wei, G., Rodermund, L., Almeida, M., Coker, H., Carter, E.J., Kadaster, A., Brockdorff, N., 2022. *Xist*-mediated silencing requires additive functions of SPEN and Polycomb together with differentiation-dependent recruitment of SmcHD1. *Cell Rep.* 39, 110830. <https://doi.org/10.1016/j.celrep.2022.110830>
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., Young, R.A., 2005. Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell* 122, 947–956. <https://doi.org/10.1016/j.cell.2005.08.020>
- Boyle, S., Flyamer, I.M., Williamson, I., Sengupta, D., Bickmore, W.A., Illingworth, R.S., 2020. A central role for canonical PRC1 in shaping the 3D nuclear landscape. *Genes Dev.* 34, 931–949. <https://doi.org/10.1101/gad.336487.120>
- Bredenkamp, N., Stirparo, G.G., Nichols, J., Smith, A., Guo, G., 2019a. The Cell-Surface Marker Sushi Containing Domain 2 Facilitates Establishment of Human Naive Pluripotent Stem Cells. *Stem Cell Rep.* 12, 1212–1222. <https://doi.org/10.1016/j.stemcr.2019.03.014>
- Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D., Wu, C., Rostovskaya, M., Eminli-Meissner, S., Smith, A., Guo, G., 2019b. Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. *Stem Cell Rep.* 13, 1083–1098. <https://doi.org/10.1016/j.stemcr.2019.10.009>
- Briggs, R., King, T.J., 1952. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc. Natl. Acad. Sci.* 38, 455–463. <https://doi.org/10.1073/pnas.38.5.455>
- Brockdorff, N., 2017. Polycomb complexes in X chromosome inactivation. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20170021. <https://doi.org/10.1098/rstb.2017.0021>
- Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., Rastan, S., 1991. Conservation of position and exclusive expression of mouse *Xist* from the inactive X chromosome. *Nature* 351, 329–331. <https://doi.org/10.1038/351329a0>
- Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., Rastan, S., 1992. The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* 71, 515–526. [https://doi.org/10.1016/0092-8674\(92\)90519-I](https://doi.org/10.1016/0092-8674(92)90519-I)
- Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., Vallier, L., 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191–195. <https://doi.org/10.1038/nature05950>

- Brookes, E., de Santiago, I., Hebenstreit, D., Morris, K.J., Carroll, T., Xie, S.Q., Stock, J.K., Heidemann, M., Eick, D., Nozaki, N., Kimura, H., Ragoussis, J., Teichmann, S.A., Pombo, A., 2012. Polycomb Associates Genome-wide with a Specific RNA Polymerase II Variant, and Regulates Metabolic Genes in ESCs. *Cell Stem Cell* 10, 157–170. <https://doi.org/10.1016/j.stem.2011.12.017>
- Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., Willard, H.F., 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349, 38–44. <https://doi.org/10.1038/349038a0>
- Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C.H.-H., Lim, B., Vardy, L., Dunn, N.R., Trotter, M., Pedersen, R., Vallier, L., 2011. Activin/Nodal Signaling Controls Divergent Transcriptional Networks in Human Embryonic Stem Cells and in Endoderm Progenitors. *Stem Cells* 29, 1176–1185. <https://doi.org/10.1002/stem.666>
- Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., Sixma, T.K., 2006. Structure and E3-ligase activity of the Ring–Ring complex of Polycomb proteins Bmi1 and Ring1b. *EMBO J.* 25, 2465–2474. <https://doi.org/10.1038/sj.emboj.7601144>
- Bunker, C.A., Kingston, R.E., 1994. Transcriptional repression by Drosophila and mammalian Polycomb group proteins in transfected mammalian cells. *Mol. Cell. Biol.* 14, 1721–1732. <https://doi.org/10.1128/MCB.14.3.1721>
- Burdon, T., Stracey, C., Chambers, I., Nichols, J., Smith, A., 1999. Suppression of SHP-2 and ERK Signalling Promotes Self-Renewal of Mouse Embryonic Stem Cells. *Dev. Biol.* 210, 30–43. <https://doi.org/10.1006/dbio.1999.9265>
- Campagne, A., Lee, M.-K., Zielinski, D., Michaud, A., Le Corre, S., Dingli, F., Chen, H., Shahidian, L.Z., Vassilev, I., Servant, N., Loew, D., Pasmant, E., Postel-Vinay, S., Wassef, M., Margueron, R., 2019. BAP1 complex promotes transcription by opposing PRC1-mediated H2A ubiquitylation. *Nat. Commun.* 10, 348. <https://doi.org/10.1038/s41467-018-08255-x>
- Campbell, K.H.S., McWhir, J., Ritchie, W.A., Wilmut, I., 1996. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* 380, 64–66. <https://doi.org/10.1038/380064a0>
- Cao, R., Tsukada, Y., Zhang, Y., 2005. Role of Bmi-1 and Ring1A in H2A Ubiquitylation and Hox Gene Silencing. *Mol. Cell* 20, 845–854. <https://doi.org/10.1016/j.molcel.2005.12.002>
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., Zhang, Y., 2002. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science* 298, 1039–1043. <https://doi.org/10.1126/science.1076997>
- Cao, R., Zhang, Y., 2004. SUZ12 Is Required for Both the Histone Methyltransferase Activity and the Silencing Function of the EED-EZH2 Complex. *Mol. Cell* 15, 57–67. <https://doi.org/10.1016/j.molcel.2004.06.020>
- Castel, G., Meistermann, D., Bretin, B., Firmin, J., Blin, J., Loubersac, S., Bruneau, A., Chevolleau, S., Kilens, S., Chariou, C., Gaignerie, A., Francheteau, Q., Kagawa, H., Charpentier, E., Flippe, L., François-Campion, V., Haider, S., Dietrich, B., Knöfler, M., Arima, T., Bourdon, J., Rivron, N., Masson, D., Fournier, T., Okae, H., Fréour, T., David, L., 2020. Induction of Human Trophoblast Stem Cells from Somatic Cells and Pluripotent Stem Cells. *Cell Rep.* 33, 108419. <https://doi.org/10.1016/j.celrep.2020.108419>
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., Smith, A., 2003. Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells. *Cell* 113, 643–655. [https://doi.org/10.1016/S0092-8674\(03\)00392-1](https://doi.org/10.1016/S0092-8674(03)00392-1)
- Chan, Y.-S., Göke, J., Lu, X., Venkatesan, N., Feng, B., Su, I.-H., Ng, H.-H., 2013a. A PRC2-Dependent Repressive Role of PRDM14 in Human Embryonic Stem Cells and Induced Pluripotent Stem Cell Reprogramming. *STEM CELLS* 31, 682–692. <https://doi.org/10.1002/stem.1307>

- Chan, Y.-S., Göke, J., Ng, J.-H., Lu, X., Gonzales, K.A.U., Tan, C.-P., Tng, W.-Q., Hong, Z.-Z., Lim, Y.-S., Ng, H.-H., 2013b. Induction of a Human Pluripotent State with Distinct Regulatory Circuitry that Resembles Preimplantation Epiblast. *Cell Stem Cell* 13, 663–675. <https://doi.org/10.1016/j.stem.2013.11.015>
- Chen, H., Aksoy, I., Gonnot, F., Osteil, P., Aubry, M., Hamela, C., Rognard, C., Hochard, A., Voisin, S., Fontaine, E., Mure, M., Afanassieff, M., Cleroux, E., Guibert, S., Chen, J., Vallot, C., Acloque, H., Genthon, C., Donnadiou, C., De Vos, J., Sanlaville, D., Guérin, J.-F., Weber, M., Stanton, L.W., Rougeulle, C., Pain, B., Bourillot, P.-Y., Savatier, P., 2015. Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat. Commun.* 6, 7095. <https://doi.org/10.1038/ncomms8095>
- Chen, Z., Djekidel, M.N., Zhang, Y., 2021. Distinct dynamics and functions of H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. *Nat. Genet.* 53, 551–563. <https://doi.org/10.1038/s41588-021-00821-2>
- Chia, N.-Y., Chan, Y.-S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.-S., Huss, M., Soh, B.-S., Kraus, P., Li, P., Lufkin, T., Lim, B., Clarke, N.D., Bard, F., Ng, H.-H., 2010. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* 468, 316–320. <https://doi.org/10.1038/nature09531>
- Chittock, E.C., Latwiel, S., Miller, T.C.R., Müller, C.W., 2017. Molecular architecture of polycomb repressive complexes. *Biochem. Soc. Trans.* 45, 193–205. <https://doi.org/10.1042/BST20160173>
- Choi, H.W., Joo, J.Y., Hong, Y.J., Kim, J.S., Song, H., Lee, J.W., Wu, G., Schöler, H.R., Do, J.T., 2016. Distinct Enhancer Activity of Oct4 in Naive and Primed Mouse Pluripotency. *Stem Cell Rep.* 7, 911–926. <https://doi.org/10.1016/j.stemcr.2016.09.012>
- Chovanec, P., Collier, A.J., Krueger, C., Várnai, C., Semprich, C.I., Schoenfelder, S., Corcoran, A.E., Rugg-Gunn, P.J., 2021. Widespread reorganisation of pluripotent factor binding and gene regulatory interactions between human pluripotent states. *Nat. Commun.* 12, 2098. <https://doi.org/10.1038/s41467-021-22201-4>
- Cieślak-Pobuda, A., Knoflach, V., Ringh, M.V., Stark, J., Likus, W., Siemianowicz, K., Ghavami, S., Hudecki, A., Green, J.L., Łos, M.J., 2017. Transdifferentiation and reprogramming: Overview of the processes, their similarities and differences. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1864, 1359–1369. <https://doi.org/10.1016/j.bbamcr.2017.04.017>
- Cinkornpumin, J.K., Kwon, S.Y., Guo, Y., Hossain, I., Sirois, J., Russett, C.S., Tseng, H.-W., Okae, H., Arima, T., Duchaine, T.F., Liu, W., Pastor, W.A., 2020. Naive Human Embryonic Stem Cells Can Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. *Stem Cell Rep.* 15, 198–213. <https://doi.org/10.1016/j.stemcr.2020.06.003>
- Clemson, C.M., McNeil, J.A., Willard, H.F., Lawrence, J.B., 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* 132, 259–275. <https://doi.org/10.1083/jcb.132.3.259>
- Cohen, I., Zhao, D., Menon, G., Nakayama, M., Koseki, H., Zheng, D., Ezhkova, E., 2019. PRC1 preserves epidermal tissue integrity independently of PRC2. *Genes Dev.* 33, 55–60. <https://doi.org/10.1101/gad.319939.118>
- Collier, A.J., Bendall, A., Fabian, C., Malcolm, A.A., Tilgner, K., Semprich, C.I., Wojdyla, K., Nisi, P.S., Kishore, K., Roamio Franklin, V.N., Mirshekar-Syahkal, B., D'Santos, C., Plath, K., Yusa, K., Rugg-Gunn, P.J., 2022. Genome-wide screening identifies Polycomb repressive complex 1.3 as an essential regulator of human naïve pluripotent cell reprogramming. *Sci. Adv.* 8, eabk0013. <https://doi.org/10.1126/sciadv.abk0013>
- Collier, A.J., Panula, S.P., Schell, J.P., Chovanec, P., Plaza Reyes, A., Petropoulos, S., Corcoran, A.E., Walker, R., Douagi, I., Lanner, F., Rugg-Gunn, P.J., 2017.

- Comprehensive Cell Surface Protein Profiling Identifies Specific Markers of Human Naive and Primed Pluripotent States. *Cell Stem Cell* 20, 874-890.e7. <https://doi.org/10.1016/j.stem.2017.02.014>
- Collier, A.J., Rugg-Gunn, P.J., 2018. Identifying Human Naïve Pluripotent Stem Cells – Evaluating State-Specific Reporter Lines and Cell-Surface Markers. *BioEssays* 40, 1700239. <https://doi.org/10.1002/bies.201700239>
- Collinson, A., Collier, A.J., Morgan, N.P., Sienerth, A.R., Chandra, T., Andrews, S., Rugg-Gunn, P.J., 2016. Deletion of the Polycomb-Group Protein EZH2 Leads to Compromised Self-Renewal and Differentiation Defects in Human Embryonic Stem Cells. *Cell Rep.* 17, 2700–2714. <https://doi.org/10.1016/j.celrep.2016.11.032>
- Conway, E., Rossi, F., Fernandez-Perez, D., Ponzo, E., Ferrari, K.J., Zanotti, M., Manganaro, D., Rodighiero, S., Tamburri, S., Pasini, D., 2021. BAP1 enhances Polycomb repression by counteracting widespread H2AK119ub1 deposition and chromatin condensation. *Mol. Cell* 81, 3526-3541.e8. <https://doi.org/10.1016/j.molcel.2021.06.020>
- Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N.P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., Heger, A., Brockdorff, N., 2014. Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep.* 7, 1456–1470. <https://doi.org/10.1016/j.celrep.2014.04.012>
- Cooper, S., Grijzenhout, A., Underwood, E., Ancelin, K., Zhang, T., Nesterova, T.B., Anil-Kirmizitas, B., Bassett, A., Kooistra, S.M., Agger, K., Helin, K., Heard, E., Brockdorff, N., 2016. Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat. Commun.* 7, 13661. <https://doi.org/10.1038/ncomms13661>
- Costanzi, C., Pehrson, J.R., 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* 393, 599–601. <https://doi.org/10.1038/31275>
- Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., Pirrotta, V., 2002. Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity that Marks Chromosomal Polycomb Sites. *Cell* 111, 185–196. [https://doi.org/10.1016/S0092-8674\(02\)00975-3](https://doi.org/10.1016/S0092-8674(02)00975-3)
- Dahéron, L., Opitz, S.L., Zaehres, H., Lensch, W.M., Andrews, P.W., Itskovitz-Eldor, J., Daley, G.Q., 2004. LIF/STAT3 Signaling Fails to Maintain Self-Renewal of Human Embryonic Stem Cells. *STEM CELLS* 22, 770–778. <https://doi.org/10.1634/stemcells.22-5-770>
- Daled, S., Willems, S., Van Puyvelde, B., Corveleyn, L., Verhelst, S., De Clerck, L., Deforce, D., Dhaenens, M., 2021. Histone Sample Preparation for Bottom-Up Mass Spectrometry: A Roadmap to Informed Decisions. *Proteomes* 9, 17. <https://doi.org/10.3390/proteomes9020017>
- Davidson, K.C., Mason, E.A., Pera, M.F., 2015. The pluripotent state in mouse and human. *Development* 142, 3090–3099. <https://doi.org/10.1242/dev.116061>
- De Clerck, L., Taelman, J., Popovic, M., Willems, S., Van der Jeught, M., Heindryckx, B., De Sutter, P., Marks, H., Deforce, D., Dhaenens, M., 2019. Untargeted histone profiling during naive conversion uncovers conserved modification markers between mouse and human. *Sci. Rep.* 9, 17240. <https://doi.org/10.1038/s41598-019-53681-6>
- de Jaime-Soguero, A., Abreu de Oliveira, W., Lluís, F., 2018. The Pleiotropic Effects of the Canonical Wnt Pathway in Early Development and Pluripotency. *Genes* 9, 93. <https://doi.org/10.3390/genes9020093>
- De Lucia, F., Crevillen, P., Jones, A.M.E., Greb, T., Dean, C., 2008. A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc. Natl. Acad. Sci.* 105, 16831–16836. <https://doi.org/10.1073/pnas.0808687105>
- de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., Koseki, H., Brockdorff, N., 2004. Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene

