ROLES OF TRANSCRIPTION FACTOR EOMESODERMIN IN THE TROPHECTODERM AND DEFINITIVE ENDODERM LINEAGES OF THE MURINE EMBRYO



Alexandra Maria Bisia
Sir William Dunn School of Pathology
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Brasenose College
Michaelmas Term 2022

DECLARATION

I declare that the work presented in this thesis is my own unless otherwise stated.

ACKNOWLEDGEMENTS

I want to thank Liz Robertson and Liz Bikoff for agreeing to supervise my DPhil. This past 4-and-a-half years, they gave me the opportunity to explore a vast array of topics and experimental techniques, and for all that I have learned thanks to their deep knowledge and expertise, I am very grateful.

My DPhil projects would not have progressed a fraction as much as they did without the substantial help and guidance of the Robertson lab members. Whether it was offering experimental advice, helping me out with bioinformatics issues, or just sharing my excitement over interesting experimental results, the intellectual and emotional support of Ita Costello and Luke Harland has been irreplaceable. Ita Costello's experimental contributions to the TSC project were also indispensable. Maria-Eleni Xypolita, in addition to being a wonderful colleague and a great advisor on all things 'wet lab,' has been instrumental in carrying out the mouse breeding, embryo manipulation and imaging that is described in this thesis, along with Ita Costello. I want to thank Anna Senft, who helped me get my bearings when I first joined the lab and is an example of what a curious lifelong learner should aspire to. Finally, I sincerely thank Philipp Kurbel for his help in the lab, particularly the dozens of western blots and the marathon genotyping sessions that he undertook.

I am very grateful to the Wellcome Trust for funding for my DPhil project. For their experimental assistance and advice I thank Robert Hedley and Vasiliki Tsioligka (Don Mason flow facility), Joey Riepsaame (Genome Engineering Oxford), Simon Engledow (WTCHG), Amanda Williams (Department of Biology), Timothy Rostron (WIMM), and Joseph Caesar (Dunn School). I also want to thank Bertie Göttgens and his students Mai-Linh Ton and Bart Theeuwes at the University of Cambridge for their collaboration and help with the DE project.

My brother Dimitri, my dad Christos, my stepdad Hans Thoolen, but my mum Elena Pavlidou in particular, have encouraged my academic curiosity from the very beginning, and for letting me choose what and where to study and backing my choices unconditionally, I am endlessly thankful. I would not have reached this point without their support, emotional or otherwise.

Finally, the person who has been with me through every high and every low, who sat through countless presentation practices, read thesis chapter drafts, helped me over every tearful roadblock, celebrated every success, gave me great scientific and experimental advice: thank you, Brendan Largey, for sharing our DPhil adventure. I can't wait for the next one!

TABLE OF CONTENTS

Declaration....ii
Acknowledgements....iii
Table of contents....v
Abstract....xi
List of Figures....xiii
List of Tables....xvii
Abbreviations....xviii
Chapter 1: Introduction....1
1.1: Pre-implantation development....1
1.2: Isolation and in vitro maintenance of pre-implantation blastocyst lineages....4
1.3: Eomes and the T-box family....7
1.4: Role of Eomes in the trophoblast lineage....8
1.5: Additional essential components of the TSC gene regulatory network....10
1.5.1: Tead4....10
1.5.2: Cdx2....10
1.5.3: Ets2....11
1.5.4: Elf5....11
1.5.5: Fgf4/Fgfr2 axis....11
1.6: Post-implantation murine development....13
1.6.1: Gastrulation....15
1.6.2: Cell fate allocation during gastrulation....17
1.6.3: Signalling in DE specification....19
1.6.4: EMT and gastrulation....19

1.7: Roles of Eomes in post-implantation mouse embryonic development.....20
1.7.1: Eomes in DE specification.....22
1.7.2: Eomes regulation during DE formation.....22
1.7.3: Eomes activities in the DE lineage.....24
1.7.4: Essential components of DE specification gene regulatory network.....26
1.7.4.i: FoxA2.....26
1.7.4.ii: Sox17.....26
1.7.4.iii: Mixl1.....27
1.7.4.iv: GATA transcription factors.....28
1.7.4.v: FoxA2 and Sox17 regulation.....29
1.8: Tools and techniques.....30
1.8.1: Embryoid body culture.....30
1.8.2: Genome editing techniques.....32
1.8.3: The dTAG degron system.....33
1.8.4: Genome-wide techniques.....35
1.8.5: Using ESCs to generate novel genetically altered mouse strains.37
1.9: Overview and aims of this DPhil thesis.....39
Chapter 2: Generation of an inducible Eomes degradation system *in vivo* for embryo analysis.....40
2.1: Introduction.....40
2.2: Allele design and generation of a degron-tagged Eomes mouse line.....43
2.3: *Ex vivo/ in vitro* manipulation of Eomes-degron-expressing embryos.....49
2.3.1: *Ex vivo* culture of pre-implantation Eomes$^{deg/deg}$ blastocysts.....49
2.3.2: *Ex vivo* culture of post-implantation Eomes$^{deg/deg}$ embryos.....51
2.4: *In vivo/ in utero* manipulation of Eomes-degron-expressing embryos.....53
2.4.1: E5.5 Eomes$^{deg/deg}$ embryos do not show Eomes degradation after *in vivo* dTAG administration.....53

2.4.2: E6.5 and E7.5 Eomes$^{deg/deg}$ embryos show variable depletion of Eomes protein upon *in vivo* dTAG treatments.....55
2.4.3: *In vivo* dTAG treatment does not result in discernible phenotypic defects of Eomes$^{deg/deg}$ embryos.....59
2.5: Deletion of putative Eomes super-enhancer does not impact formation of CXCR4^{+}/cKit^{+} cells in DE differentiation culture.....61
2.6: Summary of findings.....65
2.6.1: Generation of an inducibly degradable Eomes allele.....65
2.6.2: *Ex vivo* dTAG treatment of Eomes$^{deg/deg}$ pre- and post-implantation embryos.....65
2.6.3: *In vivo* dTAG treatment of post-implantation embryos.....66
2.6.4: Deletion of putative Eomes super-enhancer region does not affect EB differentiation.....66
2.7: Appendix.....67
Chapter 3: Investigating the roles of Eomes in the trophectoderm lineage.....69
3.1: Introduction.....69
3.2: Derivation of Eomes$^{deg/deg}$ cell lines.....71
3.3: Transcriptional variability of Eomes$^{deg/deg}$ cell lines.....73
3.4: Eomes protein depletion results in aberrant expression of TSC-related signalling pathway and motility components.....78
3.4.1: Early-dysregulated genes.....79
3.4.2: Differentially expressed genes 3 days after dTAG-13 addition...84
3.4.3: Persistently dysregulated genes.....87
3.5: Eomes protein binds TSC-, TB-, and pluripotency-associated genes.....89
3.5.1: Investigating genes both differentially expressed and Eomes-bound.....93
3.6: Eomes and Tead4 may collaborate to promote TSC-specific gene expression.....94
3.7: Summary of findings.....100
3.7.1: Generation of Eomes$^{deg/deg}$ TSC lines.....100
3.7.2: Eomes depletion results in dysregulation of TSC GRN components and cell motility/cell contacts/cytoskeleton interactors.....100

3.7.3: Eomes binds TSC-, TB-, and pluripotency-expressed genes.....101
3.7.4: Tead4 co-binds many Eomes-bound genes.....101
3.8: Appendix.....102
Chapter 4: Investigating Eomes functions in the specification of DE and midline progenitors.....112
4.1: Introduction.....112
4.2: Optimisation of a DE progenitor EB differentiation protocol in mESCs.....114
4.3: CXCR4$^+$/cKit$^+$ cells in DE EB cultures express DE lineage markers.....120
4.4: Eomes protein ablation in DE EB cultures results in specific loss of CXCR4$^+$/cKit$^+$ double-positive cell population.....125
4.5: Eomes degradation results in dysregulation of endodermal, APS, and mesodermal markers.....130
4.6: Eomes protein loss results in changes to chromatin accessibility near APS/DE-related genes.....137
4.7: The timing of Eomes depletion affects anterior-posterior gene expression within the CXCR4$^+$/cKit$^+$ cell population.....143
4.8: Deletion of Sox17- and FoxA2-associated differentially accessible regions does not significantly impact CXCR4$^+$/cKit$^+$ cell specification.....145
4.9: Single-cell RNA- and ATAC-sequencing reveal specific loss of DE and notochord cells in Eomes-depleted DE EB cultures.....150
4.10: Summary of findings.....153
4.10.1: Optimisation of an EB differentiation protocol promoting APS/DE identity.....153
4.10.2: Eomes is essential for specification of the CXCR4$^+$/cKit$^+$ population.....153
4.10.3: Eomes depletion results in dysregulation of lineage-specific markers.....154
4.10.4: Eomes depletion results in chromatin accessibility changes...154
4.10.5: Single-cell multi-ome experiment in EBs may aid in parsing Eomes effects in APS lineages.....155
4.11: Appendix.....156

Chapter 5: Final Discussion.....163

5.1: Challenges, potential solutions, and future work in mouse degron allele application.....163

5.1.1: Discussion.....163

5.1.2: Precautions when employing degron systems.....164

5.1.3: Future experiments.....165

5.2: Eomes in the TE lineage.....166

5.2.1: Discussion.....166

5.2.2: Limitations of current experiments.....167

5.2.3: Future experiments.....168

5.3: Eomes roles in the APS and its daughter lineages.....170

5.3.1: Discussion.....170

5.3.2: Single-cell multi-ome.....173

5.3.3: Future experiments.....175

5.4: General principles of Eomes regulatory functions between TSC and APS lineages.....176

Chapter 6: Materials and Methods.....178

6.1: Methods.....178

6.1.1: Stem cell maintenance and differentiation.....178

6.1.2: Generation of Eomes-degron-P2A-mCherry repair template....179

6.1.3: Cas9-mediated mESC modification.....180

6.1.4: Mouse strains and genotyping.....181

6.1.5: Immunostaining.....181

6.1.6: Western blots.....181

6.1.7: Flow cytometry and FACS.....182

6.1.8: RNA isolation and qPCR.....183

6.1.9: RNA-seq.....183

6.1.10: RNA-seq analysis.....183

6.1.11: ATAC-seq.....183

6.1.12: ATAC-seq analysis.....184

6.1.13: CUT&RUN.....184

6.1.14: CUT&RUN analysis.....185

6.1.15: 10X multiome.....185

6.2: Materials.....186

6.2.1: Antibodies.....186

6.2.2: Cas9 targeting reagents.....187

6.2.3: Genotyping primers.....189
References.....190

ABSTRACT

Eomesodermin (Eomes), a member of the T-box family of transcription factors, is required for the specification of a multitude of lineages during mouse gastrulation. Initially it is expressed in the trophectoderm (TE) of the pre-implantation blastocyst, where it is essential for peri-implantation development. Eomes is also expressed in the post-implantation embryo, where it has a critical role in the specification of several lineages during gastrulation, including the definitive endoderm (DE). The DE is fated to give rise to the gut tube and all gut-associated organs of the developing embryo. Although Eomes is necessary in both, the molecular mechanisms by which it acts to maintain or specify these two diverse cell lineages remain largely unknown. The aim of my thesis was to investigate how Eomes functions at a cellular level to generate these distinct lineages.

In this thesis, I describe the generation of an Eomes-degron allele that permits the inducible depletion of Eomes protein in Eomes-expressing cells, and the subsequent generation of a mouse line harbouring this allele. I have taken advantage of this allele, in combination with genome-wide RNA and chromatin accessibility techniques, to examine the functions of Eomes protein within the TE and DE lineages in a finely time-resolved manner.

Eomes$^{deg/deg}$ mice are viable and fertile. Although the Eomes-degron system is functional and highly efficient *in vitro*, we encountered setbacks in its implementation *in vivo*, resulting in inconsistent depletion of Eomes protein in post-implantation embryos. We suggest that degron-tagged mouse lines may therefore not always be the best approach for the study of protein functionality, and rigorous validation of the system should always be carried out when it is adapted to a gene or protein of interest.

In trophoblast stem cells (TSCs), the *in vitro* equivalent of the TE, Eomes depletion followed by RNA-seq suggests that in addition to regulating components of the core TSC gene regulatory network, Eomes appears to promote the expression of a suite of genes

associated with cell motility, extracellular matrix components, and cell-cell interactions. Eomes may therefore have an important role in modulating the interactions of the TE with its environment, thus promoting successful attachment and invasion of the uterus by the blastocyst.

Using an *in vitro* embryonic stem cell-derived embryoid body model of anterior primitive streak (APS) development, which includes the DE and node/notochord lineages, I demonstrate that Eomes depletion affects the expression of many known DE markers, including the DE master regulators FoxA2 and Sox17. Eomes protein loss also results in altered chromatin accessibility, suggesting a possible mechanism for the regulation of DE genes by Eomes. In collaboration with the Göttgens group at the University of Cambridge, we have profiled at single-cell resolution the transcriptome and chromatin accessibility of APS-directed EBs. Ongoing analyses of this data set which will further elucidate the roles of Eomes in the specification of the APS progenitors and the resulting daughter lineages.

LIST OF FIGURES

Figure 1.1: Mouse pre-implantation development.....3
Figure 1.2: The T-box transcription factor family.....7
Figure 1.3: Eomes expression in the pre-implantation mouse embryo.....8
Figure 1.4: The GRN of mouse TSC maintenance and differentiation.....12
Figure 1.5: Peri-implantation mouse development.....14
Figure 1.6: Mouse gastrulation.....16
Figure 1.7: Cell fate allocation during gastrulation.....18
Figure 1.8: Eomes expression at the onset of gastrulation.....21
Figure 1.9: Regulatory regions proximal to the Eomes locus.....23
Figure 1.10: Summary of putative Eomes regulatory interactions in DE lineage.....25
Figure 1.11: Embryoid bodies.....31
Figure 1.12: Cas9 DNA targeting mechanism.....33
Figure 1.13: Schematic of the dTAG degron system.....35
Figure 1.14: Schematic of steps to make a genetically modified mouse line.....38
Figure 1.15: Summary of my DPhil's central questions.....39
Figure 2.1: Generation and characterisation of mESC clones harbouring Eomes^{deg} alleles.....45
Figure 2.2: Eomes$^{deg/deg}$ DE EB differentiation.....46
Figure 2.3: Flow cytometry of day 5.0 EBs made from weaned FD Eomes-degron-tagged mESC clones trialled for appropriate DE differentiation and degron tag functionality.....48
Figure 2.4: E3.5 pre-implantation Eomes$^{deg/deg}$ blastocysts lose Eomes protein and phenocopy Eomes-null blastocysts upon dTAG addition (500nM) in *ex vivo* cultures.....50
Figure 2.5: *Ex vivo* culture, dTAG treatment, and washout recovery of E6.5 Eomes$^{deg/deg}$ embryos.....52
Figure 2.6: *In vivo* injection of dTAG-13 in pregnant mice (from a homozygous Eomes^{deg} cross) on E5.5 does not result in Eomes-degron protein degradation..54
Figure 2.7: E6.5 Eomes$^{deg/deg}$ embryos show variable depletion of Eomes-degron protein after *in vivo* dTAG treatment.....56

Figure 2.8: Selected images from E7.5 dTAG-treated Eomes $Eomes^{deg/deg}$ x $Eomes^{deg/+}$ intercrosses show variable Eomes protein degradation.....58
Figure 2.9: Selected images of $Eomes^{deg/deg}$ embryos treated in utero with dTAG-13 and harvested 24 or 48h later.....60
Figure 2.10: Deletion of putative superenhancer region regulating Eomes expression.....62
Figure 2.11: $CXCR4^{+}$/$cKit^{+}$ cells appear to be specified normally in Eomes enhancer deletion clones.....64
Figure 2.12: $Eomes^{deg/deg}$ E6.5 embryos show variable depletion of Eomes protein after 4h treatment upon in vivo dTAG-13 injection.....67
Figure 3.1: Independent TSC line derivation from $Eomes^{deg/deg}$ mouse blastocysts.....72
Figure 3.2: $Eomes^{deg/deg}$ TSC transcriptomes exhibit clone-to-clone variability..74
Figure 3.3: Transcriptional variability between $Eomes^{deg/deg}$ clones and between dTAG treatment timepoints.....76
Figure 3.4: Known TSC markers are expressed variably between $Eomes^{deg/deg}$ TSC clones.....77
Figure 3.5: Volcano plots of differentially expressed genes upon Eomes depletion.....81
Figure 3.6: Bar chart of enriched GO terms in differentially expressed genes 3 days after dTAG-13 treatment.....86
Figure 3.7: Eomes protein binds enhancer regions near genes associated with developmental progression and trophoblast formation.....90
Figure 3.8: Eomes binds genes associated with TSC maintenance, TSC/TB differentiation, as well as pluripotency.....92
Figure 3.9: Tead4 binds regions associated with TSC-related genes, and co-binds with Eomes a significant proportion of genomic loci.....95
Figure 3.10: Schematic depiction of overlap between differentially expressed, Eomes-bound, and Tead4-bound genes in Eomes-depleted TSCs.....96
Figure 3.11: Selected genomic tracks of Eomes/Tead4-co-bound genes differentially expressed upon Eomes depletion.....98
Figure 3.12: Schematic of proposed models of Eomes functions in TSCs and Eomes-Tead4 collaboration in TSCs.....99

Figure 3.13: Sashimi plot of Eomes transcript splicing in Eomes$^{deg/deg}$ clone TSC3.....102
Figure 3.14: Quality control of RNA-seq data.....103
Figure 3.15: Heatmap of log_2[fold change(timepoint/0h)] of genes detected in more than one differentiation timepoint.....104
Figure 3.16: Eomes binds near genes associated with TSCs and TSC differentiation.....105
Figure 3.17: Eomes and Tead4 co-bind mostly to genes more highly expressed in untreated TSCs.....107
Figure 4.1: Optimisation of an EB-based DE differentiation protocol in mESCs.116
Figure 4.2: Variable effects of graded CHIR and LDN concentrations in CXCR4$^+$/cKit$^+$ cell specification efficiency.....118
Figure 4.3: mCherry reporter is enriched in CXCR4$^+$/cKit$^+$ populations of EBs from both CCE and E14-RV genetic backgrounds.....119
Figure 4.4: CXCR4+/cKit+ transcriptomes are distinct from those of undifferentiated ESCs.....121
Figure 4.5: CXCR4$^+$/cKit$^+$ cells from day 5.0 DE EBs express DE and APS markers and downregulate pluripotency markers.....124
Figure 4.6: Dynamic Eomes expression in DE EB cultures.....126
Figure 4.7: Time-dependent effect of Eomes protein degradation on DE population in EB cultures.....129
Figure 4.8: Experimental layout for RNA- and ATAC-seq experiments in DE EBs upon Eomes degradation.....130
Figure 4.9: Eomes depletion results in a distinct differentiation trajectory and gene expression signature in EBs.....132
Figure 4.10: Eomes depletion results in rapid downregulation of PS, DE, and anterior mesodermal markers.....136
Figure 4.11: Foxa2 is associated with one of the earliest-arising differentially accessible sites between Eomes-depleted and Eomes-non-depleted DE EBs.....138
Figure 4.12: Differentially accessible genomic sites at +24h from dTAG-13 addition are found near genes related to APS-derived lineages, as well as rostral-caudal patterning.....140

Figure 4.13: Differentially accessible genomic sites at +24h are enriched for binding motifs of DE-specific TFs and intracellular signalling effectors.....142
Figure 4.14: Depletion of Eomes protein at different timepoints of the DE EB differentiation protocol affects expression of A-P DE markers Sox2 and Cdx2..144
Figure 4.15: FoxA2-associated region identified as differentially accessible in ATAC-seq experiments is Tcf3/ß-catenin/Smad2/3-bound in DE differentiation.....146
Figure 4.16: Sox17-associated differentially accessible region is homologous to a human region that is differentially methylated specifically in DE.....148
Figure 4.17: Deletion of FoxA2- and Sox17-associated differentially accessible regions does not result in reduction of the percentage of $CXCR4^{+}$/$cKit^{+}$ cells in DE EB cultures.....149
Figure 4.18: Single-cell multiome (RNA- & ATAC-seq) on cells from DE EBs shows that Eomes depletion results in loss of anterior primitive streak, DE, and notochord cells.....152
Figure 4.19: Previous flow cytometry experiments carried out to optimise conditions for APS/DE EB differentiation.....156
Figure 4.20: Quality control for RNA-seq of undifferentiated mESCs and day 5.0 $CXCR4^{+}$/cKit+ EB cells.....159
Figure 4.21: Quality control for $Eomes^{deg/deg}$ G9 clone EB differentiation timecourse RNA-seq.....160

LIST OF TABLES

Table 2.1: *In vivo* Eomes protein depletion in Eomes$^{deg/deg}$ embryos varies depending on developmental stage and choice of small molecule.....68
Table 3.1: Summary statistics of differentially expressed genes in Eomes$^{deg/deg}$ TSC lines.....79
Table 3.2: Selected genes significantly differentially expressed at different timepoints of dTAG-13 treatment (upregulated in Eomes+ TSCs).....83
Table 3.3: Selected classes of differentially expressed genes 3 days after Eomes protein depletion.....84
Table 3.4: Selected differentially expressed genes 3 days after dTAG-13 treatment.....85
Table 3.5: Top 5 persistently differentially expressed genes upregulated in Eomes+ and Eomes– TSCs with known function or expression in the TE lineage.87
Table 3.6: Summary statistics of genes both differentially expressed upon Eomes depletion, and bound by Eomes.....93
Table 3.7: Summary statistics of differentially expressed genes and Eomes/Tead4-bound genes across all timepoints of Eomes depletion time course experiment.....97
Table 3.8: Differentially expressed genes from Eomes protein depletion timecourse experiment that are also Eomes-bound or Eomes/Tead4-co-bound (merged across timepoints).....108
Table 4.1: Summary statistics of differential accessibility analysis in Eomes degradation timecourse experiment carried out in DE EBs.....137
Table 4.2: Positive correlation between chromatin accessibility and increase in expression level at the +24h dTAG-13 addition timepoint.....139
Table 4.3: List of genes associated with Eomes+/- –differentially accessible genomic sites at +12h of dTAG/VC treatment.....161
Table 4.4: List of genes associated with Eomes+/- –differentially accessible genomic sites at +24h of dTAG/VC treatment.....162
Table 6.1: List of antibodies.....170
Table 6.2: List of reagents for Cas9-mediated genome editing.....171
Table 6.3: Primers used for genotyping and RT-qPCR.....173

ABBREVIATIONS

ActA: activin A

CTB: cytotrophoblast

DE: definitive endoderm

EB: embryoid body

EMT: epithelial-to-mesenchymal transition

EPC: ectoplacental cone

Epi: epiblast

(m/h)ESC: (mouse/human) embryonic stem cell

(E)VT: (extra)villous trophoblast

ExE: extraembryonic ectoderm

ExM: extraembryonic mesoderm

GO: gene ontology

ICM: inner cell mass

KD: knock-down

KI: knock-in

KO: knock-out

MET: mesenchymal-to-epithelial transition

PCA: principal component analysis

PrE: primitive endoderm

(A)PS: (anterior) primitive streak

TB: trophoblast

(p/m)TE: (polar/mural) trophectoderm

TF: transcription factor

TGC: trophoblast giant cell

TSC: trophoblast stem cell

TSS: transcriptional start site

VC: vehicle control

(A/D)VE: (anterior/distal) visceral endoderm

WB: western blot

WT: wild type

CHAPTER 1: INTRODUCTION

1.1: PRE-IMPLANTATION DEVELOPMENT

At the onset of mouse embryonic development, the zygote undergoes sequential cleavages to produce the morula (Figure 1.1a). Early-expressed transcription factors (TFs), such as Oct4 (octamer-binding transcription factor 4), Nanog, and Cdx2 (caudal type homeobox 2) initially co-localise in cells, but through mutually repressive activities between Cdx2 and Nanog, as well as Cdx2 and Oct4, their expression domains become distinct. At the 16-cell stage, around embryonic day 3.0 (E3.0), the morula undergoes compaction, a process by which cells increase the contacts between them, in part through increased expression of adhesion molecules. This results in organisation of the cells into non-polar inner cells fated to give rise to the inner cell mass (ICM), expressing Nanog and Oct4, and polarised outer cells fated to give rise to the trophectoderm (TE), in contact with the surrounding environment and defined by their Cdx2 expression (Posfai et al., 2017). This constitutes the first cell fate bifurcation in mammalian development (Figure 1.1b). The pluripotent ICM eventually gives rise to the embryo proper, as well as the parietal and visceral extraembryonic endoderm lineages (Chen et al., 2009; Niwa et al., 2005). The TE generates the majority of the placenta, with the exception of maternal-origin vasculature (Adamson et al., 2002; Watson & Cross, 2005).

By E3.5, the embryo, now called a blastocyst, has developed a fluid filled cavity, the blastocoel, and the ICM is only attached to a small portion of the TE, known as the polar trophectoderm (pTE) (Figure 1.1c). Under the influence of FGF4 (fibroblast growth factor 4) signalling from the ICM, relayed via FGFR2 (FGF receptor 2), pTE cells continue dividing, while, in absence of these signals, mural TE (mTE) cells differentiate into primary TGCs. In a reciprocal manner, the pTE expresses BMP4 (bone morphogenetic protein 4), which promotes appropriate growth of the ICM (Murohashi et al., 2010). The pTE compartment expands to produce the extraembryonic ectoderm (ExE), which contains the trophoblast stem cell (TSC) pool that generates the ectoplacental cone

(EPC). The EPC is responsible for invasion of the embryo into the maternal stroma and gives rise to a diverse array of specialised trophoblast (TB) cell types that organise to form the embryonic placental tissues, a process which continues for much of gestation. At the same time, ICM cells initially co-expressing Nanog and Gata6 again make a choice between expression of these the two TFs, with Gata6+ (GATA-binding protein 6) cells migrating to the blastocoel-facing surface of the ICM and giving rise to the primitive endoderm (PrE), which further diversifies to give rise to the parietal endoderm, which contributes to the Reichert's membrane; and the visceral endoderm, which generates the outer endodermal layer of the yolk sac. Pluripotent Nanog+ cells give rise to the epiblast, which gives rise to the embryo proper and the primordial germ cells (Chambers et al., 2003; Mitsui et al., 2003; Saiz & Plusa, 2013; Schrode et al., 2014).

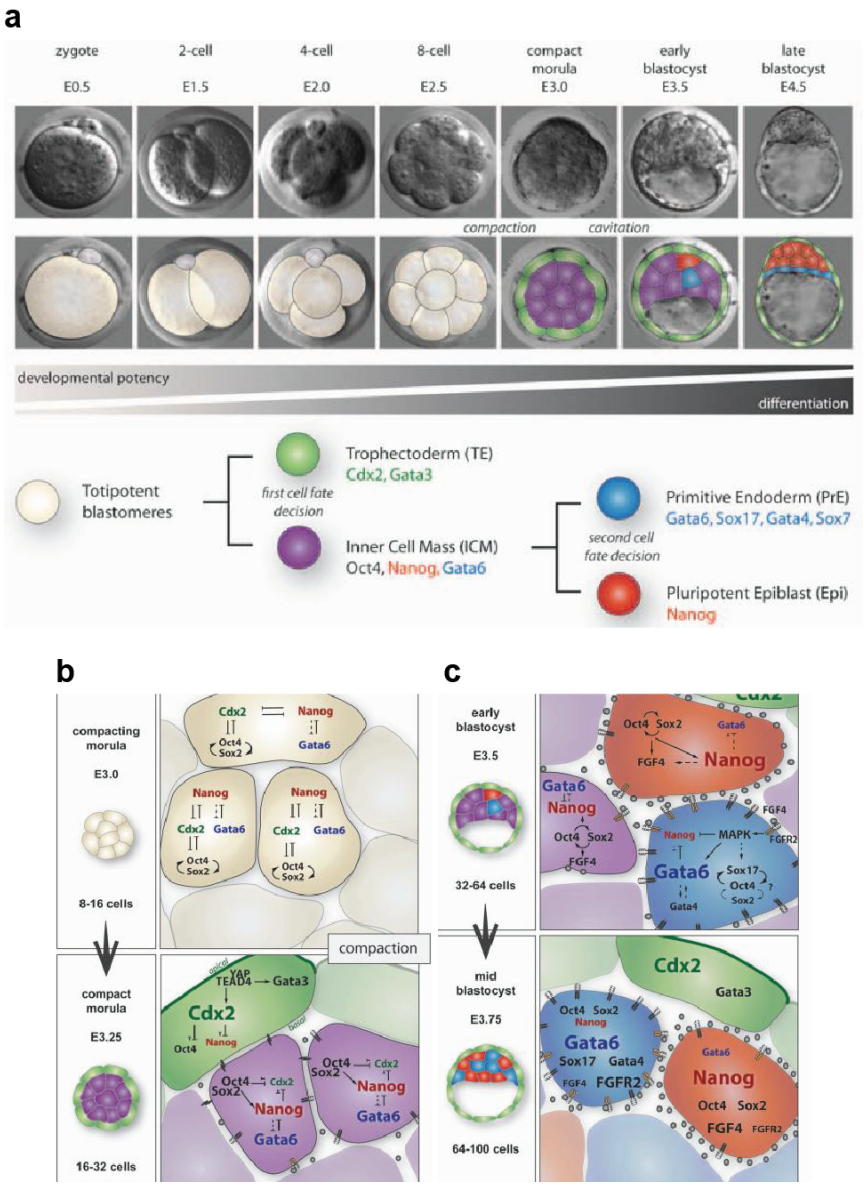


Figure 1.1: Mouse pre-implantation development.
Figure panels adapted from Schrode et al. (2013).
(a) Totipotent blastomeres give rise to the TE and the ICM in the first lineage segregation event of mouse embryonic development. The ICM further gives rise to the PrE and the pluripotent Epi.
(b) Blastomeres co-express Nanog, Oct4, and Cdx2, but there become lineage-restricted through mutually-repressive interactions, thus giving rise to the TE and the ICM.
(c) Cells of the ICM co-express Nanog and Gata6, which similarly have mutually repressive effects on each other, thus resulting in the segregation of the epiblast and PrE lineages.

1.2: Isolation and *in vitro* maintenance of pre-implantation blastocyst lineages

The three cell types found in the pre-implantation blastocyst (TE, epiblast, and PrE) have all been successfully isolated *in vitro*. It has further been demonstrated that these *in vitro* counterpart cells can be maintained in perpetuity in appropriate culture media.

The first of the three cell types to be isolated *in vitro* was the epiblast, in the form of embryonic stem cells (ESCs). It was demonstrated that, like their *in vivo* counterpart, ESCs are pluripotent, i.e. they retain the ability to differentiate into cells of all three germ layers (ectoderm, mesoderm, and endoderm), as well as the germ line (Martin, 1981). They can thus contribute to teratomas when injected subcutaneously into mice, generate embryoid bodies (EBs) containing cells of all three germ layers when differentiated *in vitro*, and contribute to the formation of a chimeric animal when injected into host blastocysts (Evans & Kaufman, 1981). ESCs can be maintained indefinitely in leukemia inhibitory factor (LIF)-containing medium (Smith et al., 1988).

Pluripotent stem cells can be isolated from embryos both at pre-implantation (ESCs) as well as at post-implantation stages (epiblast stem cells, EpiSCs) (Tesar et al., 2007; Vallier et al., 2009). EpiSCs can contribute to all three germ layers as well as primordial germ cells when grafted into the epiblast of a post-implantation embryo (Huang et al., 2012). Notably, ESCs cannot successfully engraft into the post-implantation epiblast, nor can EpiSCs successfully engraft into the pre-implantation epiblast (Huang et al., 2012). Additionally, although both ESCs and EpiSCs demonstrate activity of pluripotency markers Oct4, Sox2, and Nanog (Kojima et al., 2014), EpiSCs are dependent on Activin and FGF for *in vitro* maintenance (Vallier et al., 2009), unlike ESCs. These data, along with the extensive epigenetic differences between ESCs and EpiSCs (Takahashi et al., 2018), indicate that these pluripotency states are distinct, and they have been dubbed ‘naïve’ and ‘primed’ respectively. Notably, these pluripotent states are also interchangeable,

based on the medium of culture of the cells, indicating a degree of plasticity between the two states (Bao et al., 2009).

Naïve pluripotency *in vitro* is maintained by culturing ESCs in a LIF- and '2i' (2-inhibitor)-containing medium. The two inhibitors in question are CHIR99021 ('CHIR') and PD0325901 ('PD') (Ying et al., 2008). CHIR is an inhibitor of the enzyme GSK-3 (Murray et al., 2004), which normally phosphorylates the Wnt signalling intracellular effector ß-catenin, resulting in its destabilisation and degradation (Aberle et al., 1997; Munemitsu et al., 1995; Rubinfeld et al., 1996). CHIR thus acts as a Wnt signalling agonist, a pathway which promotes naïve ESC self-renewal (Merrill, 2012). PD, in contrast, is an inhibitor of MAPK kinase-1 (MKK1) (Bain et al., 2007), a component of the MEK/ERK signalling pathway which inhibits self-renewal and promotes lineage commitment in ESCs (Kunath et al., 2007).

As mentioned above, ESC and EpiSC are interconvertible states, with FGF + Activin treatment promoting the transition to a post-implantation epiblast-like pluripotency. During this transition, cells acquire a transitional epiblast-like cell (EpiLS) identity (thought to be representative of the *in vivo* epiblast's 'formative state') that permits induction of germ cells (Hayashi et al., 2011). It should be noted that naïve ESCs have been demonstrated not to be directly responsive to germ layer induction in hESC models, as they first need to undergo a process termed 'capacitation,' a process requiring Wnt signalling inhibition among other conditions, in order to become responsive to germ layer induction (Kalkan & Smith, 2014; Rostovskaya et al., 2019). The transition to primed pluripotency then permits the specification of cell lineages corresponding to the embryonic germ layers (Stavridis et al., 2007).