- Silencing and X Inactivation. *Dev. Cell* 7, 663–676.
<https://doi.org/10.1016/j.devcel.2004.10.005>
- del Mar Lorente, M., Marcos-Gutierrez, C., Perez, C., Schoorlemmer, J., Ramirez, A., Magin, T., Vidal, M., 2000. Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development* 127, 5093–5100.
<https://doi.org/10.1242/dev.127.23.5093>
- Denell, R.E., Frederick, R.D., 1983. Homoeosis in *Drosophila*: A description of the polycomb lethal syndrome. *Dev. Biol.* 97, 34–47. [https://doi.org/10.1016/0012-1606\(83\)90061-1](https://doi.org/10.1016/0012-1606(83)90061-1)
- Desai, D., Khanna, A., Pethe, P., 2020. Inhibition of RING1B alters lineage specificity in human embryonic stem cells. *Cell Biol. Int.* 44, 1299–1311.
<https://doi.org/10.1002/cbin.11325>
- Dhami, P., Saffrey, P., Bruce, A.W., Dillon, S.C., Chiang, K., Bonhoure, N., Koch, C.M., Bye, J., James, K., Foad, N.S., Ellis, P., Watkins, N.A., Ouwehand, W.H., Langford, C., Andrews, R.M., Dunham, I., Vetrie, D., 2010. Complex Exon-Intron Marking by Histone Modifications Is Not Determined Solely by Nucleosome Distribution. *PLoS ONE* 5, e12339. <https://doi.org/10.1371/journal.pone.0012339>
- Di Stefano, B., Di Stefano, B., Hochedlinger, K., 2018. Reduced MEK inhibition preserves genomic stability in naïve human ES cells. *Protoc. Exch.*
<https://doi.org/10.1038/protex.2018.062>
- Diamante, L., Martello, G., 2022. Metabolic regulation in pluripotent stem cells. *Curr. Opin. Genet. Dev.* 75, 101923. <https://doi.org/10.1016/j.gde.2022.101923>
- Disteche, C.M., 2012. Dosage Compensation of the Sex Chromosomes. *Annu. Rev. Genet.* 46, 537–560. <https://doi.org/10.1146/annurev-genet-110711-155454>
- Dobrinić, P., Szczurek, A.T., Klose, R.J., 2021. PRC1 drives Polycomb-mediated gene repression by controlling transcription initiation and burst frequency. *Nat. Struct. Mol. Biol.* 28, 811–824. <https://doi.org/10.1038/s41594-021-00661-y>
- Dobrovolsky, P.L., Bess, D., 2011. Optimized PCR-based Detection of Mycoplasma. *J. Vis. Exp.* 3057. <https://doi.org/10.3791/3057>
- Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park, K., Yoon, E.-J., Xing, X., Kommagani, R., Wang, T., Solnica-Krezel, L., Theunissen, T.W., 2020. Derivation of trophoblast stem cells from naïve human pluripotent stem cells. *eLife* 9, e52504. <https://doi.org/10.7554/eLife.52504>
- Du, Z., Zheng, H., Kawamura, Y.K., Zhang, K., Gassler, J., Powell, S., Xu, Q., Lin, Z., Xu, K., Zhou, Q., Ozonov, E.A., Véron, N., Huang, B., Li, L., Yu, G., Liu, L., Au Yeung, W.K., Wang, P., Chang, L., Wang, Q., He, A., Sun, Y., Na, J., Sun, Q., Sasaki, H., Tachibana, K., Peters, A.H.F.M., Xie, W., 2020. Polycomb Group Proteins Regulate Chromatin Architecture in Mouse Oocytes and Early Embryos. *Mol. Cell* 77, 825–839.e7. <https://doi.org/10.1016/j.molcel.2019.11.011>
- Duan-Arnold, Y., Gyurdieva, A., Johnson, A., Uveges, T.E., Jacobstein, D.A., Danilkovitch, A., 2015. Retention of Endogenous Viable Cells Enhances the Anti-Inflammatory Activity of Cryopreserved Amnion. *Adv. Wound Care* 4, 523–533.
<https://doi.org/10.1089/wound.2015.0636>
- Duggal, G., Warrier, S., Ghimire, S., Broekaert, D., Van der Jeught, M., Lierman, S., Deroo, T., Peelman, L., Van Soom, A., Cornelissen, R., Menten, B., Mestdagh, P., Vandesompele, J., Roost, M., Sliker, R.C., Heijmans, B.T., Deforce, D., De Sutter, P., De Sousa Lopes, S.C., Heindryckx, B., 2015. Alternative Routes to Induce Naïve Pluripotency in Human Embryonic Stem Cells. *Stem Cells* 33, 2686–2698.
<https://doi.org/10.1002/stem.2071>
- Dutta, D., Ray, S., Home, P., Larson, M., Wolfe, M.W., Paul, S., 2011. Self-Renewal Versus Lineage Commitment of Embryonic Stem Cells: Protein Kinase C Signaling Shifts the Balance. *Stem Cells* 29, 618–628. <https://doi.org/10.1002/stem.605>
- Ebrahimi, B., 2015. Reprogramming barriers and enhancers: strategies to enhance the efficiency and kinetics of induced pluripotency. *Cell Regen.* 4, 4:10.
<https://doi.org/10.1186/s13619-015-0024-9>

- Elderkin, S., Maertens, G.N., Endoh, M., Mallery, D.L., Morrice, N., Koseki, H., Peters, G., Brockdorff, N., Hiom, K., 2007. A Phosphorylated Form of Mel-18 Targets the Ring1B Histone H2A Ubiquitin Ligase to Chromatin. *Mol. Cell* 28, 107–120. <https://doi.org/10.1016/j.molcel.2007.08.009>
- Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M., Koseki, H., 2008. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* 135, 1513–1524. <https://doi.org/10.1242/dev.014340>
- Endoh, M., Endo, T.A., Endoh, T., Isono, K., Sharif, J., Ohara, O., Toyoda, T., Ito, T., Eskeland, R., Bickmore, W.A., Vidal, M., Bernstein, B.E., Koseki, H., 2012. Histone H2A Mono-Ubiquitination Is a Crucial Step to Mediate PRC1-Dependent Repression of Developmental Genes to Maintain ES Cell Identity. *PLoS Genet.* 8, e1002774. <https://doi.org/10.1371/journal.pgen.1002774>
- Eskeland, R., Leeb, M., Grimes, G.R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultschi, A.I., Wutz, A., Bickmore, W.A., 2010. Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. *Mol. Cell* 38, 452–464. <https://doi.org/10.1016/j.molcel.2010.02.032>
- Evans, M.J., Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156. <https://doi.org/10.1038/292154a0>
- Farcas, A.M., Blackledge, N.P., Sudbery, I., Long, H.K., McGouran, J.F., Rose, N.R., Lee, S., Sims, D., Cerase, A., Sheahan, T.W., Koseki, H., Brockdorff, N., Ponting, C.P., Kessler, B.M., Klose, R.J., 2012. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLife* 1, e00205. <https://doi.org/10.7554/eLife.00205>
- Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T., Pasini, D., 2014. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol. Cell* 53, 49–62. <https://doi.org/10.1016/j.molcel.2013.10.030>
- Fischle, W., 2003. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* 17, 1870–1881. <https://doi.org/10.1101/gad.1110503>
- Fisher, R.D., Wang, B., Alam, S.L., Higginson, D.S., Robinson, H., Sundquist, W.I., Hill, C.P., 2003. Structure and Ubiquitin Binding of the Ubiquitin-interacting Motif. *J. Biol. Chem.* 278, 28976–28984. <https://doi.org/10.1074/jbc.M302596200>
- Fitz-James, M.H., Cavalli, G., 2022. Molecular mechanisms of transgenerational epigenetic inheritance. *Nat. Rev. Genet.* 23, 325–341. <https://doi.org/10.1038/s41576-021-00438-5>
- Fogarty, N.M.E., McCarthy, A., Snijders, K.E., Powell, B.E., Kubikova, N., Blakeley, P., Lea, R., Elder, K., Wamaitha, S.E., Kim, D., Maciulyte, V., Kleinjung, J., Kim, J.-S., Wells, D., Vallier, L., Bertero, A., Turner, J.M.A., Niakan, K.K., 2017. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature* 550, 67–73. <https://doi.org/10.1038/nature24033>
- Fournier, T., 2020. Human chorionic gonadotropin, in: 100 Years of Human Chorionic Gonadotropin. Elsevier, pp. 31–43. <https://doi.org/10.1016/B978-0-12-820050-6.00004-7>
- Francis, N.J., 2004. Chromatin Compaction by a Polycomb Group Protein Complex. *Science* 306, 1574–1577. <https://doi.org/10.1126/science.1100576>
- Francis, N.J., Saurin, A.J., Shao, Z., Kingston, R.E., 2001. Reconstitution of a Functional Core Polycomb Repressive Complex. *Mol. Cell* 8, 545–556. [https://doi.org/10.1016/S1097-2765\(01\)00316-1](https://doi.org/10.1016/S1097-2765(01)00316-1)
- Fritsch, C., Beuchle, D., Müller, J., 2003. Molecular and genetic analysis of the Polycomb group gene *Sex combs extra/Ring* in *Drosophila*. *Mech. Dev.* 120, 949–954. [https://doi.org/10.1016/S0925-4773\(03\)00083-2](https://doi.org/10.1016/S0925-4773(03)00083-2)
- Fursova, N.A., Blackledge, N.P., Nakayama, M., Ito, S., Koseki, Y., Farcas, A.M., King, H.W., Koseki, H., Klose, R.J., 2019. Synergy between Variant PRC1 Complexes

- Defines Polycomb-Mediated Gene Repression. *Mol. Cell* 74, 1020–1036.e8. <https://doi.org/10.1016/j.molcel.2019.03.024>
- Fursova, N.A., Turberfield, A.H., Blackledge, N.P., Findlater, E.L., Lastuvkova, A., Huseyin, M.K., Dobrinić, P., Klose, R.J., 2021. BAP1 constrains pervasive H2AK119ub1 to control the transcriptional potential of the genome. *Genes Dev.* 35, 749–770. <https://doi.org/10.1101/gad.347005.120>
- Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., Rais, Y., Shipony, Z., Mukamel, Z., Krupalnik, V., Zerbib, M., Geula, S., Caspi, I., Schneir, D., Shwartz, T., Gilad, S., Amann-Zalcenstein, D., Benjamin, S., Amit, I., Tanay, A., Massarwa, R., Novershtern, N., Hanna, J.H., 2013. Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504, 282–286. <https://doi.org/10.1038/nature12745>
- Gao, Z., Lee, P., Stafford, J.M., von Schimmelmann, M., Schaefer, A., Reinberg, D., 2014. An AUTS2–Polycomb complex activates gene expression in the CNS. *Nature* 516, 349–354. <https://doi.org/10.1038/nature13921>
- Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., Reinberg, D., 2012. PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol. Cell* 45, 344–356. <https://doi.org/10.1016/j.molcel.2012.01.002>
- Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., Bardwell, V.J., 2006. Polycomb Group and SCF Ubiquitin Ligases Are Found in a Novel BCOR Complex That Is Recruited to BCL6 Targets. *Mol. Cell. Biol.* 26, 6880–6889. <https://doi.org/10.1128/MCB.00630-06>
- Gerri, C., McCarthy, A., Alanis-Lobato, G., Demtschenko, A., Bruneau, A., Loubersac, S., Fogarty, N.M.E., Hampshire, D., Elder, K., Snell, P., Christie, L., David, L., Van de Velde, H., Fouladi-Nashta, A.A., Niakan, K.K., 2020. Initiation of a conserved trophoderm program in human, cow and mouse embryos. *Nature* 587, 443–447. <https://doi.org/10.1038/s41586-020-2759-x>
- Ghimire, S., Mantziou, V., Moris, N., Martinez Arias, A., 2021. Human gastrulation: The embryo and its models. *Dev. Biol.* 474, 100–108. <https://doi.org/10.1016/j.ydbio.2021.01.006>
- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., Qi, L.S., Kampmann, M., Weissman, J.S., 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647–661. <https://doi.org/10.1016/j.cell.2014.09.029>
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442–451. <https://doi.org/10.1016/j.cell.2013.06.044>
- Giulitti, S., Pellegrini, M., Zorzan, I., Martini, P., Gagliano, O., Mutarelli, M., Ziller, M.J., Cacchiarelli, D., Romualdi, C., Elvassore, N., Martello, G., 2019. Direct generation of human naive induced pluripotent stem cells from somatic cells in microfluidics. *Nat. Cell Biol.* 21, 275–286. <https://doi.org/10.1038/s41556-018-0254-5>
- Godwin, S., Ward, D., Pedone, E., Homer, M., Fletcher, A.G., Marucci, L., 2017. An extended model for culture-dependent heterogeneous gene expression and proliferation dynamics in mouse embryonic stem cells. *Npj Syst. Biol. Appl.* 3, 19. <https://doi.org/10.1038/s41540-017-0020-5>
- Gong, X., Du, D., Deng, Y., Zhou, Y., Sun, L., Yuan, S., 2020. The structure and regulation of the E3 ubiquitin ligase HUWE1 and its biological functions in cancer. *Invest. New Drugs* 38, 515–524. <https://doi.org/10.1007/s10637-020-00894-6>
- Goodwin, J., Laslett, A.L., Rugg-Gunn, P.J., 2020. The application of cell surface markers to demarcate distinct human pluripotent states. *Exp. Cell Res.* 387, 111749. <https://doi.org/10.1016/j.yexcr.2019.111749>
- Grant, C.E., Bailey, T.L., Noble, W.S., 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>

- Grau, D.J., Chapman, B.A., Garlick, J.D., Borowsky, M., Francis, N.J., Kingston, R.E., 2011. Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev.* 25, 2210–2221. <https://doi.org/10.1101/gad.17288211>
- Greber, B., Coulon, P., Zhang, M., Moritz, S., Frank, S., Müller-Molina, A.J., Araújo-Bravo, M.J., Han, D.W., Pape, H.-C., Schöler, H.R., 2011. FGF signalling inhibits neural induction in human embryonic stem cells: FGF inhibits neural induction in hESCs. *EMBO J.* 30, 4874–4884. <https://doi.org/10.1038/emboj.2011.407>
- Greber, B., Lehrach, H., Adjaye, J., 2007. Fibroblast Growth Factor 2 Modulates Transforming Growth Factor β Signaling in Mouse Embryonic Fibroblasts and Human ESCs (hESCs) to Support hESC Self-Renewal. *Stem Cells* 25, 455–464. <https://doi.org/10.1634/stemcells.2006-0476>
- Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M., Szumska, D., McGouran, J.F., Bhattacharya, S., Kessler, B.M., Brockdorff, N., Cooper, S., 2016. Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ES cells. *Development dev.* 123935. <https://doi.org/10.1242/dev.123935>
- Groh, B.S., Yan, F., Smith, M.D., Yu, Y., Chen, X., Xiong, Y., 2016. The antiobesity factor WDTC 1 suppresses adipogenesis via the CRL 4^{WDTC 1} E3 ligase. *EMBO Rep.* 17, 638–647. <https://doi.org/10.15252/embr.201540500>
- Guo, G., Stirparo, G.G., Strawbridge, S., Spindlow, D., Yang, J., Clarke, J., Dattani, A., Yanagida, A., Li, M.A., Myers, S., Özel, B.N., Nichols, J., Smith, A., 2020a. Human Naïve Epiblast Cells Possess Unrestricted Lineage Potential (preprint). *Developmental Biology*. <https://doi.org/10.1101/2020.02.04.933812>
- Guo, G., Stirparo, G.G., Strawbridge, S., Yang, J., Clarke, J., Li, M.A., Myers, S., Özel, B.N., Nichols, J., Smith, A., 2020b. Trophoctoderm Potency is Retained Exclusively in Human Naïve Cells (preprint). *Developmental Biology*. <https://doi.org/10.1101/2020.02.04.933812>
- Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., Plath, K., Smith, A., 2017. Epigenetic resetting of human pluripotency. *Development* 144, 2748–2763. <https://doi.org/10.1242/dev.146811>
- Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., Nichols, J., 2016a. Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Rep.* 6, 437–446. <https://doi.org/10.1016/j.stemcr.2016.02.005>
- Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., Nichols, J., 2016b. Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Rep.* 6, 437–446. <https://doi.org/10.1016/j.stemcr.2016.02.005>
- Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., Smith, A., 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136, 1063–1069. <https://doi.org/10.1242/dev.030957>
- Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., Jin, X., Shi, X., Liu, P., Wang, X., Wang, W., Wei, Y., Li, X., Guo, F., Wu, X., Fan, X., Yong, J., Wen, L., Xie, S.X., Tang, F., Qiao, J., 2014. The DNA methylation landscape of human early embryos. *Nature* 511, 606–610. <https://doi.org/10.1038/nature13544>
- Gurdon, J.B., 1960. The Developmental Capacity of Nuclei Taken from Differentiating Endoderm Cells of *Xenopus laevis*. *Development* 8, 505–526. <https://doi.org/10.1242/dev.8.4.505>
- Hackett, J.A., Surani, M.A., 2014. Regulatory Principles of Pluripotency: From the Ground State Up. *Cell Stem Cell* 15, 416–430. <https://doi.org/10.1016/j.stem.2014.09.015>
- Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., Jaenisch, R., 2010. Human embryonic stem cells with biological and

- epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl. Acad. Sci.* 107, 9222–9227. <https://doi.org/10.1073/pnas.1004584107>
- Hauri, S., Comoglio, F., Seimiya, M., Gerstung, M., Glatter, T., Hansen, K., Aebersold, R., Paro, R., Gstaiger, M., Beisel, C., 2016. A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Rep.* 17, 583–595. <https://doi.org/10.1016/j.celrep.2016.08.096>
- Hay, D.C., Sutherland, L., Clark, J., Burdon, T., 2004. Oct-4 Knockdown Induces Similar Patterns of Endoderm and Trophoblast Differentiation Markers in Human and Mouse Embryonic Stem Cells. *Stem Cells* 22, 225–235. <https://doi.org/10.1634/stemcells.22-2-225>
- Hayashi, H., Kubo, Y., Izumida, M., Matsuyama, T., 2020. Efficient viral delivery of Cas9 into human safe harbor. *Sci. Rep.* 10, 21474. <https://doi.org/10.1038/s41598-020-78450-8>
- Hayashi, K., Neve, J., Hirose, M., Kuboki, A., Shimada, Y., Kepinski, S., Nozaki, H., 2012. Rational Design of an Auxin Antagonist of the SCF^{TIR1} Auxin Receptor Complex. *ACS Chem. Biol.* 7, 590–598. <https://doi.org/10.1021/cb200404c>
- Healy, E., Mucha, M., Glancy, E., Fitzpatrick, D.J., Conway, E., Neikes, H.K., Monger, C., Van Mierlo, G., Baltissen, M.P., Koseki, Y., Vermeulen, M., Koseki, H., Bracken, A.P., 2019. PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation. *Mol. Cell* 76, 437–452.e6. <https://doi.org/10.1016/j.molcel.2019.08.012>
- Healy, L., Ruban, L., 2015. *Atlas of Human Pluripotent Stem Cells in Culture*. Springer US, Boston, MA. <https://doi.org/10.1007/978-1-4899-7507-2>
- Henderson, J.K., Draper, J.S., Baillie, H.S., Fishel, S., Thomson, J.A., Moore, H., Andrews, P.W., 2002. Preimplantation Human Embryos and Embryonic Stem Cells Show Comparable Expression of Stage-Specific Embryonic Antigens. *STEM CELLS* 20, 329–337. <https://doi.org/10.1634/stemcells.20-4-329>
- Hickey, G.J., Wike, C.L., Nie, X., Guo, Y., Tan, M., Murphy, P.J., Cairns, B.R., 2022. Establishment of developmental gene silencing by ordered polycomb complex recruitment in early zebrafish embryos. *eLife* 11, e67738. <https://doi.org/10.7554/eLife.67738>
- Hirai, H., Karian, P., Kikyo, N., 2011. Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor. *Biochem. J.* 438, 11–23. <https://doi.org/10.1042/BJ20102152>
- Hitoshi, N., Ken-ichi, Y., Jun-ichi, M., 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108, 193–199. [https://doi.org/10.1016/0378-1119\(91\)90434-D](https://doi.org/10.1016/0378-1119(91)90434-D)
- Holoch, D., Margueron, R., 2017. Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. *Trends Biochem. Sci.* 42, 531–542. <https://doi.org/10.1016/j.tibs.2017.04.003>
- Hori, K., Nagai, T., Shan, W., Sakamoto, A., Taya, S., Hashimoto, R., Hayashi, T., Abe, M., Yamazaki, M., Nakao, K., Nishioka, T., Sakimura, K., Yamada, K., Kaibuchi, K., Hoshino, M., 2014. Cytoskeletal Regulation by AUTS2 in Neuronal Migration and Neuritogenesis. *Cell Rep.* 9, 2166–2179. <https://doi.org/10.1016/j.celrep.2014.11.045>
- Huang, K., Maruyama, T., Fan, G., 2014. The Naive State of Human Pluripotent Stem Cells: A Synthesis of Stem Cell and Preimplantation Embryo Transcriptome Analyses. *Cell Stem Cell* 15, 410–415. <https://doi.org/10.1016/j.stem.2014.09.014>
- Huang, S.-M.A., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., Hild, M., Shi, X., Wilson, C.J., Mickanin, C., Myer, V., Fazal, A., Tomlinson, R., Serluca, F., Shao, W., Cheng, H., Shultz, M., Rau, C., Schirle, M., Schlegl, J., Ghidelli, S., Fawell, S., Lu, C., Curtis, D., Kirschner, M.W., Lengauer, C., Finan, P.M., Tallarico, J.A., Bouwmeester, T., Porter, J.A., Bauer, A., Cong, F., 2009. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* 461, 614–620. <https://doi.org/10.1038/nature08356>

- Huseyin, M.K., Klose, R.J., 2020. Live-cell single particle tracking of PRC1 reveals a highly dynamic system with low target site occupancy (preprint). *Molecular Biology*. <https://doi.org/10.1101/2020.04.25.061358>
- Hwang, B., Lee, J.H., Bang, D., 2018. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* 50, 1–14. <https://doi.org/10.1038/s12276-018-0071-8>
- Hyslop, L., Stojkovic, M., Armstrong, L., Walter, T., Stojkovic, P., Przyborski, S., Herbert, M., Murdoch, A., Strachan, T., Lako, M., 2005. Downregulation of NANOG Induces Differentiation of Human Embryonic Stem Cells to Extraembryonic Lineages. *STEM CELLS* 23, 1035–1043. <https://doi.org/10.1634/stemcells.2005-0080>
- Illingworth, R.S., Moffat, M., Mann, A.R., Read, D., Hunter, C.J., Pradeepa, M.M., Adams, I.R., Bickmore, W.A., 2015. The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development. *Genes Dev.* 29, 1897–1902. <https://doi.org/10.1101/gad.268151.115>
- Ingham, P.W., 1985. Genetic Control of the Spatial Pattern of Selector Gene Expression in *Drosophila*. *Cold Spring Harb. Symp. Quant. Biol.* 50, 201–208. <https://doi.org/10.1101/SQB.1985.050.01.026>
- Io, S., Kabata, M., Iemura, Y., Semi, K., Morone, N., Minagawa, A., Wang, B., Okamoto, I., Nakamura, T., Kojima, Y., Iwatani, C., Tsuchiya, H., Kaswandy, B., Kondoh, E., Kaneko, S., Woltjen, K., Saitou, M., Yamamoto, T., Mandai, M., Takashima, Y., 2021. Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell* 28, 1023–1039.e13. <https://doi.org/10.1016/j.stem.2021.03.013>
- Ishihara, A., Hou, Y., Jacobson, K., 1987. The Thy-1 antigen exhibits rapid lateral diffusion in the plasma membrane of rodent lymphoid cells and fibroblasts. *Proc. Natl. Acad. Sci.* 84, 1290–1293. <https://doi.org/10.1073/pnas.84.5.1290>
- Ismail, I.H., McDonald, D., Strickfaden, H., Xu, Z., Hendzel, M.J., 2013. A Small Molecule Inhibitor of Polycomb Repressive Complex 1 Inhibits Ubiquitin Signaling at DNA Double-strand Breaks. *J. Biol. Chem.* 288, 26944–26954. <https://doi.org/10.1074/jbc.M113.461699>
- Isono, K., Endo, T.A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B.E., Koseki, H., 2013. SAM Domain Polymerization Links Subnuclear Clustering of PRC1 to Gene Silencing. *Dev. Cell* 26, 565–577. <https://doi.org/10.1016/j.devcel.2013.08.016>
- Jadhav, U., Manieri, E., Nalapareddy, K., Madha, S., Chakrabarti, S., Wucherpfennig, K., Barefoot, M., Shivdasani, R.A., 2020. Replicational Dilution of H3K27me3 in Mammalian Cells and the Role of Poised Promoters. *Mol. Cell* 78, 141–151.e5. <https://doi.org/10.1016/j.molcel.2020.01.017>
- James, D., Levine, A.J., Besser, D., Hemmati-Brivanlou, A., 2005. TGF β /activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* 132, 1273–1282. <https://doi.org/10.1242/dev.01706>
- Jürgens, G., 1985. A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* 316, 153–155. <https://doi.org/10.1038/316153a0>
- Kagawa, H., Javali, A., Khoei, H.H., Sommer, T., Sestini, G., Novatchkova, M., Reimer, Y.S., Rivron, N., 2022. Human blastoids model blastocyst development and implantation (preprint). *Protocol Exchange*. <https://doi.org/10.21203/rs.3.pex-1639/v1>
- Kalb, R., Latwiel, S., Baymaz, H.I., Jansen, P.W.T.C., Müller, C.W., Vermeulen, M., Müller, J., 2014. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* 21, 569–571. <https://doi.org/10.1038/nsmb.2833>
- Kanki, Y., Muramatsu, M., Miyamura, Y., Kikuchi, K., Higashijima, Y., Nakaki, R., Suehiro, J., Sasaki, Y., Kubota, Y., Koseki, H., Morioka, H., Kodama, T., Nakao, M., Kurotaki, D., Aburatani, H., Minami, T., 2022. Bivalent-histone-marked immediate-early gene regulation is vital for VEGF-responsive angiogenesis. *Cell Rep.* 38, 110332. <https://doi.org/10.1016/j.celrep.2022.110332>