TE cells have also successfully been captured and cultured *in vitro* (Tanaka et al., 1998), both from pre- and post-implantation embryos. So-termed trophoblast stem cells (TSCs) can successfully contribute to chimeras, where they give rise to TGCs, the EPC, and the

ExE (but not the epiblast or the PrE), and support normal embryonic development (Tanaka et al., 1998). TSCs were initially perpetuated *in vitro* through culture in serum-containing, heparin- and FGF4-supplemented media (Tanaka, 2006; Tanaka et al., 1998). The identification of TGF-ß (transforming growth factor ß) as the active ingredient of the serum permitting TSC maintenance (Erlebacher et al., 2004) more recently led to the successful design of a defined TSC medium (Kubaczka et al., 2014).

PrE cells have also been isolated and cultured *in vitro*. So-called e**x**traembryonic **en**doderm (XEN) cells can contribute cells to blastocyst chimeras, exclusively to the PrE lineage (Kunath et al., 2005). XEN cells can be isolated through ESC or TSC derivation protocols, with slight modifications in media or timing of disaggregation (Niakan et al., 2013), and can also be generated by conversion of ESCs and differentiated cell types through forced expression of Gata4 or Gata6 (Fujikura et al., 2002; Shimosato et al., 2007; Wamaitha et al., 2015) or by ESC exposure to growth factors retinoic acid and activin A (Cho et al., 2012). Sox17 (SRY-box transcription factor 17) is also a central factor in activating the PrE gene regulatory network (Niakan et al., 2010). Notably, XEN cells can also be isolated from pre- as well as post-implantation embryos (Lin et al., 2016).

1.3: Eomes and the T-box TF family

The T-box TF family is an ancient group of TFs found in all bilaterians. They contain a conserved DNA binding motif (the T-box) of 180-200 amino acid residues. Most members of the family are involved in essential aspects of vertebrate embryonic development, often in more than one context, ranging from regulating gastrulation, to branching morphogenesis and organogenesis, to regulating limb development. Additional roles in adult stem cell maintenance have since emerged. Eomes is a T-box family member that belongs to the Tbr1 (t-box brain transcription factor 1) clade of the T-box family, and is involved in the specification and development of multiple embryonic cell lineages, as well as adult cell populations, including neural stem cells and natural killer T-cells (Papaioannou, 2014).

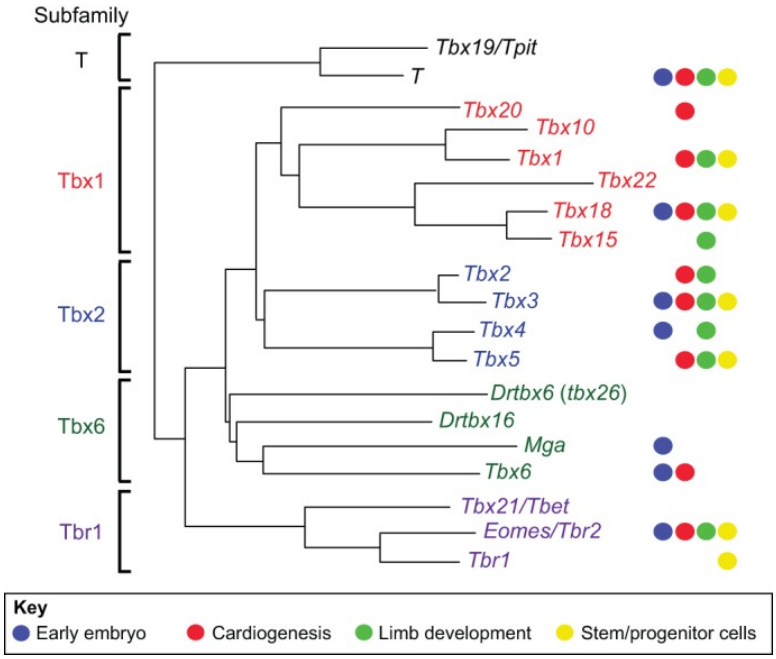


Figure 1.2: The T-box transcription factor family.
Eomes belongs to the Tbr1 branch of the T-box TF family, with Tbr1 being its most closely-related paralogue. Eomes is also known as Tbr2. Figure panel adapted from Papaioannou (2014).

1.4: Roles of Eomes in the trophoblast lineage

Eomes is one of the earliest-expressed members of the T-box TF family during mouse development. It is first detectable in the TE lineage at the blastocyst stage (Figure 1.3), activated by TE lineage markers Cdx2 and Tead4 (TEA domain transcription factor 4) (Nishioka et al., 2008; Strumpf et al., 2005; Yagi et al., 2007). It is essential for maintenance of the TSC pool and in its absence the TE compartment, although initially specified, fails to expand (Russ et al., 2000). Homozygous Eomes-null embryos arrest at the peri-implantation stage and exhibit a significantly diminished TE compartment, as cells stop dividing and rapidly differentiate into TGCs. After implantation, as the embryo develops, Eomes continues to be expressed throughout the ExE. By E9.5, expression is confined to the developing chorion, the most distal layer of the developing placenta, and is then down-regulated by E10.5 (Kuales et al., 2015; Latos et al., 2015b).

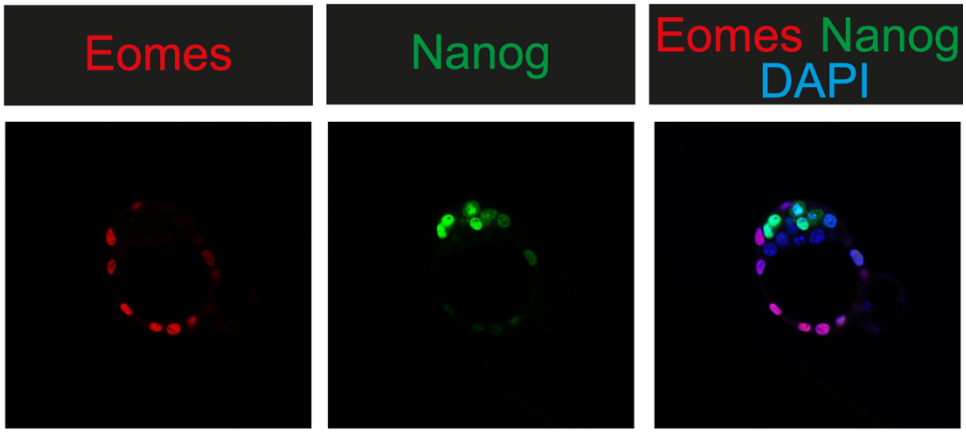


Figure 1.3: Eomes expression in the pre-implantation mouse embryo.
Immunofluorescence of E3.5 pre-implantation mouse blastocyst showing Eomes expression in the TE. The ICM is distinguishable by its expression of pluripotency marker Nanog. DAPI is a nuclear marker. Images taken by Dr Ita Costello.

Eomes is also essential for TSC maintenance in culture, as RNA interference (RNAi) knock-down (KD) results in loss of self-renewal and expression of TSC-specific markers (Kidder & Palmer, 2010). Experiments performed in cultured TSCs have shown Eomes occupancy at several TSC-associated genes, including Elf5 (E74-like ETS transcription factor 5), Tfap2c (transcription factor AP-2γ), components of FGF signalling Fgfr2/4, as well as the

Eomes locus itself (Kidder & Palmer, 2010; Lee et al., 2019). However, a more recent re-examination of the Kidder and Palmer (2010) ChIP-seq data by members of our lab (unpublished) may call for reservations with these conclusions.

It has also been demonstrated that forced expression of four TSC TFs, namely Eomes, Tfap2c, Ets2 (ETS proto-oncogene 2 transcription factor), and Gata3, enables the direct conversion of mouse fibroblasts to so-termed "induced TSCs" (iTSCs) (Kubaczka et al., 2015). The authors note that it was not possible, in their hands, to generate iTSCs with fewer than these four factors, suggesting Eomes is necessary for this process. Another reported combination of factors permitting fibroblast-to-iTSC conversion is Gata3, Eomes, and Tfap2c, with or without Myc (MYC proto-oncogene, bHLH transcription factor) (Benchetrit et al., 2015). By contrast, the conversion efficiency of mESCs to iTSCs was much lower. The authors attributed this difference to potentially greater epigenetic similarity between fibroblasts and TSCs compared to ESCs and TSCs. The authors do not comment on attempts to generate iTSCs with fewer than those three factors. Although neither publication addresses the specific role of Eomes within iTSC reprogramming, the fact that Eomes is reported by both as necessary for the process suggests an essential role for Eomes during reprogramming.

In TSCs, Eomes has been proposed to act in a trimeric protein complex with Elf5 and Tfap2c to maintain a TSC identity by occupying "stemness-related" genes (Latos et al., 2015b). In contrast, during TSC differentiation, when Eomes levels fall and Tfap2c levels increase, an Elf5-Tfap2c dimeric complex occupies sites with a stronger Tfap2c binding site signature and which are associated with differentiation genes (Latos et al., 2015b). This hypothesis may also match the *in vivo* situation, with TB cells more proximal to the Epi receiving FGF4 signals via FGFR2, which permits them to maintain their stem-like status, while more distal EPC cells are no longer exposed to FGF signalling, resulting in reduced Eomes and increased Tfap2c expression, therefore tipping the balance toward differentiation (Latos et al., 2015b).

However, although Eomes is known to be required for TSC maintenance, its exact roles in regulation of components of the TSC gene regulatory network (GRN) remain to be explored.

1.5: Additional essential components of the TSC gene regulatory network

1.5.1: Tead4

TEA domain transcription factor 4 (Tead4) is the earliest-expressed determinant of the TE lineage. In Tead4-null mouse embryos, TE lineage markers such as Cdx2, Eomes, and Fgfr2 are not activated, while pluripotency markers such as Oct4 persist in their expression, thereby hampering the first lineage segregation event of the mouse embryo (Figure 1.4a). The activity of Tead4 depends on a cell's position within the morula. Inner cells, where the Hippo signalling pathway is active, do not exhibit Tead4-mediated transcription. On the contrary, in outer cells where the Hippo pathway is inactive, the transcriptional partner of Tead4, Yap, localises to the nucleus, resulting in Tead4-mediated transcriptional activity (Figure 1.4b) (Nishioka et al., 2009). Tead4 is additionally required for blastocoel cavity formation, as well as for the establishment of TSCs, but not of ESCs. Finally, Tead4 is exclusively required for pre-implantation development, as a conditional epiblast deletion after implantation does not hamper embryonic development (Home et al., 2012; Nishioka et al., 2008; Yagi et al., 2007).

1.5.2: Cdx2

Caudal type homeobox 2 (Cdx2) is an early lineage determinant essential in the segregation of the ICM and TE lineages through mutually repressive activity with the pluripotency factors Nanog and Oct4 (Chen et al., 2009; Niwa et al., 2005) (Figure 1.4a). Cdx2-null embryos fail to downregulate Oct4 and Nanog, and fail to implant, resulting in

embryonic lethality. Cdx2-null embryos also express lower levels of Eomes protein (Strumpf et al., 2005).

1.5.3: Ets2

E26 transformation-specific 2 (Ets2) is expressed early in the mouse and is also involved in TSC self-renewal (Figure 1.4a). Ets2-null cells exhibit reduced expression of Cdx2 and Eomes (Wen et al., 2007). Additionally, Ets2 and Elf5 appear to have partially epistatic and synergistic roles in the proliferative ExE through regulation of partially overlapping downstream targets, and loss of Elf5 can be partially compensated for by Ets2 overexpression (Donnison et al., 2015).

1.5.4: Elf5

As discussed above, E74-like ETS transcription factor 5 (Elf5) is also involved in the TE lineage in partially epistatic manner with Ets2 (Figure 1.4a). Elf5-null embryos implant and form an EPC, however they fail to produce an ExE as the TSC population cannot be maintained (Donnison et al., 2005). Also as previously discussed, Elf5 is central in the balance between TSC-specific gene expression as member of the Elf5-Eomes-Tfap2c protein complex, and the TB differentiation gene expression as member of the Elf5-Tfap2c dimer (Latos et al., 2015b).

1.5.5: Fgf4/Fgfr2 axis

Fibroblast growth factor (FGF) signalling is essential for TSC maintenance, *in vivo* as well as *in vitro* (Figure 1.4b). FGF4 signals released from the ICM/Epi are transduced through FGF receptor 2 (FGFR2), which is found in the surface of TE cells. Fgf4- or Fgfr2-null embryos show similar peri-implantation lethality phenotypes, due to trophoblast defects as well as stunted ICM development (Arman et al., 1998, 1999; Feldman et al., 1995; Murohashi et al., 2010).

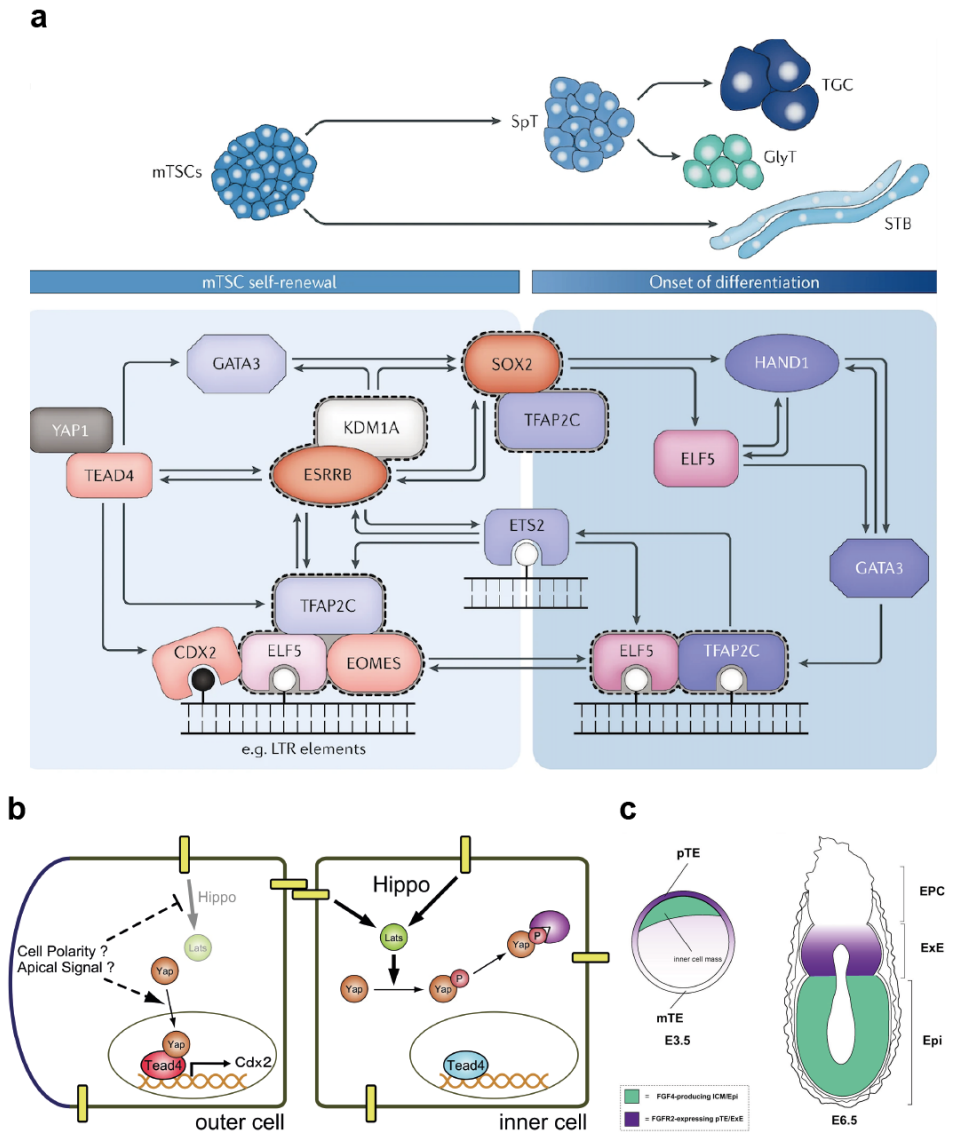


Figure 1.4: The GRN of mouse TSC maintenance and differentiation.
(a) A network of TFs including among others Cdx2, Tead4, Eomes, Elf5, Tfap2c, Gata3, and Ets2 are involved in TSC maintenance and proliferation. Conversely, a partially overlapping network of TFs, including Elf5, Tafp2c, Ets2, Hand1, and Gata3 are implicated in TSC differentiation. Figure panel adapted from Hemberger et al. (2020).
(b) Differential activity of the Hippo pathway in inner and outer morula cells results in activation of Tead4-mediated transcription only in outer cells, which become fated to the TE. Through a mechanism that has not yet been fully elucidated, the Hippo pathway is selectively active in inner cells. This results in phosphorylation of Yap, the transcriptional partner of Tead4, by Lats, a member of the Hippo pathway. Yap is thus excluded from the nucleus. In contrast, outer cells don't exhibit Hippo activity and Yap remains unphosphorylated. It can then enter the nucleus and partner with Tead4 to promote transcription of Tead4 target loci. Figure panel adapted from Nishioka et al. (2009).
(c) Epi-secreted FGF4 promotes TSC maintenance by signalling through TE/ExE-expressed FGFR2. Figure panel adapted from Ralston and Rossant (2006).

1.6: Post-implantation murine development

By E4.5 the three first cell lineages of the embryo (TE, Epi, PrE) have been established, and the embryo implants in the uterus (Theiler, 1989). Post-implantation development is associated with a rapid increase in cell proliferation rates and expansion of all three primary lineages (Ciemerych & Sicinski, 2005). The epiblast elongates and expands in size, and develops a new fluid-filled cavity (E5.5-6.25), the proamniotic cavity (Kim et al., 2021; Theiler, 1989). This gives the embryo a cup-like appearance, giving rise to the name of this stage, the "egg cylinder" (Figure 1.5, first panel). At the same time, the pTE expands to form the cup-shaped ExE, which secretes BMP4 signals – these are detected more strongly by more proximal regions of the egg cylinder (Lawson et al., 1999). The PrE expands down the inner wall of the blastocyst and gives rise to the visceral endoderm (VE), which remains associated with the epiblast and ExE, and the parietal endoderm, which migrates away from the egg cylinder to form a largely acellular sac-like structure, the so-termed Reichert's membrane, forming a barrier between the embryo and the uterine stroma (Gardner & Papaioannou, 1975; Hogan et al., 1980).

The VE also acts as a signalling centre to impart pattern to the developing epiblast, with the induction of the so-termed distal visceral endoderm (DVE) at the distal tip of the egg cylinder (Figure 1.5, first panel) (Thomas & Beddington, 1996). The DVE, in response to low-level Nodal signalling from the proximal (i.e. nearest to the ExE) epiblast expresses Lefty1 (left-right determinator factor 1), a Nodal inhibitor (Mesnard et al., 2006; Rodriguez et al., 2005; Takaoka et al., 2017). The DVE expands and migrates towards the prospective anterior side of the epiblast, while recruiting additional cells to form the anterior VE (AVE) (Rodriguez et al., 2001; Srinivas et al., 2004). The AVE, also expressing Lefty1 as well as the TF Hex (haematopoietically-expressed homeobox) and BMP/Wnt antagonists Cer1 (Cerberus 1) and Dkk (Dickkopf), promotes specification of the future anterior structures of the embryo (as its name suggests) (Miura & Mishina, 2007). Conversely, on the opposite side of the proximal epiblast, Nodal signalling acts in

conjunction with BMP4 and Wnt3 in a positive-feedback loop to promote posterior fates (Figure 1.5, second and third panels) (Ben-Haim et al., 2006). Maximal Nodal/BMP/Wnt signalling identifies the future posterior of the embryo, induces the epiblast to form nascent mesoderm and marks the site of the prospective primitive streak (PS) (Conlon et al., 1994; Takaoka & Hamada, 2012; Zhang & Hiiragi, 2018).

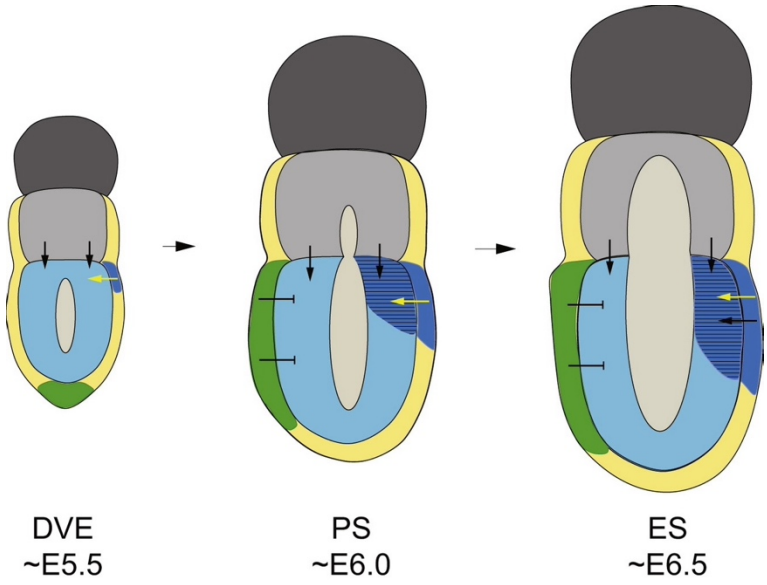


Figure 1.5: Peri-implantation mouse development.
At peri-implantation (E5.5), the ICM has given rise to the Epi (light blue), the PrE has produced the VE (yellow) and parietal endoderm (not shown in this image), and the TE has given rise to the ExE (light grey) and the EPC (dark grey). BMP4 signalling from the ExE is shown in black arrows, and Wnt3 signalling from the VE (and later, the Epi itself) is shown in yellow arrows. A specialised subset of the VE, the DVE, is shown in green. The DVE migrates to the future anterior of the embryo (E6.0 onwards), concomitantly recruiting the AVE. The DVE and AVE secrete Nodal, BMP, and Wnt signalling antagonists (shown in black blunt-headed arrows), while the future posterior of the embryo experiences a positive-feedback loop between BMP4, Nodal, and Wnt3. This pinpoints the position of the PS (marked by expression of T in hatches). Figure adapted from Yoon et al. (2015).

1.6.1: Gastrulation

The PS is a structure through which epiblast cells will ingress in a process known as gastrulation (Figure 1.6). During gastrulation, which begins around E6.5, Epi cells undergo EMT, which is thought to be the result of a switch from E- to N-cadherin, enabling nascent mesoderm to delaminate from the epiblast and ingress through the PS, between the epiblast and the VE on the posterior side of the embryo (Basilicata et al., 2016). The PS gradually elongates, eventually reaching the distalmost point of the egg cylinder, where a flattened epithelial structure called the node appears. This process results in formation of the three primary germ layers, with ingressing cells giving rise to the mesoderm and definitive endoderm (DE), and the cells remaining in the epiblast generating the ectoderm. The ectoderm will give rise to the epidermis and neurectoderm, while the mesoderm will give rise to the progenitors of multiple cell types including the heart, blood, skeleton and musculature, bone, and connective tissue (Bardot & Hadjantonakis, 2020). Finally, the DE specified during gastrulation will generate the entire gut tube and all gut-associated organs, including the lungs, liver, pancreas, and bile duct, as well as contributing to portions of the urinary system (Liaw et al., 2018; Wells & Melton, 1999).

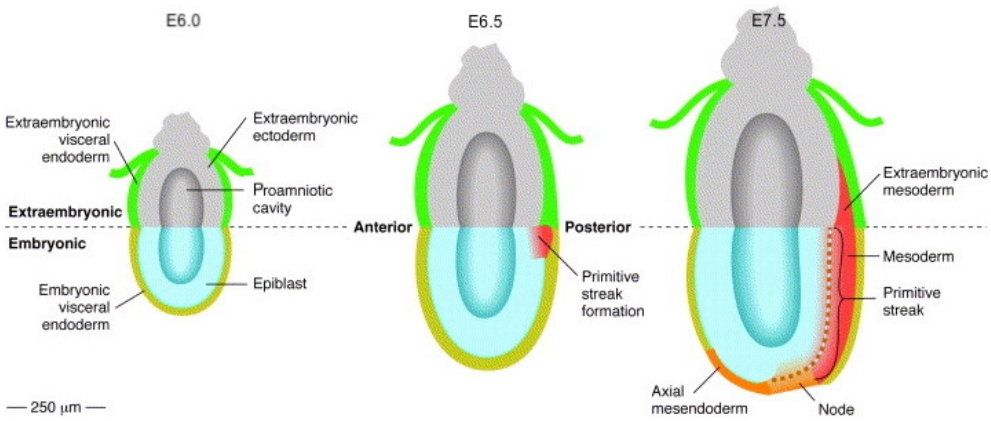


Figure 1.6: Mouse gastrulation.
The site of the PS specifies the posterior of the embryo. During gastrulation, cells from the Epi ingress through the PS. Cells ingressing through the posterior PS and migrate proximally and enter the extraembryonic portion of the egg cylinder, becoming ExM. Cells ingressing through the anterior PS become axial mesoderm (which will give rise to the node and notochord), cardiac and cranial mesoderm, and DE. Cells ingressing through more medial regions of the PS will give rise to other mesodermal lineages, paraxial, intermediate, and lateral plate. Figure panel adapted from Beddington and Robertson (1998).

The signalling pathways discussed above do not only function to specify the initial A-P axis of the embryo, but also have critical roles during gastrulation itself. Nodal is key in promoting anterior PS (APS) derivative lineages, namely the DE, the node and midline, cranial/cardiac mesoderm, and paraxial mesoderm (Tam & Loebel, 2007). Wnt signalling has been shown to be essential in the initiation and maintenance of the PS (Gadue et al., 2006; Liu et al., 1999). Furthermore, loss of ß-catenin, an intracellular signal transducer in the canonical Wnt signalling pathway, results in a block of PS formation (Huelsken et al., 2000). Epi-specific deletion of Wnt3, however, does allow PS specification, but results in slight developmental delay, shorter and reduced expression of PS markers, and a reduction in Wnt3 expression in the overlying VE, suggesting a positive feedback loop of Wnt on its own expression (Arkell et al., 2013).

As gastrulation proceeds and the egg cylinder elongates, BMP signalling from the ExE remains robust in the adjacent proximal epiblast (Lawson et al., 1999). In contrast, due to expression of antagonists in the AVE, Nodal signalling becomes confined to the proximal

epiblast prior to PS formation, and after formation of an overt PS, becomes confined to the nascent mesoderm entering the streak, before becoming downregulated (Collignon et al., 1996).

On E8.0 (the late streak stage), the PS begins to recede, with cells from the node trailing toward the posterior end of the embryo and giving rise to the midline (axial) mesoderm. By E9.5, gastrulation is complete. Overall, although before gastrulation epiblast cells are pluripotent, after they have traversed the PS they become fate-restricted, and their capacity for fate 'switching' is greatly diminished, limited generally to within their assumed germ layer (Tam et al., 1997).

1.6.2: Cell fate allocation during gastrulation

Throughout gastrulation, the embryo undergoes a considerable expansion in cell numbers, accompanied by morphogenesis and elongation of the egg cylinder, resulting in a highly dynamic signalling environment encountered by cells traversing the lengthening PS. At early streak stages, the first mesodermal lineages to be specified as cells begin to ingress through the PS are the extraembryonic mesoderm (ExM), followed by the cranial/cardiac mesodermal progenitors. The ExM arises from cells migrating through the proximal PS toward the extraembryonic portion of the egg cylinder to form the amnion and lining of the emerging yolk sac. The cranial/cardiac mesodermal progenitors, which ingresses more proximally and migrate anteriorly. At later streak stages cells give rise to the paraxial mesoderm, which gives rise to somites, and the lateral plate mesoderm, which forms the body wall. Cells ingressing through the anteriormost portion of the streak give rise to progenitors of the DE, node, and midline of the embryo (Figure 1.7).

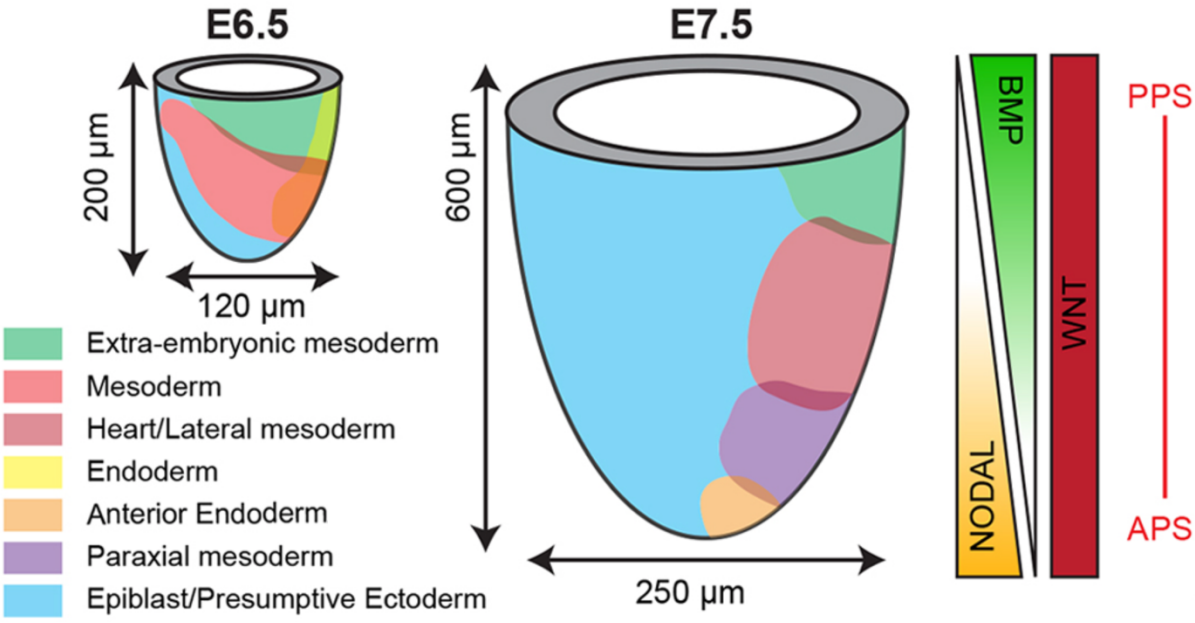


Figure 1.7: Cell fate allocation during gastrulation.
Cell fate allocation depends on the region of the PS through which cells ingress, and by extension the signals to which the cells are exposed during gastrulation. Cells ingressing through the APS (shown in orange) are exposed to high levels of Nodal signalling and low levels of BMP, and contribute to the node, cardiac/cranial mesoderm, and DE. Cells ingressing through the middle portions of the streak give rise to paraxial mesoderm (in purple) and intermediate and lateral plate mesoderm (in red), while cells ingressing through the posteriormost portion of the streak (in green) are exposed to high levels of BMP and will contribute to ExM lineages. Figure panel adapted from Martyn et al. (2019).

1.6.3: Signalling in DE specification

DE cells ingress through the anteriormost region of the PS. As mentioned earlier, sustained Nodal signalling is essential for APS-derived lineages. In the mouse, Nodal is capable of inducing both mesodermal and endodermal lineages, but endodermal and midline lineages fail to be specified in embryos carrying a Nodal hypomorphic allele, indicating a requirement for robust and sustained Nodal signalling during gastrulation for correct specification of APS derivatives (Vincent et al., 2003).

The intracellular effectors of Nodal signalling, Smad2/3, are integral for DE specification, as Smad2-deficient cells in mouse embryonic chimeras fail to colonise the foregut and hindgut, while compound heterozygous Smad2 and Smad3 mice exhibit variably penetrant anterior midline defects, including a reduction in DE and node/midline progenitors (Dunn et al., 2005; Tremblay et al., 2000; Vincent et al., 2003).

1.6.4: EMT and gastrulation

Although EMT is considered to be a defining hallmark of gastrulation, recent findings paint a more complex picture of the role of EMT role in DE specification and morphogenetic movement. The DE itself is an epithelial tissue (Tam & Beddington, 1992), and until the mid-to-late 2000s it was widely accepted that during gastrulation, future DE cells undergo a round of EMT as they ingress through the PS, followed by a process of MET, during which cells of the newly-specified DE emerge onto the outer surface of the embryo and actively displace the overlying VE proximally into the extraembryonic yolk sac region of the egg cylinder (Lawson & Pedersen, 1987). More recent findings indicate that DE cells intercalate with VE cells, and indeed that VE cells contribute to the embryonic gut tube of the murine embryo at mid-gestation (Kwon et al., 2008; Nowotschin et al., 2019). It remains to be determined if VE cells persist in the gut tube of the neonate and adult mouse. Additionally, during gastrulation, cells ingressing through the APS appear to undergo a partial EMT (Scheibner et al., 2021), independent of the classical EMT TF

Snai1, with Foxa2 playing a key role in preventing complete EMT by suppressing expression of Snai1 and other EMT genes, and upregulating EMT suppressors such as Grhl2, Ovol2, Pofut2, and Elf3 (Kwon et al., 2008; Nowotschin et al., 2019).

1.7: Roles of Eomes in post-implantation mouse embryonic development

At E5.5, Eomes begins to be expressed in the VE (Arnold et al., 2009), where it functions to maintain the DVE and recruit additional Hex-expressing migratory cells that contribute to the AVE (Nowotschin et al., 2013). VE-specific KO of Eomes results initially in universal thickening of the VE, followed by thickening specifically at the distal tip of the VE (the latter a phenotype generally associated with failure of AVE migration). Additionally, AVE markers normally found on the future anterior side of the embryo remain confined to the distal tip of the embryo (Nowotschin et al., 2013). Additional evidence from overexpression and ChIP-seq experiments *in vitro* suggest that AVE recruitment is in part due to the Eomes activation of Lhx1, a TF essential for AVE specification and function (Nowotschin et al., 2013).

Eomes also begins to be expressed at the site of the forming PS, the future posterior of the embryo, around E6.0-6.5, with expression maintained in the PS during most of gastrulation (Figure 1.8) (Ciruna & Rossant, 1999; Hancock et al., 1999). Eomes is essential for promoting EMT at the onset of gastrulation (Arnold et al., 2008). An epiblast-specific deletion of Eomes based on a Sox2Cre system results in embryos that survive beyond peri-implantation (Arnold et al., 2008; Hayashi et al., 2002). Nascent mesoderm induced in these embryos fails to delaminate, resulting in a profound block of gastrulation, which is visible as a build-up of cells in the site of the putative PS and a failure to properly allocate cells to form the three germ layers (Arnold et al., 2008). In the absence of Eomes, E-cadherin (E-cad) is not downregulated as it normally is during EMT (Arnold et al., 2008). However, expression of the known regulators of E-cad

repression, namely Fgf4 and Snail are unperturbed, so at present it remains unclear as to how Eomes prevents E-cad being correctly downregulated (Arnold et al., 2008).

Previous experiments from our lab have established that Eomes expression in the prospective ExM of the yolk sac governs haemogenic competency and is required for the specification of the first extraembryonic waves of blood formation (Harland et al., 2021). Eomes activity is also essential for specification of the cardiac mesoderm progenitors. This reflects a requirement for Eomes binding to T-box enhancer elements regulating the induction of Mesp1/2 (mesoderm posterior bHLH transcription factor 1/2), known master regulators of the cardiac lineage (Costello et al., 2011).

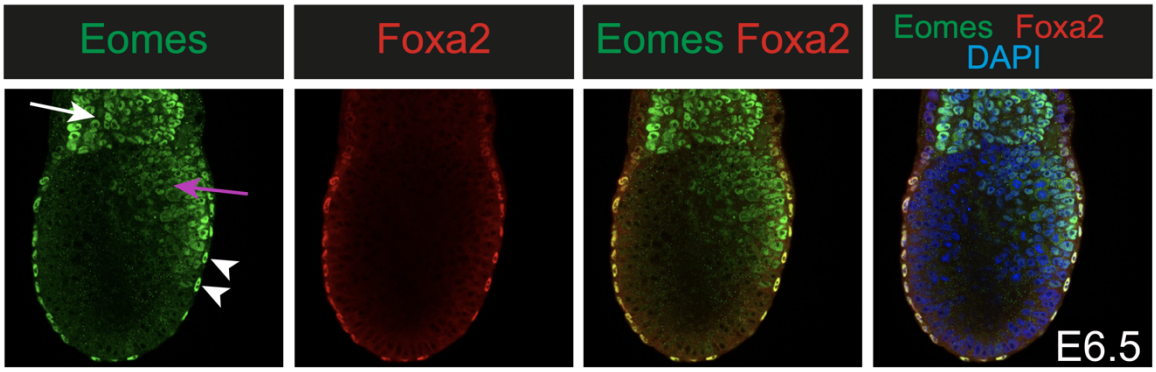


Figure 1.8: Eomes expression at the onset of gastrulation.
Eomes (green) shows three distinct expression domains at the onset of gastrulation: Eomes expression in the ExE (white arrow) is sustained from its preimplantation TE expression; Eomes expression in the VE (white arrowheads); and Eomes expression in the PS (magenta arrow). Foxa2 (red) is a marker of VE at the onset of gastrulation, while DAPI (blue) is a nuclear marker. Images taken by Dr Ita Costello.

1.7.1: Eomes in DE specification

In vivo fate mapping experiments employing an $Eomes^{Cre}$ allele have shown that, in addition to the yolk sac mesoderm and developing heart, Eomes-expressing cells in the PS also give rise to the DE, node, and midline (Costello et al., 2011; Harland et al., 2021). In keeping with this finding, epiblast-mutant-Eomes embryos exhibit complete loss of DE marker expression, including Foxa2 and Cer1 (Arnold et al., 2008). Similarly, Eomes-deficient ESCs fail to colonise the DE and midline in chimeric embryos (Costello et al., 2011). Collectively, these findings point to an essential requirement for Eomes in specification of the discrete APS progenitor during mouse gastrulation.

1.7.2: Eomes regulation in the epiblast and PS

During gastrulation, Nodal and Wnt signalling pathways synergise to promote Eomes expression. In ESC-derived mouse embryoid body (EB) cultures, Smad2/3, the intracellular effector of Nodal (and Activin A, a Nodal surrogate) signalling, binds enhancers upstream (–10, –7 kb) and downstream (+8 kb) of the Eomes promoter (Simon et al., 2017; Wang et al., 2017). Additionally, the Wnt intracellular effector ß-catenin, in complex with Tcf3, binds the Eomes –10 kb and +8 kb enhancers as well, in a p53/73-dependent manner (Wang et al., 2017). However, from *in vivo* deletion experiments, it appears that *in vivo* only the downstream (+8 kb) enhancer is required for appropriate Eomes expression in the PS (Simon et al., 2017). ß-catenin KD in ActA-treated ESCs also results in significantly reduced Eomes transcript levels (Funa et al., 2015) (Figure 1.9). ATAC-seq experiments, in conjunction with Next Generation Capture-C approaches, have identified multiple additional putative enhancers associated with Smad2/3 occupancy that become accessible during EB differentiation (Simon et al., 2017). However, the contributions of those enhancer regions in regulating Eomes activation in the DE remains to be determined.

Nodal and Wnt signalling are similarly required for Eomes expression during human DE specification. EOMES is, as a result of ActA/Nodal signalling, bound by SMAD2/3 and directly upregulated by it. Knock-down of SMAD2 results in a reduction in Eomes expression, a result which is not replicated by knock-down of SMAD3 (Brown et al., 2011). Eomes subsequently physically interacts with SMAD2/3 to promote activation of the DE transcriptional network (Teo et al., 2011).

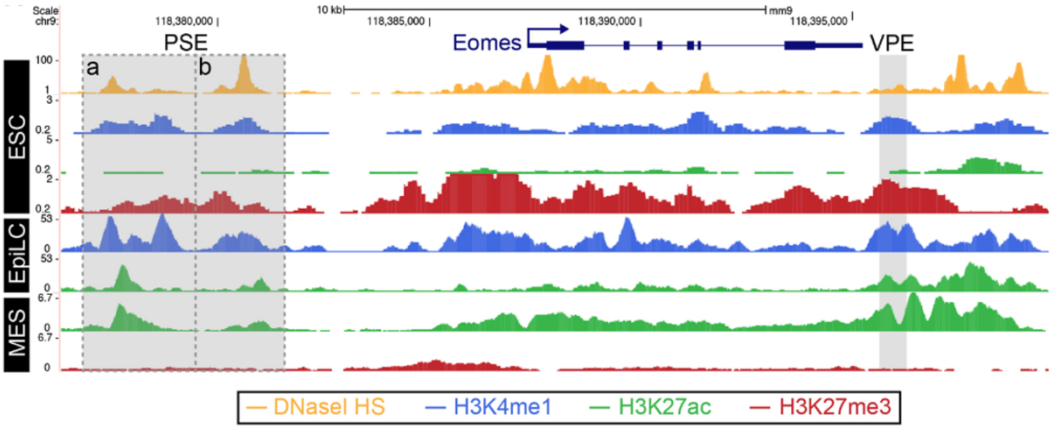


Figure 1.9: Regulatory regions proximal to the Eomes locus.
Eomes enhancer regions (denoted PSE_a/b, VPE in the figure) are bound by Nodal and Wnt signalling intracellular effectors (Smad2/3 and ß-cat/Tcf3 respectively), which in turn promote Eomes expression in the PS and during DE specification. Image adapted from Simon et al. (2017).

1.7.3: Eomes activities in the DE lineage

As discussed above, Eomes is essential for specification of the DE, as in its absence, in addition to a profound block in gastrulation and failure of epiblast cells to delaminate, there is a complete absence of definitive endoderm markers (Arnold et al., 2008).

Eomes appears to regulate expression of various endodermal genes, such as Mixl1, as demonstrated by Eomes KD experiments in DE-directed mESC cultures (Izumi et al., 2007). An epiblast-specific Eomes KO additionally results in complete loss of DE marker expression, such as Foxa2 and Cer1, from the mouse embryo (Arnold et al., 2008). Finally, in Eomes-null mESC-derived EBs cultured in presence of a high ActA concentration, Sox17 also shows significantly reduced expression (Tosic et al., 2019).

In parallel to DE gene activation, Eomes is additionally involved in repression of genes associated with pluripotency and the neurectoderm. Eomes and Brachyury (T) are jointly involved in specifying all mesodermal and DE lineages of the embryo, and double KO of these genes results in sustained expression of pluripotency markers (e.g. Nanog, Sox2, Oct4, Esrrb) and NE markers (e.g. Olig3, Gbx2, Neurod1). Eomes single KO also results in upregulation of mesodermal fate regulators, such as Foxc1 and Prrx1, highlighting its function as a regulator of the mesoderm vs. DE fate specification (Tosic et al., 2019) (Figure 1.10).

In summary, although we are beginning to appreciate the molecular mechanisms of Eomes activation in the PS/DE progenitors, we only have a limited understanding of the identity of key Eomes downstream targets required for DE and midline specification. Additionally, it remains unclear how Eomes controls the specification of such distinct lineages as the haemotopoietic, cardiac, and DE and midline within the very short developmental window of gastrulation.

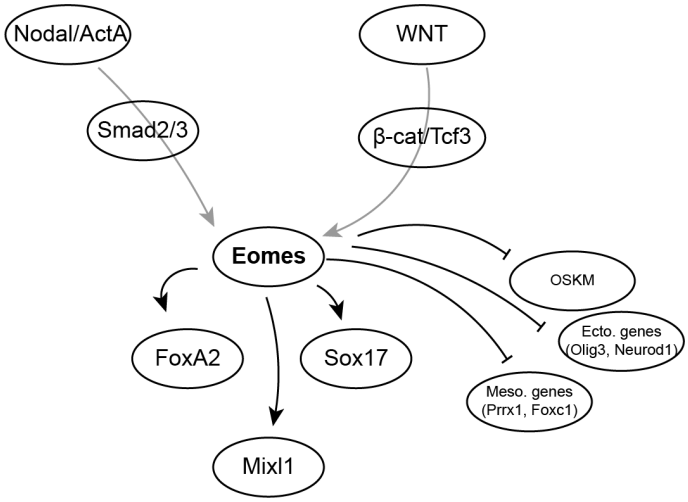


Figure 1.10: Summary of putative Eomes regulatory interactions in DE lineage.
Eomes expression in the PS is promoted by Wnt and Nodal signalling. Eomes appears to control expression of DE master regulators and other DE markers. Eomes may also downregulate a suite of ectodermal and mesodermal genes, as well as pluripotency markers Oct4, Sox2, Klf4, and Myc (OSKM).

1.7.4: Essential components of DE specification gene regulatory network

1.7.4.i: Foxa2

TF forkhead box 2 (Foxa2) (previously known as HNF3β) is essential for DE specification. It is expressed in the embryonic VE by day E5.5 (Kimura-Yoshida et al., 2007), as well as in the anterior portion of the PS, in the node and anterior midline, including the notochord, as well as the overlying neural plate the mid-streak stage (approximately E7.0) (Sasaki & Hogan, 1993). Expression continues throughout gastrulation and early organogenesis along the midline and in the fore- and hindgut (Sasaki & Hogan, 1993). Loss of Foxa2 function results in two distinct defects: (i) failure of VE specification, which consequently leads to failure of PS/axial elongation (Dufort et al., 1998; Kimura-Yoshida et al., 2007), and (ii) defects in tissues of the embryo proper, specifically in the anterior midline (Dufort et al., 1998). Tetraploid complementation experiments with a WT Foxa2 in the visceral endoderm (thus circumventing the early VE defect) showed failure of specification of the node and notochord, as well as failure to specify the fore-and midgut (but not hindgut) within the DE lineage (Ang et al., 1993; Dufort et al., 1998; Sasaki & Hogan, 1993; Weinstein et al., 1994).

1.7.4.ii: Sox17

SRY-box transcription factor 17 (Sox17) is a critical component of endodermal specification specific to vertebrates (Alexander & Stainier, 1999; Clements & Woodland, 2000; Hudson et al., 1997; Kanai-Azuma et al., 2002). Similarly to Foxa2, Sox17 is initially expressed in the extraembryonic VE at pre-streak stages (Kanai-Azuma et al., 2002). At mid-streak, a novel expression domain appears at the APS, coinciding with the domain of DE ingression through the streak. At late streak (E7.5), Sox17 expression is also found in the parietal endoderm (PE) (Kanai-Azuma et al., 2002). Prior to the early head fold (EHF) stage (approx. E7.5-8.0), Sox17 expression is restricted to the VE and foregut, whereas by the EHF stage, expression in those domains has become reduced and by the 7-8-somite

stage (approx. E8.5) become intensified in the mid- and hindgut (Kanai-Azuma et al., 2002).

Although Sox17 is expressed in the PE and extraembryonic VE during gastrulation and early organogenesis, Sox17-null embryos do not exhibit early developmental defects, possibly due to co-expression of, and possible functional compensation by, the closely-related Sox7 in this tissue (Kanai-Azuma et al., 2002). Later, however, Sox17-null embryos exhibit significantly reduced numbers of DE cells, with a FoxA1 (endodermal marker) domain restricted to a more medial region than normal. Concomitantly, more lateral regions of the embryonic endoderm seem to be colonised by VE cells. Subsequently the foregut is significantly restricted in size, and the mid- and hindgut degenerate into a cord-like structure (Kanai-Azuma et al., 2002). Cell apoptosis is significantly increased after the late neural plate stage (approx. E7.5-8.0), explaining the reduced gut tube size. Finally, Sox17-null cells introduced into WT embryos to form chimeras are shown to be able to contribute to foregut structures, albeit in limited numbers, but did not contribute to mid- and hindgut structures in the resultant chimeric embryos (Kanai-Azuma et al., 2002).

1.7.4.iii: Mixl1

Mix paired-like homeobox (Mixl1) is expressed initially in the VE and PS. Although Mixl1-null embryos do express DE markers Cer1 and Sox17, their domain is significantly reduced, and embryos exhibit gross morphological abnormalities in their anterior midline and DE structures. Additionally, Mixl1-null cells in embryonic chimera experiments show a very minor contribution to the foregut, and almost none at all to the mid- and hindgut structures (Hart et al., 2002). Similarly, Mixl1-KD mESCs subjected to a DE differentiation protocol show a notable reduction in the generation of Gsc^{+}/$Sox17^{+}$ DE cells (Izumi et al., 2007).

Data from Smad2/3-null mouse teratomas and EBs suggest that they regulate Mixl1, as Mixl1 levels are significantly reduced relative to the WT (Aragón et al., 2019; Liu et al.,

2016). Mixl1 is additionally found downstream of Eomes in the DE gene regulatory hierarchy, as Eomes KD in mESCs results in reduced Mixl1 expression levels, but the reverse effect (i.e. a reduction in Eomes expression levels upon Mixl1 KD) is not observed (Izumi et al., 2007).

1.7.4.iv: GATA transcription factors

Gata TFs are also essential for endodermal marker expression and appropriate endodermal tissue morphogenesis. Gata4 and Gata6 are initially expressed early in mouse embryonic development, prior to PS formation, with Gata4 appearing in the visceral and parietal endoderm at E5.5, and Gata6 being expressed a day later in both (Morrisey et al., 1998). Interestingly, Gata6-null embryos die soon after expected expression of Gata6, appearing severely growth-restricted and with Gata4 and extraembryonic endoderm marker HNF4 (hepatocyte nuclear factor 4) expression also significantly downregulated (Morrisey et al., 1998).

EBs generated with Gata6-null mESCs show significantly reduced expression of Foxa2. Additionally, Gata6-null cells introduced into WT blastocysts fail to contribute to bronchial epithelium, a DE-derived tissue (Morrisey et al., 1998). In human pluripotent stem cells, GATA6 directly binds to and activates DE genes, and knock-down of GATA6 in DE-directed differentiation of hESCs results in large-scale apoptosis and reduction of DE gene expression (Fisher et al., 2017).

In human ESC differentiation towards DE, GATA6 is regulated through the interaction of a lncRNA, GATA6-AS1 ("GATA6 antisense1") and SMAD2/3, with SMAD2/3 binding the GATA6 promoter in a GATA6-AS1-dependent manner (Brown et al., 2011; Yang et al., 2020).

During DE specification and regionalisation, Gata TFs proceed to activate endoderm-specific markers collaboratively with Foxa2, which displays pioneer factor characteristics (Cernilogar et al., 2019).

Notably, forced expression of Gata4 or Gata6 in mESCs results in transdifferentiation into XEN cells (Niakan et al., 2013; Shimosato et al., 2007). However, Gata4 is dispensable for PrE specification *in vivo* (Kuo et al., 1997; Molkentin et al., 1997).

1.7.4.v: Foxa2 and Sox17 regulation

Mixl1 appears to upregulate PS markers, such as Gsc, as well as DE marker Foxa2, as demonstrated through generation of inducible-Mixl1 mESCs. Upon Mixl1 induction in differentiating cells (where LIF has been removed from the medium), Gsc and Foxa2 exhibited premature expression relative to non-Mixl1 induced mESCs (Zhang et al., 2009). Additionally, mESCs constitutively expressing Mixl1 contain a higher proportion of Foxa2/Sox17/E-cadherin-positive epithelial cells, consistent with increased definitive endoderm production (Lim et al., 2009).

In a human ESC DE differentiation model, SMAD2/3 directly binds, along with EOMES, to both SOX17 and FOXA2 genes. WNT/ERK signalling also results in upregulation of both genes (Kim et al., 2011; Teo et al., 2011; Yang & Jiang, 2020). Loss of Eomes results in downregulation of both SOX17 and FOXA2, along with MIXL1 (Teo et al., 2011).

Conditional deletion of ß-catenin in mouse embryos indicates dependence of Sox17 expression on WNT signalling, as loss of ß-catenin results in loss of Sox17 expression and resulting defects in mid- and hindgut (as well as additional defects in the VE and organiser) (Engert et al., 2013).

Wnt3a/ActA signalling regulates global histone acetylation and subsequent gene expression via a set of miRNAs, including regulation of Sox17 and Foxa2 in DE-differentiated mESCs. Administration of HDAC inhibitors results in significant increase in DE specification (Fu et al., 2011).

During endodermal differentiation of human ESCs, FOXA2 additionally appears to be regulated by the methylation of three CpG islands found near its promoter region. This CpG methylation is accompanied by histone 3 lysine 27 trimethylation (H3K27me3), which

is classically associated with gene silencing. However, in the case of FOXA2 during DE differentiation, this seems to be directly associated with FOXA2 activation, as inhibition of DNA methyltransferase action resulted in significant reduction in FOXA2 expression levels. This effect was not shared by SOX17 and DE cell surface marker CXCR4 (Bahar Halpern et al., 2014).

Sox17 directly regulates Foxa2 in *Xenopus laevis* embryos during endodermal specification (Sinner et al., 2004). Similarly, in Sox17-null EBs produced from mESCs, Foxa2, as well as Gata6 and Gata4, are downregulated (Niakan et al., 2010).

GATA6 also regulates expression of FOXA2 and SOX17 in human ESC. DE differentiation using a line which produces only a truncated GATA6 protein shows a marked reduction in expression of both DE markers, which is rescued upon lentiviral transduction of the cells with a full-length GATA6 transgene (Tiyaboonchai et al., 2017).

1.8: Tools and techniques

1.8.1: Embryoid body culture

The study of embryonic development sees various limitations imposed on it, both physical as well as ethical. The size and limited cell number of embryos, especially at the earliest stages of development, makes them difficult to access, manipulate, and extract sufficient material from. Additionally, from an ethical standpoint it is highly desirable to reduce to a minimum the number of pregnant mice and embryos sacrificed in pursuit of scientific research. The isolation of pluripotent ESCs has therefore been revolutionary in our ability to model *in vitro* cell type specification and tissues formation. ESCs have the ability to contribute to all somatic cell types upon blastocyst injection, including the germ line, and can also give rise to cells of all three germ layers *in vitro* by forming embryoid bodies (EBs) (Evans & Kaufman, 1981; Itskovitz-Eldor et al., 2000) (Figure 1.11). Besides mice

and humans, ESCs have been isolated from several other species, including rat, pig, rabbit, and cow (Betts & Tobias, 2015; Choi et al., 2019).

In the intervening decades, highly elaborate protocols have been developed to manipulate EBs and preferentially guide their cells to differentiate toward a lineage/cell type of interest (Coucouvanis & Martin, 1995; Martin & Evans, 1975). If left to develop in the absence of LIF, EBs will generally differentiate to form an aggregate comprised of various cell types, based on interactions and autocrine/paracrine signals released by the cells themselves. Addition of signalling proteins/molecules, or indeed signalling inhibitors, which mimic the position within the Epi that cells find themselves in, will result in preferential activation of specific transcriptional networks – for example, addition of exogenous Wnt to the culture medium will result in generation of a PS-like cell population, while suppression of Wnt signalling will result in preferential specification of a neurectodermal cell population. High levels of ActA will result in higher levels of DE cell specification (Brickman & Serup, 2017). Cells generated within differentiating EBs are highly similar to cells produced during embryonic development, and EB differentiation protocols have been developed that permit production of terminally-differentiated cell types, such as cardiomyocytes, pancreatic cells, and blood cells (Boheler et al., 2002; Harland et al., 2021; Shahjalal et al., 2018).

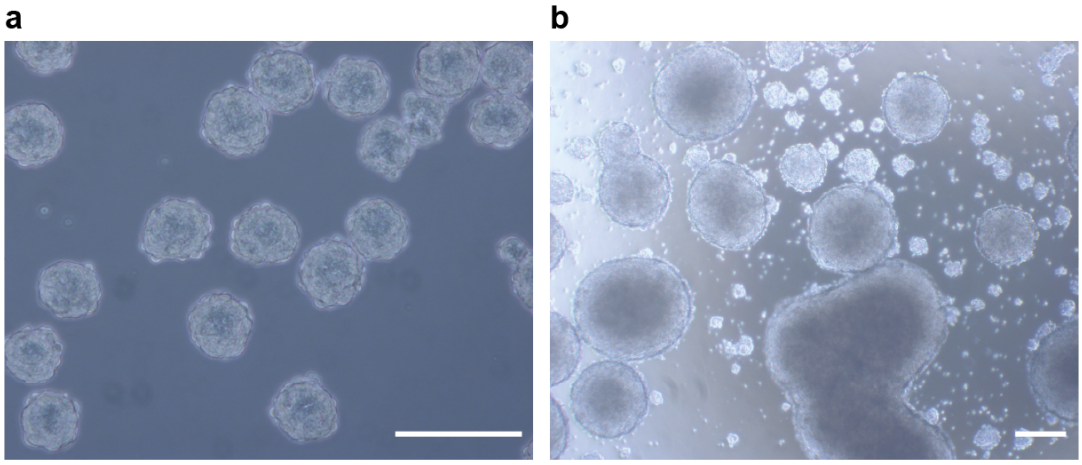


Figure 1.11: Embryoid bodies.
EBs after 1 (a) and 2 (b) days of mouse ESC differentiation. Scale bar = 500 μm.

1.8.2: Genome editing techniques

Genetic modification technologies have advanced greatly over the past decades, especially so since the discovery of bacterial anti-viral defence system CRISPR-Cas and its application to genome editing. The earliest tools employed to achieve genome editing include meganucleases (dsDNA-cleaving proteins with very long and specific recognition sites) (Epinat et al., 2003; Rouet et al., 1994), zinc-finger nucleases (ZFNs) (Bibikova et al., 2002; Porteus & Baltimore, 2003), and transcription activator-like effector nucleases (TALENs) (Boch et al., 2009; Christian et al., 2010). The drawback to these tools is that targeting specific genomic sites requires re-engineering of the protein sequence itself, which is time- and resource-consuming.

This is where the Cas protein family offers a revolutionary alternative. Bacterial Cas9, and other closely related Cas proteins, can cut both strands of DNA in a sequence-specific manner – however, the sequence specificity is not conferred by the protein sequence itself, which remains constant, but by a variable RNA sequence known as the crispr RNA (crRNA) or single guide RNA (sgRNA) (Figure 1.12) (Cong et al., 2013; Gasiunas et al., 2012). The ability to confer sequence specificity via changing an RNA sequence rather than the amino acid residue sequence of the nuclease makes the process of generating the desired genome-editing nucleases much easier, faster, and cheaper. Cas9-based genome editing also possesses great specificity, especially as computational tools have been developed to assist researchers with designing sgRNAs with very low off-target editing rates. The ability to target near any desired genomic region with a double-stranded break can be combined with providing a DNA repair template containing homologous 5' and 3' arms to the targeted region. Double-stranded breaks can thus be exploited to generate knock-in (KI) cell lines. These lines are stably modified to contain a specific sequence of interest in the desired locus, permitting endogenous expression of modified peptide sequences or co-expression of markers (Anzalone et al., 2020; Khalil, 2020).

More recently, Cas9 proteins have been modified to expand their functions, for example to use them as transcriptional activators or silencers by inactivating their nuclease domains and fusing them to transcriptional activators or repressors, to use them as "nickases" by only inactivating one of the two nuclease domains, or to use them as base "editors" to generate specific point mutations (Anzalone et al., 2020; Khalil, 2020).

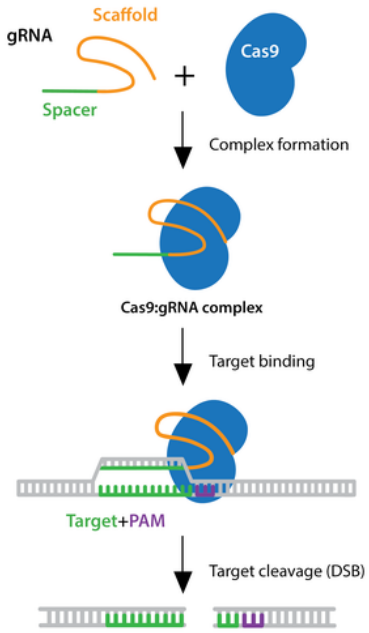


Figure 1.12: Cas9 DNA targeting mechanism.
Cas9 can be directed to target any genomic region through modification of the spacer (green) sequence of the sgRNA. After double-stranded cleavage, a provided repair template can be provided to KI a sequence of interest. Figure adapted from Addgene's CRISPR guide (https://www.addgene.org/guides/crispr/).

1.8.3: The dTAG degron system

The function of specific genes and proteins within the context of a biological process has been studied for decades through KO (stable or conditional genomic deletion) or KD (targeting gene transcripts for silencing or degradation), as well as overexpression systems. These enable researchers to determine what effects the loss of a particular gene or protein might have in a particular system, or conversely what the outcome of ectopic or excessive expression of a gene or protein might be. These studies have contributed enormously to our knowledge and understanding of biological systems, specifically so in

the context of developmental and stem cell biology (Zimmer et al., 2019). However, they may present various drawbacks which make them inappropriate tools to answer all research questions. For instance, conditional gene deletions (for example, via the Cre-loxP system) are sometimes incomplete (Bao et al., 2013; Turlo et al., 2010), as is often the case with gene knock-downs. Additionally, conditional deletion of a gene may only result in loss of its protein product several hours if not days later, making the system unworkable for the study of immediate, direct effects of protein loss. Additionally, compensatory mechanisms in cells and secondary effects from disruption of a protein's downstream targets or protein partners may further confound the results of the study. Finally, the deletions are irreversible, as the gene of interest is permanently excised from the genome.

These issues can be addressed through the use of degron systems, which have been developed in the past decade. A "degron tag" is a term for a peptide moiety attached usually at the N- or C-terminus of a protein of interest (POI), which permits its rapid and specific degradation upon administration of a degradation-inducing molecule. This degron tag is introduced through genomic editing, allowing stable expression of the degron-tagged POI. Various such systems have been developed: GFP tag-based protein degradation (Caussinus et al., 2012), auxin-inducible degradation (Holland et al., 2012; Morawska & Ulrich, 2013; Nishimura et al., 2009), and degradation based on the dTAG family of synthetic small molecules (Nabet et al., 2018, 2020) (Figure 1.13).

The dTAG system was developed on the basis of previous work redesigning the ligand binding interface of endogenous protein FKBP12. A single amino acid residue modification (F36V) results in an additional cavity that allows synthetic ligands to exclusively bind the modified protein (Clackson et al., 1998). Thanks to this property, the modified FKBP12^{F36V} has since been employed as a 'degron tag' on proteins of interest (POIs). Nabet et al. (2018, 2020) designed heterobifunctional small molecules (dTAG-13 and dTAG^{V}-1) that mediate an interaction between the degron tag and an E3 ubiquitin ligase complex (CRL4-

CRBN in the case of dTAG-13, and VHL in the case of dTAG^{V}-1). Addition of one of these small molecules to the culture medium results in the rapid and specific polyubiquitination and proteasomal degradation of the tagged POI. This process occurs on the scale of minutes-to-hours, such that a protein may be fully degraded within 30-60 minutes.

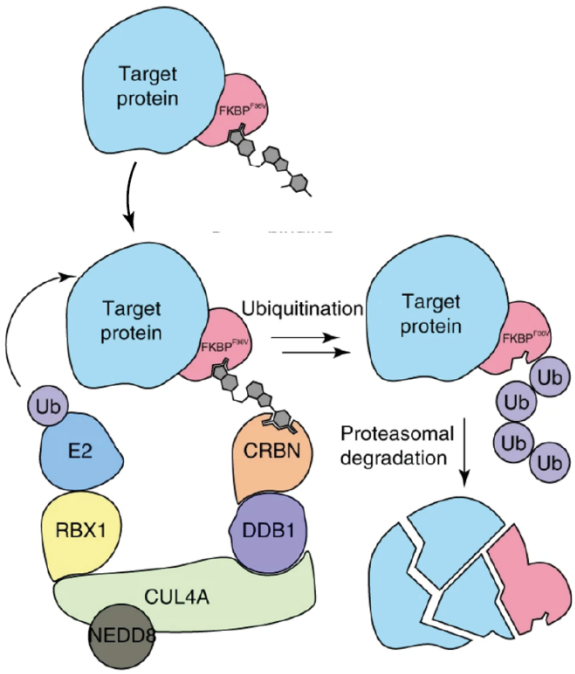


Figure 1.13: Schematic of the dTAG degron system.
The protein of interest is tagged by insertion of an FKBP12^{F36V} ("degron tag") coding sequence at the POI's endogenous locus. The expressed chimeric protein can then be inducibly degraded by addition of small molecule dTAG-13 or dTAG^{V}-1. The dTAG molecule contains a moiety that binds the degron tag, and a moiety that interacts with an E3 ubiquitin ligase. This results in ubiquitination of the POI and rapid proteasomal degradation. Figure adapted from Nabet et al. (2018).

1.8.4: Genome-wide techniques

With the development and commercial availability of next-generation sequencing (NGS), it has become possible to survey biological systems on a genome-wide scale. RNA-seq makes it possible to assess the full transcriptome of a group of cells at a given moment (Stark et al., 2019). Assay for transposase-accessible chromatin (ATAC)-seq enables the researcher to survey chromatin accessibility across the genome (Buenrostro et al., 2013).

Although these techniques are powerful, their resolution can be greatly enhanced by applying them to single cells. scRNA- and scATAC-seq enable researchers to disentangle the “average” signal arising from a population of cells and assess the state of individual cells (Haque et al., 2017; Xu et al., 2021).

These techniques have more recently been leveraged in combination with each other in order to probe the state of individual cells across different *omics* modalities. Countless combinations of single-cell multi-omics methods have been developed recently, probing various combinations of single cell transcriptome, chromatin accessibility, chromatin occupancy of proteins, epigenomic markers, and chromatin methylation state (extensively reviewed in Vandereyken et al., 2023).

These techniques have proven very useful in developmental biology, as they enable researchers to examine *bona fide* embryonic cells, with low cell numbers no longer posing the barrier that they once did. Several groups have recently published scRNA-seq atlases of various developmental stages characterising not just normal development and cell fate acquisition (Chan et al., 2019; Cheng et al., 2019; He et al., 2020; Ibarra-Soria et al., 2018; Mittnenzweig et al., 2021; Pijuan-Sala et al., 2019), but also developmental defects of knock-out models at the single-cell level (Grosswendt et al., 2020; Lescroart et al., 2018). In addition to providing a detailed view of the transcriptional state of each cell type generated during embryonic development, these atlases can be used to “map” transcriptional data from other cells based on their transcriptional similarity to the cells used to generate the atlas, in order to determine their identity and the developmental stage they correspond to, for instance as Mittnenzweig et al. (2021) did with the mouse gastrulation atlas from Pijuan-Sala et al. (2019).

Multi-ome technologies have similarly been employed to assess wild-type and mutant development in more than one modalities (Argelaguet et al., 2019; Clark et al., 2022; Guo et al., 2017; Jiang et al., 2023; Wang et al., 2021). For example, single-cell multi-

modal readouts in a recent RNA/chromatin accessibility atlas of early mouse organogenesis have been used to generate predictive models on transcription factor binding behaviour in cells. This was used to predict the effect of TF perturbations *in silico*, and experimental validation demonstrated the requirement of TF Brachyury for differentiation of neuromesodermal progenitors into somitic mesoderm (Argelaguet et al., 2022).

1.8.5: Using ESCs to generate novel genetically altered mouse strains

In order to assess the role of a gene or protein in a biological process, often the gold standard has been to manipulate the GOI/POI in live animals (Costantini & Lacy, 1981; Gordon et al., 1980).

As introduced earlier in this Chapter, pluripotent ESCs can readily be isolated from pre-implantation mouse blastocysts (Evans & Kaufman, 1981). In addition to their numerous uses in *in vitro* systems, ESCs can be manipulated by Cas9-based approaches, and the resulting genetically altered (GA) ESCs used to generate novel mouse strains (Wang et al., 2013). In brief, ESCs can be injected into mouse blastocysts and these subsequently transferred into pseudopregnant surrogate females (Dewar, 1957; Hasegawa et al., 2017). ESC derivatives contribute to all somatic tissues in the resulting live born chimeric animals, and in a high proportion of these, the ESC will also contribute to the primordial germ cell lineage (Du et al., 2019). If the GA ESCs contribute to the germline, then crossing chimeric animals with WT partners results in heterozygous carried offspring that can then be intercrossed to generate embryos and adults that are homozygous for the genetic alteration (Du et al., 2019) (Figure 1.14).