- Karakaidos, P., Karagiannis, D., Rampias, T., 2020. Resolving DNA Damage: Epigenetic Regulation of DNA Repair. *Molecules* 25, 2496. <https://doi.org/10.3390/molecules25112496>
- Kassis, J.A., Brown, J.L., 2013. Polycomb Group Response Elements in *Drosophila* and Vertebrates, in: *Advances in Genetics*. Elsevier, pp. 83–118. <https://doi.org/10.1016/B978-0-12-407677-8.00003-8>
- Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., Henikoff, S., 2019. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* 10, 1930. <https://doi.org/10.1038/s41467-019-09982-5>
- Khan, S.A., Park, K., Fischer, L.A., Dong, C., Lungjangwa, T., Jimenez, M., Casalena, D., Chew, B., Dietmann, S., Auld, D.S., Jaenisch, R., Theunissen, T.W., 2021. Probing the signaling requirements for naive human pluripotency by high-throughput chemical screening. *Cell Rep.* 35, 109233. <https://doi.org/10.1016/j.celrep.2021.109233>
- Kheir, T.B., Lund, A.H., 2010. Epigenetic dynamics across the cell cycle. *Essays Biochem.* 48, 107–120. <https://doi.org/10.1042/bse0480107>
- Kimura, H., 2013. Histone modifications for human epigenome analysis. *J. Hum. Genet.* 58, 439–445. <https://doi.org/10.1038/jhg.2013.66>
- King, H.W., Fursova, N.A., Blackledge, N.P., Klose, R.J., 2018. Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes. *Genome Res.* 28, 1494–1507. <https://doi.org/10.1101/gr.237180.118>
- Kinoshita, M., Barber, M., Mansfield, W., Cui, Y., Spindlow, D., Stirparo, G.G., Dietmann, S., Nichols, J., Smith, A., 2021. Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* 28, 453–471.e8. <https://doi.org/10.1016/j.stem.2020.11.005>
- Knöfler, M., Haider, S., Saleh, L., Pollheimer, J., Gamage, T.K.J.B., James, J., 2019. Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell. Mol. Life Sci.* 76, 3479–3496. <https://doi.org/10.1007/s00018-019-03104-6>
- Kondrychyn, I., Robra, L., Thirumalai, V., 2017. Transcriptional Complexity and Distinct Expression Patterns of *auts2* Paralogs in *Danio rerio*. *G3* 7, 2577–2593. <https://doi.org/10.1534/g3.117.042622>
- Konze, K.D., Ma, A., Li, F., Barsyte-Lovejoy, D., Parton, T., MacNevin, C.J., Liu, F., Gao, C., Huang, X.-P., Kuznetsova, E., Rougie, M., Jiang, A., Pattenden, S.G., Norris, J.L., James, L.I., Roth, B.L., Brown, P.J., Frye, S.V., Arrowsmith, C.H., Hahn, K.M., Wang, G.G., Vedadi, M., Jin, J., 2013. An Orally Bioavailable Chemical Probe of the Lysine Methyltransferases EZH2 and EZH1. *ACS Chem. Biol.* 8, 1324–1334. <https://doi.org/10.1021/cb400133j>
- Kuhn, R.M., Haussler, D., Kent, W.J., 2013. The UCSC genome browser and associated tools. *Brief. Bioinform.* 14, 144–161. <https://doi.org/10.1093/bib/bbs038>
- Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., McDermott, M.G., Monteiro, C.D., Gundersen, G.W., Ma'ayan, A., 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44, W90–W97. <https://doi.org/10.1093/nar/gkw377>
- Kumar, B., Navarro, C., Winblad, N., Schell, J.P., Zhao, C., Weltner, J., Baqué-Vidal, L., Salazar Mantero, A., Petropoulos, S., Lanner, F., Elsässer, S.J., 2022. Polycomb repressive complex 2 shields naïve human pluripotent cells from trophectoderm differentiation. *Nat. Cell Biol.* <https://doi.org/10.1038/s41556-022-00916-w>
- Kunath, T., Saba-El-Leil, M.K., Almousaileakh, M., Wray, J., Meloche, S., Smith, A., 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* 134, 2895–2902. <https://doi.org/10.1242/dev.02880>
- Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J.T., Sadreyev, R.I., Dekker, J., Kingston, R.E., 2017. Polycomb Repressive Complex 1 Generates Discrete Compacted Domains

- that Change during Differentiation. *Mol. Cell* 65, 432–446.e5. <https://doi.org/10.1016/j.molcel.2017.01.009>
- Lagarou, A., Mohd-Sarip, A., Moshkin, Y.M., Chalkley, G.E., Bezstarosti, K., Demmers, J.A.A., Verrijzer, C.P., 2008. dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev.* 22, 2799–2810. <https://doi.org/10.1101/gad.484208>
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lau, M.S., Schwartz, M.G., Kundu, S., Savol, A.J., Wang, P.I., Marr, S.K., Grau, D.J., Schorderet, P., Sadreyev, R.I., Tabin, C.J., Kingston, R.E., 2017. Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science* 355, 1081–1084. <https://doi.org/10.1126/science.aah5403>
- Lea, R.A., McCarthy, A., Boeing, S., Fallesen, T., Elder, K., Snell, P., Christie, L., Adkins, S., Shaikly, V., Taranissi, M., Niakan, K.K., 2021. KLF17 promotes human naïve pluripotency but is not required for its establishment. *Development* 148, dev199378. <https://doi.org/10.1242/dev.199378>
- Lee, H.-G., Kahn, T.G., Simcox, A., Schwartz, Y.B., Pirrotta, V., 2015. Genome-wide activities of Polycomb complexes control pervasive transcription. *Genome Res.* 25, 1170–1181. <https://doi.org/10.1101/gr.188920.114>
- Lee, J.T., Jaenisch, R., 1997. Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature* 386, 275–279. <https://doi.org/10.1038/386275a0>
- Legrand, J.M.D., Chan, A.-L., La, H.M., Rossello, F.J., Änkö, M.-L., Fuller-Pace, F.V., Hobbs, R.M., 2019. DDX5 plays essential transcriptional and post-transcriptional roles in the maintenance and function of spermatogonia. *Nat. Commun.* 10, 2278. <https://doi.org/10.1038/s41467-019-09972-7>
- Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., Kingston, R.E., 2002. The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans. *Mol. Cell. Biol.* 22, 6070–6078. <https://doi.org/10.1128/MCB.22.17.6070-6078.2002>
- Lewis, E.B., 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565–570. <https://doi.org/10.1038/276565a0>
- Lewis: New mutants report - Google Scholar [WWW Document], n.d. URL https://scholar.google.com/scholar_lookup?title=New%20mutants%20report&author=P.H.%20Lewis&publication_year=1947 (accessed 6.10.22).
- Li, D., Hsu, S., Purushotham, D., Sears, R.L., Wang, T., 2019. WashU Epigenome Browser update 2019. *Nucleic Acids Res.* 47, W158–W165. <https://doi.org/10.1093/nar/gkz348>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, H., Lai, P., Jia, J., Song, Y., Xia, Q., Huang, K., He, N., Ping, W., Chen, J., Yang, Z., Li, J., Yao, M., Dong, X., Zhao, J., Hou, C., Esteban, M.A., Gao, S., Pei, D., Hutchins, A.P., Yao, H., 2017. RNA Helicase DDX5 Inhibits Reprogramming to Pluripotency by miRNA-Based Repression of RYBP and its PRC1-Dependent and -Independent Functions. *Cell Stem Cell* 20, 462–477.e6. <https://doi.org/10.1016/j.stem.2016.12.002>
- Li, K., Wang, Z., 2021. Histone crotonylation-centric gene regulation. *Epigenetics Chromatin* 14, 10. <https://doi.org/10.1186/s13072-021-00385-9>
- Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., Liu, X.S., 2014. MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* 15, 554. <https://doi.org/10.1186/s13059-014-0554-4>
- Linneberg-Agerholm, M., Wong, Y.F., Romero Herrera, J.A., Monteiro, R.S., Anderson, K.G.V., Brickman, J.M., 2019. Naïve human pluripotent stem cells respond to Wnt,

- Nodal and LIF signalling to produce expandable naïve extra-embryonic endoderm. *Development* 146, dev180620. <https://doi.org/10.1242/dev.180620>
- Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S., O'Brien, C.M., Huang, C., Mishra, K., Hodgson-Garms, M., Jansz, N., Williams, S.M., Blewitt, M.E., Nilsson, S.K., Schittenhelm, R.B., Laslett, A.L., Lister, R., Polo, J.M., 2017. Comprehensive characterization of distinct states of human naïve pluripotency generated by reprogramming. *Nat. Methods* 14, 1055–1062. <https://doi.org/10.1038/nmeth.4436>
- Liu, X., Ouyang, J.F., Rossello, F.J., Tan, J.P., Davidson, K.C., Valdes, D.S., Schröder, J., Sun, Y.B.Y., Chen, J., Knaupp, A.S., Sun, G., Chy, H.S., Huang, Z., Pflueger, J., Firas, J., Tano, V., Buckberry, S., Paynter, J.M., Larcombe, M.R., Poppe, D., Choo, X.Y., O'Brien, C.M., Pastor, W.A., Chen, D., Leichter, A.L., Naeem, H., Tripathi, P., Das, P.P., Grubman, A., Powell, D.R., Laslett, A.L., David, L., Nilsson, S.K., Clark, A.T., Lister, R., Nefzger, C.M., Martelotto, L.G., Rackham, O.J.L., Polo, J.M., 2020. Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature* 586, 101–107. <https://doi.org/10.1038/s41586-020-2734-6>
- Liu, X., Tan, J.P., Schröder, J., Aberkane, A., Ouyang, J.F., Mohenska, M., Lim, S.M., Sun, Y.B.Y., Chen, J., Sun, G., Zhou, Y., Poppe, D., Lister, R., Clark, A.T., Rackham, O.J.L., Zenker, J., Polo, J.M., 2021. Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* 591, 627–632. <https://doi.org/10.1038/s41586-021-03372-y>
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lynch, C.J., Bernad, R., Calvo, I., Serrano, M., 2020. Manipulating the Mediator complex to induce naïve pluripotency. *Exp. Cell Res.* 395, 112215. <https://doi.org/10.1016/j.yexcr.2020.112215>
- Lyon, M.F., 1961. Gene Action in the X-chromosome of the Mouse (*Mus musculus* L.). *Nature* 190, 372–373. <https://doi.org/10.1038/190372a0>
- Ma, X., Chen, H., Chen, L., 2016. A dual role of Erk signaling in embryonic stem cells. *Exp. Hematol.* 44, 151–156. <https://doi.org/10.1016/j.exphem.2015.12.008>
- Maan, Z.N., Rennert, R.C., Koob, T.J., Januszyk, M., Li, W.W., Gurtner, G.C., 2015. Cell recruitment by amnion chorion grafts promotes neovascularization. *J. Surg. Res.* 193, 953–962. <https://doi.org/10.1016/j.jss.2014.08.045>
- Mackinlay, K.M., Weatherbee, B.A., Souza Rosa, V., Handford, C.E., Hudson, G., Coorens, T., Pereira, L.V., Behjati, S., Vallier, L., Shahbazi, M.N., Zernicka-Goetz, M., 2021. An in vitro stem cell model of human epiblast and yolk sac interaction. *eLife* 10, e63930. <https://doi.org/10.7554/eLife.63930>
- Maezawa, S., Hasegawa, K., Yukawa, M., Sakashita, A., Alavattam, K.G., Andreassen, P.R., Vidal, M., Koseki, H., Barski, A., Namekawa, S.H., 2017. Polycomb directs timely activation of germline genes in spermatogenesis. *Genes Dev.* 31, 1693–1703. <https://doi.org/10.1101/gad.302000.117>
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-Guided Human Genome Engineering via Cas9. *Science* 339, 823–826. <https://doi.org/10.1126/science.1232033>
- Mamede, A.C., Carvalho, M.J., Abrantes, A.M., Laranjo, M., Maia, C.J., Botelho, M.F., 2012. Amniotic membrane: from structure and functions to clinical applications. *Cell Tissue Res.* 349, 447–458. <https://doi.org/10.1007/s00441-012-1424-6>
- Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I., Judge, L.M., Gordon, D.E., Eskildsen, T.V., Villalta, J.E., Horlbeck, M.A., Gilbert, L.A., Krogan, N.J., Sheikh, S.P., Weissman, J.S., Qi, L.S., So, P.-L., Conklin, B.R., 2016. CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell* 18, 541–553. <https://doi.org/10.1016/j.stem.2016.01.022>

- Marahrens, Y., Loring, J., Jaenisch, R., 1998. Role of the Xist Gene in X Chromosome Choosing. *Cell* 92, 657–664. [https://doi.org/10.1016/S0092-8674\(00\)81133-2](https://doi.org/10.1016/S0092-8674(00)81133-2)
- Martello, G., Bertone, P., Smith, A., 2013. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *EMBO J.* 32, 2561–2574. <https://doi.org/10.1038/emboj.2013.177>
- Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci.* 78, 7634–7638. <https://doi.org/10.1073/pnas.78.12.7634>
- McLeay, R.C., Bailey, T.L., 2010. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11, 165. <https://doi.org/10.1186/1471-2105-11-165>
- Meers, M.P., Bryson, T.D., Henikoff, J.G., Henikoff, S., 2019. Improved CUT&RUN chromatin profiling tools. *eLife* 8, e46314. <https://doi.org/10.7554/eLife.46314>
- Mei, H., Kozuka, C., Hayashi, R., Kumon, M., Koseki, H., Inoue, A., 2021. H2AK119ub1 guides maternal inheritance and zygotic deposition of H3K27me3 in mouse embryos. *Nat. Genet.* 53, 539–550. <https://doi.org/10.1038/s41588-021-00820-3>
- Messmer, S., Franke, A., Paro, R., 1992. Analysis of the functional role of the Polycomb chromo domain in *Drosophila melanogaster*. *Genes Dev.* 6, 1241–1254. <https://doi.org/10.1101/gad.6.7.1241>
- Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.L., Marioni, J.C., Reik, W., 2019. Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. *Cell Rep.* 26, 815-824.e4. <https://doi.org/10.1016/j.celrep.2018.12.099>
- Mierlo, G. van, Veenstra, G.J.C., Vermeulen, M., Marks, H., 2019. The Complexity of PRC2 Subcomplexes. *Trends Cell Biol.* 29, 660–671. <https://doi.org/10.1016/j.tcb.2019.05.004>
- Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., Shilatifard, A., 2001. COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci.* 98, 12902–12907. <https://doi.org/10.1073/pnas.231473398>
- Min, J., 2003. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* 17, 1823–1828. <https://doi.org/10.1101/gad.269603>
- Minkovsky, A., Patel, S., Plath, K., 2012. Concise Review: Pluripotency and the Transcriptional Inactivation of the Female Mammalian X Chromosome. *STEM CELLS* 30, 48–54. <https://doi.org/10.1002/stem.755>
- Mohammed, H., Taylor, C., Brown, G.D., Papachristou, E.K., Carroll, J.S., D'Santos, C.S., 2016. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* 11, 316–326. <https://doi.org/10.1038/nprot.2016.020>
- Mohan, M., Herz, H.-M., Smith, E.R., Zhang, Y., Jackson, J., Washburn, M.P., Florens, L., Eissenberg, J.C., Shilatifard, A., 2011. The COMPASS Family of H3K4 Methylases in *Drosophila*. *Mol. Cell. Biol.* 31, 4310–4318. <https://doi.org/10.1128/MCB.06092-11>
- Molè, M.A., Coorens, T.H.H., Shahbazi, M.N., Weberling, A., Weatherbee, B.A.T., Gantner, C.W., Sancho-Serra, C., Richardson, L., Drinkwater, A., Syed, N., Engley, S., Snell, P., Christie, L., Elder, K., Campbell, A., Fishel, S., Behjati, S., Vento-Tormo, R., Zernicka-Goetz, M., 2021. A single cell characterisation of human embryogenesis identifies pluripotency transitions and putative anterior hypoblast centre. *Nat. Commun.* 12, 3679. <https://doi.org/10.1038/s41467-021-23758-w>
- Monderer-Rothkoff, G., Tal, N., Risman, M., Shani, O., Nissim-Rafinia, M., Malki-Feldman, L., Medvedeva, V., Groszer, M., Meshorer, E., Shifman, S., 2019. AUTS2 isoforms control neuronal differentiation. *Mol. Psychiatry*. <https://doi.org/10.1038/s41380-019-0409-1>
- Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., Magnuson, T., 2005. The Murine Polycomb Group Protein Eed Is Required for Global