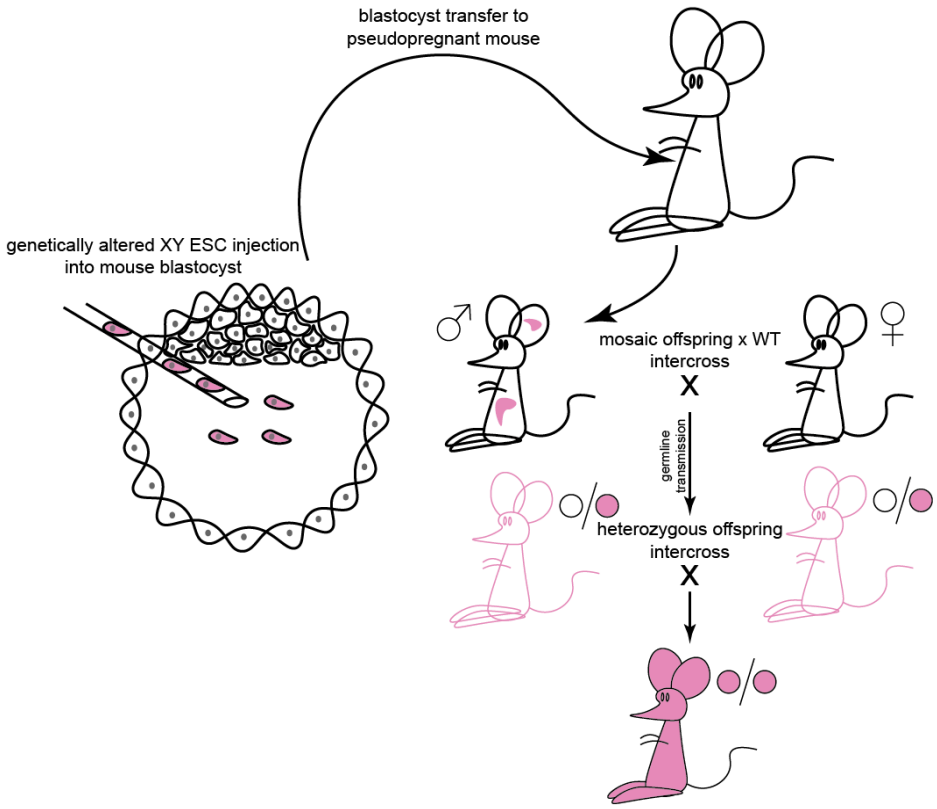


Figure 1.14: Schematic of steps to make a genetically modified mouse line.
Genetically modified ESCs are injected into a blastocyst, which is transferred into a pseudopregnant mouse. The resulting mosaic offspring are crossed to WT individuals, and if the GM cells have contributed to the germline, these crosses can result in heterozygous offspring. These offspring can be further crossed to generate homozygous individuals.

1.9: Overview and aims of this DPhil thesis

Although Eomes is essential both for the maintenance of the TSC lineage as well as the specification of the DE, the molecular mechanisms by which these processes are achieved remain unclear.

In order to further elucidate these questions, over the course of my DPhil I have constructed an Eomes-degron allele in mESCs using the recently-developed dTAG system, permitting me to inducibly and rapidly degrade Eomes protein. This genetic modification was also introduced to a mouse line, in order to extend the study *in vivo*. These tools, as well as a range of techniques, have allowed me to study the functions of Eomes at a high temporal resolution. The main questions I aim to answer within this thesis are (Figure 1.15):

1) How does Eomes contribute to the maintenance of the trophoblast stem cell lineage of the mouse embryo?
2) How does Eomes function to specify definitive endoderm progenitors of the mouse embryo during gastrulation?
3) More broadly, how does Eomes function within the context of such different embryonic lineages?

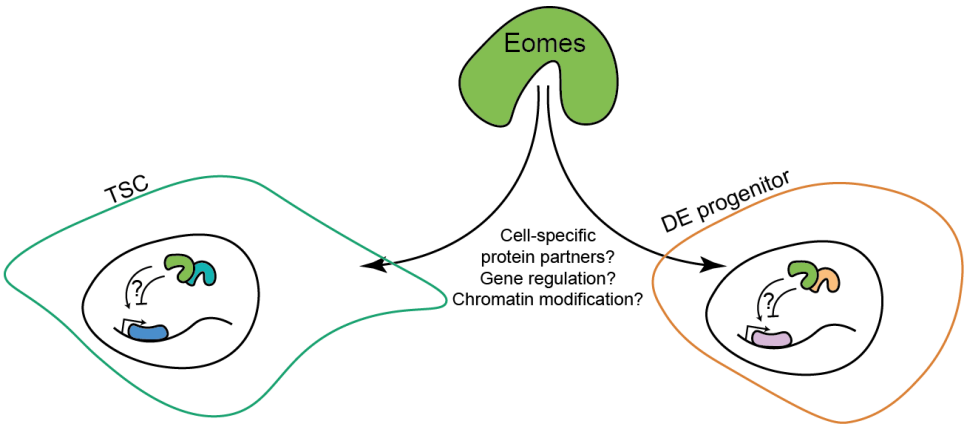


Figure 1.15: Summary of my DPhil's central questions.
How does Eomes perform such different functions as promoting the maintenance of a stem cell lineage (TSCs) and the specification of a transient progenitor population (DE)?

CHAPTER 2: GENERATION OF AN INDUCIBLE EOMES DEGRADATION SYSTEM *IN VIVO* FOR EMBRYO ANALYSIS

2.1 INTRODUCTION

The study of embryonic development is not without its challenges, as researchers must allow embryos to develop with as little perturbation as possible, while still collecting the data they require. This is easier said than done particularly in the case of mammals, where the embryo develops in a largely inaccessible uterus. However, modern developmental biology techniques offer an explosive array of investigative options to overcome these difficulties, from high-powered microscopes that permit observation of 3D structures on the multi-µm scale, to fluorescent markers, to multi-day *ex vivo* embryo culture that permits normal development (Aguilera-Castrejon et al., 2021; McDole et al., 2018). Combining these techniques has led to extraordinary insight into development at unprecedented temporal resolution and at the single-cell scale.

Eomes, the T-box TF at the core of my DPhil projects, is initially expressed in the mouse embryo in the TE of the pre-implantation blastocyst (Russ et al., 2000). Its necessity for embryonic survival becomes apparent at the peri-implantation stage, as it is essential for the maintenance of the TE lineage stem cell pool (the TSCs), and Eomes-null embryos die at peri-implantation. However, Eomes is also required for the specification of several embryonic lineages during post-implantation development, including the cardiac lineage (Costello et al., 2011), the anterior mesendoderm (Costello et al., 2015), the definitive endoderm (DE) (Arnold et al., 2008), and the yolk sac haemogenic endothelium (Harland et al., 2021), among others. This means that for much of Eomes' window of action, the mouse embryo is not easily accessible.

Eomes' requirement in post-implantation embryos has been examined, among others, through an epiblast (Epi)-specific deletion allele (Arnold et al., 2008). A Cre protein expressed from an epiblast-expressed Sox2 regulatory element (Hayashi et al., 2002)

allows conditional Eomes alleles ($Eomes^{CA}$) ("floxed," i.e. with LoxP sites flanking exons 2-5) to be rendered functionally null (Arnold et al., 2008; Sauer, 1998) exclusively in the Epi. Therefore, Sox2.Cre; $Eomes^{null/+}$ x $Eomes^{CA/CA}$ mouse crosses permit the generation of embryos that circumvent the pre-implantation Eomes-null defect in extraembryonic tissues, thereby permitting investigation of Eomes' Epi functions. This study demonstrated for the first time that Eomes is essential in specification of the DE lineage and proper gastrulation (Arnold et al., 2008). However, Cre-Lox systems have been demonstrated repeatedly to delete floxed alleles incompletely (Bao et al., 2013; Turlo et al., 2010). Additionally, the options for the timing of allele deletion are rigid, as they depend on the timing and localisation of other genes' expression (in this case, Sox2). There do exist inducible Cre-Lox systems (e.g. tamoxifen-inducible $CreER^{T2}$ systems), but they have been reported to be leaky, activating floxed site excision even in absence of the inducer (Álvarez-Aznar et al., 2020). Finally, Cre-mediated deletion is permanent, with the excised sequence irretrievably degraded by the cell's vast array of nucleases. This is inconvenient, but even worse, it can also give misleading results in those occasions when the excised sequence is retained in the cell, potentially continuing to be expressed, particularly in quiescent cells (Leonhard et al., 2008; Srivastava & Ow, 2003; Turlo et al., 2010).

In order to overcome these limitations, and to enable us to further study the roles of Eomes in the developing mouse embryo, we decided to adapt the recently-developed $FKBP12^{F36V}$ degron tag method (the "dTAG system") (Nabet et al., 2018). We designed an allele for inducibly-degradable Eomes, which would permit precise elimination of Eomes protein at any chosen moment, following exposure to small molecule of the dTAG family (dTAG-13 or $dTAG^{V}$-1). The dTAG system has recently been demonstrated to be functional in mouse embryos, both via *ex vivo* culture and *in utero* dTAG injections, in a system tagging endogenous Nelfb (negative elongation factor b), a negative regulator of RNAPII-mediated transcription (Abuhashem et al., 2022a,b). Our aim was to improve upon the

Cre-based method, by adapting a system that permits inducible, and complete, protein degradation; is reversible; can be employed *in utero* without physical perturbation of the embryo; and permits examination of downstream effects in a time-sensitive manner.

Here, I describe my design of a degron-tagged Eomes allele, the Cas9-based targeting of mESCs with it, and the *in vitro* validation of the tag's functionality. I also describe the generation of a mouse line harbouring the Eomes^{deg} allele, and the *in vitro* and *in vivo* assays to investigate the effects of Eomes protein loss on embryonic development. We compared the performance of small molecule dTAG-13 with dTAG^{V}-1, a more recent addition to the FKBP12^{F36V} degron system. Whereas dTAG-13 mediates specific interactions between E3 ubiquitin ligase CRBN (cereblon) and a degron-tagged protein of interest, dTAG^{V}-1 does so between the VHL (Von Hippel-Landau) E3 ubiquitin ligase complex and a degron-tagged protein, thereby permitting degradation of proteins recalcitrant to CRBN-mediated degradation (Nabet et al., 2020). dTAG^{V}-1 was also reported to show improved pharmacokinetic and pharmacodynamic properties relative to dTAG-13, having a greater half-life and exposure (drug levels achieved in the body) than dTAG-13 (Nabet et al., 2020).

Finally, I discuss the problems and limitations encountered in the validation of this system. Although *ex vivo* degradation of Eomes is successful and Eomes degradation in blastocysts results in a phenocopy of the Eomes-null phenotype previously reported (Russ et al., 2000), *in vivo* degradation gives inconsistent results which we hypothesise may be due to inconsistent ability of dTAG-13 to reach embryos at early post-implantation/early gastrulation stages, perhaps since trophoblast invasion of the endometrium has only recently begun and is far from complete, making the (maternal) blood supply to the embryo scant (Hemberger et al., 2003). Alternatively (or additionally), a powerful feedback mechanism may drastically upregulate Eomes expression upon the protein's degradation in the system; or the sheer amount of Eomes expressed in the cells of the

early gastrula is too great to consistently deplete. Caution and methodical validation are therefore required when making use of systems such as the degron one in one's own research.

2.2 Allele design and generation of a degron-tagged Eomes mouse line

I designed and implemented a Cas9-based targeting system in order to insert the coding sequence for an XTEN peptide linker, the FKBP12^{F36V} degron tag, and an mCherry fluorescent reporter in the C-terminus of the Eomes coding sequence (Figure 2.1a). There were previous attempts in our lab to tag the N terminus of Eomes with a V5 tag. However, a western blot against N-V5-Eomes-C-expressing EBs with anti-V5 antibody did not recognise all three peptide species that Eomes typically produces, which made us concerned that some Eomes protein was not tagged, perhaps due to alternative splicing (unpublished). This informed our decision to target the C terminus of the coding sequence. I targeted both feeder-free (FF) and feeder-dependent (FD) mESCs with this strategy, picked clonal colonies of targeted cells, and screened them via PCR to identify homozygously targeted clones (Figure 2.1b). The FF cells were used for *in vitro* differentiation studies, while the FD cells were used for construction of the mouse line. I subsequently carried out EB differentiations in order to confirm that the Eomes-degron chimeric protein was expressed, functional, and amenable to degradation upon addition of dTAG-13.

I initially carried out the haemogenic endothelium (HE) EB differentiation previously described in Harland et al. (2021) (Figure 2.1c). I added 100nM dTAG-13 to day 3.0 HE EBs and assessed protein presence at various timepoints after the addition via western blot (WB). Within 1 hour of dTAG-13 addition, Eomes protein was undetectable via WB, an effect which persisted over several hours (the EBs were maintained in dTAG during that time) (Figure 2.1d). Luke Harland carried out intracellular flow cytometry for Eomes

protein in day 4.0, after two days of culturing EBs in dTAG-13, whereupon Eomes protein was also undetectable (Figure 2.1e). He additionally carried out live cell flow cytometry for mesodermal markers Flk-1 and PdgfRa (Figure 2.1f). In EB cultures with dTAG-13 addition, the Eomes$^{deg/deg}$ EBs phenocopy the Eomes-null EBs in failing to express Flk-1. In contrast, when DMSO vehicle control is added, the Eomes$^{deg/deg}$ EBs are indistinguishable from the WT EBs. Notably, mCherry reporter is also shown to be expressed is the Eomes$^{deg/deg}$ flow cytometry panels of Figure 2.1f, and expression persists even after dTAG-13 addition and Eomes-degron protein degradation. mCherry therefore also acts as a short-term lineage tracer indicating historic expression of Eomes in those cells.

I also carried out an EB differentiation optimised for generation of DE cells, following the protocol described in Chapter 5 of this thesis (Figure 2.2a). Addition of dTAG-13 on day 3.0 of DE EB differentiation results in Eomes protein being undetectable 30 minutes later, an effect which, as in the HE EB differentiation, lasts for multiple hours (Figure 2.2b). Moreover, adding dTAG-13 to DE EB cultures on day 2.5 of differentiation resulted in complete loss of the cKit^{+}/CXCR4^{+} population as judged by flow cytometry on day 5.0 (Figure 2.2c). Therefore, Eomes-degron protein is efficiently depleted from dTAG-13-treated EBs regardless of the differentiation protocol that the cells are subjected to.

During the *in vitro* assessment of the Eomes^{deg} allele's functionality, I did not carry out experiments using dTAG^{V}-1, as dTAG-13 appeared to be working very efficiently. As will be discussed in later sections of this Chapter, dTAG-13 was shown not to be fully functional during *in vivo* Eomes$^{deg/deg}$ embryo treatment via IP injections, as judged by the incomplete depletion of Eomes protein in those embryos. We therefore utilised dTAG^{V}-1 to assess whether *in vivo* Eomes depletion would be more efficient with this small molecule (which was ultimately not the case). However, it may be interesting to assess the relative efficiency of dTAG-13 and dTAG^{V}-1 *in vitro* as well, to establish if dTAG-13 still exhibits superior efficiency.

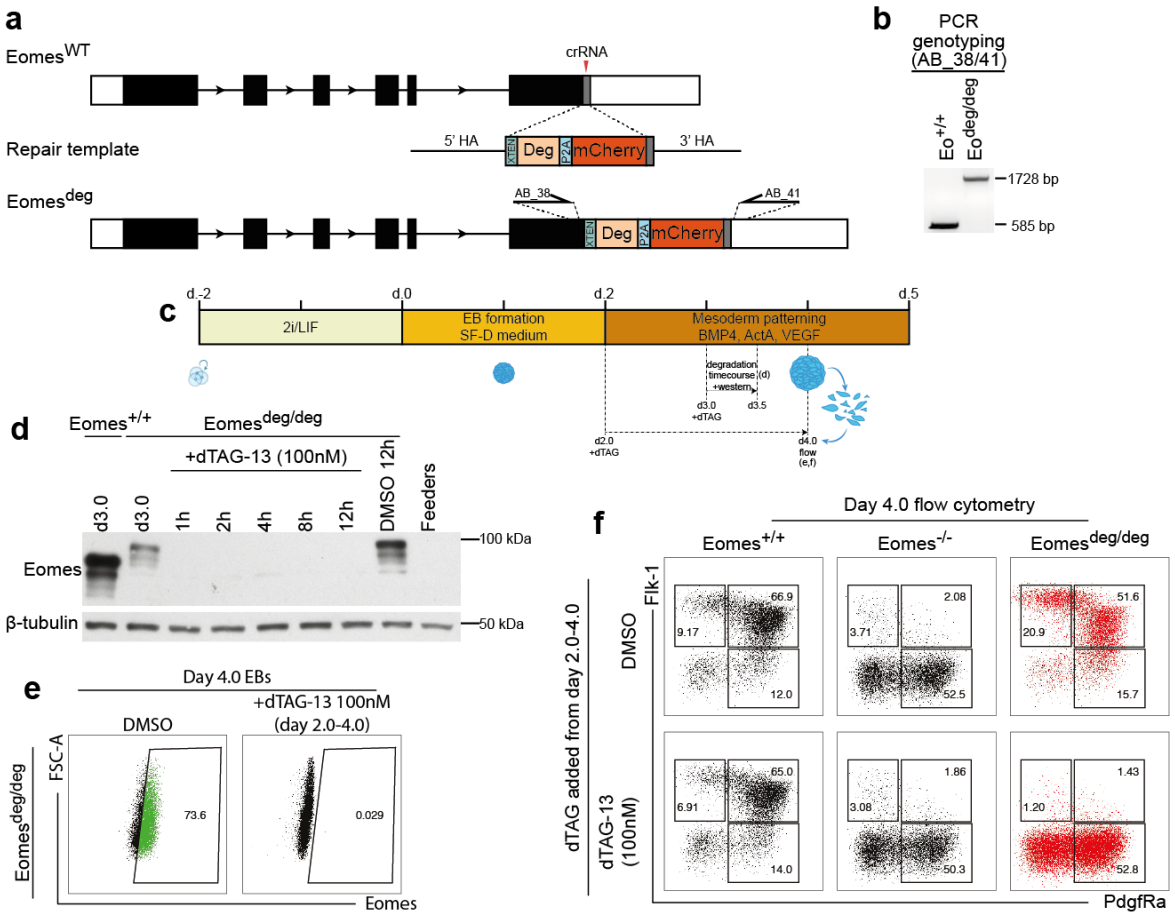


Figure 2.1: Generation and characterisation of mESC clones harbouring Eomes^{deg} alleles.

(a) Cas9-based targeting strategy for KI of Eomes^{deg} allele. A crRNA targeting the stop codon (in grey) of the endogenous Eomes locus was used along with a ssDNA repair template containing the sequence for the degron (FKBP12^{F36V}) tag, a 2A self-cleaving peptide, and an mCherry reporter. (HA: homology arm).

(b) PCR genotyping of WT and homozygously tagged clones with primers denoted in (a).

(c) Schematic of the haematoendothelial (HE) EB differentiation protocol described in Harland et al. (2021). Also indicated are the experimental procedures for the results presented in (d), (e), and (f).

(d) A western blot for Eomes protein after a degradation timecourse experiment in day 3 EBs demonstrates rapid loss of Eomes protein within 1h of dTAG-13 addition, with protein loss persisting after several hours. Note the size shift of the protein due to presence of the degron tag compared to Eomes$^{+/+}$ EBs.

(e) Intracellular flow cytometry in day 4.0 Eomes$^{deg/deg}$ EBs that have had dTAG-13 added to the culture medium on day 2.0 demonstrate loss of Eomes protein.

(f) Addition of dTAG-13 to homozygously Eomes-degron-expressing EB cultures results in a phenocopy of Eomes$^{-/-}$ embryoid bodies (EBs). dTAG-13 was added on day 2.0 of EB differentiation and maintained in the culture medium until live cell flow cytometry on day 4 for mesodermal/HE markers Flk-1 and PdgfRa. Eomes$^{deg/deg}$ EBs with dTAG-13 added fail to upregulate Flk-1. Cells in red are expressing mCherry reporter, indicating that mCherry continues to be expressed even after Eomes protein degradation.

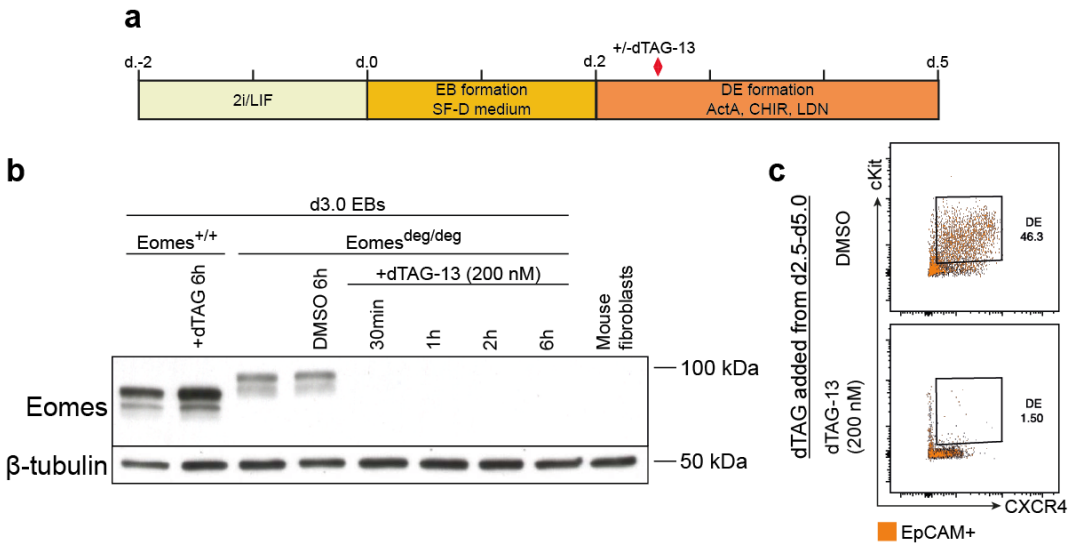


Figure 2.2: Eomes$^{deg/deg}$ DE EB differentiation.
(a) Schematic representation of DE EB differentiation protocol. The red diamond denotes the timepoint of dTAG-13 addition which results in complete loss of the cKit^{+}/CXCR4^{+} DE population, as shown in (c).
(b) Western blot for Eomes protein after a degradation timecourse experiment in day 3 EBs demonstrates rapid loss of Eomes protein within 30min of dTAG-13 addition, with protein loss persisting after several hours. Note the size shift of the protein due to presence of the degron tag compared to Eomes$^{+/+}$ EBs.
(c) Flow cytometry of day 5.0 DE EB cells shows a complete loss of cKit^{+}/CXCR4^{+} cells when dTAG-13 is added on day 2.5 of the differentiation protocol.

As mentioned earlier, the above differentiation experiments were carried out using FF ESCs. To test that the FD ESCs were also functionally degron-tagged, I weaned them from feeders by passaging them 3-4 times on feeder-free gelatinised plates, and carried out the DE differentiation protocol to confirm that the cells are behaving in a similar way to the previously-tested FF degron-tagged clones. Several clones differentiated as previously characterised, showing high, if variable, cKit$^+$/CXCR4$^+$/EpCAM$^+$ specification efficiency (Figure 2.3). Upon dTAG-13 addition on day 2.5, these EBs also lost their double-positive population, as expected. However, one clone (H3) did not specify a double-positive population in the DMSO control differentiation condition, suggesting that the degron tag repair template may have been integrated incorrectly, resulting in a non-functional allele. Based on these results, I chose two clones, B8 and G7 for blastocyst injection and mouse line generation.

Liz Robertson performed the blastocyst injections and transfer to pseudopregnant mice, and resulting male chimeric offspring were crossed with wild type females until heterozygous mice for the Eomes^{deg} allele were produced. These were then intercrossed to obtain homozygous individuals. The resulting mice are viable and fertile, transmit the Eomes^{deg} allele with Mendelian ratios, and, to the best of our knowledge, are normal and healthy.

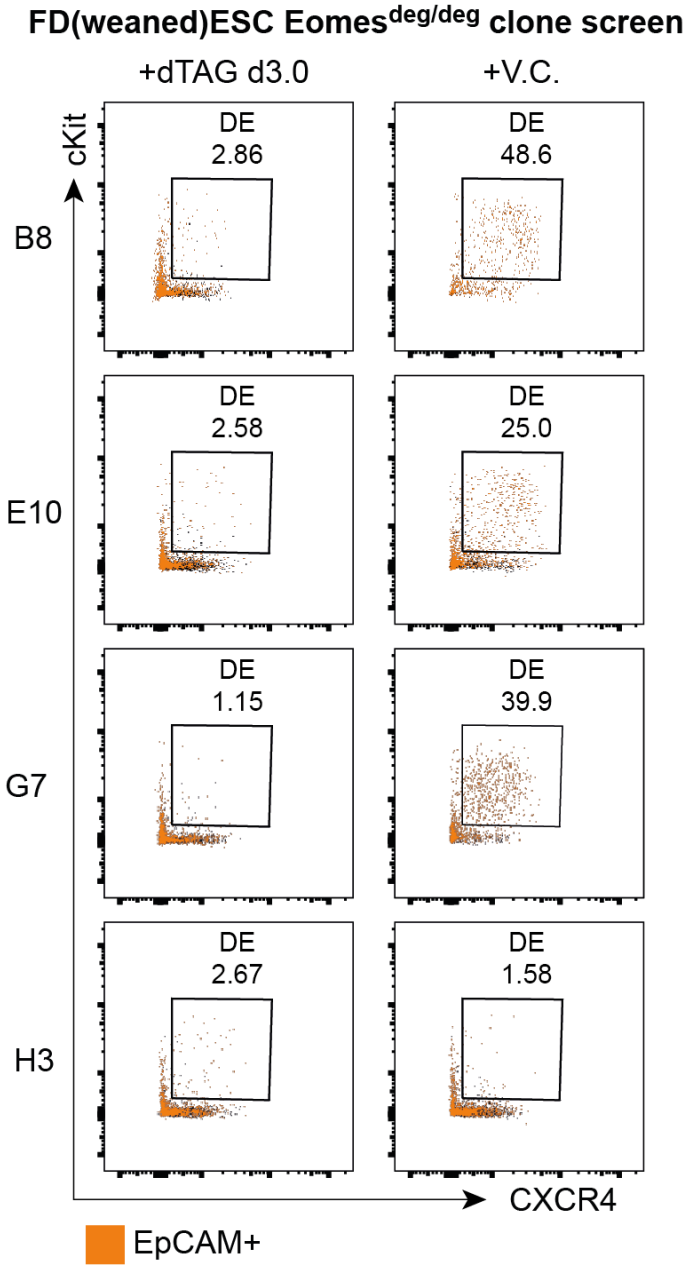


Figure 2.3: Flow cytometry of day 5.0 EBs made from weaned FD Eomes-degron-tagged mESC clones trialled for appropriate DE differentiation and degron tag functionality.

FD degron-tagged mESC clones were weaned from feeder layers and, after 3-4 passages on gelatin-coated plates, they underwent the DE EB differentiation protocol (Figure 2.2a). Although clones B8, E10, and G7 differentiated appropriately, showed efficient cKit$^+$/CXCR4$^+$ cell generation by day 5 of differentiation, and successfully lost their double-positive population upon day 2.5 dTAG-13 administration, one (H3) did not, and was therefore excluded from downstream use.

2.3 *Ex vivo/in vitro* manipulation of Eomes-degron-expressing embryos

Mouse intercrosses and immunofluorescence experiments were carried out by Maria-Eleni Xypolita and Ita Costello.

2.3.1: Ex vivo culture of pre-implantation Eomes$^{deg/deg}$ blastocysts

Homozygous Eomes$^{deg/deg}$ mice were crossed to obtain Eomes$^{deg/deg}$ blastocysts on E3.5 (the day of identification of a vaginal plug was considered as E0.5). These were isolated and treated *ex vivo* with dTAG-13, as well as dTAG^{V}-1 for 2h (Figure 2.4a). Both dTAG molecules resulted in rapid loss of Eomes protein from blastocysts' TE compartment, as it was undetectable by immunofluorescence (IF), in contrast to DMSO-treated blastocysts (Figure 2.4b). Cell IF was further quantified and showed notable decrease in the case of both dTAG treatments (Figure 2.4c). Furthermore, a 72h blastocyst outgrowth of E3.5 blastocysts was carried out both in DMSO control and dTAG treatment conditions (Figure 2.4d). DMSO-treated blastocysts showed attachment and normal trophoblast outgrowth. In contrast, this did not occur in dTAG-13- or dTAG^{V}-1-treated blastocysts, which did not attach or outgrow. This phenocopies Eomes-null embryos, which also show non-expansion of the TE (Russ et al., 2000).

Although this assay for assessment of trophoblast cell expansion is well-established (Cole & Paul, 1965; Russ et al., 2000; Sutherland et al., 1988), it can't be definitively determined without further assessment that the lack of expansion is solely due to a defect of the blastocyst's trophoblast compartment. Therefore, the result of the blastocyst outgrowth experiment could be further characterised by staining for trophoblast (stem and differentiation) markers to precisely determine the identity of the outgrowing cells, as well as perhaps assessing the area, cell numbers, strength of attachment or other characteristics of the outgrowth. Additionally, it may be useful to assess developmental defects or trophoblast marker expression after longer-term (24-48h) Eomes depletion to characterise Eomes' effect on TSC GRN members in more detail.

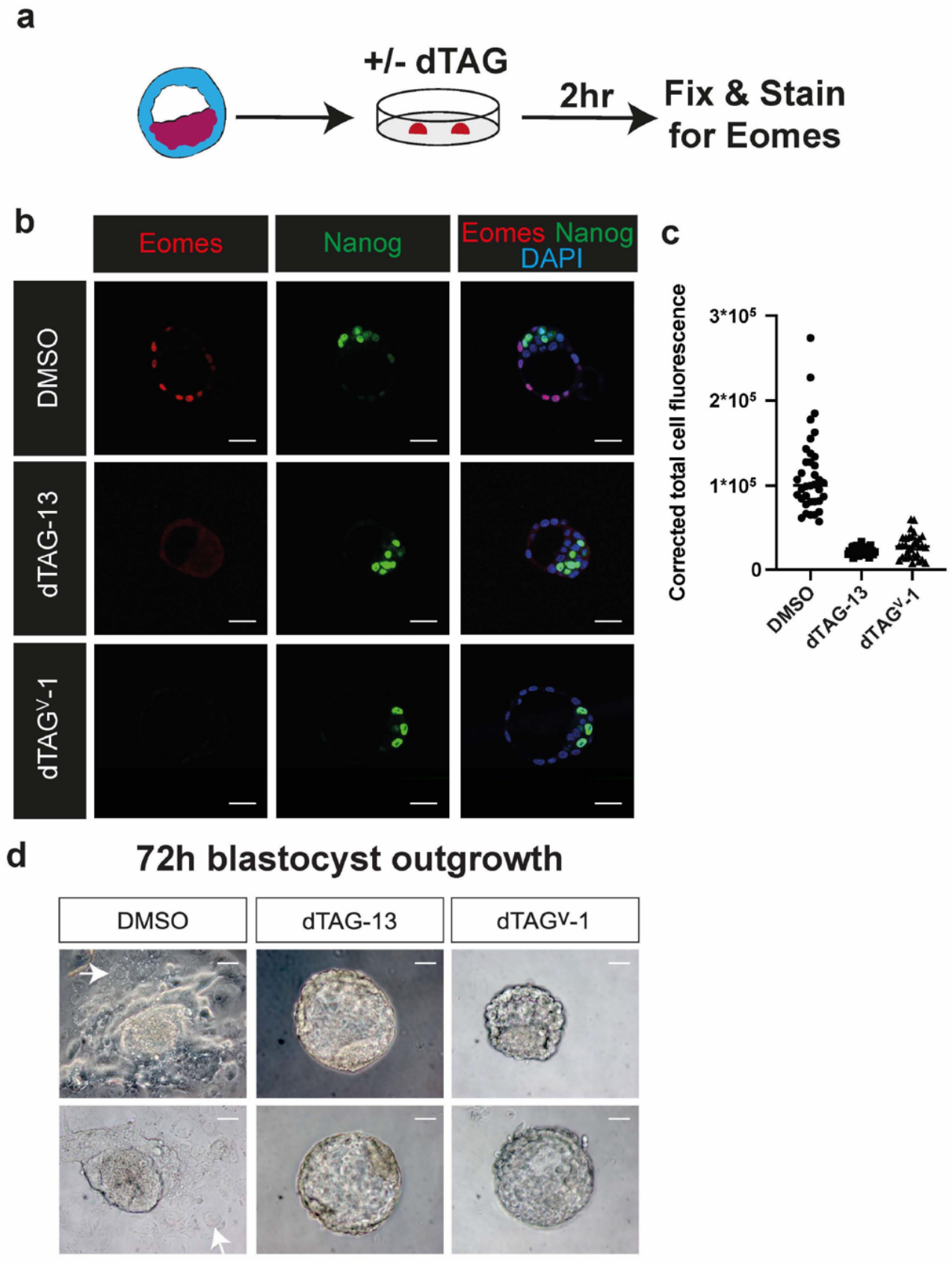


Figure 2.4 (legend on next page)

Figure 2.4: E3.5 pre-implantation Eomes$^{deg/deg}$ blastocysts lose Eomes protein and phenocopy Eomes-null blastocysts upon dTAG addition (500nM) in *ex vivo* cultures.
(a) Schematic representation of experimental design to test Eomes-degron degradation in E3.5 blastocysts. The TE is depicted in blue while the ICM is depicted in plum.
(b) Eomes and Nanog staining of DMSO-, dTAG-13- or dTAG^{V}-1 treated E3.5 pre-implantation blastocyst embryos. Eomes expression in the trophectoderm is lost in dTAG-treated embryos but not in control DMSO-treated embryos. Nanog marks the pluripotent epiblast and DAPI is a nuclear stain (n=5 DMSO, n=6 dTAG-13, n=5 dTAG^{V}-1 embryos). Scale bars = 50 μm.
(c) Quantification of Eomes expression in Eomes-degron blastocysts treated with DMSO, dTAG-13 or dTAG^{V}-1. Corrected total cell fluorescence is calculated as the Integrated Density – (Area of selected cell)*(Mean fluorescence of background readings) (n=5 DMSO, n=6 dTAG-13, n=5 dTAG^{V}-1 embryos).
(d) Cultured blastocysts from Eomes$^{deg/deg}$ homozygous intercrosses treated with DMSO, dTAG-13 or dTAG^{V}-1 for 72 hours. Eomes$^{deg/deg}$ blastocysts treated with DMSO only developed trophoblast outgrowths with typical giant cells (arrows). Eomes-degron blastocysts treated with dTAG-13 or dTAGv-1 did not attach and outgrow as controls, phenocopying Eomes-null embryos (n=6 DMSO, n=7 dTAG-13, n=8 dTAG^{V}-1 embryos). Scale bars = 50 μm.

2.3.2: Ex vivo culture of post-implantation Eomes$^{deg/deg}$ embryos

Subsequently, Eomes$^{deg/deg}$ mice were crossed to obtain post-implantation embryos. E6.5 embryos were cultured *ex vivo* in dTAG-13 and dTAG^{V}-1, and assessed for Eomes protein presence via IF microscopy (Figure 2.5). After 2h of dTAG treatment, Eomes is virtually undetectable by IF (a few cells in dTAG^{V}-1-treated embryos may still appear faintly Eomes-positive), whereas the DMSO control treatment shows a strong nuclear Eomes signal in all its known expression of domains at the early streak stage: the ExE, the VE, and the PS (Figure 2.5a). Washout experiments were also carried out in order to determine the ability of Eomes-degron protein to recover after being ablated. This was indeed the case, with both dTAG-13- and dTAG^{V}-1-treated embryos recovering Eomes protein expression 2h after dTAG washout (Figure 2.5b). dTAG-13-treated embryos generally showed lower/slower recovery levels compared to dTAG^{V}-1-treated ones.

As this experiment was only carried out once by the time of this thesis' submission, repeating the experiment would strengthen the conclusions discussed above.

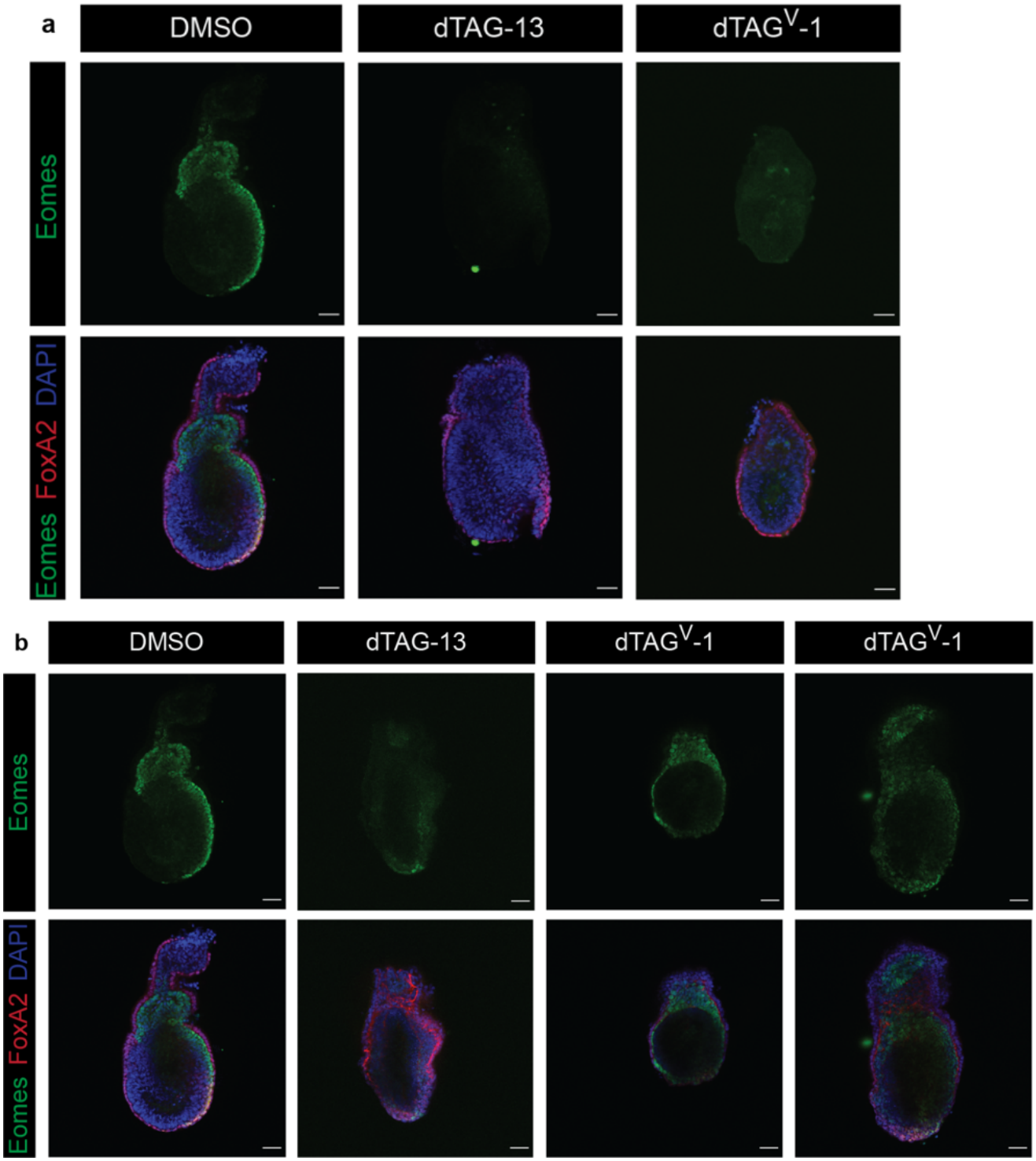


Figure 2.5 (legend on next page)

Figure 2.5: *Ex vivo* culture, dTAG treatment, and washout recovery of E6.5 Eomes$^{deg/deg}$ embryos.
(a) Representative images of E6.5 embryos isolated and cultured *ex vivo* in DMSO- or dTAG-containing (500nM) medium for 2h. While DMSO-treated embryos show an expected pattern of Eomes expression for early streak stage embryos (ExE, VE, and PS), in dTAG-13- and dTAG^{V}-1-treated embryos Eomes protein was undetectable by IF microscopy. Eomes is shown in green, FoxA2, an endoderm marker, is in red, and DAPI staining is shown in blue. Scale bars: 50μm.
(b) Representative images of E6.5 embryos cultured *ex vivo* in DMSO or dTAG for 2h, followed by a 2h washout. Embryos were washed several times in order to fully remove dTAG from the medium, after which they were further cultured in dTAG-free medium. dTAG-treated embryos show partial recovery of Eomes protein levels after 2h. Scale bars: 50μm.
n=6 DMSO; 9 dTAG-13; 4 dTAG^{V}-1; 7 dTAG-13 + 2 hr washout; 6 dTAGV-1 + 2 hr washout.

2.4 *In vivo/in utero* manipulation of Eomes-degron-expressing embryos

2.4.1: E5.5 Eomes$^{deg/deg}$ embryos do not show Eomes degradation after in vivo dTAG administration

The degradation of Eomes was then assessed *in vivo* by intraperitoneal (IP) dTAG-13 injections of pregnant mice using a similar strategy to the one described in Abuhashem et al. (2022a, b). Mice were injected on the morning of E5.5 and two hours later mice were sacrificed and embryos collected and stained.

As a control, Eomes heterozygous (+/-) intercrosses were carried out to allow a comparison to the Eomes-null embryo phenotype (Figure 2.6a). Eomes-null embryos showed severely stunted growth, with the TE/ExE compartment unexpanded and the Epi also severely underdeveloped, consistent with Eomes' role in TE maintenance and the TE's role in promoting appropriate Epi expansion and development.

Embryos from Eomes^{deg} homozygous intercrosses were treated with DMSO control or dTAG-13, and embryos were isolated after 2 or 4h (Figure 2.6b). Surprisingly, dTAG-13-treated mice did not show noticeably reduced Eomes levels, with expression both in the VE and ExE indistinguishable from control embryos, both after 2 or 4h. This may be attributed to the fact that at this stage of embryonic development, when implantation is

still in progress or has only recently occurred, the maternal and embryonic circulatory systems have not yet become connected, or at least not to the degree that dTAG can successfully reach the whole, or any, of the embryo.

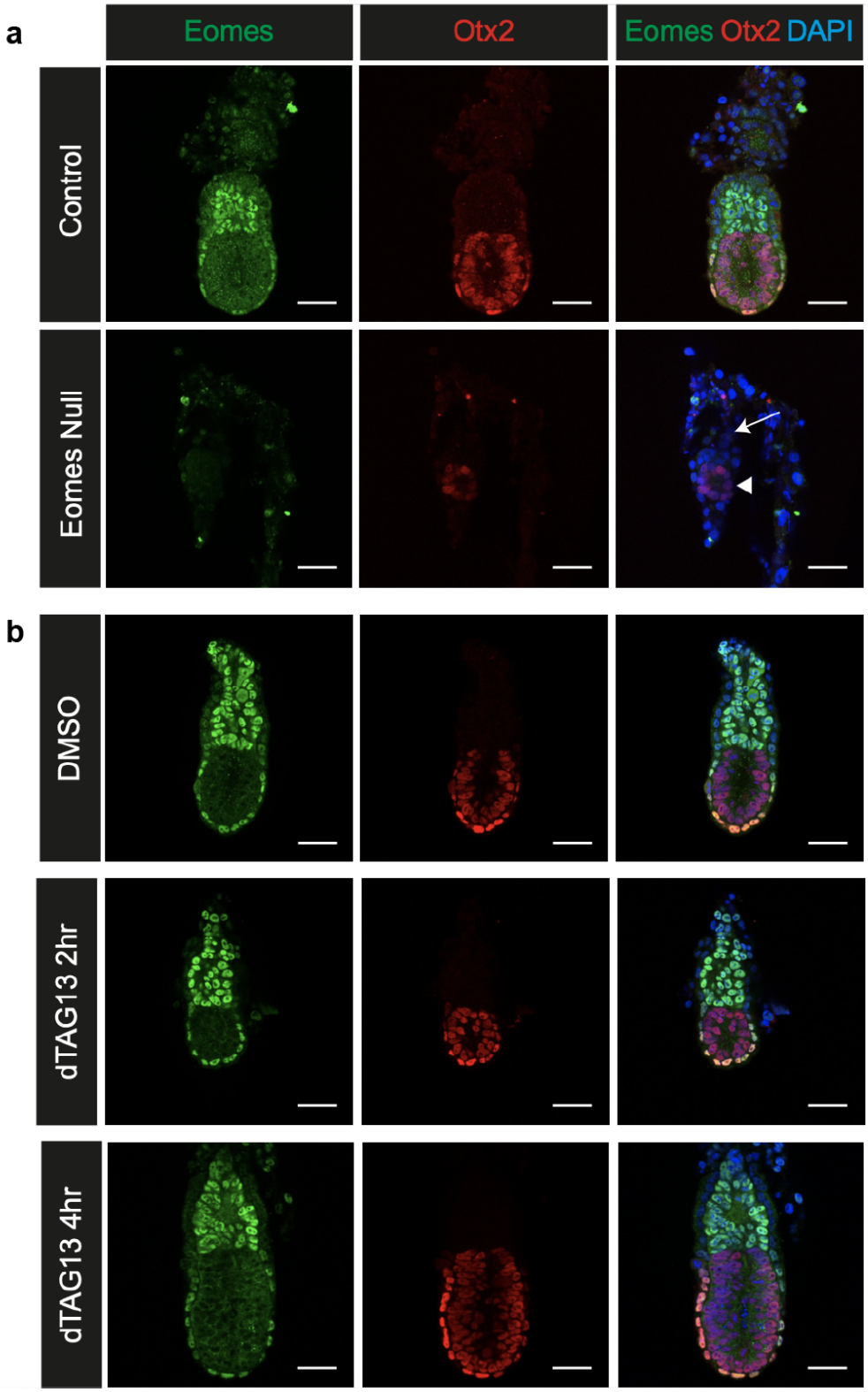


Figure 2.6 (legend on next page)

Figure 2.6: *In vivo* injection of dTAG-13 in pregnant mice (from a homozygous Eomes^{deg} cross) on E5.5 does not result in Eomes-degron protein degradation.
(a) E5.5 embryos from an Eomes$^{+/-}$ x Eomes$^{+/-}$ cross were stained for Eomes (green) (in the ExE and VE at E5.5), Otx2 [a ubiquitously transcribed epiblast and VE marker (Zakin et al., 2000)] (red), and DAPI (blue). Eomes$^{+/+}$ and Eomes$^{+/-}$ embryos showed normal expression of both markers, whereas Eomes$^{-/-}$ embryos showed no Eomes expression at all and significantly reduced Otx2, and only in the Epi (white arrowhead). The white arrow indicates the location of the ExE, which has failed to expand due to absence of Eomes. The Epi is stunted due to the absence of functional ExE. Scale bars = 50 μm.
(b) Eomes$^{deg/deg}$ homozygous embryos treated *in utero* with DMSO or dTAG-13 for 2 or 4h and fixed and stained for Eomes and Otx2. Eomes expression is still evident in both the ExE and VE in treated embryos. Otx2 expression is also unperturbed in dTAG-13-treated embryos (n=12 across three litters, detailed in Appendix Table 2.1a). Scale bars = 50 μm.

2.4.2: E6.5 and E7.5 Eomes$^{deg/deg}$ embryos show variable depletion of Eomes protein upon in vivo dTAG treatments

We then assessed the effect of *in vivo* dTAG treatment via IP injection on E6.5 Eomes-degron-expressing embryos. 4h post-injection, embryos were isolated and stained for Eomes, as well as FoxA2, which by now is not only expressed in the VE but also emerging DE cells. dTAG injections did result in loss of Eomes protein, however this effect was variable (Figure 2.7a). dTAG-13-treated embryos ranged from those where Eomes protein was completely undetectable, to those where Eomes was only depleted from the VE, or from the VE and partially from the ExE (see Appendix Figure 2.13). The variability in the efficiency of Eomes protein depletion was also apparent within litter-matched homozygous embryos, and therefore does not appear to be related to the efficiency of individual injections.

In contrast, dTAG^{V}-1 treatment did not result in complete Eomes loss (Figure 2.7a). In an attempt to see if splitting the dose over two injections would promote more complete protein degradation, pregnant mothers were injected twice within three hours, at variable intervals, however this too was insufficient to result in full Eomes protein depletion (Figure 2.7b). The embryos showed loss of Eomes expression in the VE, and perhaps a reduction in the ExE, but PS expression of Eomes was not impacted.

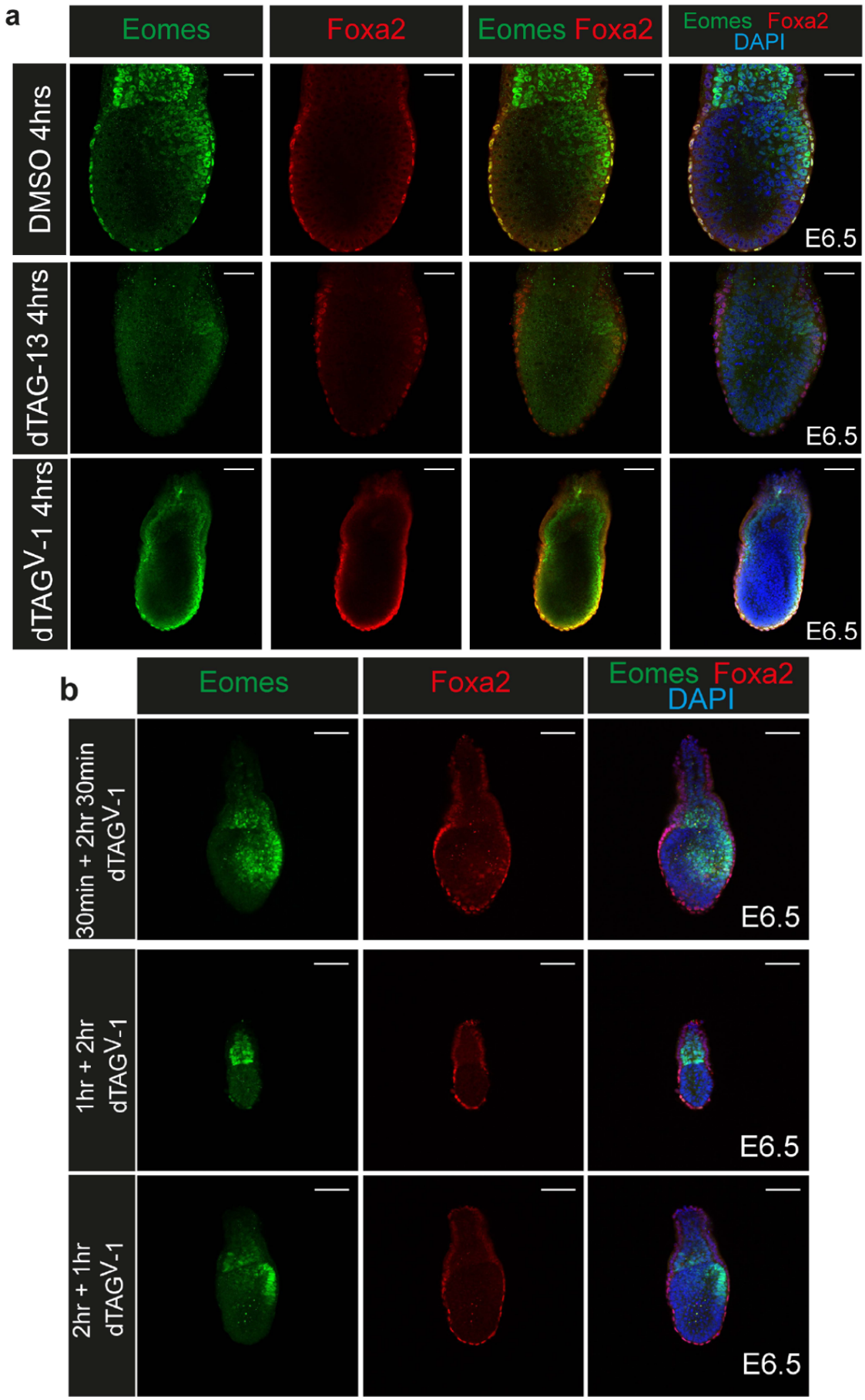


Figure 2.7 (legend on next page)

Figure 2.7: E6.5 Eomes$^{deg/deg}$ embryos show variable depletion of Eomes-degron protein after *in vivo* dTAG treatment.
(a) Selected images of E6.5 embryos treated *in utero* with dTAG-13 or dTAG^{V}-1. Although some dTAG-13-treated embryos showed full Eomes protein depletion, some did not (see Figure 2.8). dTAG^{V}-1-treated embryos did not show a great reduction in Eomes protein depletion (n=50 across 10 litters, detailed breakdown in Appendix Table 2.1b).
(b) Selected images of dTAG^{V}-1-treated embryos via two consecutive injections spaced over the course of 3h did not fully deplete Eomes protein.

In E7.5 embryos, there was similar variability in the degree of degradation of Eomes in Eomes$^{deg/deg}$ embryos, with some degrading Eomes-degron protein to a lesser degree (Figure 2.8, second row, where there is still visible nuclear signal) than others (Figure 2.8, third row). Once again, this variability was also noted within litter-matched embryos. Eomes$^{deg/deg}$ x Eomes$^{deg/+}$ intercrosses also permitted the intra-litter comparison of the degree of Eomes degradation relative to heterozygous Eomes$^{deg/+}$ offspring (Figure 2.8, first row).

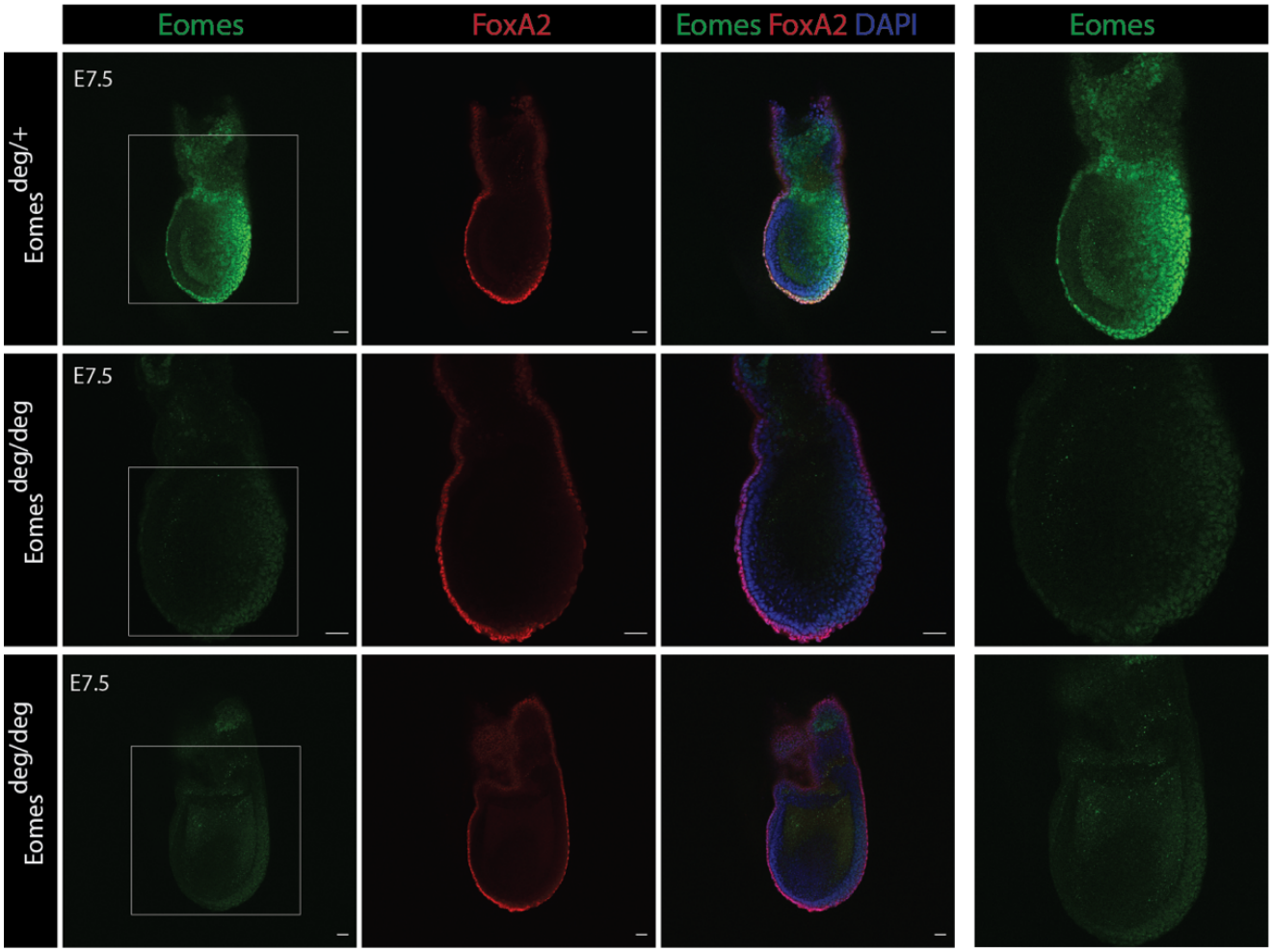


Figure 2.8: Selected images from E7.5 dTAG-13-treated Eomes $Eomes^{deg/deg}$ x $Eomes^{deg/+}$ intercrosses show variable Eomes protein degradation.
There was similar variability observed after 2 and 4h of treatment. The last panel of each row shows a magnified view of the region in the white box of the corresponding first panel. Scale bar: 50 μm (n=41 across 5 litters, detailed breakdown in Appendix Table 2.1c).

2.4.3: In vivo dTAG treatment does not result in discernible phenotypic defects of Eomes$^{deg/deg}$ embryos

As Eomes *in vivo* protein degradation in post-implantation Eomes$^{deg/deg}$ embryos was variable, we sought to investigate whether it would have an impact on embryo development after treatment. To this end, dTAG injections were followed by harvesting/dissection at 24 or 28h. This, however, did not appear to impact the normal developmental progression of embryos compared to their Eomes$^{deg/+}$ littermates, which showed both Eomes and FoxA2 expression in the APS (Figure 2.9, first row, arrow), as well as appropriate formation of the node, which also showed Eomes/FoxA2 expression (Figure 2.9, second row, asterisk). Additionally, Eomes levels appeared to be fully recovered, suggesting that dTAG had been metabolised or cleared from the vicinity of the embryo within the 24 or 28h of treatment. These data are representative of only one litter per condition (24 or 28h post-injection harvesting). Therefore, additional repeats of these experiments will be required to fully assess the consequences of *in utero* treatment and (temporarily) complete, or nearly complete, loss of Eomes protein.

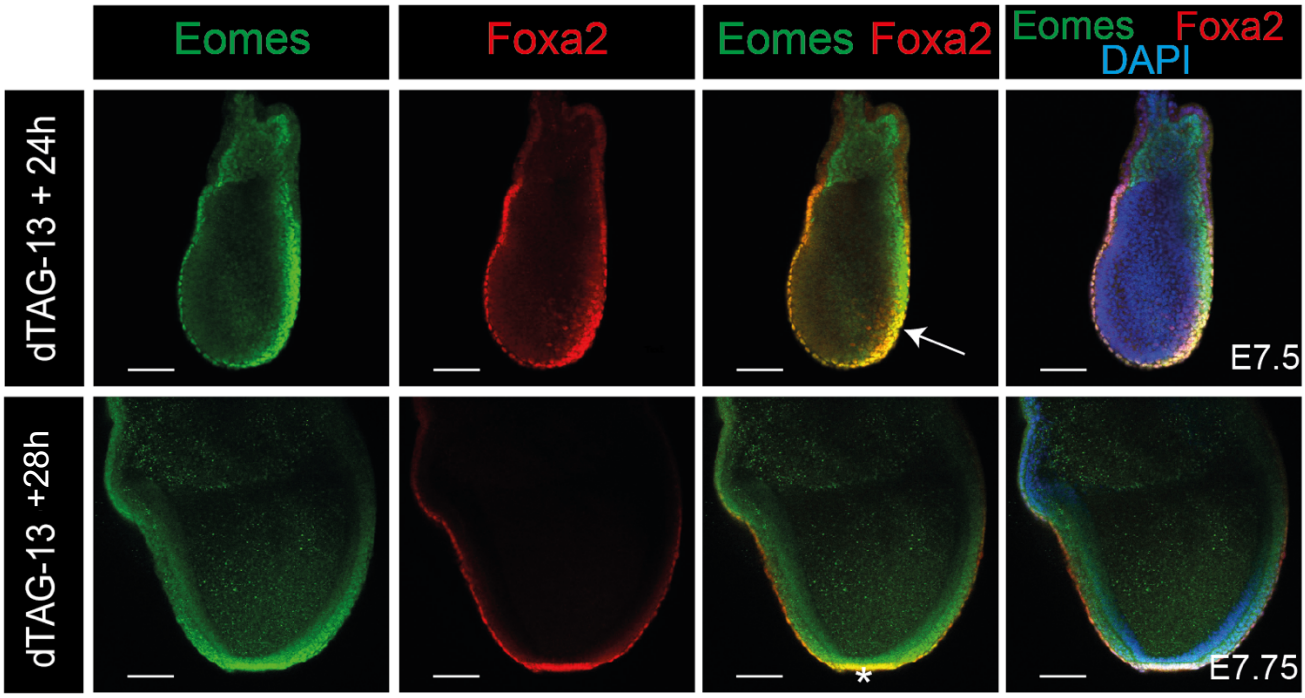


Figure 2.9: Selected images of Eomes$^{deg/deg}$ embryos treated *in utero* with dTAG-13 and harvested 24 or 28h later.
There do not appear to be any visible developmental defects with respect to appearance of Eomes-expressing lineages such as the APS (indicated by arrow) and the node (indicated by asterisk), as they both express Eomes and Foxa2, and do not appear morphologically different from litter-matched heterozygous embryos (n=7 E7.5 embryos, n=7 E7.75 embryos). Scale bars = 100 μm.

2.5 Deletion of putative Eomes super-enhancer does not impact formation of CXCR4+/cKit+ cells in DE differentiation culture

A previous study showed degradation of a widely-expressed transcriptional regulator, Nelfb, in gastrulating mouse embryos following IP injection of dTAG, proving that complete degradation of protein using the degron system is possible *in utero* at E6.5 and E7.5 (Abuhashem et al., 2022b). However, in contrast to the steady-state protein levels of Nelfb, Eomes expression is tightly controlled, spatially and temporally, in early development. Eomes protein levels appear to recover soon after washout of dTAG following treatment (Figure 2.5b), and *in vivo* dTAG injections does not consistently result in full Eomes degradation, which suggests rapid Eomes protein recovery.

We proposed that Eomes may be under the control of a tightly-monitored feedback system. Hence, we sought to identify novel enhancer regions within the *Eomes* locus that may control *Eomes* expression. Previous studies have identified *Eomes* enhancers (Simon et al., 2017), but these enhancers only have mild to moderate effects on Eomes expression. We identified a region approximately 85 kb upstream of the Eomes promoter. This region may act as a super-enhancer in TSCs, as it is bound by super-enhancer-associated protein p300 (Lee et al., 2019), DNA demethylase Tet1 (Chrysanthou et al., 2018), as well as Eomes and Tead4 in our CUT&RUN datasets (Figure 2.10a). Additionally, it has been shown via promoter capture Hi-C (PCHi-C) to strongly interact with the Eomes promoter in TSCs – an interaction not found in ESCs, where Eomes is not expressed (Schoenfelder et al., 2018) (Figure 2.10b). Eomes also binds this region during haemogenic endothelium differentiation (Harland et al., 2021). We therefore sought to investigate if this region also plays a role in regulating Eomes expression during gastrulation, and specifically in DE specification.



Figure 2.10: Deletion of putative super-enhancer region regulating Eomes expression. (a) Genomic track view of regions upstream of the Eomes promoter bound by TSC-specific TFs, including Eomes, Tead4, and Tet1, as well as super-enhancer-associated protein p300. Also indicated are locations targeted by Cas9 and the resulting deletion.
(b) View of the same genomic region highlighting the direct promoter interactions with distal genomic regions. Note the highlighted region that specifically interacts with the Eomes promoter in TSCs, and which contains the region targeted via Cas9-mediated deletion. Figure reproduced from Schoenfelder et al. (2018).

I designed gRNAs to delete an approximately 10kb region containing these coinciding Eomes/Tead4/Tet1/p300 sites, and after targeting ESCs, I isolated two independent clones with a homozygous deletion of the region. I then assessed the effect of this deletion on the generation of Eomes-dependent CXCR4+/cKit+ population in the DE EB differentiation protocol described in this thesis. However, the population was still specified in the enhancer deletion clones as assessed by flow cytometry on day 5.0 of differentiation, and there did not appear to be a significant difference in the percentage of double-positive cells between the parental cell line and the putative super-enhancer-deleted clones (Figure 2.11a,b). It has been demonstrated that Eomes levels need to be reduced by 75-80% for a loss-of-function effect to become apparent (Simon et al., 2017). Additional investigation into Eomes expression levels via intracellular flow cytometry or WB will provide a clearer picture on whether this is the case in the enhancer deletion clones. However, the lack of a distinctive reduction in the double-positive population suggests that Eomes expression is not severely enough affected to impede DE specification, and its expression during DE specification is likely controlled by a multitude of cis- and trans-regulatory factors, this enhancer region perhaps only being one of them.

It may be worth further investigating the function of this putative super-enhancer region in TSCs to determine if it is essential for sufficient Eomes expression in TSCs, and thus for TSC maintenance. It would be possible to flank the region of interest in TSCs with LoxP sites, in order to be able to inducibly excise the region and observe whether the cells sustain divisions and maintain a typical TS phenotype, or whether they begin to differentiate. Alternatively, it may be possible to assess the ESC-to-iTSC transdifferentiation efficiency of 'super-enhancer'-deleted ESCs through one of the protocols discussed in section 1.4 of the Introduction (Benchetrit et al., 2015; Kubaczka et al., 2015), although the inherently low efficiency of the ESC-to-iTSC transdifferentiation (Kubaczka et al., 2015) may render this an unideal assay.

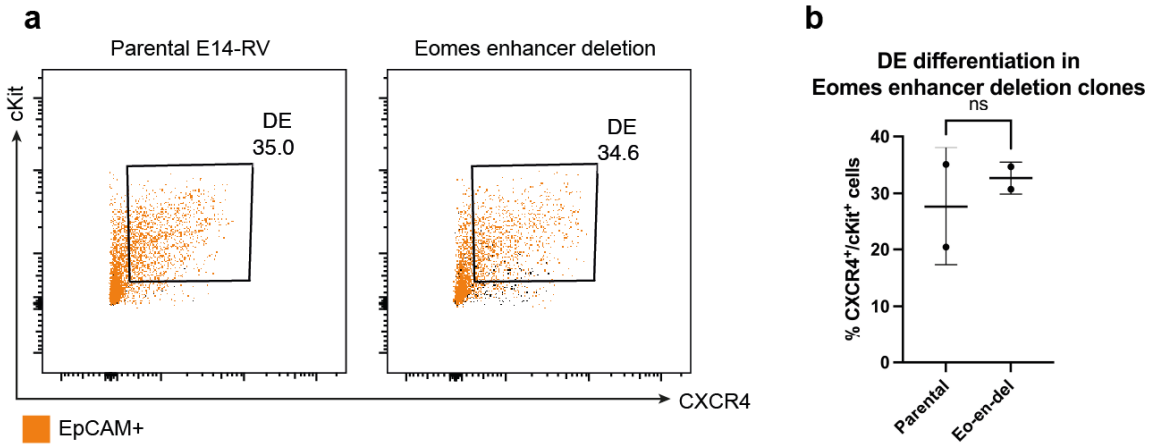


Figure 2.11: CXCR4⁺/cKit⁺ cells appear to be specified normally in Eomes enhancer deletion clones.

(a) Flow cytometry plot of CXCR4⁺ and cKit⁺ cells on day 5.0 of DE EB differentiation shows no apparent difference in the specification of the CXCR4⁺/cKit⁺ population between the parental and Eomes enhancer deletion daughter clone.

(b) There is no statistically significant difference between the percentage of double-positive cells between the parental and two Eomes deletion daughter clones. The unusually big difference between the two parental replicates is probably attributable to technical error, as this differentiation is usually very consistent, as described in chapter 3 of this thesis (n=2 for each of parental and Eo-en-del conditions).