- Histone H3 Lysine-27 Methylation. *Curr. Biol.* 15, 942–947.
<https://doi.org/10.1016/j.cub.2005.04.051>
- Morikawa, M., Koinuma, D., Mizutani, A., Kawasaki, N., Holmborn, K., Sundqvist, A., Tsutsumi, S., Watabe, T., Aburatani, H., Heldin, C.-H., Miyazono, K., 2016. BMP Sustains Embryonic Stem Cell Self-Renewal through Distinct Functions of Different Krüppel-like Factors. *Stem Cell Rep.* 6, 64–73.
<https://doi.org/10.1016/j.stemcr.2015.12.004>
- Moris, N., Anlas, K., van den Brink, S.C., Alemany, A., Schröder, J., Ghimire, S., Balayo, T., van Oudenaarden, A., Martinez Arias, A., 2020. An in vitro model of early anteroposterior organization during human development. *Nature* 582, 410–415.
<https://doi.org/10.1038/s41586-020-2383-9>
- Müller, F.-J., Schuld, B.M., Williams, R., Mason, D., Altun, G., Papapetrou, E.P., Danner, S., Goldmann, J.E., Herbst, A., Schmidt, N.O., Aldenhoff, J.B., Laurent, L.C., Loring, J.F., 2011. A bioinformatic assay for pluripotency in human cells. *Nat. Methods* 8, 315–317. <https://doi.org/10.1038/nmeth.1580>
- Müller, J., 1995. Transcriptional silencing by the Polycomb protein in *Drosophila* embryos. *EMBO J.* 14, 1209–1220. <https://doi.org/10.1002/j.1460-2075.1995.tb07104.x>
- Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., Simon, J.A., 2002. Histone Methyltransferase Activity of a *Drosophila* Polycomb Group Repressor Complex. *Cell* 111, 197–208.
[https://doi.org/10.1016/S0092-8674\(02\)00976-5](https://doi.org/10.1016/S0092-8674(02)00976-5)
- Murr, R., 2010. Interplay Between Different Epigenetic Modifications and Mechanisms, in: *Advances in Genetics*. Elsevier, pp. 101–141. <https://doi.org/10.1016/B978-0-12-380866-0.60005-8>
- Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., Scott, T.G., Vittori, S., Perry, J.A., Qi, J., Winter, G.E., Wong, K.-K., Gray, N.S., Bradner, J.E., 2018. The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* 14, 431–441.
<https://doi.org/10.1038/s41589-018-0021-8>
- Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., Roder, J.C., 1993. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci.* 90, 8424–8428. <https://doi.org/10.1073/pnas.90.18.8424>
- Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., Saitou, M., 2016. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* 537, 57–62.
<https://doi.org/10.1038/nature19096>
- Natsume, T., Kiyomitsu, T., Saga, Y., Kanemaki, M.T., 2016. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep.* 15, 210–218. <https://doi.org/10.1016/j.celrep.2016.03.001>
- Nesterova, T.B., Wei, G., Coker, H., Pintacuda, G., Bowness, J.S., Zhang, T., Almeida, M., Bloechl, B., Moindrot, B., Carter, E.J., Alvarez Rodrigo, I., Pan, Q., Bi, Y., Song, C.-X., Brockdorff, N., 2019. Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat. Commun.* 10, 3129.
<https://doi.org/10.1038/s41467-019-11171-3>
- Ng, J., Hart, C.M., Morgan, K., Simon, J.A., 2000. A *Drosophila* ESC-E(Z) Protein Complex Is Distinct from Other Polycomb Group Complexes and Contains Covalently Modified ESC. *Mol. Cell. Biol.* 20, 3069–3078. <https://doi.org/10.1128/MCB.20.9.3069-3078.2000>
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., Smith, A., 1998. Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell* 95, 379–391.
[https://doi.org/10.1016/S0092-8674\(00\)81769-9](https://doi.org/10.1016/S0092-8674(00)81769-9)
- Nickel, B.E., Davie, J.R., 1989. Structure of polyubiquitinated histone H2A. *Biochemistry* 28, 964–968. <https://doi.org/10.1021/bi00429a007>

- Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., Kanemaki, M., 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* 6, 917–922. <https://doi.org/10.1038/nmeth.1401>
- Niwa, H., 2001. Molecular Mechanism to Maintain Stem Cell Renewal of ES Cells. *Cell Struct. Funct.* 26, 137–148. <https://doi.org/10.1247/csf.26.137>
- Niwa, H., Burdon, T., Chambers, I., Smith, A., 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* 12, 2048–2060. <https://doi.org/10.1101/gad.12.13.2048>
- Noble, D., 2015. Conrad Waddington and the origin of epigenetics. *J. Exp. Biol.* 218, 816–818. <https://doi.org/10.1242/jeb.120071>
- O'Brien, C.M., Chy, H.S., Zhou, Q., Blumenfeld, S., Lamshead, J.W., Liu, X., Kie, J., Capaldo, B.D., Chung, T.-L., Adams, T.E., Phan, T., Bentley, J.D., McKinstry, W.J., Oliva, K., McMurrick, P.J., Wang, Y.-C., Rossello, F.J., Lindeman, G.J., Chen, D., Jarde, T., Clark, A.T., Abud, H.E., Visvader, J.E., Nefzger, C.M., Polo, J.M., Loring, J.F., Laslett, A.L., 2017. New Monoclonal Antibodies to Defined Cell Surface Proteins on Human Pluripotent Stem Cells. *Stem Cells* 35, 626–640. <https://doi.org/10.1002/stem.2558>
- O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., Jenuwein, T., 2001. The Polycomb-Group Gene *Ezh2* Is Required for Early Mouse Development. *Mol. Cell Biol.* 21, 4330–4336. <https://doi.org/10.1128/MCB.21.13.4330-4336.2001>
- Ogiyama, Y., Schuettengruber, B., Papadopoulos, G.L., Chang, J.-M., Cavalli, G., 2018. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during *Drosophila* Development. *Mol. Cell* 71, 73-88.e5. <https://doi.org/10.1016/j.molcel.2018.05.032>
- Okabe, H., Chiba, H., Hiura, H., Hamada, H., Sato, A., Utsunomiya, T., Kikuchi, H., Yoshida, H., Tanaka, A., Suyama, M., Arima, T., 2014. Genome-Wide Analysis of DNA Methylation Dynamics during Early Human Development. *PLoS Genet.* 10, e1004868. <https://doi.org/10.1371/journal.pgen.1004868>
- Okabe, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M., Sasaki, H., Arima, T., 2018. Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* 22, 50-63.e6. <https://doi.org/10.1016/j.stem.2017.11.004>
- Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., Heard, E., 2004. Epigenetic Dynamics of Imprinted X Inactivation During Early Mouse Development. *Science* 303, 644–649. <https://doi.org/10.1126/science.1092727>
- Okamoto, I., Patrat, C., Thépot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.-P., Renard, J.-P., Duranthon, V., Heard, E., 2011. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* 472, 370–374. <https://doi.org/10.1038/nature09872>
- Oksenberg, N., Ahituv, N., 2013. The role of *AUTS2* in neurodevelopment and human evolution. *Trends Genet.* 29, 600–608. <https://doi.org/10.1016/j.tig.2013.08.001>
- Oksenberg, N., Haliburton, G.D.E., Eckalbar, W.L., Oren, I., Nishizaki, S., Murphy, K., Pollard, K.S., Birnbaum, R.Y., Ahituv, N., 2014. Genome-wide distribution of *Auts2* binding localizes with active neurodevelopmental genes. *Transl. Psychiatry* 4, e431–e431. <https://doi.org/10.1038/tp.2014.78>
- Oksenberg, N., Stevison, L., Wall, J.D., Ahituv, N., 2013. Function and Regulation of *AUTS2*, a Gene Implicated in Autism and Human Evolution. *PLoS Genet.* 9, e1003221. <https://doi.org/10.1371/journal.pgen.1003221>
- Ortega, N.M., Winblad, N., Plaza Reyes, A., Lanner, F., 2018. Functional genetics of early human development. *Curr. Opin. Genet. Dev.* 52, 1–6. <https://doi.org/10.1016/j.gde.2018.04.005>
- Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., Melton, D.A., 2008. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat. Biotechnol.* 26, 313–315. <https://doi.org/10.1038/nbt1383>

- Osnato, A., Brown, S., Krueger, C., Andrews, S., Collier, A.J., Nakanoh, S., Quiroga Londoño, M., Wesley, B.T., Muraro, D., Brumm, A.S., Niakan, K.K., Vallier, L., Ortmann, D., Rugg-Gunn, P.J., 2021. TGF β signalling is required to maintain pluripotency of human naïve pluripotent stem cells. *eLife* 10, e67259. <https://doi.org/10.7554/eLife.67259>
- Oudet, P., Gross-Bellard, M., Chambon, P., 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4, 281–300. [https://doi.org/10.1016/0092-8674\(75\)90149-X](https://doi.org/10.1016/0092-8674(75)90149-X)
- Panova, A.V., Bogomazova, A.N., Lagarkova, M.A., Kiselev, S.L., 2018. Epigenetic reprogramming by naïve conditions establishes an irreversible state of partial X chromosome reactivation in female stem cells. *Oncotarget* 9, 25136–25147. <https://doi.org/10.18632/oncotarget.25353>
- Papachristou, E.K., Kishore, K., Holding, A.N., Harvey, K., Roumeliotis, T.I., Chilamakuri, C.S.R., Omarjee, S., Chia, K.M., Swarbrick, A., Lim, E., Markowitz, F., Eldridge, M., Siersbaek, R., D'Santos, C.S., Carroll, J.S., 2018. A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. *Nat. Commun.* 9, 2311. <https://doi.org/10.1038/s41467-018-04619-5>
- Papp, B., Plath, K., 2011. Reprogramming to pluripotency: stepwise resetting of the epigenetic landscape. *Cell Res.* 21, 486–501. <https://doi.org/10.1038/cr.2011.28>
- Park, I.-H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., Daley, G.Q., 2008. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451, 141–146. <https://doi.org/10.1038/nature06534>
- Paro, R., Hogness, D.S., 1991. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci.* 88, 263–267. <https://doi.org/10.1073/pnas.88.1.263>
- Parolini, O., Soncini, M., Evangelista, M., Schmidt, D., 2009. Amniotic membrane and amniotic fluid-derived cells: potential tools for regenerative medicine? *Regen. Med.* 4, 275–291. <https://doi.org/10.2217/17460751.4.2.275>
- Pasini, D., Bracken, A.P., Jensen, M.R., Denchi, E.L., Helin, K., 2004. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* 23, 4061–4071. <https://doi.org/10.1038/sj.emboj.7600402>
- Pastor, W.A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S.E., Clark, A.T., 2016. Naïve Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* 18, 323–329. <https://doi.org/10.1016/j.stem.2016.01.019>
- Pauklin, S., Vallier, L., 2015. Activin/Nodal signalling in stem cells. *Development* 142, 607–619. <https://doi.org/10.1242/dev.091769>
- Paulson, J.R., Hudson, D.F., Cisneros-Soberanis, F., Earnshaw, W.C., 2021. Mitotic chromosomes. *Semin. Cell Dev. Biol.* 117, 7–29. <https://doi.org/10.1016/j.semcdb.2021.03.014>
- Pengelly, A.R., Kalb, R., Finkl, K., Müller, J., 2015. Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes Dev.* 29, 1487–1492. <https://doi.org/10.1101/gad.265439.115>
- Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., Brockdorff, N., 1996. Requirement for Xist in X chromosome inactivation. *Nature* 379, 131–137. <https://doi.org/10.1038/379131a0>
- Petropoulos, S., Edsgård, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., Lanner, F., 2016. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* 165, 1012–1026. <https://doi.org/10.1016/j.cell.2016.03.023>
- Pickart, C.M., 2001. Mechanisms Underlying Ubiquitination. *Annu. Rev. Biochem.* 70, 503–533. <https://doi.org/10.1146/annurev.biochem.70.1.503>
- Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B.A., Solcan, N., Cerase, A., Castello, A., Mohammed, S., Moindrot, B., Nesterova, T.B., Brockdorff, N., 2017. hnRNPK

- Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Mol. Cell* 68, 955-969.e10.
<https://doi.org/10.1016/j.molcel.2017.11.013>
- Piunti, A., Shilatifard, A., 2022. Author Correction: The roles of Polycomb repressive complexes in mammalian development and cancer. *Nat. Rev. Mol. Cell Biol.* 23, 444–444. <https://doi.org/10.1038/s41580-022-00495-6>
- Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., Zhang, Y., 2003. Role of Histone H3 Lysine 27 Methylation in X Inactivation. *Science* 300, 131–135.
<https://doi.org/10.1126/science.1084274>
- Pollheimer, J., Vondra, S., Baltayeva, J., Beristain, A.G., Knöfler, M., 2018. Regulation of Placental Extravillous Trophoblasts by the Maternal Uterine Environment. *Front. Immunol.* 9, 2597. <https://doi.org/10.3389/fimmu.2018.02597>
- Prakash, S., Robbins, P.W., 1998. Cloning and Analysis of the cDNA for Human Fibrosin, a Novel Fibrogenic Lymphokine. *DNA Cell Biol.* 17, 879–884.
<https://doi.org/10.1089/dna.1998.17.879>
- Prokhorova, T.A., Harkness, L.M., Frandsen, U., Ditzel, N., Schröder, H.D., Burns, J.S., Kassem, M., 2009. Teratoma Formation by Human Embryonic Stem Cells Is Site Dependent and Enhanced by the Presence of Matrigel. *Stem Cells Dev.* 18, 47–54.
<https://doi.org/10.1089/scd.2007.0266>
- Przyborski, S.A., 2005. Differentiation of Human Embryonic Stem Cells After Transplantation in Immune-Deficient Mice. *STEM CELLS* 23, 1242–1250.
<https://doi.org/10.1634/stemcells.2005-0014>
- Qin, H., Hejna, M., Liu, Y., Percharde, M., Wossidlo, M., Blouin, L., Durruthy-Durruthy, J., Wong, P., Qi, Z., Yu, J., Qi, L.S., Sebastiano, V., Song, J.S., Ramalho-Santos, M., 2016. YAP Induces Human Naïve Pluripotency. *Cell Rep.* 14, 2301–2312.
<https://doi.org/10.1016/j.celrep.2016.02.036>
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., Manke, T., 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165.
<https://doi.org/10.1093/nar/gkw257>
- Reijnen, M.J., Hamer, K.M., den Blaauwen, J.L., Lambrechts, C., Schoneveld, I., van Driel, R., Otte, A.P., 1995. Polycomb and bmi-1 homologs are expressed in overlapping patterns in *Xenopus* embryos and are able to interact with each other. *Mech. Dev.* 53, 35–46. [https://doi.org/10.1016/0925-4773\(95\)00422-X](https://doi.org/10.1016/0925-4773(95)00422-X)
- Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.-C., Korving, J., Vivié, J., Truckenmüller, R.K., van Oudenaarden, A., van Blitterswijk, C.A., Geijsen, N., 2018. Blastocyst-like structures generated solely from stem cells. *Nature* 557, 106–111.
<https://doi.org/10.1038/s41586-018-0051-0>
- Robzyk, K., 2000. Rad6-Dependent Ubiquitination of Histone H2B in Yeast. *Science* 287, 501–504. <https://doi.org/10.1126/science.287.5452.501>
- Rose, N.R., King, H.W., Blackledge, N.P., Fursova, N.A., Ember, K.J., Fischer, R., Kessler, B.M., Klose, R.J., 2016. RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. *eLife* 5, e18591.
<https://doi.org/10.7554/eLife.18591>
- Ross, C., Boroviak, T.E., 2020. Origin and function of the yolk sac in primate embryogenesis. *Nat. Commun.* 11, 3760. <https://doi.org/10.1038/s41467-020-17575-w>
- Rossant, J., Tam, P.P.L., 2017. New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. *Cell Stem Cell* 20, 18–28.
<https://doi.org/10.1016/j.stem.2016.12.004>
- Rostovskaya, M., Andrews, S., Reik, W., Rugg-Gunn, P.J., 2022. Amniogenesis occurs in two independent waves in primates. *Cell Stem Cell* 29, 744-759.e6.
<https://doi.org/10.1016/j.stem.2022.03.014>

- Rostovskaya, M., Stirparo, G.G., Smith, A., 2019. Capacitation of human naïve pluripotent stem cells for multi-lineage differentiation. *Development* 146, dev172916. <https://doi.org/10.1242/dev.172916>
- Rugg-Gunn, P.J., 2022a. Flow Cytometry Analysis of to Identify Human Naïve Pluripotent Stem Cells, in: Rugg-Gunn, P. (Ed.), *Human Naïve Pluripotent Stem Cells, Methods in Molecular Biology*. Springer US, New York, NY, pp. 257–265. https://doi.org/10.1007/978-1-0716-1908-7_16
- Rugg-Gunn, P.J., 2022b. Induction of Human Naïve Pluripotency Using, in: Rugg-Gunn, P. (Ed.), *Human Naïve Pluripotent Stem Cells, Methods in Molecular Biology*. Springer US, New York, NY, pp. 29–37. https://doi.org/10.1007/978-1-0716-1908-7_3
- Ruiz-Cañada, C., Bernabé-García, Á., Liarte, S., Rodríguez-Valiente, M., Nicolás, F.J., 2021. Chronic Wound Healing by Amniotic Membrane: TGF- β and EGF Signaling Modulation in Re-epithelialization. *Front. Bioeng. Biotechnol.* 9, 689328. <https://doi.org/10.3389/fbioe.2021.689328>
- Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E., Langerman, J., Clark, A.T., Jaenisch, R., Plath, K., 2017. Human Naïve Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell* 20, 87–101. <https://doi.org/10.1016/j.stem.2016.10.006>
- Sánchez-Herrero, E., Vernós, I., Marco, R., Morata, G., 1985. Genetic organization of *Drosophila bithorax* complex. *Nature* 313, 108–113. <https://doi.org/10.1038/313108a0>
- Sankar, A., Mohammad, F., Sundaramurthy, A.K., Wang, H., Lerdrup, M., Tatar, T., Helin, K., 2022. Histone editing elucidates the functional roles of H3K27 methylation and acetylation in mammals. *Nat. Genet.* 54, 754–760. <https://doi.org/10.1038/s41588-022-01091-2>
- Sarel-Gallily, R., Benvenisty, N., 2022. Large-Scale Analysis of X Inactivation Variations between Primed and Naïve Human Embryonic Stem Cells. *Cells* 11, 1729. <https://doi.org/10.3390/cells11111729>
- Satijn, D.P., Gunster, M.J., van der Vlag, J., Hamer, K.M., Schul, W., Alkema, M.J., Saurin, A.J., Freemont, P.S., van Driel, R., Otte, A.P., 1997. RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. *Mol. Cell. Biol.* 17, 4105–4113. <https://doi.org/10.1128/MCB.17.7.4105>
- Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., Brivanlou, A.H., 2004. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat. Med.* 10, 55–63. <https://doi.org/10.1038/nm979>
- Saurin, A.J., Shao, Z., Erdjument-Bromage, H., Tempst, P., Kingston, R.E., 2001. A *Drosophila* Polycomb group complex includes Zeste and dTAFII proteins. *Nature* 412, 655–660. <https://doi.org/10.1038/35088096>
- Scelfo, A., Fernández-Pérez, D., Tamburri, S., Zanotti, M., Lavarone, E., Soldi, M., Bonaldi, T., Ferrari, K.J., Pasini, D., 2019. Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities. *Mol. Cell* 74, 1037–1052.e7. <https://doi.org/10.1016/j.molcel.2019.04.002>
- Schmöle, A.-C., Brennfürher, A., Karapetyan, G., Jaster, R., Pews-Davtyan, A., Hübner, R., Ortinau, S., Beller, M., Rolfs, A., Frech, M.J., 2010. Novel indolylmaleimide acts as GSK-3 β inhibitor in human neural progenitor cells. *Bioorg. Med. Chem.* 18, 6785–6795. <https://doi.org/10.1016/j.bmc.2010.07.045>
- Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., Wutz, A., 2006. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* 25, 3110–3122. <https://doi.org/10.1038/sj.emboj.7601187>
- Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.-M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., Segonds-Pichon, A., Jurkowski, W., Wingett, S.W., Tabbada, K., Andrews, S., Herman, B., LeProust, E., Osborne, C.S., Koseki, H., Fraser, P., Luscombe, N.M., Elderkin, S.,

2015. Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* 47, 1179–1186. <https://doi.org/10.1038/ng.3393>
- Schuettengruber, B., Bourbon, H.-M., Di Croce, L., Cavalli, G., 2017. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* 171, 34–57. <https://doi.org/10.1016/j.cell.2017.08.002>
- Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D.A., Benvenisty, N., 2000. Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. *Proc. Natl. Acad. Sci.* 97, 11307–11312. <https://doi.org/10.1073/pnas.97.21.11307>
- Schumacher, A., Faust, C., Magnuson, T., 1996. Positional cloning of a global regulator of anterior–posterior patterning in mice. *Nature* 383, 250–253. <https://doi.org/10.1038/383250a0>
- Sebolt-Leopold, J.S., Herrera, R., 2004. Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nat. Rev. Cancer* 4, 937–947. <https://doi.org/10.1038/nrc1503>
- Sellers, R.A., Robertson, D.L., Tassabehji, M., 2020. Ancestry of the AUTS2 family—A novel group of polycomb-complex proteins involved in human neurological disease. *PLOS ONE* 15, e0232101. <https://doi.org/10.1371/journal.pone.0232101>
- Sewalt, R.G.A.B., van der Vlag, J., Gunster, M.J., Hamer, K.M., den Blaauwen, J.L., Satijn, D.P.E., Hendrix, T., van Driel, R., Otte, A.P., 1998. Characterization of Interactions between the Mammalian Polycomb-Group Proteins Enx1/EZH2 and EED Suggests the Existence of Different Mammalian Polycomb-Group Protein Complexes. *Mol. Cell. Biol.* 18, 3586–3595. <https://doi.org/10.1128/MCB.18.6.3586>
- Seydoux, G., Braun, R.E., 2006. Pathway to Totipotency: Lessons from Germ Cells. *Cell* 127, 891–904. <https://doi.org/10.1016/j.cell.2006.11.016>
- Shakiba, N., White, C.A., Lipsitz, Y.Y., Yachie-Kinoshita, A., Tonge, P.D., Hussein, S.M.I., Puri, M.C., Elbaz, J., Morrissey-Scoot, J., Li, M., Munoz, J., Benevento, M., Rogers, I.M., Hanna, J.H., Heck, A.J.R., Wollscheid, B., Nagy, A., Zandstra, P.W., 2015. CD24 tracks divergent pluripotent states in mouse and human cells. *Nat. Commun.* 6, 7329. <https://doi.org/10.1038/ncomms8329>
- Shan, Y., Liang, Z., Xing, Q., Zhang, T., Wang, B., Tian, S., Huang, W., Zhang, Y., Yao, J., Zhu, Y., Huang, K., Liu, Y., Wang, X., Chen, Q., Zhang, J., Shang, B., Li, S., Shi, X., Liao, B., Zhang, C., Lai, K., Zhong, X., Shu, X., Wang, J., Yao, H., Chen, J., Pei, D., Pan, G., 2017. PRC2 specifies ectoderm lineages and maintains pluripotency in primed but not naïve ESCs. *Nat. Commun.* 8, 672. <https://doi.org/10.1038/s41467-017-00668-4>
- Shan, Y., Ma, L., Zhang, C., Zhang, Y., Zhang, J., Xiao, L., Wei, Y., Yu, Y., 2020. Generation of two RNF2 homozygous knockout human embryonic stem cell lines by CRISPR/Cas9 system. *Stem Cell Res.* 47, 101885. <https://doi.org/10.1016/j.scr.2020.101885>
- Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C., Bender, W., Kingston, R.E., 1999. Stabilization of Chromatin Structure by PRC1, a Polycomb Complex. *Cell* 98, 37–46. [https://doi.org/10.1016/S0092-8674\(00\)80604-2](https://doi.org/10.1016/S0092-8674(00)80604-2)
- Shaytan, A.K., Landsman, D., Panchenko, A.R., 2015. Nucleosome adaptability conferred by sequence and structural variations in histone H2A–H2B dimers. *Curr. Opin. Struct. Biol.* 32, 48–57. <https://doi.org/10.1016/j.sbi.2015.02.004>
- Shukla, S., Ying, W., Gray, F., Yao, Y., Simes, M.L., Zhao, Q., Miao, H., Cho, H.J., González-Alonso, P., Winkler, A., Lund, G., Purohit, T., Kim, E., Zhang, X., Ray, J.M., He, S., Nikolaidis, C., Ndoj, J., Wang, J., Jaremko, Ł., Jaremko, M., Ryan, R.J.H., Guzman, M.L., Grembecka, J., Cierpicki, T., 2021. Small-molecule inhibitors targeting Polycomb repressive complex 1 RING domain. *Nat. Chem. Biol.* 17, 784–793. <https://doi.org/10.1038/s41589-021-00815-5>
- Sierra, R.A., Hoverter, N.P., Ramirez, R.N., Vuong, L.M., Mortazavi, A., Merrill, B.J., Waterman, M.L., Donovan, P.J., 2018. *TCF7L1* suppresses primitive streak gene

- expression to support human embryonic stem cell pluripotency. *Development* 161, 161075. <https://doi.org/10.1242/dev.161075>
- Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., Smith, A., 2008. Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition. *PLoS Biol.* 6, e253. <https://doi.org/10.1371/journal.pbio.0060253>
- Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H.F.M., Jenuwein, T., Otte, A.P., Brockdorff, N., 2003. Establishment of Histone H3 Methylation on the Inactive X Chromosome Requires Transient Recruitment of Eed-Enx1 Polycomb Group Complexes. *Dev. Cell* 4, 481–495. [https://doi.org/10.1016/S1534-5807\(03\)00068-6](https://doi.org/10.1016/S1534-5807(03)00068-6)
- Singh, P.P., Arora, J., Isambert, H., 2015. Identification of Ohnolog Genes Originating from Whole Genome Duplication in Early Vertebrates, Based on Synteny Comparison across Multiple Genomes. *PLOS Comput. Biol.* 11, e1004394. <https://doi.org/10.1371/journal.pcbi.1004394>
- Skene, P.J., Henikoff, S., 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* 6, e21856. <https://doi.org/10.7554/eLife.21856>
- Smith, Z.D., Meissner, A., 2013. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* 14, 204–220. <https://doi.org/10.1038/nrg3354>
- Soncin, F., Morey, R., Bui, T., Requena, D.F., Cheung, V.C., Kallol, S., Kittle, R., Jackson, M.G., Farah, O., Dumdie, J., Meads, M., Pizzo, D., Horii, M., Fisch, K.M., Parast, M.M., 2022. Derivation of functional trophoblast stem cells from primed human pluripotent stem cells. *Stem Cell Rep.* 17, 1303–1317. <https://doi.org/10.1016/j.stemcr.2022.04.013>
- Stewart-Morgan, K.R., Petryk, N., Groth, A., 2020. Chromatin replication and epigenetic cell memory. *Nat. Cell Biol.* 22, 361–371. <https://doi.org/10.1038/s41556-020-0487-y>
- Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., Bertone, P., 2018. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* 145, dev158501. <https://doi.org/10.1242/dev.158501>
- Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., Pombo, A., 2007. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat. Cell Biol.* 9, 1428–1435. <https://doi.org/10.1038/ncb1663>
- Struhl, G., Akam, M., 1985. Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of *Drosophila*. *EMBO J.* 4, 3259–3264. <https://doi.org/10.1002/j.1460-2075.1985.tb04075.x>
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* 102, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>
- Sultana, R., Yu, C.-E., Yu, J., Munson, J., Chen, D., Hua, W., Estes, A., Cortes, F., de la Barra, F., Yu, D., Haider, S.T., Trask, B.J., Green, E.D., Raskind, W.H., Distech, C.M., Wijsman, E., Dawson, G., Storm, D.R., Schellenberg, G.D., Villacres, E.C., 2002. Identification of a Novel Gene on Chromosome 7q11.2 Interrupted by a Translocation Breakpoint in a Pair of Autistic Twins. *Genomics* 80, 129–134. <https://doi.org/10.1006/geno.2002.6810>
- Surmacz, B., Noisa, P., Risner-Janiczek, J.R., Hui, K., Ungless, M., Cui, W., Li, M., 2012. DLK1 Promotes Neurogenesis of Human and Mouse Pluripotent Stem Cell-Derived Neural Progenitors Via Modulating Notch and BMP Signalling. *Stem Cell Rev. Rep.* 8, 459–471. <https://doi.org/10.1007/s12015-011-9298-7>
- Sybirna, A., Tang, W.W.C., Pierson Smela, M., Dietmann, S., Gruhn, W.H., Brosh, R., Surani, M.A., 2020. A critical role of PRDM14 in human primordial germ cell fate revealed by inducible degrons. *Nat. Commun.* 11, 1282. <https://doi.org/10.1038/s41467-020-15042-0>