2.6 Summary of findings

2.6.1: Generation of an inducibly degradable Eomes allele

In this Chapter, I describe the design of an allele for inducibly degradable Eomes protein that co-expresses an mCherry fluorescent reporter from the endogenous Eomes locus. The allele was used to target FF and FD mESCs, and its functionality was established by differentiating FF ESCs into Eomes-expressing EBs through addition of small molecule dTAG-13 or dTAG^{V}-1. Eomes protein was undetectable via western blot after a 1h incubation in dTAG-13. Moreover, Eomes depletion phenocopies the Eomes-null allele in the HE differentiation protocol, while in the DE differentiation, Eomes depletion results in loss of a CXCR4^{+}/cKit^{+} putative DE progenitor population.

2.6.2: Ex vivo dTAG treatment of Eomes$^{deg/deg}$ pre- and post-implantation embryos

The FD ESCs were also assessed for functionality of the knock-in allele via EB differentiation, and were subsequently used to generate a mouse line harbouring the allele. The homozygous Eomes$^{deg/deg}$ mice are viable and fertile. We isolated E3.5 Eomes$^{deg/deg}$ blastocysts and cultured them in dTAG-containing medium for 2h. Eomes protein was undetectable via IF, and dTAG-treated blastocysts did not successfully generate an outgrowth in a blastocyst outgrowth assay, phenocopying Eomes-null blastocysts.

Post-implantation E6.5 Eomes$^{deg/deg}$ embryos were isolated and treated with dTAG *ex vivo*, and similarly to the blastocysts showed undetectable Eomes by IF after a 2h treatment. They additionally showed robust Eomes protein recovery 2h after dTAG washout from the medium.

It was also observed that treatment with dTAG-13 was faster and more effective at fully depleting Eomes protein from the embryos compared to dTAG^{V}-1, and washout Eomes recovery was slower.

2.6.3: In vivo dTAG treatment of post-implantation embryos

In vivo dTAG treatment results were less consistent. E5.5 Eomes$^{deg/deg}$ embryos isolated two or four hours after IP dTAG injection of the pregnant mother did not result in loss of Eomes protein. E6.5 and E7.5 Eomes$^{deg/deg}$ embryos similarly treated *in vivo* showed variable degrees of Eomes loss, with some showing no Eomes IF signal, but most showing variable degrees of Eomes depletion when compared with Eomes$^{deg/+}$ littermates. Eomes expression in the VE was the first to be lost, followed by the ExE, and finally in some cases also the PS. We hypothesise this may have to do with the variable Eomes expression dynamics in these three tissues of post-implantation embryos. Varying the *in vivo* treatment protocol by splitting the dTAG dose across two injections similarly did not produce more consistent Eomes protein depletion. Finally, assessment of the phenotype of Eomes$^{deg/deg}$ embryos 24 or 28h after IP injection did not indicate an obvious developmental defect as assessed by presence of the FoxA2-positive APS and morphological node. Therefore, either incomplete depletion of Eomes protein after dTAG treatment *in vivo*, or its loss and recovery, were insufficient in producing a visible phenotypic defect at E7.5.

2.6.4: Deletion of a putative Eomes super-enhancer region does not affect EB differentiation

As Eomes recovery in the PS after depletion via dTAG treatment is rapid and robust (both after *ex vivo* washout as well as during *in vivo* dTAG treatment), we decided to delete an approximately 10 kb region known to act as an Eomes-regulating super-enhancer in trophoblast stem cells (Schoenfelder et al., 2018) in order to investigate whether this may produce a loss-of-function phenotype. However, EB differentiation toward a DE identity did not appear affected, however, suggesting that this region does not regulate Eomes expression in this lineage, or that its effect on Eomes expression does not have the magnitude required to produce an obvious phenotype, which has previously been shown to be a 75-80% reduction in Eomes levels (Simon et al., 2017).

2.7 APPENDIX

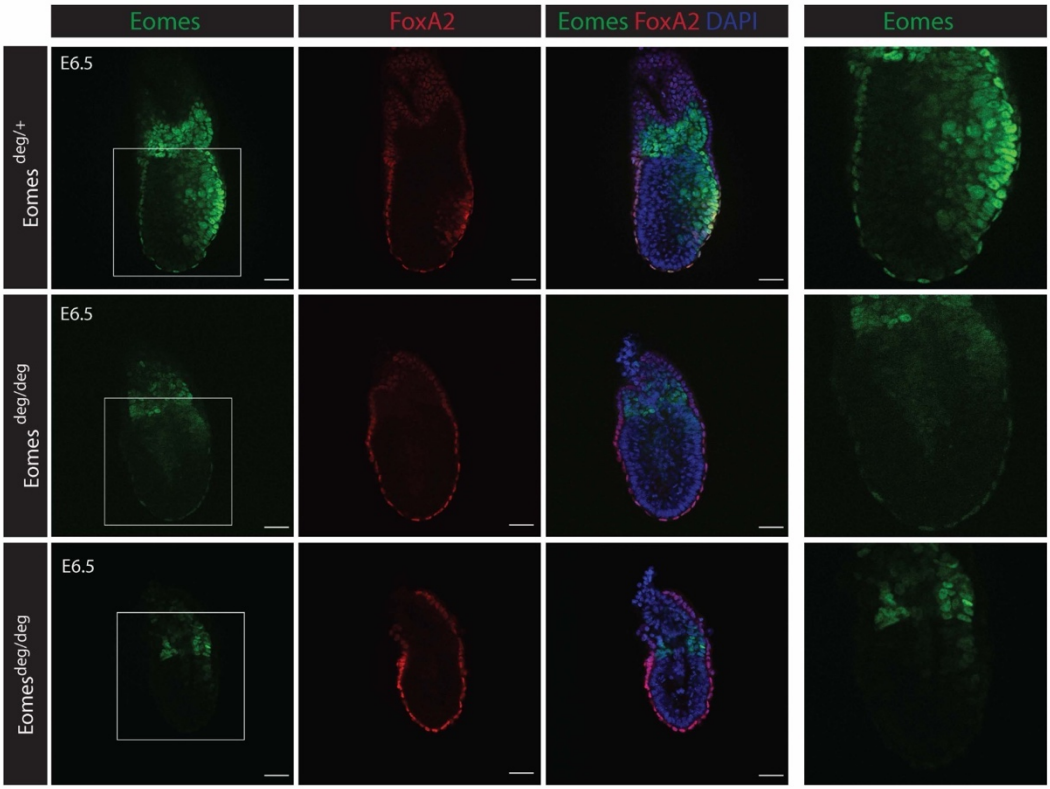
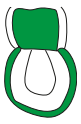
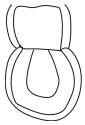
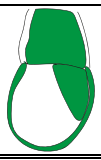
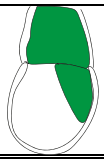
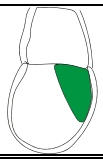
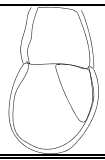


Figure 2.12: Eomes$^{deg/deg}$ E6.5 embryos show variable depletion of Eomes protein after 4h treatment upon in vivo dTAG-13 injection.
E6.5 Eomes$^{deg/+}$ heterozygous embryos exhibit Eomes expression in ExE, VE, and the PS. E6.5 Eomes$^{deg/deg}$ homozygous embryos showed variable reduction, with some losing expression in the VE (second row), and some both in the VE and most of the ExE (third row). Embryos' genotype was confirmed by PCR after staining. The last panel of each row shows a magnified view of the region in the white box of the corresponding first panel. Scale bar: 50μm.

a

Stage	dTAG	Duration	Litter size		
E5.5	dTAG-13	4hr	3	3	
E5.5	dTAG-13	4hr	6	6	
				100%	
E5.5	dTAG^{V}-1	4hr	3	3	
				100%	

b

Stage	dTAG	Duration	Litter size				
E6.5	dTAG-13	2hr	4				4
E6.5	dTAG-13	2hr	6				6
E6.5	dTAG-13	2hr	6		4		2
					25%		75%
E6.5	dTAG^{V}-1	2hr	6		3		3
E6.5	dTAG^{V}-1	2hr	3				3
E6.5	dTAG^{V}-1	2hr	5			2	3
E6.5	dTAG^{V}-1	3hr	7		5	2	
E6.5	dTAG^{V}-1	3hr	7		5	2	
E6.5	dTAG^{V}-1	4hr	3	3			
E6.5	dTAG^{V}-1	4hr	3	2	1		
				15%	41%	18%	26%

c

Stage	dTAG	Duration	Litter size			
E7.5	dTAG-13	2hr	6	6		
E7.5	dTAG-13	2hr	12	5	5	2
				61%	28%	11%
E7.5	dTAG^{V}-1	2hr	4		2	2
E7.5	dTAG^{V}-1	2hr	14	8	6	
E7.5	dTAG^{V}-1	2hr	5	3		2
				48%	35%	17%

Table 2.1: *In vivo* Eomes protein depletion in Eomes$^{deg/deg}$ embryos varies depending on developmental stage and choice of small molecule.

Stage refers to days after fertilisation (embryonic day) and duration refers to number of hours between dTAG administration and embryo collection. Aggregate statistics are provided for each small molecule and developmental stage: (a) for E5.5, (b) for E6.5, and (c) for E7.5 embryos.

CHAPTER 3: INVESTIGATING THE ROLES OF EOMES IN THE TROPHECTODERM LINEAGE

3.1 INTRODUCTION

The trophectoderm (TE) differentiates to form the outer layer of cells in the pre-implantation blastocyst. It gives rise to most of the placenta, with the exception of the maternal vasculature. Eomes is expressed throughout the TE at the blastocyst stage and is essential in its development during peri- and post-implantation development. Although TE cells are initially specified in Eomes-null pre-implantation embryos, they appear to rapidly differentiate into primary trophoblast giant cells (TGCs) during implantation, thus impeding appropriate implantation and further embryonic development (Russ et al., 2000).

The polar TE (pTE) gives rise to the ExE once the embryo implants. The TE/ExE lineage of mouse embryos maintains its stem cell niche until approximately mid-gestation (Kuales et al., 2015), due to FGF signalling from the ICM/epiblast (pre- and post-implantation respectively). Trophoblast stem cells (TSCs) can be isolated from the pTE of pre-implantation or the ExE of post-implantation embryos, and represent the *in vitro* equivalent of the TE/ExE lineage stem cells. They can be indefinitely passaged when cultured in defined medium containing FGF4, heparin, and TGF-β1 (Erlebacher et al., 2004).

Eomes is expressed in TSCs and downregulated upon TSC differentiation induced by growth factor withdrawal. It is an essential component of the TSC maintenance gene regulatory network (GRN), as it is found in protein complexes with other core TSC TFs Elf5 and Tfap2c, and hence thought to promote expression of genes involved in self-renewal (Kidder & Palmer, 2010; Latos et al., 2015b). Additionally, RNAi knock-down of Eomes mRNA results in loss of normal TSC colony morphology and downregulation of TSC-specific genes (Kidder and Palmer, 2010). Furthermore, TSCs cannot be derived from

Eomes-null blastocysts, due to the inability of the Eomes-null TE layer to attach and outgrow on a culture plate (Russ et al., 2000).

In this Chapter, I aim to characterise in greater detail the molecular functions of Eomes in the trophoblast lineage. To this end, we generated degron-tagged Eomes TSC lines from homozygous Eomes$^{deg/deg}$ mouse blastocysts. I carried out a time course of Eomes protein degradation followed by RNA-seq to determine what the earliest-affected genes were. In addition to TSC-specific TFs, components of signalling pathways involved in regulating the balance between TSC renewal and differentiation were affected. Expression of genes associated with ECM components, adhesion molecules, cytoskeletal components and cell motility was also dramatically altered. These are integral to the appropriate function of the epithelial TE layer of the blastocyst, and later of the differentiated trophoblast (TB) lineages. TB cells need to appropriately modify the uterine environment in order to successfully implant and invade into the uterine stroma to initiate placental morphogenesis.

Dr Ita Costello, postdoc in the Robertson lab, carried out CUT&RUN (cleavage under targets and release using nuclease) experiments for Eomes and Tead4. Here we show that Eomes binds the genes of several known TSC-associated TFs, along with genes associated with differentiating TSCs. Eomes also binds the pluripotency-associated genes Nanog and Myc. Our datasets suggest multiple modes of Eomes function, including promoting TSC-associated genes, regulating cell shape, cell-cell and cell-ECM interactions, and repressing TSC differentiation genes and possibly pluripotency genes. Finally, Eomes binding sites are also shown to be enriched in Tead4 binding motifs, as well as to significantly overlap with Tead4 binding sites, strongly suggesting co-regulation (perhaps by protein-protein interaction) of these two TFs. Protein interactome experiments will further elucidate the cell type-specific functions of Eomes.

3.2 Derivation of Eomes$^{deg/deg}$ cell lines

Prof Liz Robertson derived three independent TSC lines from E3.5 Eomes$^{deg/deg}$ embryos following standardised protocols (Tanaka, 2006). The cells were initially grown on a layer of inactivated mouse fibroblasts and subsequently passaged onto Matrigel-coated plates in defined medium (Kubaczka et al., 2014). The morphology of the cells cultured in the defined medium was characteristic of TSCs, with generally tight epithelial colonies (Figure 3.1a-a''). However, clone TSC5 formed less tight colonies, with the cells not tessellating as closely (Figure 3.1a''). Expression of Eomes and functionality of the degron tag was tested via western blot (WB), which indicated that all three colonies express degron-tagged Eomes protein (note the size shift from WT Eomes) and that Eomes protein is depleted upon overnight addition of dTAG-13 to the culture medium (Figure 3.1b). However, the Eomes protein expression level was variable, with clones TSC3 and TSC4 expressing much higher levels than TSC5. This may be characteristic of the phenotypic variation between TSC lines, or may indicate that TSC5 is not a true stem cell line, but rather a transient progenitor state that does not form tight epithelial colonies. However, it should be noted that there was no noticeable change in cell proliferation rates of the cell lines for the duration of their culture.

Surprisingly, in all three clones, there appear to be two Eomes-degron protein products – one with the expected molecular weight for degron-tagged Eomes protein, and one smaller. As this unexpected product is consistent (both in its presence, as well as its molecular weight) across all three clones, it is unlikely that it is due to random heterozygous loss of a section of the locus at the chromatin level. It may be that the knock-in of the sequence coding for the degron tag and the mCherry reporter affects transcriptional and/or splicing machinery of the cell in a TSC-specific manner, as this dual protein product was not observed in WBs of ESC lines differentiated into Eomes-producing EBs (see Figures 2.1, 2.2). Importantly, all cell lines were targeted with the same Cas9-based strategy, as the Eomes$^{deg/deg}$ ESC lines were used to generate the

Eomes^{deg} mouse line from which the Eomes$^{deg/deg}$ TSCs were derived. It is noteworthy that this truncated product also disappears after dTAG-13 incubation, suggesting it also has a functional degron tag, although the functionality of this peptide product is unknown. It may be useful to carry out qPCR to examine the relative expression levels of each exon of the Eomes gene, to determine if a 'fragmentary' Eomes mRNA is being expressed and translated. A Sashimi plot of exon splicing of the Eomes gene in one of the three Eo$^{deg/deg}$ clones (TSC3) did not show any obvious exon skipping (Appendix Figure 3.13a), but it did contain shorter Eomes mRNAs that did not include the first few exons (Appendix Figure 3.13b). However, a thorough comparison against RNA from a WT TSC clone should be carried out, in order to make note of significant differences between the two.

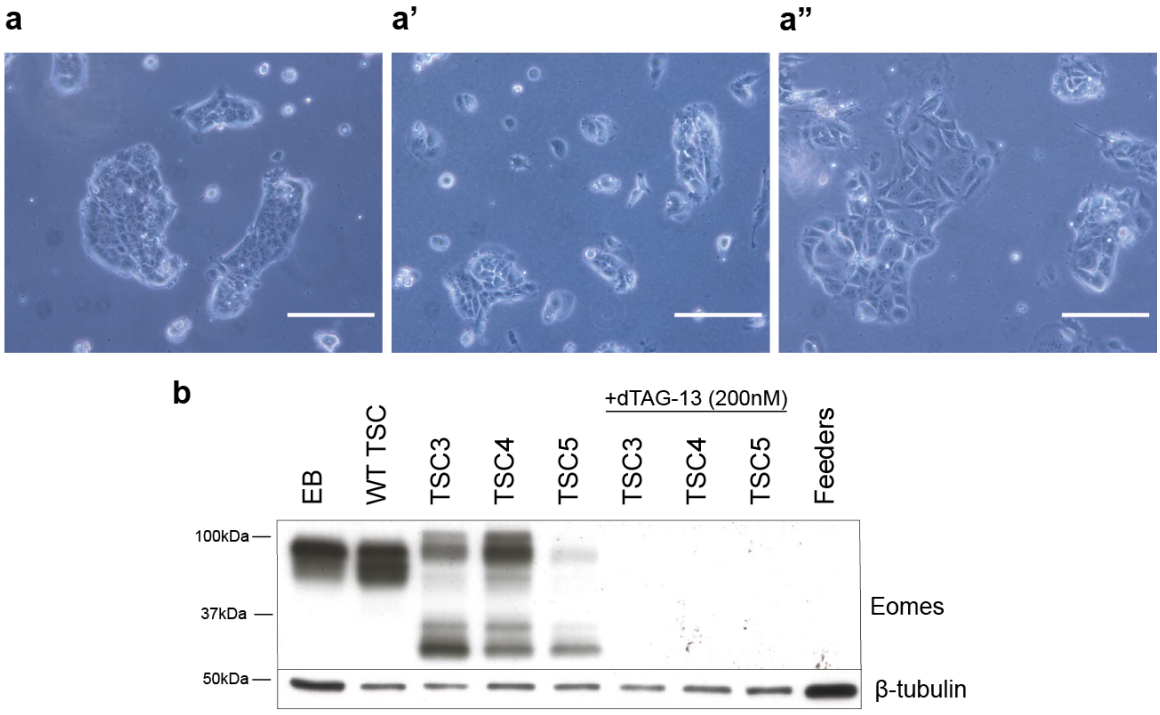


Figure 3.1: Independent TSC line derivation from Eomes$^{deg/deg}$ mouse blastocysts.
(a-a'') Phase-contrast images of Clones 3, 4, and 5 in (a), (a'), and (a''), respectively. Eomes$^{deg/deg}$ clones showed typical TSC morphology. Scale bar = 100μm.
(b) WB of dTAG- and DMSO-only treated Eomes$^{deg/deg}$ TSC lines. WT Eomes-expressing EBs and TSCs are shown as a positive control. Feeders are shown as a negative control. Eomes$^{deg/deg}$ TSC clones express a size-shifted degron-tagged Eomes protein, as well as a smaller peptide fragment detectable by the α-Eomes antibody. This effect is consistent across the three independently-derived clones. ß-tubulin was used as a loading control.

3.3 Transcriptional variability of Eomes$^{deg/deg}$ cell lines

Eomes$^{deg/deg}$ TSCs were then subjected to an Eomes protein degradation time course, followed by RNA-seq at various timepoints, in order to determine the short-term, direct effects of Eomes loss on TSC transcriptomes. Transcriptomes were assessed 6, 12, 24h, and 3 days after dTAG-13 addition. Transcriptomes were also assessed in untreated TSCs, and TSCs subjected to 3 days of DMSO control treatment. Only uniquely-mapping reads were assessed during the RNA-seq analysis. To assess library quality, the fraction of reads in genes and exons was assessed, with a higher proportion indicating a library with little contamination and little unspliced mRNA. Additionally, contamination from rRNA and reads aligning to mitochondrial DNA was quantified (Appendix Figure 3.14a). Finally, a QQ plot was generated to assess if there were any significant differences in probe quantitation between samples (Appendix Figure 3.14b). Clustering and dimensionality reduction were carried out by the SeqMonk software.

PCA (Figure 3.2a) indicated that most of the variance between samples is accounted for by the two first components (52% and 32%, respectively) (Figure 3.2b). Samples were also clustered by Pearson correlation, which showed a higher degree of correlation between samples of the same TSC clone, rather than clustering by timepoint of the degradation (Figure 3.2c). This suggests notable differences between the transcriptomes of the clones.

The samples are distributed across PC1 based on duration of dTAG-13 incubation (Figure 3.2a), with the untreated cells and earlier treatment timepoints (6, 12, 24h) occupying one end of the axis, and the TSCs treated for 3 days on the other end. However, interestingly, the 3 day DMSO-treated cells, although one might expect them to be close to the untreated TSCs, fall nearer the 3 day dTAG-13-treated TSCs. This is likely because after 3 days of DMSO treatment, the TSCs were much denser than the untreated “day 0” TSCs, which was also the case for the 3 day dTAG-13 treated cells. This, in turn, may be

accounting for similar density-dependent transcriptional changes. PC2, on the other hand, appears to be segregating the samples by clone.

Since there appears to be a density-dependent effect on TSC transcriptomes, this indicates that a DMSO control should have been implemented for each separate timepoint of the experiment. However, as the doubling time of TSCs in the defined medium that I used is approximately 24h (Kubaczka et al., 2014), the density effect should not be as much of a contributing factor in samples treated for ≤24h.

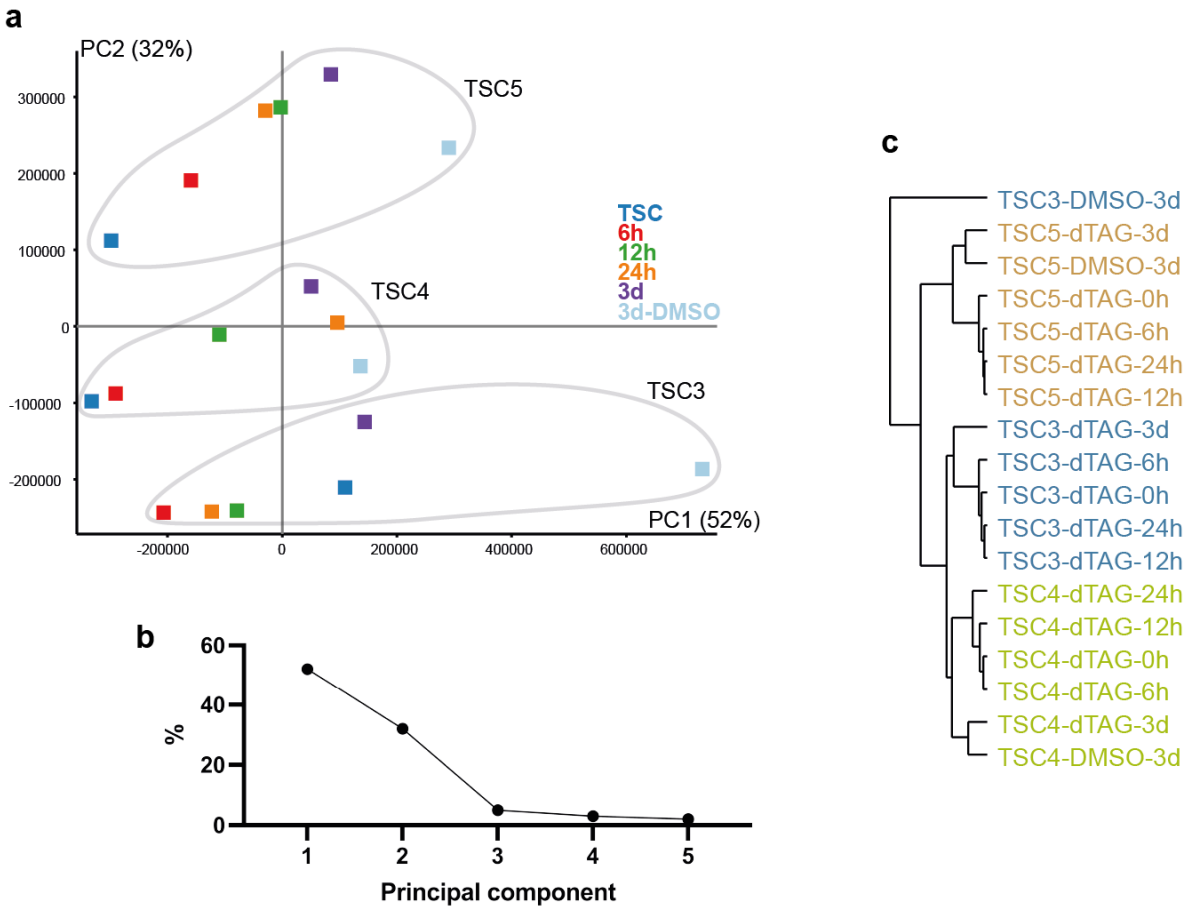


Figure 3.2: Eomes$^{deg/deg}$ TSC transcriptomes exhibit clone-to-clone variability.
(a) PCA of the three Eomes$^{deg/deg}$ clone transcriptomes during an Eomes protein degradation time course experiment. Samples belonging to the same clone are encircled and demarcated.
(b) Scatterplot of percent contribution of each principal component (PC) to inter-sample variability. PC1 and PC2 account for most of the variability between transcriptomes.
(c) Pearson correlation and clustering of the RNA-seq samples demonstrates that samples cluster by clone.

To investigate the genes accounting for the two major principal components, the 200 most highly contributing genes to each of PC1 and 2 were examined by GO analysis (Figure 3.3). PC1 contains two groups of genes, those more highly expressed in untreated and early-treated TSCs (+6, +12, +24h dTAG-13), and those more highly expressed in the day 3 dTAG-13- and control-treated TSCs. Several of the genes encode ECM components, proteolytic, cytoskeleton or cytoskeleton-associated proteins, important factors in the migration to and invasion of the maternal uterus by trophoblast cells during establishment of the placenta.

PC2 contains genes that are variably expressed by clone. Specifically, PC2-b genes, which include genes with a central role in TSC maintenance, such as Eomes and Esrrb, and TSC markers like Plet1 (Placenta expressed transcript 1) (Cambuli et al., 2014; Latos et al., 2015a; Murray et al., 2016). Notably, among the three clones, these genes are most highly expressed in clone TSC3 in untreated ("0h") TSCs, are slightly lower in TSC4-0h, and considerably lower in TSC5-0h. This may reflect a stronger activity of the TSC maintenance GRN in TSC3 compared to the other two clones, and would indeed be consistent with the less tight epithelial morphology and significantly lower Eomes protein levels of TSC5 compared to TSC3 and TSC4 (Figure 3.1). This line of evidence, thus, also suggests that TSC5 may constitute a transitional phase between bona fide TSCs and differentiated TB cells.

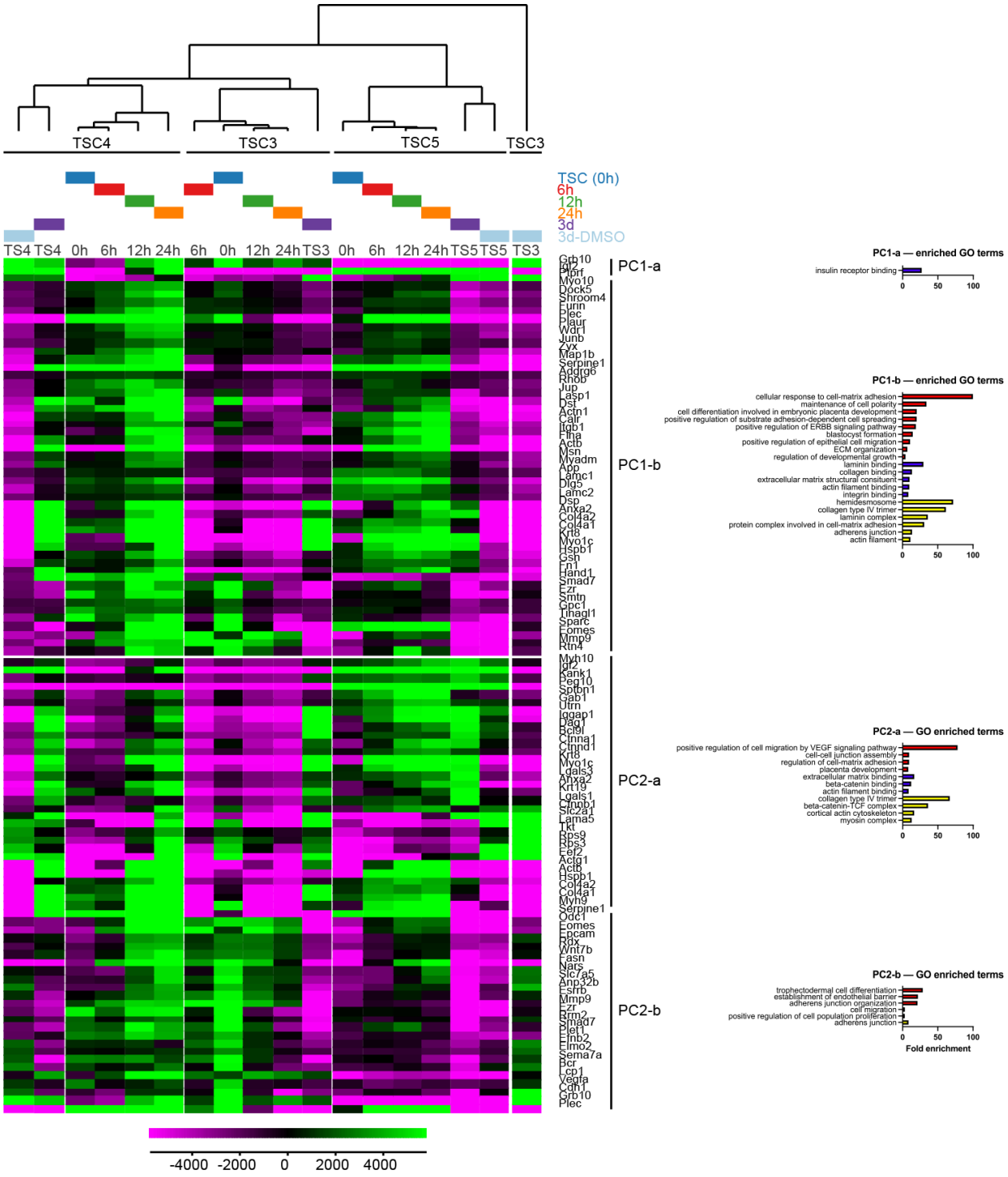


Figure 3.3: Transcriptional variability between Eomes$^{deg/deg}$ clones and between dTAG treatment timepoints.

Hierarchical clustering of genes showing highest contributions to PC1 and PC2. PC1 distributes samples by dTAG treatment timepoint, whereas PC2 distributes samples by clone (see Figure 3.2a). Enriched terms identified by GO analysis are presented next to the corresponding gene set. Terms in red correspond to biological process, those in blue to molecular function, and those in yellow to cellular component.

A more comprehensive examination of known TSC markers confirms that these are indeed more highly expressed in TSC3 and TSC4 (Figure 3.4). Interestingly, in all three clones, Eomes transcript levels initially increase in the hours after dTAG-13 addition, before beginning to fall after 24h. This may point to a potential compensatory mechanism at the transcriptional level, attempting to rectify the rapidly falling levels of Eomes protein.

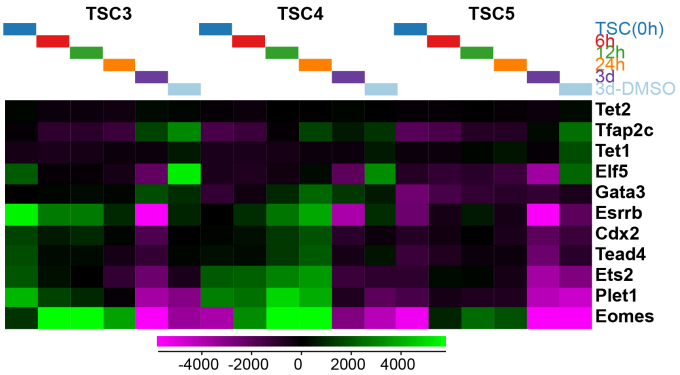


Figure 3.4: Known TSC markers are expressed variably between Eomes$^{deg/deg}$ TSC clones.
Hierarchical clustering of known TSC markers in Eomes$^{deg/deg}$ TSC clones. Higher levels of expression in untreated TSCs in clones 3 and 4 relative to clone 5 are apparent.

3.4 EOMES PROTEIN DEPLETION RESULTS IN ABERRANT EXPRESSION OF TSC-RELATED SIGNALLING PATHWAY AND MOTILITY COMPONENTS

Differential expression analysis of the three TSC Eomes$^{deg/deg}$ clones indicates that as early as 6h after dTAG-13 addition, there are already 59 differentially expressed genes, with the number increasing at each consecutive timepoint (Figure 3.5a). At each timepoint, there is a greater number of genes more highly expressed in the untreated TSCs than in the dTAG-treated cells, suggesting that Eomes may mostly be functioning as a gene activator rather than a repressor in multipotent TSCs.

Although 3 days after dTAG addition to the TSCs there were 3162 differentially expressed genes compared to the untreated TSCs, there were also many differentially expressed genes between the TSCs and the 3 day control (DMSO)-treated TSCs (3513 – more, in fact, than in the dTAG-treated cells). This indicated that TSC proliferation and increased cell density may account for much of this differential expression. In an effort to parse the effects of Eomes depletion via dTAG from the effects of increased cell density, I narrowed down the 3162 “TSC-vs-3d-dTAG" differentially expressed genes by removing those that were also differentially expressed in the “TSC-vs-3d-DMSO" group. The remaining genes are recorded in parentheses and are the ones that I retained for use in downstream analyses.

Differentially expressed genes (≥1.5-fold change)				
	Upregulated in untreated TSCs ←	**Total**	**Upregulated in dTAG-treated TSCs →**	
TSC (0h)	50	59	9	6h dTAG
	111	127	16	12h dTAG
	236	277	41	24h dTAG
	1692 (828)	3162 (1420)	1470 (592)	3d dTAG
	1664	3513	1849	3d DMSO
3d dTAG	1222	2016	794	3d DMSO

Table 3.1: Summary statistics of differentially expressed genes in Eomes$^{deg/deg}$ TSC lines.
Differential expression was assessed with DESeq in SeqMonk and genes were narrowed down to include those with at least 1.5-fold difference in expression levels. In the TSC-vs-3d-dTAG group, I have also recorded the number of genes differentially expressed between TSC-vs-3d-dTAG *but not* between TSC-vs-3d-DMSO (symbolised by "green – blue = yellow").

3.4.1: Early-dysregulated genes

Volcano plots of the log-fold change of gene expression against the false discovery rate (FDR) permit a closer inspection of the genes differentially expressed between untreated and dTAG-treated TSCs (Figure 3.5). The role of a selection of these genes in the TE lineage is detailed in Table 3.2.

At 6h, there is already a significant reduction in the expression levels of genes known to be expressed in mouse or human TSCs or TB-derived tissues, such as Cd24a, Wnt6, Six2, and ECM components Lama3 and Col27a1 (Figure 3.5a, Table 3.2).

12h after Eomes depletion, a larger group of TSC-expressed genes are downregulated, including Gata4, Efnb3, Duoxa1, Cd44, and Cdh22 (Figure 3.5b, Table 3.2). In contrast, Capn6, which encodes a microtubule-stabilising protein (Tonami et al., 2007) expressed in TSCs is expressed more strongly 12h after Eomes protein depletion. Although it has previously been reported to become downregulated upon differentiation into TGCs (Ullah et al., 2020), it has also been reported to be expressed in the mid-to-late gestation murine placenta (Dear & Boehm, 1999). The Eomes transcript is also expressed more

strongly at +12h, suggesting the presence of a compensation mechanism at the transcript level, attempting to restore Eomes to its expected levels.

24h after Eomes depletion, members of several classes of signalling molecules are downregulated: FGF pathway members Fgf3, Fgfr1, and Fgfr4; WNT pathway components Wnt6, Wnt5b, and Fzd1/7/10; signalling receptor Notch1; BMP pathway components Bmp2, Bmp2k (BMP2-inducible kinase), and Bambi (BMP and Activin Membrane-Bound Inhibitor). Additionally, groups of downregulated genes include Cluster of Differentiation (CD) genes Cd24a, Cd40, Cd44, and Cd47; ECM and cytoskeleton-associated genes, including Elmo1, Ezr, Mmp17, and Tinagl1; and cell-cell adhesion molecules Cdh22 and Pcdh12. Conversely, 24h after Eomes depletion, transcription factor Id3, TSC differentiation gene Ascl2, and ECM component Col18a1 are all upregulated (Figure 3.5c, Table 3.2).

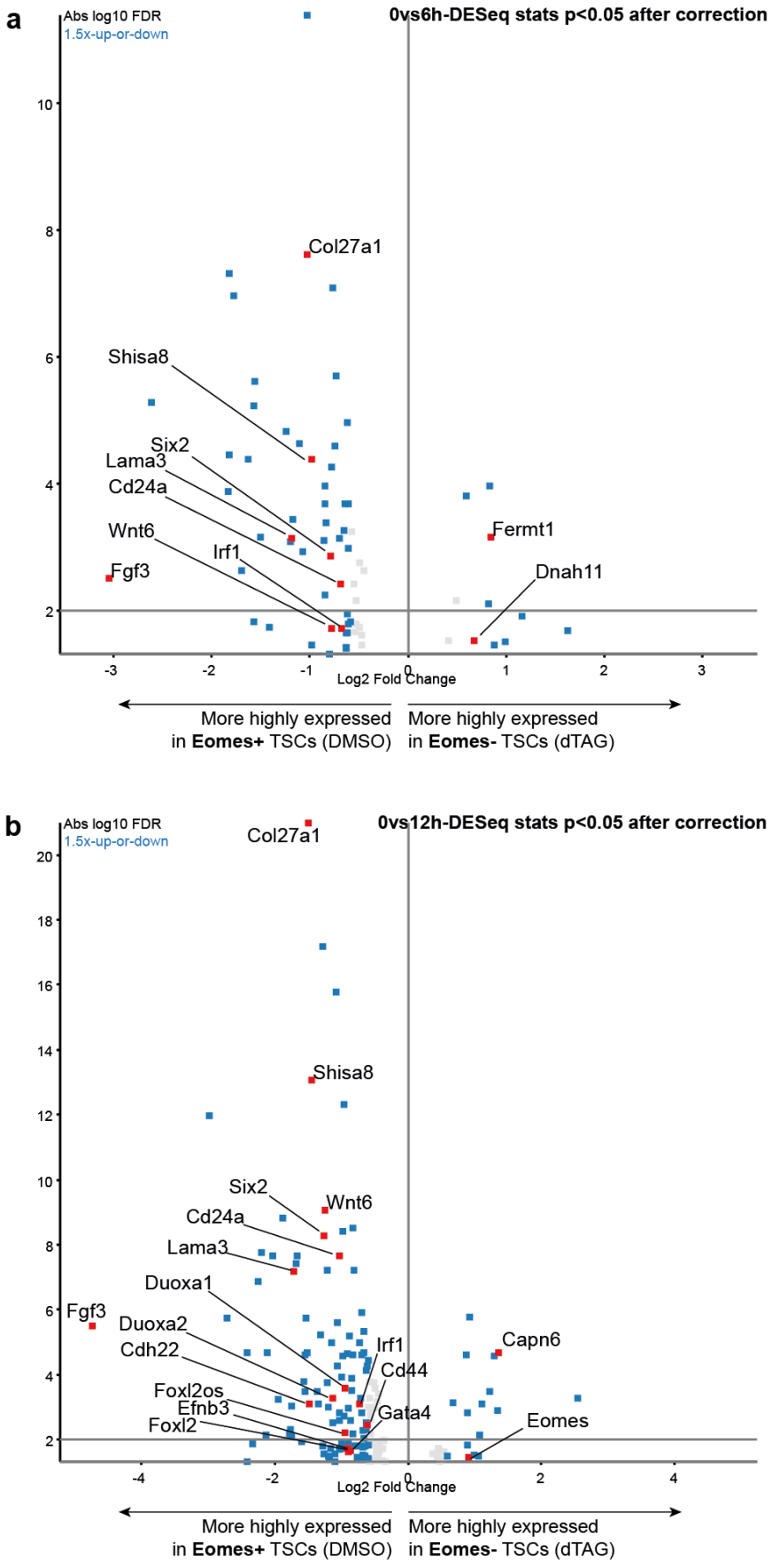


Figure 3.5 (continued on next page)

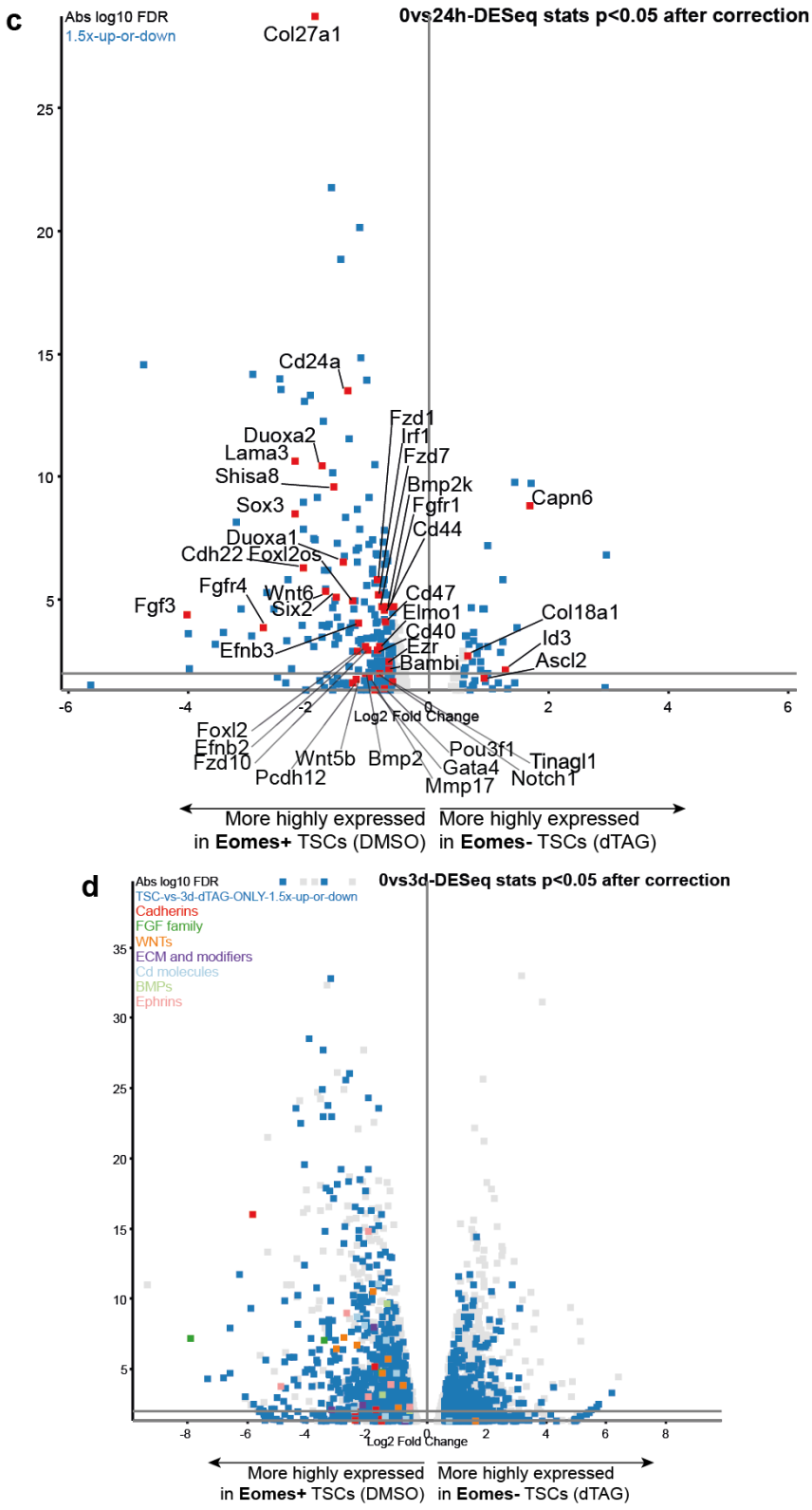


Figure 3.5: Volcano plots of differentially expressed genes upon Eomes depletion. Differentially expressed genes are shown in grey dots. Genes with a ≥1.5-fold change in expression are shown in blue dots. Denoted genes are highlighted in red.
(a-d) Differentially expressed genes in untreated TSCs vs 6h (a), 12h (b), 24h (c), and 3 days (d). In (d), only those differentially expressed genes NOT present in the TSC-vs-3d-DMSO differential expression comparison were included. Categories of differentially expressed genes are denoted in the legend and highlighted in the corresponding colour.

	Gene	Details in TE lineage	Reference
Expression reduced by +6h dTAG	**Cd24a**	Mucin-like adhesion molecule with reduced expression in human pre-eclamptic placentae	Nagy et al., 2008; Sferruzzi-Perri et al., 2009
	Wnt6	Low expression in human 1st trimester cytotrophoblasts (CTBs)	Knöfler & Pollheimer, 2013
	Six2	In ESCs, Six2/Six3 both have inactive H3k27me3 histone marks and involved in long-range, mutually repressive interaction. Both long-range interaction and repressive marks absent from Six2 in TSCs	Schoenfelder et al., 2018
	Lama3	ECM component of smooth chorion CTBs	Marsh et al., 2022
	Col27a1	Increased expression in differentiating human BeWo TB cells	Renaud et al., 2015
	Irf1	Interferon regulatory factor 1 – Expressed in fetal tissues and myometrium, has pro-inflammatory role in human parturition	Lim et al., 2016
	Shisa8	Transmembrane protein of Shisa family of WNT and FGF signalling modulators; expressed during mouse embryonic fibroblast reprogramming to iTSCs by forced expression of GETM (Gata3, Eomes, Tfap2c, Myc)	Jaber et al., 2022
	Fgf3	Little is known about its role in TB lineage, although one report suggests its misregulation in parthenote embryos may affect the balance between proliferation and differentiation of embryonic cells	Chen & Yu, 2012
Expression reduced by +12h dTAG	**Gata4**	Not known to be essential in mouse TSCs, but impacts progenitor state of human TB progenitors	Genbacev et al., 2016
	Efnb3	Transcriptional target of human early TB marker DLX5	Zadora et al., 2017
	Duoxa1	NADPH oxidase downregulated upon mouse TSC differentiation	Ullah et al., 2020
	Cd44	Present on blastocysts, involved in successful uterine attachment and implantation. Dysregulation associated with preeclampsia	Berneau et al., 2019; Todd et al., 2021
	Cdh22	Expressed more strongly in human TE than ICM	Adjaye et al., 2005
Expression reduced by +24h	**Fgfr1**	Regulates TE development and facilitates blastocyst implantation	Kurowski et al., 2019
	Fgfr4	Expressed in pre-implantation blastocyst, exclusively in TE	Ozawa et al., 2013
	Wnt5b	Low expression in human 1st trimester cytotrophoblasts (CTBs)	Knöfler & Pollheimer, 2013; Marsh & Blelloch, 2020
	Fzd1/7/10	Expressed in human CTBs	Knöfler and Pollheimer, 2013
	Notch1	Important in proliferation and survival of human EVT precursors	Dietrich et al., 2022
	Bmp2	Important in invasive differentiation of human and mouse TBs, dysregulation involved in early-onset preeclampsia	Yi et al., 2021; You et al., 2021
	Bmp2k	Mouse TSC-expressed gene, induced by Gata3	Ralston et al., 2010
	Bambi	Involved in human TB migration and invasion	Zhou et al., 2022
	Cd40	TSC marker, exclusive to TE at E5.5	Rugg-Gunn et al., 2012
	Cd44	TE-expressed, involved in human EVT invasion	Takahashi et al., 2014
	Elmo1	Engulfment and cell motility 1 – TSC-expressed	Ralston et al., 2010
	Ezr	Involved in epithelial integrity, hypomethylated in TSCs vs. ESCs	Cambuli et al., 2014
	Tinagl1	Hypomethylated in TSCs vs. ESCs	Cambuli et al., 2014
	Pcdh12	Expressed in mouse TB, but null mice viable and fertile	Rampon et al., 2005
Expression up by +24h	**Id3**	Bmp2-activated gene in human EVTs	You et al., 2021
	Ascl2	Human CTB-expressed; required for glycogen cell specification in mouse placenta	Bogutz et al., 2018; Knöfler et al., 2019
	Col18a1	Highly expressed in Dazu black goat (*Capra hircus*) placental tissues	Luo et al., 2021

Table 3.2: Selected genes significantly differentially expressed at different timepoints of dTAG-13 treatment (upregulated in Eomes+ TSCs).

Known expression or function in the TE lineage is detailed, along with the corresponding reference.

3.4.2: Differentially expressed genes 3 days after dTAG-13 addition

3 days after Eomes depletion, consistent with earlier timepoints, a suite of genes associated with various signalling pathways was strongly downregulated (Figure 3.5d, Table 3.3), including components of the FGF signalling pathway, essential in TSC maintenance, and the Wnt signalling pathway, which is also known to regulate TSC differentiation in the mouse (Bao et al., 2020).

Pathway/function/family	Genes
FGF	Fgf2, Fgf3, Fgfr4
WNT	Wnt3, Wnt5b, Wnt11, Wnt16*, Fzd3, Fzd7, Fzd10, Sfrp1, Tcf7
BMP	Bmp6, Bmp8b, Bambi, Smad6
ECM and ECM modifiers	Mmp15, Mmp16, Mmp17, Col6a4, Adamts2, Adamts15, Adam1b*, Adam4*, Adam8*, Adam19*, Adam22*, Elmo2, Elmo3
Cadherins	Cdh4, Cdh6, Cdh22, Pcdh10, Pcdh12, Pcdh17, Pcdh19
Cell-cell contacts	Efnb2, Efnb3, Ephb2, Ephb3, Epha5, Epha7
CD molecules	Cd24a, Cd34, Cd40, Cd44, Cd47, Cd59a, Cd93, Cd274*, Cd300a

Table 3.3: Selected classes of differentially expressed genes 3 days after Eomes protein depletion.
All genes are more highly expressed in Eomes+ cells, except those denoted with an asterisk, which are more highly expressed in dTAG-13-treated cells.

In addition to these classes of genes, there are further genes with known functions in the TE lineage misregulated after 3 days of dTAG-13 treatment, including Elf5, Tfap2a, Vefgc/Kdr, Tet3, Cxcl12/Cxcl14. A summary of their functions in the TE lineage is given in Table 3.4. Surprisingly, although Tfap2a and Tet3 appear to promote a TS-like identity, they are more strongly expressed in Eomes-depleted TSCs.

	Gene(s)	Function	Reference
Expression reduced by +3d	Elf5	Essential in TSC maintenance	Hemberger et al., 2010
	Vegfc, Kdr	VEGF signalling (Kdr is VEGF receptor) induces differentiation into TGCs	Fan et al., 2014
	Cxcl12	Expressed by TB cells, associates with CXCR4 promoting invasion of decidual stromal cells. Attracts regulatory T-cells, preventing pregnancy loss in non-obese diabetic mice	Lin et al., 2009
	Cxcl14	Inhibits TB invasion in human; inhibits TB outgrowth in early mouse pregnancy	Kuang et al., 2009a,b
Expression increased by +3d	Tfap2a	Part of TF network including Tfap2c and Gata2/3 resulting in BMP4-induced reprogramming of ESCs to TE	Krendl et al., 2017
	Tet3	Dynamically expressed in TE lineage throughout gestation, associates with RNAPII, increasing global transcription and promoting TS-like identity when expressed in ESCs	Krueger et al., 2017; Rakoczy et al., 2017

Table 3.4: Selected differentially expressed genes 3 days after dTAG-13 treatment.

In order to further investigate groups of genes that are broadly dysregulated after dTAG-13 treatment, I carried out a GO analysis on the differentially expressed genes at the 3-day timepoint (Figure 3.6). In genes more highly expressed in Eomes+ untreated TSCs, in addition to terms associated with previously-identified groups, such as the Wnt and FGF signalling pathways and cell migration and adhesion, several terms point to cation (and most specifically, calcium) regulation. Calcium is known to be involved in regulating TB proliferation, differentiation, and fusion, and several calcium channels are expressed in placental tissues throughout gestation (de Clercq et al., 2021). Accordingly, several calcium channel genes such as Cacna1a/c/i, calcium regulator Calca, and TRP (transient receptor potential) channel Trpc3 are expressed in untreated cells (Figure 3.6a), while Trpm6 is more highly expressed in 3-day dTAG-13-treated cells.

In contrast, GO terms enriched in dTAG-13-treated (Eomes–) TSCs include response to interferons (IFNs) and viruses. It is known that pregnancy constitutes a unique immunological situation, whereby a part-foreign organism (the fetus) must be tolerated by its host (the mother) while still permitting the host's defence from infection. Type I

IFNs in the placenta are known to modulate antiviral and immune responses in a manner protective to pregnancy (Kwon et al., 2018). Moreover, TBs can provide IFN-based antiviral protection in a paracrine manner (Fendereski et al., 2022). In agreement with this, several related genes show increased expression, including Ifnar2 (interferon α/β receptor subunit 2), Oas1a (an IFN-responsive gene) (Pulit-Penaloza et al., 2012), and Iigp1 (interferon-inducible GTPase 1) (Figure 3.6b).

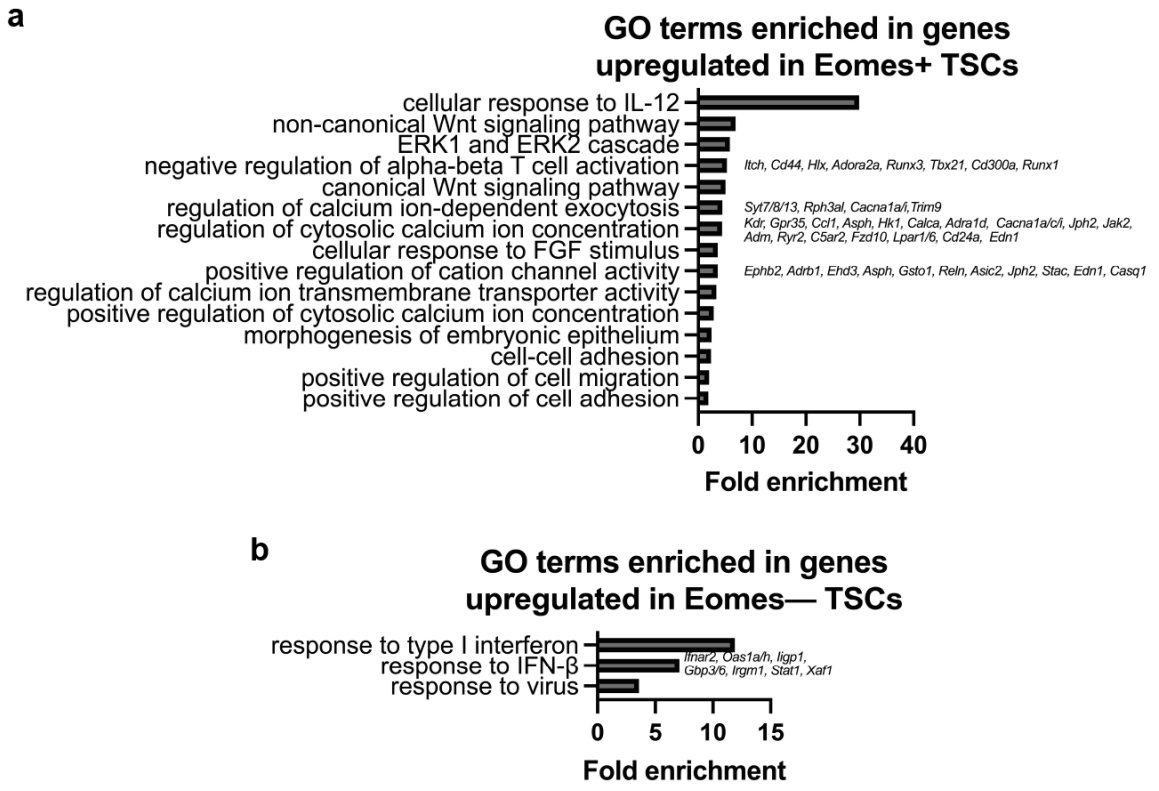


Figure 3.6: Bar chart of enriched GO terms in differentially expressed genes 3 days after dTAG-13 treatment.
Fold enrichment for each term is reported. Genes from selected terms are shown next to the corresponding bar.
(a) Enriched GO terms in genes more highly expressed in untreated (Eomes+) TSCs relative to 3 day dTAG-13 treatment.
(b) Enriched GO terms in genes upregulated in 3-day dTAG-13-treated (Eomes–) TSCs.

3.4.3: Persistently dysregulated genes

In order to investigate which genes Eomes may be most strongly/persistently regulating (and thus attempting to omit potential noise that may have led to genes appearing differentially expressed at only one Eomes degradation timepoint), I identified the genes differentially expressed in more than one timepoint within the dTAG-treated +6h, +12h, +24h, and +3d datasets, relative to untreated (0h) TSCs ("persistently" differentially expressed). The full list of genes can be found in Figure 3.15 in the Appendix. Of these genes, those most strongly downregulated upon Eomes depletion (and therefore possibly upregulated by Eomes) with known expression or function in the TE lineage are Sox3, Prss46, Calca (calcitonin/calcitonin-related polypeptide alpha), Six2, and Duoxa1. Genes that are most strongly upregulated upon Eomes depletion include Scara5, Capn6 (calpain 6), Itga1 (integrin-α1), Pcsk1n (proprotein convertase subtilisin/kexin type 1 inhibitor), and Tm6sf1 (transmembrane 6 superfamily member 1). Known details about their expression or function in the TE lineage are found in Table 3.5. Further investigation may prove some of these "persistently" differentially expressed genes as markers of TSCs or differentiating TBs.

Top persistently differentially expressed genes in Eomes-depleted TSCs					
Upregulated in Eomes+ TSCs			**Upregulated in Eomes- TSCs**		
Sox3	Elf5-upregulated, inhibits precocious TB differentiation	(Pearton et al., 2014)	**Scara5**	Ferritin receptor involved in decidualisation in human	Duncan et al., 2011; Lucas et al., 2020
Prss46	Serine protease, Elf5-upregulated	Pearton et al., 2014	**Capn6**	Microtubule-stabilising protein expressed in mid-to-late gestation placenta	Dear and Boehm, 1999; Tonami et al., 2007
Calca	Expressed in human VT/EVTs, proangiogenic function in placenta	(Dong et al., 2007)	**Itga1**	ECM interactor expressed in human EVTs	Zhou et al., 1993
Six2	Differential ESC/TSC methylation status (see Table 3.2)	Schoenfelder et al., 2018	**Pcsk1n**	Proteolytic inhibitor expressed in TSCs	Ralston et al., 2010
Duoxa1	Maturation factor for NADPH oxidase Duox1, expressed in undifferentiated TSCs	Ullah et al., 2020	**Tm6sf1**	Differentially methylated gene in human placentas based on maternal lifetime stress; found in locus where copy number variants are associated with pregnancy complications	Brunst et al., 2018; Wu et al., 2019

Table 3.5: Top 5 persistently differentially expressed genes upregulated in Eomes+ and Eomes- TSCs with known function or expression in the TE lineage.

The data from both the earlier dTAG-13 treatment timepoints, as well as the 3d long treatment, suggest that Eomes may regulate several genes known to be involved in TSC maintenance or to promote a TSC-like identity, such as Elf5, Tfap2a, and Tet3. Secondarily, Eomes may also play a role in suppressing TB differentiation-related genes, such as Ascl2, which becomes upregulated following Eomes depletion.

In addition, Eomes appears to affect the expression of a large number of genes involved in cell signalling, including ligands, receptors, and signalling modulators. Similarly, Eomes depletion affects several genes encoding ECM components, ECM modifiers, or cell-cell contact molecules. This suggests a role for Eomes in modulating the physical and signalling environment of TE cells in order to promote successful blastocyst implantation and TB cell invasion. Eomes also appears to promote genes involved in regulation of intracellular ions, which are essential for TB cell functions, and also seems to suppress expression of immunomodulatory genes that are only required later in gestation.

However, for several of the identified differentially expressed genes, little is known about them in the context of the TE lineage other than their expression; or their functions are not as well-characterised in the mouse TE as they are in the human TE. Therefore, the precise functionality of some Eomes-regulated genes in the mouse TE lineage is not yet fully clear. Further investigation into these genes may provide insight into whether they are functional in the context of TSC differentiation.

3.5 EOMES PROTEIN BINDS TSC-, TB-, AND PLURIPOTENCY-ASSOCIATED GENES

In order to identify Eomes-bound regions in the TE lineage, CUT&RUN was carried out for Eomes in WT TSCs. Peak calling revealed 924 Eomes-bound regions, the majority of which are located in genomic regions distal to gene promoters (50-500 kb from TSS), consistent with Eomes' known function as a TF (Figure 3.7a). GREAT (Genomic Regions Enrichment of Annotations Tool) (McLean et al., 2010; Tanigawa et al., 2022) analysis revealed 1391 peak-associated genes. The enriched GO terms in this group of genes include embryonic lethality and abnormal blastocyst and trophoblast layer formation (Figure 3.7b), which are consistent with Eomes functions in maintenance of the TE lineage and successful embryonic implantation (Russ et al., 2000). Known motif enrichment analysis additionally revealed that, in addition to its own consensus binding motif (a T-box binding site), Eomes-bound regions are also enriched for other TFs centrally involved in the TCS self-renewal GRN, including Tead4, Cdx2, TF AP2γ, Gata3, and Elf5 (Figure 3.7c). Sox2 and Sox3 binding motifs are also enriched. Although Sox2 is a pluripotency factor, in the context of ExE in the developing embryo, it is involved in maintaining the TSC stem cell state (along with its partially redundant family member Sox3) through activation of a very different network of genes in cooperation with different TF partners (Adachi et al., 2013; Pearton et al., 2014). Interestingly, in addition to these, the binding motif for pluripotency factor Klf4 is also enriched.

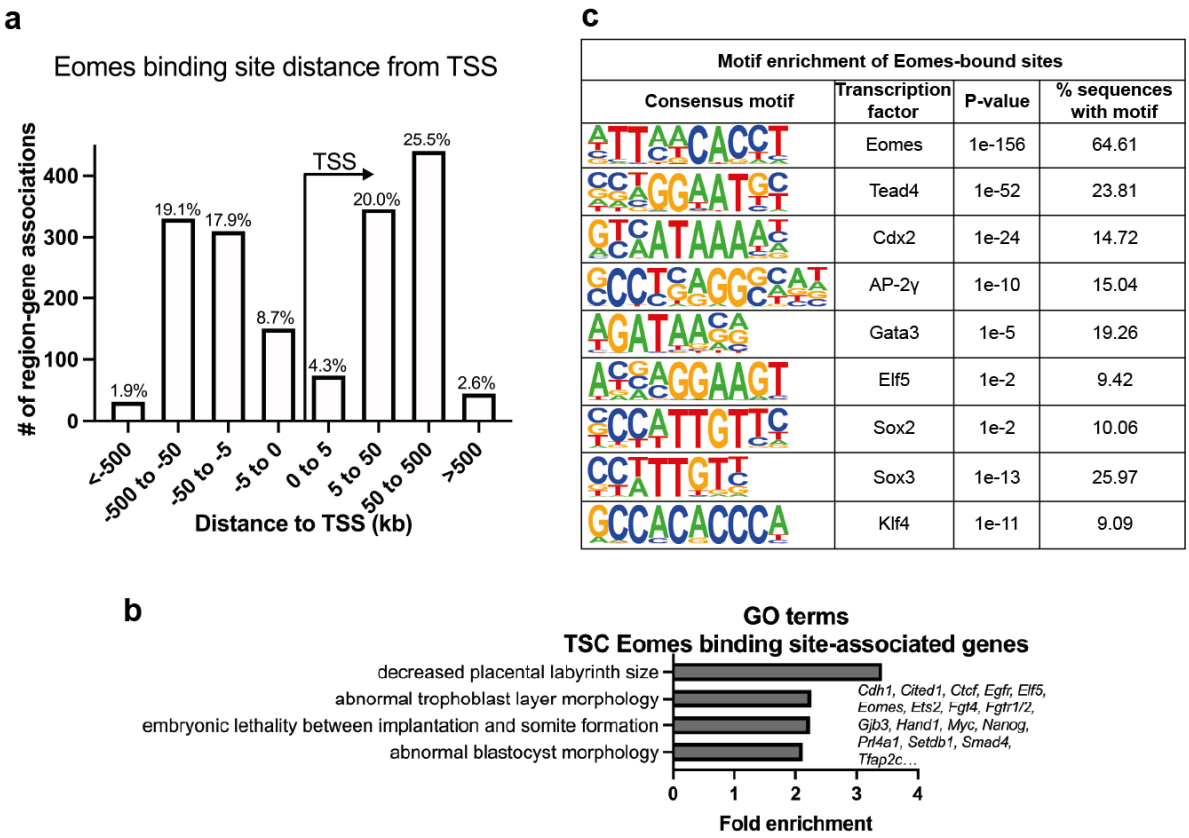


Figure 3.7: Eomes protein binds enhancer regions near genes associated with developmental progression and trophoblast formation.
(a) Distribution graph indicating the distance of Eomes binding sites from the TSS of the nearest annotated genes.
(b) Mouse phenotype single KO GO terms enriched in the gene set associated with Eomes binding sites in TSCs. A partial list of genes associated with the listed terms is provided alongside.
(c) Consensus motifs in Eomes-bound genomic regions.

Among the Eomes peak-associated genes ("Eomes-bound genes", for simplicity), there are genes associated with the TSC GRN, including Eomes itself, Elf5, Tfap2c, Fgfr2, Ets2 (Figure 3.8a); genes associated with TSC differentiation and various TB lineages, such as Id2, Hand1, Cited1/2, Rxra, Hopx, prolactin genes Prl4a1 and Prl5a1, as well as some only recently characterised within the TE lineage context by Lee et al. (Lee et al., 2019), such as Meis1 and Pou3f1 (Oct6) (Figure 3.8b). Finally, Eomes binds to pluripotency-regulating genes such as Nanog and Myc (Figure 3.8c). However, as Myc is also employed as a factor for iTSC reprogramming (due to its general function as a genome-destabilising factor) (Benchetrit et al., 2015; Felsher & Bishop, 1999), the fact that Eomes binds it may not be related to suppression of pluripotency factors, but to the cell type-independent properties of Myc.

These findings suggest multiple possible modes of action for Eomes, including promoting TSC-specific gene transcription, repressing or priming for expression TB differentiation genes, as well as repressing pluripotency gene expression. Additional genomic tracks of Eomes binding sites are supplied in Figure 3.16 (in the Appendix).

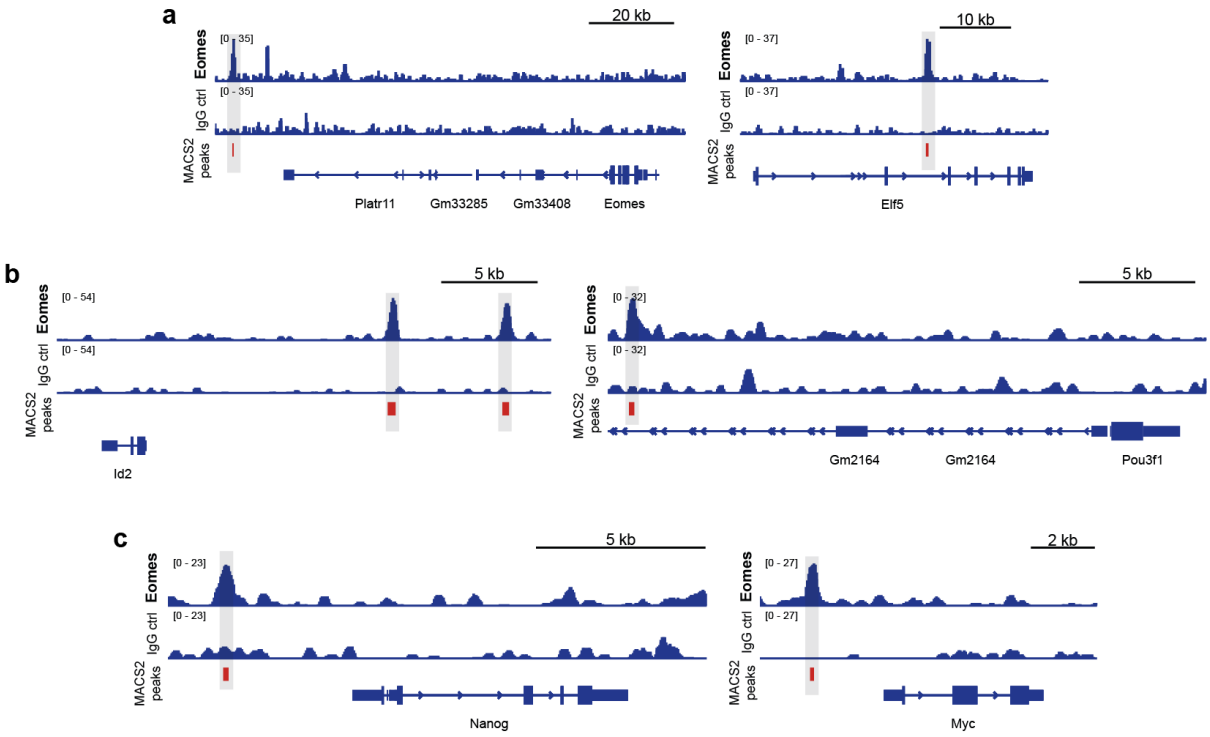


Figure 3.8: Eomes binds genes associated with TSC maintenance, TSC/TB differentiation, as well as pluripotency.
CUT&RUN tracks for Eomes, an IgG control, and Macs2-called peaks are shown. MACS2-identified peaks are highlighted on the track with a grey bar.
(a) Eomes and IgG control CUT&RUN tracks for selected Eomes-bound TSC maintenance genes. Shown here are Eomes and Elf5.
(b) Eomes and IgG control CUT&RUN tracks for selected Eomes-bound differentiation genes. Shown here are Id2 and Pou3f1.
(c) Eomes and IgG control CUT&RUN tracks for selected Eomes-bound pluripotency genes Nanog and Myc.