- Szczerbinska, I., Gonzales, K.A.U., Cukuroglu, E., Ramli, M.N.B., Lee, B.P.G., Tan, C.P., Wong, C.K., Rancati, G.I., Liang, H., Göke, J., Ng, H.-H., Chan, Y.-S., 2019. A Chemically Defined Feeder-free System for the Establishment and Maintenance of the Human Naive Pluripotent State. *Stem Cell Rep.* 13, 612–626. <https://doi.org/10.1016/j.stemcr.2019.08.005>
- T. Das, A., Tenenbaum, L., Berkhout, B., 2016. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr. Gene Ther.* 16, 156–167. <https://doi.org/10.2174/1566523216666160524144041>
- Taherbhoy, A.M., Huang, O.W., Cochran, A.G., 2015. BMI1–RING1B is an autoinhibited RING E3 ubiquitin ligase. *Nat. Commun.* 6, 7621. <https://doi.org/10.1038/ncomms8621>
- Takahashi, S., Kobayashi, S., Hiratani, I., 2018. Epigenetic differences between naïve and primed pluripotent stem cells. *Cell. Mol. Life Sci.* 75, 1191–1203. <https://doi.org/10.1007/s00018-017-2703-x>
- Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., Reik, W., Bertone, P., Smith, A., 2014. Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell* 158, 1254–1269. <https://doi.org/10.1016/j.cell.2014.08.029>
- Tamburri, S., Lavarone, E., Fernández-Pérez, D., Conway, E., Zanotti, M., Manganaro, D., Pasini, D., 2020. Histone H2AK119 Mono-Ubiquitination Is Essential for Polycomb-Mediated Transcriptional Repression. *Mol. Cell* 77, 840–856.e5. <https://doi.org/10.1016/j.molcel.2019.11.021>
- Tanay, A., O'Donnell, A.H., Damelin, M., Bestor, T.H., 2007. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc. Natl. Acad. Sci.* 104, 5521–5526. <https://doi.org/10.1073/pnas.0609746104>
- Taubenschmid-Stowers, J., Rostovskaya, M., Santos, F., Ljung, S., Argelaguet, R., Krueger, F., Nichols, J., Reik, W., 2022. 8C-like cells capture the human zygotic genome activation program in vitro. *Cell Stem Cell* 29, 449–459.e6. <https://doi.org/10.1016/j.stem.2022.01.014>
- Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., Wutz, A., Vidal, M., Elderkin, S., Brockdorff, N., 2012. RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. *Cell* 148, 664–678. <https://doi.org/10.1016/j.cell.2011.12.029>
- Taylor, G.C.A., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M.M., Bickmore, W.A., 2013. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.* 23, 2053–2065. <https://doi.org/10.1101/gr.155028.113>
- Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., McKay, R.D.G., 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448, 196–199. <https://doi.org/10.1038/nature05972>
- The International Stem Cell Initiative, 2018. Assessment of established techniques to determine developmental and malignant potential of human pluripotent stem cells. *Nat. Commun.* 9, 1925. <https://doi.org/10.1038/s41467-018-04011-3>
- The International Stem Cell Initiative, 2007. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat. Biotechnol.* 25, 803–816. <https://doi.org/10.1038/nbt1318>
- The Milieu Intérieur Consortium, Kilens, S., Meistermann, D., Moreno, D., Chariou, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., Flippe, L., Firmin, J., Song, J., Charpentier, E., Lammers, J., Donnart, A., Marec, N., Deb, W., Bihouée, A., Le Caignec, C., Pecqueur, C., Redon, R., Barrière, P., Bourdon, J., Pasque, V., Soumillon, M., Mikkelsen, T.S., Rougeulle, C., Fréour, T., David, L., 2018a. Parallel derivation of isogenic human primed and naive induced

- pluripotent stem cells. *Nat. Commun.* 9, 360. <https://doi.org/10.1038/s41467-017-02107-w>
- The Milieu Intérieur Consortium, Kilens, S., Meistermann, D., Moreno, D., Chariou, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., Flippe, L., Firmin, J., Song, J., Charpentier, E., Lammers, J., Donnart, A., Marec, N., Deb, W., Bihouée, A., Le Caignec, C., Pecqueur, C., Redon, R., Barrière, P., Bourdon, J., Pasque, V., Soumillon, M., Mikkelsen, T.S., Rougeulle, C., Fréour, T., David, L., 2018b. Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat. Commun.* 9, 360. <https://doi.org/10.1038/s41467-017-02107-w>
- Theunissen, T.W., Friedli, M., He, Y., Planet, E., O’Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., Duc, J., Cohen, M.A., Wert, K.J., Castanon, R., Zhang, Z., Huang, Y., Nery, J.R., Drotar, J., Lungjangwa, T., Trono, D., Ecker, J.R., Jaenisch, R., 2016. Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* 19, 502–515. <https://doi.org/10.1016/j.stem.2016.06.011>
- Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., Lungjangwa, T., Imsoonthornruksa, S., Stelzer, Y., Rangarajan, S., D’Alessio, A., Zhang, J., Gao, Q., Dawlaty, M.M., Young, R.A., Gray, N.S., Jaenisch, R., 2014. Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell* 15, 471–487. <https://doi.org/10.1016/j.stem.2014.07.002>
- Thomson, J.A., 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* 282, 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., Jones, J.M., 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* 282, 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>
- Tian, B., Yang, J., Brasier, A.R., 2012. Two-Step Cross-linking for Analysis of Protein–Chromatin Interactions, in: Vancura, A. (Ed.), *Transcriptional Regulation, Methods in Molecular Biology*. Springer New York, New York, NY, pp. 105–120. https://doi.org/10.1007/978-1-61779-376-9_7
- Todd, M., Ivanochko, D., Picketts, D., 2015. PHF6 Degrees of Separation: The Multifaceted Roles of a Chromatin Adaptor Protein. *Genes* 6, 325–352. <https://doi.org/10.3390/genes6020325>
- Toh, C.-X.D., Chan, J.-W., Chong, Z.-S., Wang, H.F., Guo, H.C., Satapathy, S., Ma, D., Goh, G.Y.L., Khattar, E., Yang, L., Tergaonkar, V., Chang, Y.-T., Collins, J.J., Daley, G.Q., Wee, K.B., Farran, C.A.E., Li, H., Lim, Y.-P., Bard, F.A., Loh, Y.-H., 2016. RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming. *Cell Rep.* 15, 2597–2607. <https://doi.org/10.1016/j.celrep.2016.05.049>
- Tsuboi, M., Kishi, Y., Yokozeki, W., Koseki, H., Hirabayashi, Y., Gotoh, Y., 2018. Ubiquitination-Independent Repression of PRC1 Targets during Neuronal Fate Restriction in the Developing Mouse Neocortex. *Dev. Cell* 47, 758–772.e5. <https://doi.org/10.1016/j.devcel.2018.11.018>
- Tsuneyoshi, N., Sumi, T., Onda, H., Nojima, H., Nakatsuji, N., Suemori, H., 2008. PRDM14 suppresses expression of differentiation marker genes in human embryonic stem cells. *Biochem. Biophys. Res. Commun.* 367, 899–905. <https://doi.org/10.1016/j.bbrc.2007.12.189>
- Tucker, K.L., Wang, Y., Dausman, J., Jaenisch, R., 1997. A transgenic mouse strain expressing four drug-selectable marker genes. *Nucleic Acids Res.* 25, 3745–3746. <https://doi.org/10.1093/nar/25.18.3745>
- Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., Srinivas, S., 2021. Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* 600, 285–289. <https://doi.org/10.1038/s41586-021-04158-y>
- Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W.B., Cho, C.H.-H., Martinez, A., Rugg-Gunn, P., Brons, G., Pedersen, R.A., 2009a.

- Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* 136, 1339–1349. <https://doi.org/10.1242/dev.033951>
- Vallier, L., Reynolds, D., Pedersen, R.A., 2004. Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. *Dev. Biol.* 275, 403–421. <https://doi.org/10.1016/j.ydbio.2004.08.031>
- Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter, M., Rugg-Gunn, P., Weber, A., Pedersen, R.A., 2009b. Early Cell Fate Decisions of Human Embryonic Stem Cells and Mouse Epiblast Stem Cells Are Controlled by the Same Signalling Pathways. *PLoS ONE* 4, e6082. <https://doi.org/10.1371/journal.pone.0006082>
- Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur-Griscelli, A., Duret, L., Rougeulle, C., 2013. XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. *Nat. Genet.* 45, 239–241. <https://doi.org/10.1038/ng.2530>
- Vallot, C., Ouimette, J.-F., Makhoulouf, M., Féraud, O., Pontis, J., Côme, J., Martinat, C., Bennaceur-Griscelli, A., Lalande, M., Rougeulle, C., 2015. Erosion of X Chromosome Inactivation in Human Pluripotent Cells Initiates with XACT Coating and Depends on a Specific Heterochromatin Landscape. *Cell Stem Cell* 16, 533–546. <https://doi.org/10.1016/j.stem.2015.03.016>
- Vallot, C., Patrat, C., Collier, A.J., Huret, C., Casanova, M., Liyakat Ali, T.M., Tosolini, M., Frydman, N., Heard, E., Rugg-Gunn, P.J., Rougeulle, C., 2017. XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell* 20, 102–111. <https://doi.org/10.1016/j.stem.2016.10.014>
- van Mierlo, G., Dirks, R.A.M., De Clerck, L., Brinkman, A.B., Huth, M., Kloet, S.L., Saksouk, N., Kroeze, L.I., Willems, S., Farlik, M., Bock, C., Jansen, J.H., Deforce, D., Vermeulen, M., Déjardin, J., Dhaenens, M., Marks, H., 2019a. Integrative Proteomic Profiling Reveals PRC2-Dependent Epigenetic Crosstalk Maintains Ground-State Pluripotency. *Cell Stem Cell* 24, 123–137.e8. <https://doi.org/10.1016/j.stem.2018.10.017>
- van Mierlo, G., Veenstra, G.J.C., Vermeulen, M., Marks, H., 2019b. The Complexity of PRC2 Subcomplexes. *Trends Cell Biol.* 29, 660–671. <https://doi.org/10.1016/j.tcb.2019.05.004>
- Verhelst, S., De Clerck, L., Willems, S., Van Puyvelde, B., Daled, S., Deforce, D., Dhaenens, M., 2020. Comprehensive histone epigenetics: A mass spectrometry based screening assay to measure epigenetic toxicity. *MethodsX* 7, 101055. <https://doi.org/10.1016/j.mex.2020.101055>
- Vilar, M., Murillo-Carretero, M., Mira, H., Magnusson, K., Besset, V., Ibáñez, C.F., 2006. Bex1, a novel interactor of the p75 neurotrophin receptor, links neurotrophin signaling to the cell cycle. *EMBO J.* 25, 1219–1230. <https://doi.org/10.1038/sj.emboj.7601017>
- Vitale, A.M., Matigian, N.A., Ravishankar, S., Bellette, B., Wood, S.A., Wolvetang, E.J., Mackay-Sim, A., 2012. Variability in the Generation of Induced Pluripotent Stem Cells: Importance for Disease Modeling. *Stem Cells Transl. Med.* 1, 641–650. <https://doi.org/10.5966/sctm.2012-0043>
- Viukov, S., Shani, T., Bayerl, J., Sheban, D., Stelzer, Y., Novershtern, N., Hanna, J., 2022. Human primed and naïve PSCs are both competent in differentiating into bona fide trophoblast stem cells (preprint). *Developmental Biology*. <https://doi.org/10.1101/2022.05.20.492766>
- Voncken, J.W., Roelen, B.A.J., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., van Lohuizen, M., 2003. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc. Natl. Acad. Sci.* 100, 2468–2473. <https://doi.org/10.1073/pnas.0434312100>
- Wang, D., Yi, R., 2019. Is it time to take R(epressive) out of PRC1? *Genes Dev.* 33, 4–5. <https://doi.org/10.1101/gad.322651.118>

- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., Zhang, Y., 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873–878. <https://doi.org/10.1038/nature02985>
- Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., Magnuson, T., 2001. Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat. Genet.* 28, 371–375. <https://doi.org/10.1038/ng574>
- Wang, L., Brown, J.L., Cao, R., Zhang, Y., Kassiss, J.A., Jones, R.S., 2004. Hierarchical Recruitment of Polycomb Group Silencing Complexes. *Mol. Cell* 14, 637–646. <https://doi.org/10.1016/j.molcel.2004.05.009>
- Wang, Q., Geng, Z., Gong, Y., Warren, K., Zheng, H., Imamura, Y., Gao, Z., 2018. WDR68 is essential for the transcriptional activation of the PRC1-AUTS2 complex and neuronal differentiation of mouse embryonic stem cells. *Stem Cell Res.* 33, 206–214. <https://doi.org/10.1016/j.scr.2018.10.023>
- Wang, R., Taylor, A.B., Leal, B.Z., Chadwell, L.V., Ilangoan, U., Robinson, A.K., Schirf, V., Hart, P.J., Lafer, E.M., Demeler, B., Hinck, A.P., McEwen, D.G., Kim, C.A., 2010. Polycomb Group Targeting through Different Binding Partners of RING1B C-Terminal Domain. *Structure* 18, 966–975. <https://doi.org/10.1016/j.str.2010.04.013>
- Wang, S.-H., Hao, J., Zhang, C., Duan, F.-F., Chiu, Y.-T., Shi, M., Huang, X., Yang, J., Cao, H., Wang, Y., 2022. KLF17 promotes human naive pluripotency through repressing MAPK3 and ZIC2. *Sci. China Life Sci.* <https://doi.org/10.1007/s11427-021-2076-x>
- Wang, Xiaohu, Wang, Xiaoxiao, Zhang, S., Sun, H., Li, S., Ding, H., You, Y., Zhang, X., Ye, S.-D., 2019. The transcription factor TFCEP2L1 induces expression of distinct target genes and promotes self-renewal of mouse and human embryonic stem cells. *J. Biol. Chem.* 294, 6007–6016. <https://doi.org/10.1074/jbc.RA118.006341>
- Wang, Z., Oron, E., Nelson, B., Razis, S., Ivanova, N., 2012. Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells. *Cell Stem Cell* 10, 440–454. <https://doi.org/10.1016/j.stem.2012.02.016>
- Ware, C.B., 2017. Concise Review: Lessons from Naïve Human Pluripotent Cells. *Stem Cells* 35, 35–41. <https://doi.org/10.1002/stem.2507>
- Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., Tesar, P.J., Okada, J., Margaretha, L., Sperber, H., Choi, M., Blau, C.A., Treuting, P.M., Hawkins, R.D., Cirulli, V., Ruohola-Baker, H., 2014. Derivation of naïve human embryonic stem cells. *Proc. Natl. Acad. Sci.* 111, 4484–4489. <https://doi.org/10.1073/pnas.1319738111>
- Warmflash, A., Sorre, B., Etoc, F., Siggia, E.D., Brivanlou, A.H., 2014. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat. Methods* 11, 847–854. <https://doi.org/10.1038/nmeth.3016>
- Warrier, S., Van der Jeught, M., Duggal, G., Tilleman, L., Sutherland, E., Taelman, J., Popovic, M., Lierman, S., Chuva De Sousa Lopes, S., Van Soom, A., Peelman, L., Van Nieuwerburgh, F., De Coninck, D.I.M., Menten, B., Mestdagh, P., Van de Sompele, J., Deforce, D., De Sutter, P., Heindryckx, B., 2017. Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nat. Commun.* 8, 15055. <https://doi.org/10.1038/ncomms15055>
- Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, Satomi, Nishikawa, Shin-ichi, Muguruma, K., Sasai, Y., 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* 25, 681–686. <https://doi.org/10.1038/nbt1310>
- Wei, Y., Wang, T., Ma, L., Zhang, Y., Zhao, Y., Lye, K., Xiao, L., Chen, C., Wang, Z., Ma, Y., Zhou, X., Sun, F., Li, W., Dunk, C., Li, S., Nagy, A., Yu, Y., Pan, G., Lye, S.J., Shan, Y., 2021. Efficient derivation of human trophoblast stem cells from primed pluripotent stem cells. *Sci. Adv.* 7, eabf4416. <https://doi.org/10.1126/sciadv.abf4416>
- Weinberger, L., Ayyash, M., Novershtern, N., Hanna, J.H., 2016. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* 17, 155–169. <https://doi.org/10.1038/nrm.2015.28>

- White, C.R., MacDonald, W.A., Mann, M.R.W., 2016. Conservation of DNA Methylation Programming Between Mouse and Human Gametes and Preimplantation Embryos. *Biol. Reprod.* 95, 61–61. <https://doi.org/10.1095/biolreprod.116.140319>
- Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., Gough, N.M., 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* 336, 684–687. <https://doi.org/10.1038/336684a0>
- Wobus, A.M., Boheler, K.R., 2005. Embryonic Stem Cells: Prospects for Developmental Biology and Cell Therapy. *Physiol. Rev.* 85, 635–678. <https://doi.org/10.1152/physrev.00054.2003>
- Wojdyla, K., Collier, A.J., Fabian, C., Nisi, P.S., Biggins, L., Oxley, D., Rugg-Gunn, P.J., 2020. Cell-Surface Proteomics Identifies Differences in Signaling and Adhesion Protein Expression between Naive and Primed Human Pluripotent Stem Cells. *Stem Cell Rep.* 14, 972–988. <https://doi.org/10.1016/j.stemcr.2020.03.017>
- Wray, J., Kalkan, T., Smith, A.G., 2010. The ground state of pluripotency. *Biochem. Soc. Trans.* 38, 1027–1032. <https://doi.org/10.1042/BST0381027>
- Xia, W., Xu, J., Yu, G., Yao, G., Xu, K., Ma, X., Zhang, N., Liu, B., Li, T., Lin, Z., Chen, X., Li, L., Wang, Q., Shi, D., Shi, S., Zhang, Y., Song, W., Jin, H., Hu, L., Bu, Z., Wang, Y., Na, J., Xie, W., Sun, Y.-P., 2019. Resetting histone modifications during human parental-to-zygotic transition. *Science* 365, 353–360. <https://doi.org/10.1126/science.aaw5118>
- Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K., He, J., Ren, S., Wu, D., Bai, Y., Shang, Z., Dai, X., Ji, W., Li, T., 2020. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* 577, 537–542. <https://doi.org/10.1038/s41586-019-1875-y>
- Xu, J., Shao, Z., Li, D., Xie, H., Kim, W., Huang, J., Taylor, J.E., Pinello, L., Glass, K., Jaffe, J.D., Yuan, G.-C., Orkin, S.H., 2015. Developmental Control of Polycomb Subunit Composition by GATA Factors Mediates a Switch to Non-Canonical Functions. *Mol. Cell* 57, 304–316. <https://doi.org/10.1016/j.molcel.2014.12.009>
- Yakushiji-Kaminatsui, N., Kondo, T., Endo, T.A., Koseki, Y., Kondo, K., Ohara, O., Vidal, M., Koseki, H., 2016. RING1 proteins contribute to early proximal-distal specification of the forelimb bud by restricting *Meis2* expression. *Development* 143, 276–285. <https://doi.org/10.1242/dev.127506>
- Yakushiji-Kaminatsui, N., Kondo, T., Hironaka, K., Sharif, J., Endo, T.A., Nakayama, M., Masui, O., Koseki, Y., Kondo, K., Ohara, O., Vidal, M., Morishita, Y., Koseki, H., 2018. Variant PRC1 competes with retinoic acid-related signals to repress *Meis2* in the mouse distal forelimb bud. *Development* 145, dev166348. <https://doi.org/10.1242/dev.166348>
- Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, Rong, Liu, P., Lian, Y., Zheng, X., Yan, J., Huang, J., Li, M., Wu, X., Wen, L., Lao, K., Li, Ruiqiang, Qiao, J., Tang, F., 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1131–1139. <https://doi.org/10.1038/nsmb.2660>
- Yan, Y., Zhao, W., Huang, Y., Tong, H., Xia, Y., Jiang, Q., Qin, J., 2017. Loss of Polycomb Group Protein Pcgf1 Severely Compromises Proper Differentiation of Embryonic Stem Cells. *Sci. Rep.* 7, 46276. <https://doi.org/10.1038/srep46276>
- Yanagida, A., Spindlow, D., Nichols, J., Dattani, A., Smith, A., Guo, G., 2021. Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* 28, 1016–1022.e4. <https://doi.org/10.1016/j.stem.2021.04.031>
- Yao, C., Choi, E.-A., Weng, L., Xie, X., Wan, J., Xing, Y., Moresco, J.J., Tu, P.G., Yates, J.R., Shi, Y., 2013. Overlapping and distinct functions of CstF64 and CstF64 τ in mammalian mRNA 3' processing. *RNA* 19, 1781–1790. <https://doi.org/10.1261/rna.042317.113>
- Yao, M., Zhou, X., Zhou, J., Gong, S., Hu, G., Li, J., Huang, K., Lai, P., Shi, G., Hutchins, A.P., Sun, H., Wang, H., Yao, H., 2018. PCGF5 is required for neural differentiation

- of embryonic stem cells. *Nat. Commun.* 9, 1463. <https://doi.org/10.1038/s41467-018-03781-0>
- Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., Scholer, H.R., 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* 122, 881–894. <https://doi.org/10.1242/dev.122.3.881>
- Yesbolatova, A., Natsume, T., Hayashi, K., Kanemaki, M.T., 2019. Generation of conditional auxin-inducible degron (AID) cells and tight control of degron-fused proteins using the degradation inhibitor auxinole. *Methods* 164–165, 73–80. <https://doi.org/10.1016/j.ymeth.2019.04.010>
- Yilmaz, A., Braverman-Gross, C., Bialer-Tsypin, A., Peretz, M., Benvenisty, N., 2020. Mapping Gene Circuits Essential for Germ Layer Differentiation via Loss-of-Function Screens in Haploid Human Embryonic Stem Cells. *Cell Stem Cell* 27, 679–691.e6. <https://doi.org/10.1016/j.stem.2020.06.023>
- Ying, Q.-L., Nichols, J., Chambers, I., Smith, A., 2003. BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. *Cell* 115, 281–292. [https://doi.org/10.1016/S0092-8674\(03\)00847-X](https://doi.org/10.1016/S0092-8674(03)00847-X)
- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., Smith, A., 2008. The ground state of embryonic stem cell self-renewal. *Nature* 453, 519–523. <https://doi.org/10.1038/nature06968>
- Yoo, D.H., Im, Y.S., Jo, E.H., Kim, B.Y., Jo, H.-Y., Park, M.-H., Koo, S.K., Kim, Y.-O., 2020. Simple differentiation method using FBS identifies DUSP6 as a marker for fine-tuning of FGF-ERK signaling activity in human pluripotent stem cells. *Biochem. Biophys. Res. Commun.* 521, 375–382. <https://doi.org/10.1016/j.bbrc.2019.10.081>
- Youmans, D.T., Gooding, A.R., Dowell, R.D., Cech, T.R., 2021. Competition between PRC2.1 and 2.2 subcomplexes regulates PRC2 chromatin occupancy in human stem cells. *Mol. Cell* 81, 488–501.e9. <https://doi.org/10.1016/j.molcel.2020.11.044>
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., Thomson, J.A., 2007. Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* 318, 1917–1920. <https://doi.org/10.1126/science.1151526>
- Yu, J.S.L., Yusa, K., 2019. Genome-wide CRISPR-Cas9 screening in mammalian cells. *Methods* 164–165, 29–35. <https://doi.org/10.1016/j.ymeth.2019.04.015>
- Yu, L., Wei, Y., Duan, J., Schmitz, D.A., Sakurai, M., Wang, L., Wang, K., Zhao, S., Hon, G.C., Wu, J., 2021. Blastocyst-like structures generated from human pluripotent stem cells. *Nature* 591, 620–626. <https://doi.org/10.1038/s41586-021-03356-y>
- Zhang, D., Zhou, M., Zhang, Y., Shan, Y., Pan, G., 2022. Generation of an RNF1-deficient human pluripotent stem cell line using CRISPR/Cas9 technology. *Stem Cell Res.* 62, 102809. <https://doi.org/10.1016/j.scr.2022.102809>
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>
- Zhao, C., Reyes, A.P., Schell, J.P., Weltner, J., Ortega, N.M., Zheng, Y., Björklund, Å.K., Rossant, J., Fu, J., Petropoulos, S., Lanner, F., 2021. Reprogrammed blastoids contain amnion-like cells but not trophectoderm (preprint). *Developmental Biology*. <https://doi.org/10.1101/2021.05.07.442980>
- Zhao, J., Wang, M., Chang, L., Yu, J., Song, A., Liu, C., Huang, W., Zhang, T., Wu, X., Shen, X., Zhu, B., Li, G., 2020. RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nat. Cell Biol.* 22, 439–452. <https://doi.org/10.1038/s41556-020-0484-1>
- Zhao, W., Huang, Y., Zhang, J., Liu, M., Ji, H., Wang, C., Cao, N., Li, C., Xia, Y., Jiang, Q., Qin, J., 2017. Polycomb group RING finger proteins 3/5 activate transcription via an interaction with the pluripotency factor Tex10 in embryonic stem cells. *J. Biol. Chem.* 292, 21527–21537. <https://doi.org/10.1074/jbc.M117.804054>

- Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., Ruohola-Baker, H., 2012. HIF1 α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition: Metabolic switch in ESC-to-EpiSC/hESC transition. *EMBO J.* 31, 2103–2116. <https://doi.org/10.1038/emboj.2012.71>
- Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K.A., Lozach, J., Glass, C.K., Rosenfeld, M.G., 2008. Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation. *Mol. Cell* 29, 69–80. <https://doi.org/10.1016/j.molcel.2007.11.002>
- Zhu, P., Guo, H., Ren, Y., Hou, Y., Dong, J., Li, R., Lian, Y., Fan, X., Hu, B., Gao, Y., Wang, X., Wei, Y., Liu, P., Yan, J., Ren, X., Yuan, P., Yuan, Y., Yan, Z., Wen, L., Yan, L., Qiao, J., Tang, F., 2018. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet.* 50, 12–19. <https://doi.org/10.1038/s41588-017-0007-6>
- Zhu, Y., Dong, L., Wang, C., Hao, K., Wang, J., Zhao, L., Xu, L., Xia, Y., Jiang, Q., Qin, J., 2022. Functional redundancy among Polycomb complexes in maintaining the pluripotent state of embryonic stem cells. *Stem Cell Rep.* 17, 1198–1214. <https://doi.org/10.1016/j.stemcr.2022.02.020>
- Zijlmans, D.W., Talon, I., Verhelst, S., Bendall, A., Van Nerum, K., Javali, A., Malcolm, A.A., van Knippenberg, S.S.F.A., Biggins, L., To, S.K., Janiszewski, A., Admiraal, D., Knops, R., Corthout, N., Balaton, B.P., Georgolopoulos, G., Panda, A., Bhanu, N.V., Collier, A.J., Fabian, C., Allsop, R.N., Chappell, J., Pham, T.X.A., Oberhuemer, M., Ertekin, C., Vanheer, L., Athanasouli, P., Lluís, F., Deforce, D., Jansen, J.H., Garcia, B.A., Vermeulen, M., Rivron, N., Dhaenens, M., Marks, H., Rugg-Gunn, P.J., Pasque, V., 2022. Integrated multi-omics reveal polycomb repressive complex 2 restricts human trophoblast induction. *Nat. Cell Biol.* <https://doi.org/10.1038/s41556-022-00932-w>
- Zimmerlin, L., Park, T.S., Huo, J.S., Verma, K., Pather, S.R., Talbot, C.C., Agarwal, J., Steppan, D., Zhang, Y.W., Considine, M., Guo, H., Zhong, X., Gutierrez, C., Cope, L., Canto-Soler, M.V., Friedman, A.D., Baylin, S.B., Zambidis, E.T., 2016. Tankyrase inhibition promotes a stable human naïve pluripotent state with improved functionality. *Development dev.* 138982. <https://doi.org/10.1242/dev.138982>
- Zorzan, I., Betto, R.M., Rossignoli, G., Arboit, M., Drusin, A., Martini, P., Martello, G., 2022. Chemical conversion of human conventional Pluripotent Stem Cells to Trophoblast Stem Cells (preprint). *Developmental Biology*. <https://doi.org/10.1101/2022.04.07.487453>