3.5.1: Investigating genes both differentially expressed and Eomes-bound

Next, I focussed on the genes that were both differentially expressed during the dTAG-13 Eomes depletion time course, and also bound by Eomes at the chromatin level (Table 3.6). A full list of the genes both Eomes-bound and differentially expressed upon dTAG-13 treatment can be found in Table 3.8 in the Appendix.

Differentially expressed and Eomes-bound genes					
	Timepoint				
Expression in Eomes+ TSC	**+6h**	**+12h**	**+24h**	**+3d**	**Overall**
+	16	31	61	118	**148**
-	0	2	2	28	**31**

Table 3.6: Summary statistics of genes both differentially expressed upon Eomes depletion, and bound by Eomes.
Numbers are shown for each individual timepoint, as well as overall.

Eomes mostly binds genes that are upregulated in its presence (untreated TSCs), consistent with the idea that it is mostly promoting, rather than repressing, gene expression. Among genes upregulated in untreated TSCs are TSC TF Elf5, BMP family genes (Bmp6, Bmp8b, Bambi, Smad6) and FGF family genes (Fgf3 and Fgfr1), as well as several genes previously identified involved in cell motility and associated with the cytoskeleton, such as Ezr, Elmo1, and Tinagl1. Interestingly, Eomes also binds the gene for TF Etv5, which has been reported to promote Fgfr2 expression in differentiation of EPSCs (extended potential stem cells) into TSCs (Zhu et al., 2020). The homeobox TF Hlx, which in humans has been shown to promote TB migration, but not invasion (Rajaraman et al., 2010), also displays Eomes occupancy. Interestingly, Smyd3, a histone methyltransferase which complexes with RNAPII and is known to promote lineage-specific gene expression in the early mouse blastocyst (Suzuki et al., 2015), including of Cdx2 and Eomes, is also an Eomes target.

In contrast, Eomes-bound genes that are only upregulated upon dTAG-13 addition include Oas1a (IFN-responsive gene discussed above) and the TF Pax8, which has been shown to promote human TB growth, migration, and invasion (Zang et al., 2022). Additionally,

amino acid transporter Slc38a4, whose (paternal) loss results in placental hypoplasia (Matoba et al., 2019), is bound by Eomes. Eomes also binds its own locus, and as discussed earlier, appears transiently upregulated at the +12h timepoint, indicating potential autoregulation.

Therefore, in keeping with the results of the differential expression analysis suggest, Eomes appears to be serving various roles, including promoting expression of certain core TSC GRN factors (Elf5 and possibly itself), regulating signalling pathways known to be involved in TE lineage function and differentiation, modulating the interactions between TSCs and their environment, and modifying cell permeability to various solutes. Additionally, Eomes seems to be repressing expression of genes normally found in differentiating or differentiated TBs.

3.6 Eomes and Tead4 may collaborate to promote TSC-specific gene expression

As we found that Eomes-bound genomic regions are additionally highly enriched for Tead4 binding motifs (Figure 3.7c), we analysed the genomic binding regions of Tead4 in TSCs in order to determine if they overlap with those of Eomes. Peak analysis showed 7834 Tead4-bound genomic regions. Similarly to Eomes binding site distribution with respect to gene TSSs, GREAT analysis indicated that most of the Tead4 binding sites are located distally from TSSs, at a distance of 50-500kb (Figure 3.9a). Genes associated with Tead4-bound regions are also enriched for GO terms related to the TE lineage development and stem cell maintenance (Figure 3.9b). I intersected the genomic regions bound by Eomes and Tead4, and found 527 regions bound by both TFs. This represents 57% of Eomes-bound regions, and 6.7% of Tead4-bound regions (Figure 3.9c). These co-bound regions are associated with 807 nearby genes.

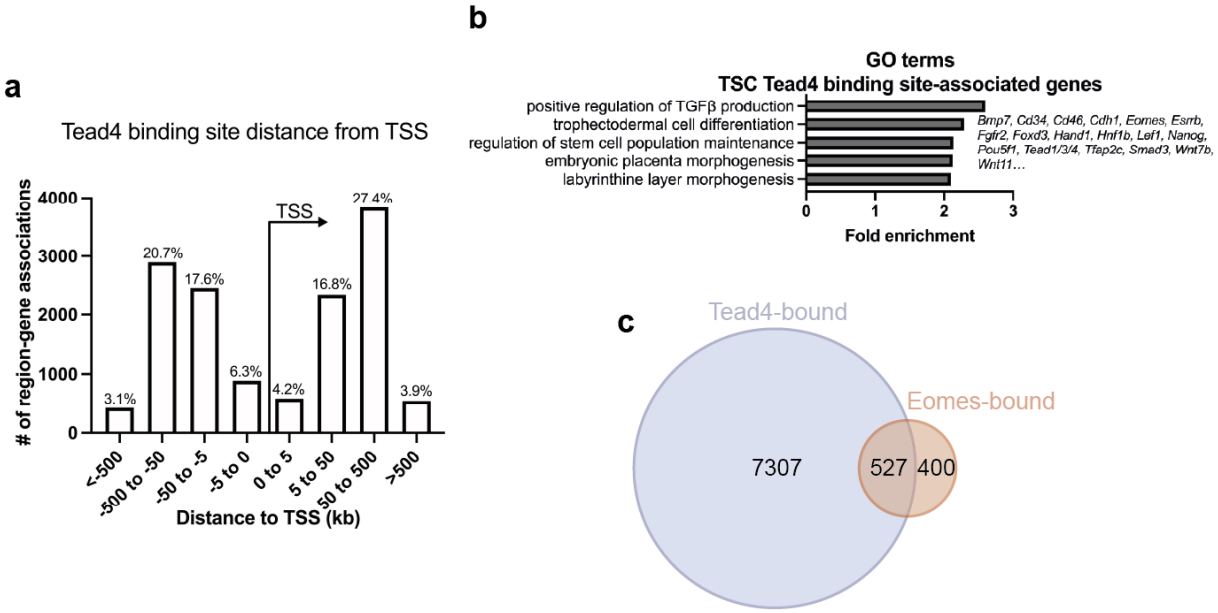


Figure 3.9: Tead4 binds regions associated with TSC-related genes, and co-binds with Eomes a significant proportion of genomic loci.
(a) Distribution histogram indicating the distance of Tead4 binding sites from the TSS of the nearest annotated genes.
(b) Biological process GO terms enriched in the gene set associated with Tead4 binding sites in TSCs. A partial list of genes associated with the listed terms is provided alongside.
(c) Overlap in Eomes- and Tead4-bound genomic regions.

In order to further elucidate Eomes' mode of action and its potential collaborative mechanisms with Tead4 in TSCs, next I identified the overlap between genes that are both differentially expressed upon Eomes protein depletion and at the same time associated with Eomes-and Tead4-bound genomic regions ("Eomes-bound genes" and "Tead4-bound genes" respectively) (Figure 3.10). As previously discussed, of the differentially expressed genes that are Eomes-bound, the majority are more highly expressed in untreated TSCs (Table 3.7). More specifically, genes more highly expressed in untreated (Eomes+) TSCs are 3 times more likely to be Eomes-bound (15% vs. 4.9%). Additionally, genes more highly expressed in untreated (Eomes+) are 4 times more likely to be Eomes-Tead4-co-bound (9.7% vs. 2.5%) than the genes upregulated upon Eomes protein depletion in TSCs.

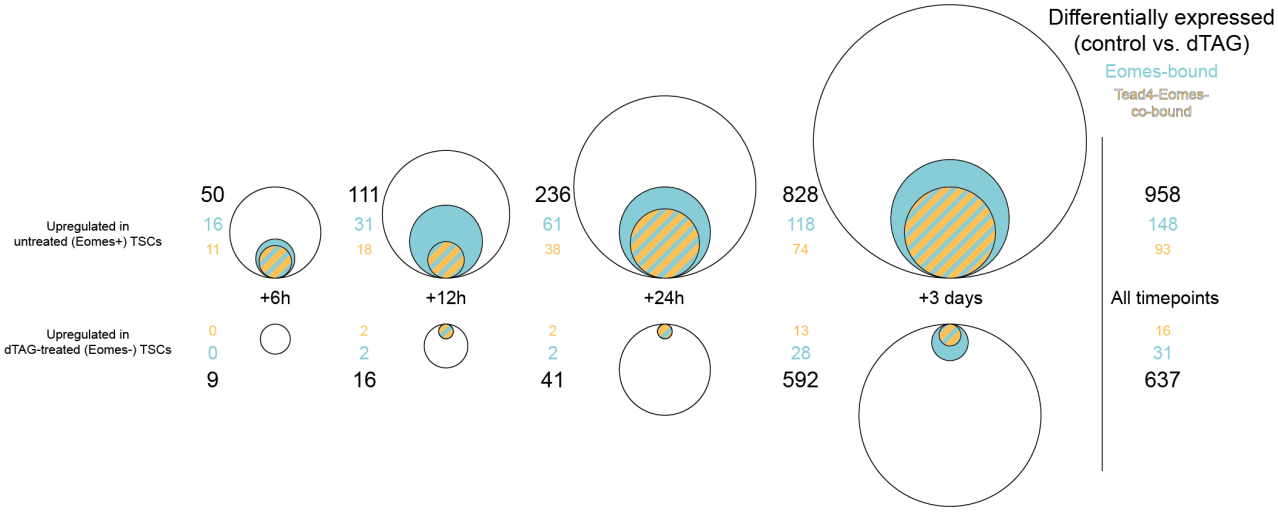


Figure 3.10: Schematic depiction of overlap between differentially expressed, Eomes-bound, and Tead4-bound genes in Eomes-depleted TSCs.
Genes misregulated upon Eomes depletion are represented by white circles, those also Eomes-bound are shown in aqua blue, and those also co-bound by Eomes and Tead4 are in striped orange circles. Circle sizes are not proportional to the number of genes they represent. Genes bound by Eomes and Tead4 are disproportionately upregulated in untreated TSCs.

Regulation in Eomes+ TSCs	# genes differentially expressed	...of which Eomes-bound (%)	...of which Eomes-Tead4-co-bound	Eomes-Tead4-co-bound as % of Eomes-bound	Eomes-Tead4-co-bound as % of upregulated genes
+	958	148 (**15%**)	93	**62.8%**	**9.7%**
-	637	31 (**4.9%**)	16	**51.6%**	**2.5%**

Table 3.7: Summary statistics of differentially expressed genes and Eomes/Tead4-bound genes across all timepoints of Eomes depletion time course experiment.
The first row (+) refers to cells significantly more highly expressed in untreated TSCs. The second row (-) refers to cells significantly more highly expressed upon Eomes protein depletion after dTAG addition.

Among the Eomes-Tead4-co-bound genes showing higher expression in Eomes+ (DMSO control) TSCs (Figure 3.11a, b) are those related to signalling pathways and their modulation (Bmp6, Bmp8b, Bambi, Smad6), transcriptional regulators (Foxo1, Meis1, Pou3f1/Oct6, Tbx21/T-bet, Zfp820, Zfpm1, Zic5), cell shape and motility regulators (Elmo1, Ezr, Dynlt1b), ECM components (Tinagl1), as well as Serpinb8, a member of the trophoblast- and placenta-expressed serpin family. Interestingly, Trp73 and Perp (a p53 transcriptional target) are also Eomes-Tead4-co-bound and more highly expressed in TSCs, suggesting that cell death regulation is an integral facet of TSC population regulation. The very few Eomes-Tead4-co-bound genes that are upregulated upon Eomes depletion include Trps1 (transcriptional repressor GATA-binding 1) and immune inhibitor Cd274 (Figure 3.11c). Additional genomic tracks for Eomes-Tead4 co-bound genes can be found in Figure 3.17 in the Appendix. The full list of differentially expressed Eomes-bound and Eomes-Tead4-co-bound genes is provided in Table 3.8 (found in the Appendix).

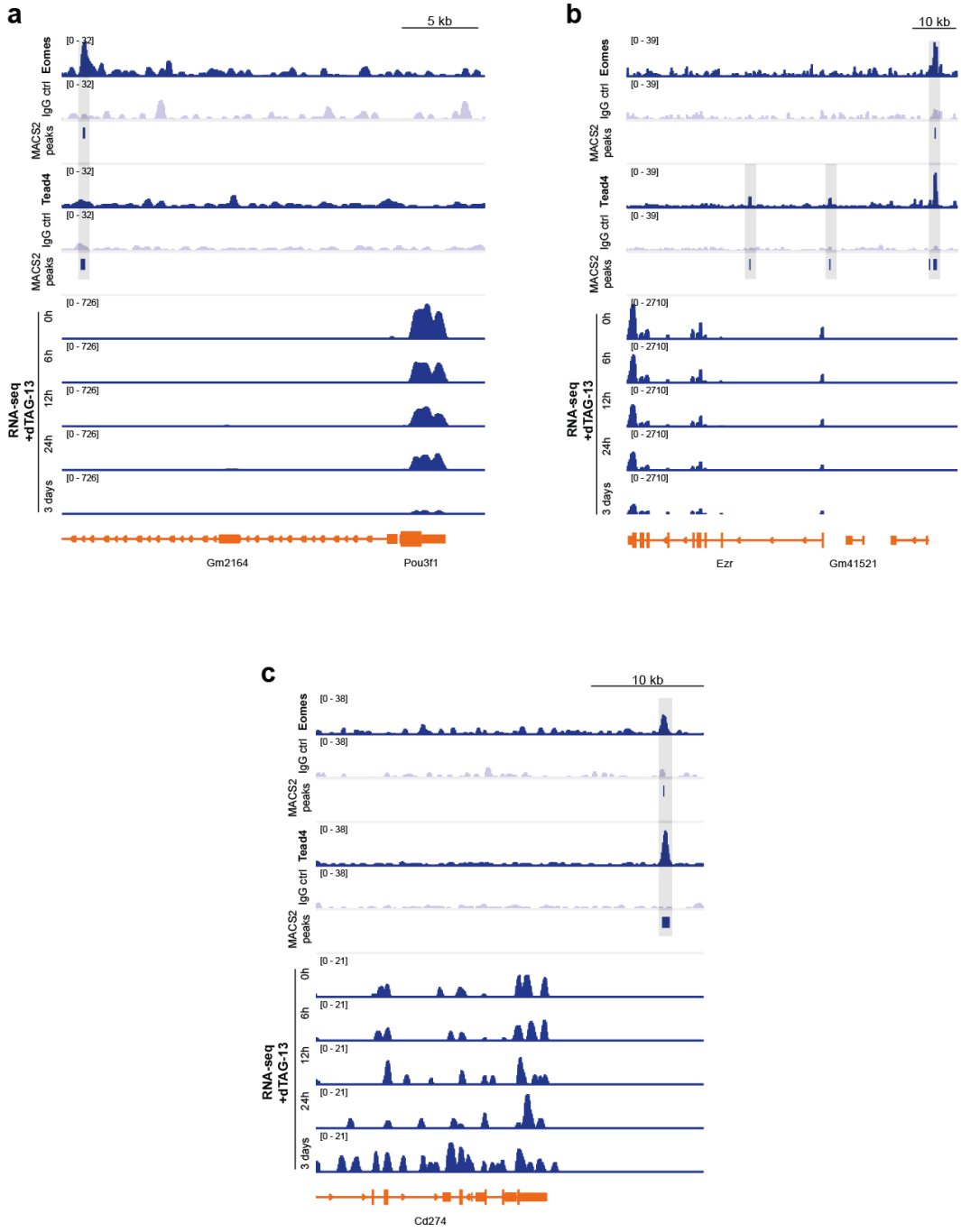


Figure 3.11: Selected genomic tracks of Eomes/Tead4-co-bound genes differentially expressed upon Eomes depletion.
(a, b) Genomic tracks for selected Eomes-Tead4-co-bound genes Pou3f1 (a) and Ezr (b) upregulated in untreated TSCs. Eomes, Tead4, and IgG control CUT&RUN tracks are shown, in addition to MACS2-called peaks, and RNA-seq tracks demonstrating the change in transcriptional levels of the genes. Grey shading highlights the Eomes- and Tead4-bound genomic regions.
(c) Genomic tracks for selected Eomes-Tead4-co-bound gene Cd274. Cd274 is one of the few Eomes-Tead4-co-bound genes that is upregulated upon dTAG-13-mediated Eomes protein depletion. Eomes, Tead4, and IgG control CUT&RUN tracks are shown, in addition to MACS2-called peaks, and RNA-seq tracks demonstrating the change in transcriptional levels of the genes. Grey shading highlights the Eomes- and Tead4-bound genomic regions.

Taken together, these data suggest that cooperative function between Eomes and Tead4 may play a role in positive regulation of gene expression in TSCs. This proposed model is presented schematically in Figure 3.12.

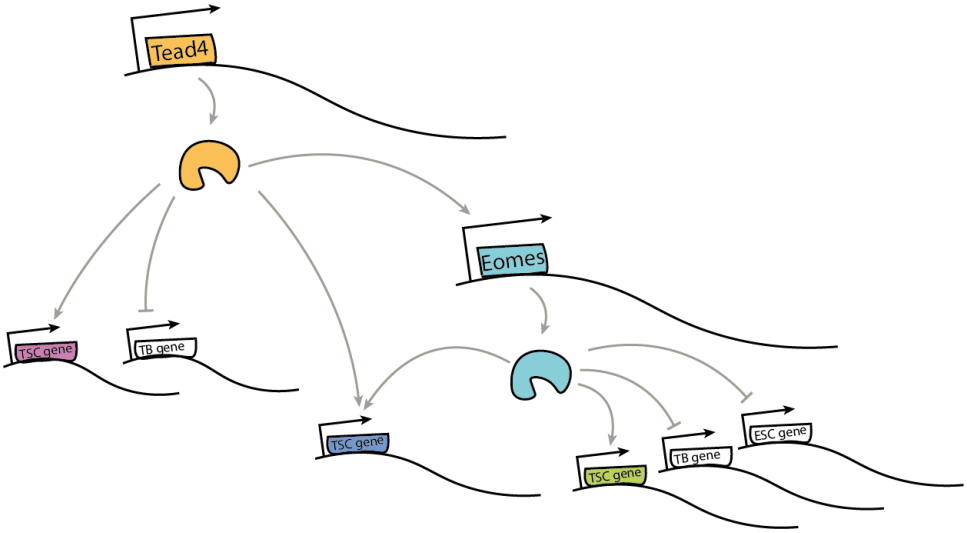


Figure 3.12: Schematic of proposed models of Eomes functions in TSCs and Eomes-Tead4 collaboration in TSCs.
Eomes and Tead4 are known to both promote and repress gene expression. However, within the specific context of TSCs, Tead4 and Eomes appear to mostly collaborate (through co-binding of genomic loci) to promote of TSC-expressed genes. Additionally, Eomes binds known TSC differentiation/TB lineage markers, as well as pluripotency genes, which it may be priming or repressing.

3.7 Summary of findings

3.7.1: Generation of Eomes$^{deg/deg}$ TSC lines

Using established protocols (Tanaka, 2006), we derived three independent TSC lines from Eomes$^{deg/deg}$ blastocysts. We established the functionality of the degron allele by dTAG-13 treatment of the cells followed by assessment of Eomes protein levels via WB. Surprisingly, a smaller peptide was additionally recognised by the Eomes antibody, which was also lost upon dTAG treatment.

Eomes$^{deg/deg}$ TSCs were then subjected to a dTAG addition time course followed by RNA-seq. The cell lines showed considerable morphological and transcriptional variability in their untreated state, notably in their expression of TSC markers.

3.7.2: Eomes depletion results in dysregulation of TSC GRN components and cell motility/cell contacts/cytoskeleton interactors

Differential expression analysis indicated that Eomes loss results in the downregulation of several TFs known to be involved in maintenance of TSC "stemness," or known to promote a TSC-like identity in cell reprogramming assays. In addition to these, many genes encoding ECM components, or involved in cell-cell or cell-ECM contacts, as well as cytoskeletal components, are also dysregulated. This raises the possibility that Eomes may be a regulator of blastocyst interaction with its environment, controlling implantation and invasion of the maternal decidua. Interestingly, many of these genes have been reported to be expressed in mouse TSC but not extensively characterised, while some genes have been reported to be expressed in human TE/placental tissues. Eomes loss also results in dysregulation of genes known to be expressed in differentiating TBs, suggesting that Eomes may also be involved in regulating expression of this class of genes.

3.7.3: Eomes binds TSC-, TB-, and pluripotency-expressed genes

CUT&RUN for Eomes indicated that it binds many of the genes associated with TSC self-renewal, as well as TSC differentiation genes and the pluripotency factor Nanog. Its binding sites are also enriched for motifs for known TSC-expressed TFs.

3.7.4: Tead4 co-binds many Eomes-bound genes

The Tead4 binding motif was enriched in Eomes-bound sites, leading us to investigate potential co-regulated targets. Over 50% of Eomes-bound sites are also Tead4-bound, suggesting there may be a high degree of gene co-regulation. Interestingly, the majority of Eomes-Tead4-co-bound genes are genes more highly expressed in Eomes+ (untreated) TSCs. Therefore, Eomes and Tead4 may be specifically co-regulating TSC-expressed genes.

3.8 Appendix

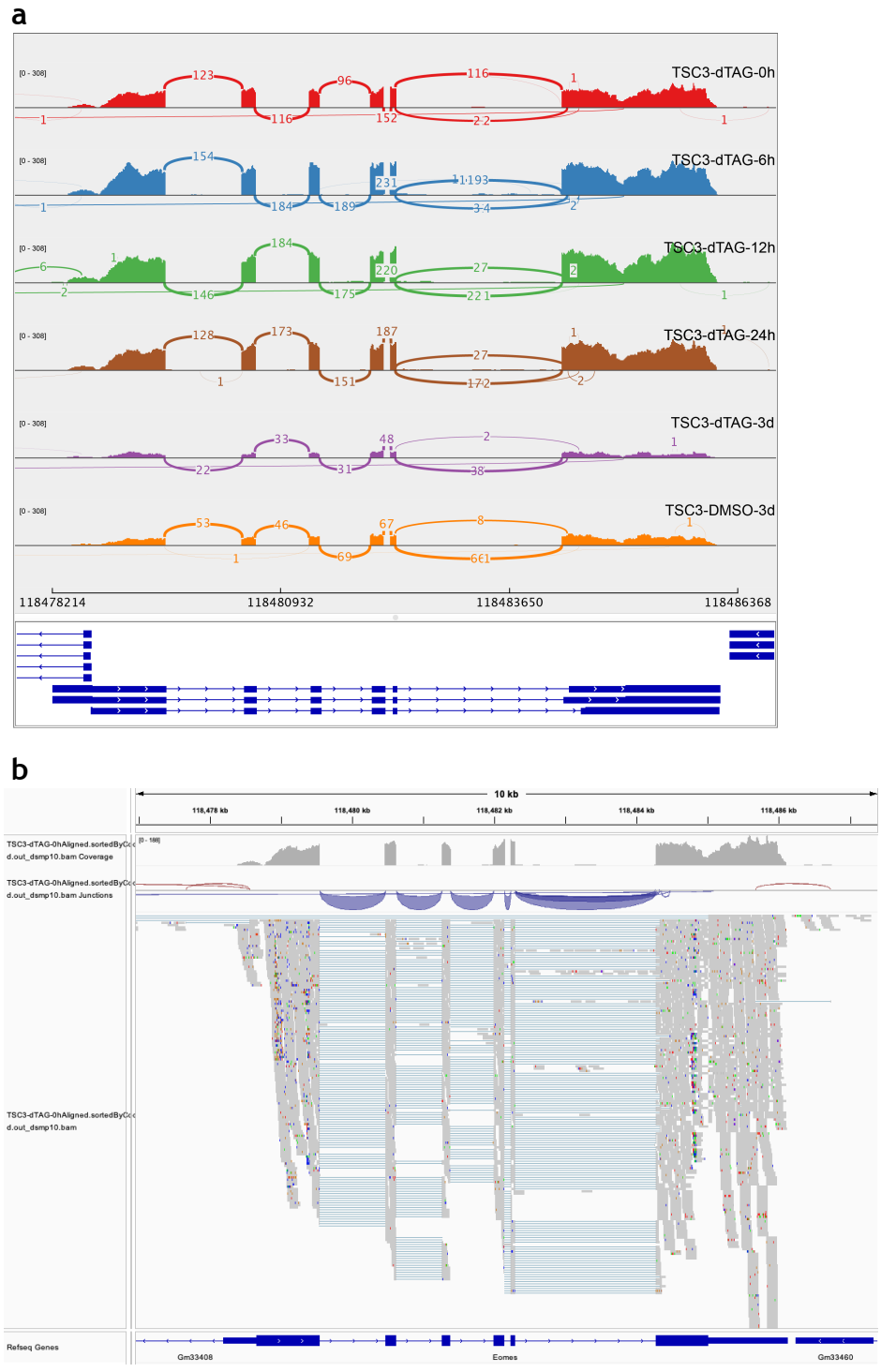


Figure 3.13: Sashimi plot of Eomes transcript splicing in Eomes$^{deg/deg}$ clone TSC3.
(a) Splicing of the Eomes mRNA does not show any obvious exon skipping.
(b) Some shorter sequenced mRNA products (not containing the first few exons) appear in clone TSC3.

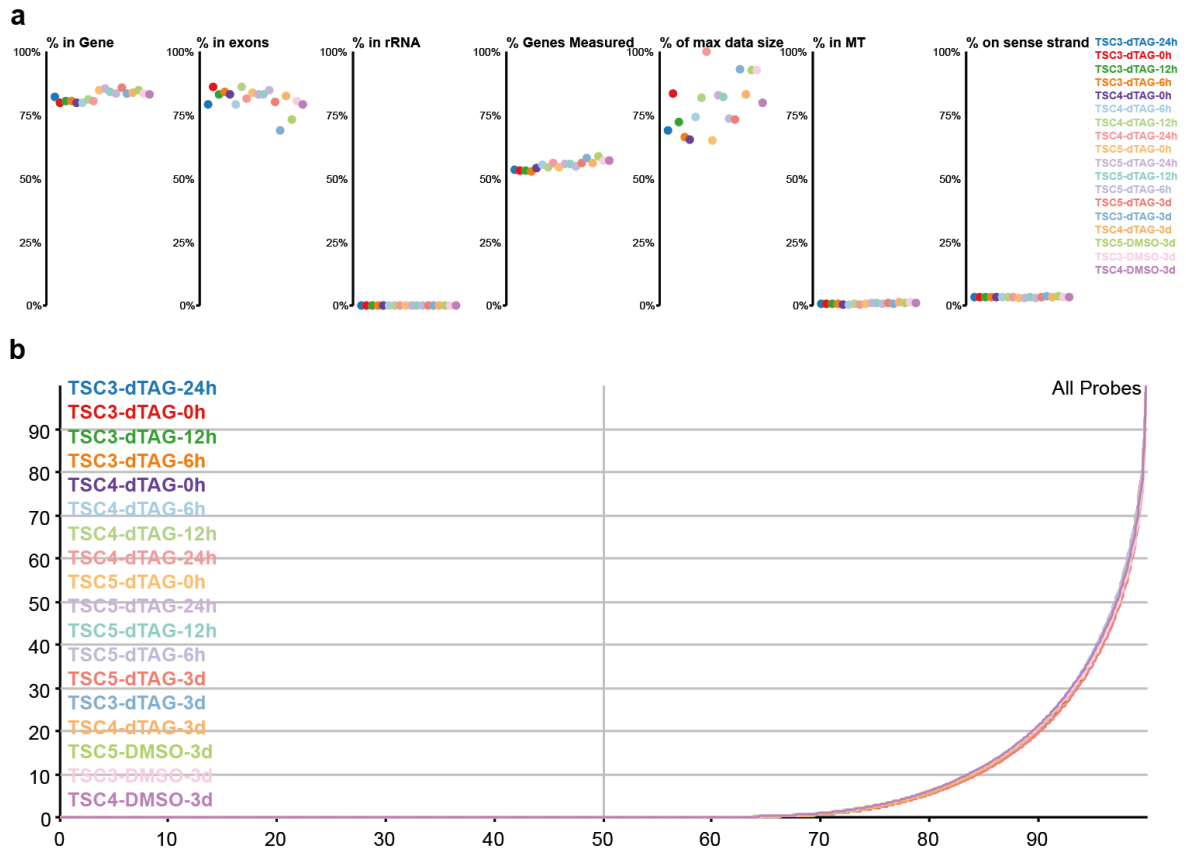


Figure 3.14: Quality control of RNA-seq data.
(a) Quality control of RNA-seq samples. Statistics shown for each sample.
(b) QQ plot for RNA-seq samples. For each sample, the sum of quantitations for each probe is calculated. Then, probes are ordered from lowest to highest quantitation. The x axis shows the percentile position of probes against the percentile quantitation until that point. Closely-aligning traces indicate that the data between samples is well-normalised.

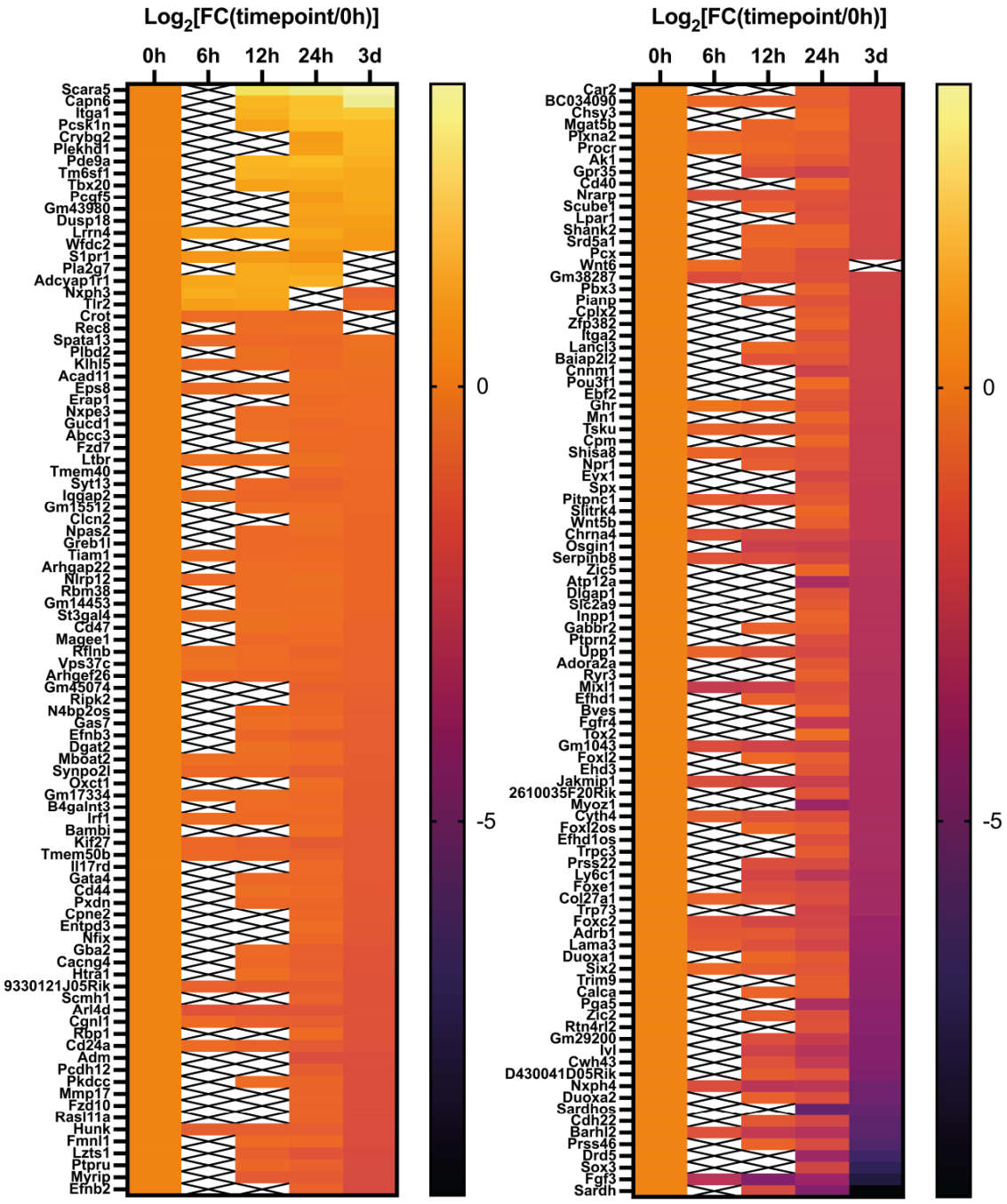


Figure 3.15: Heatmap of log_2[fold change(timepoint/0h)] of genes detected in more than one differentiation timepoint.

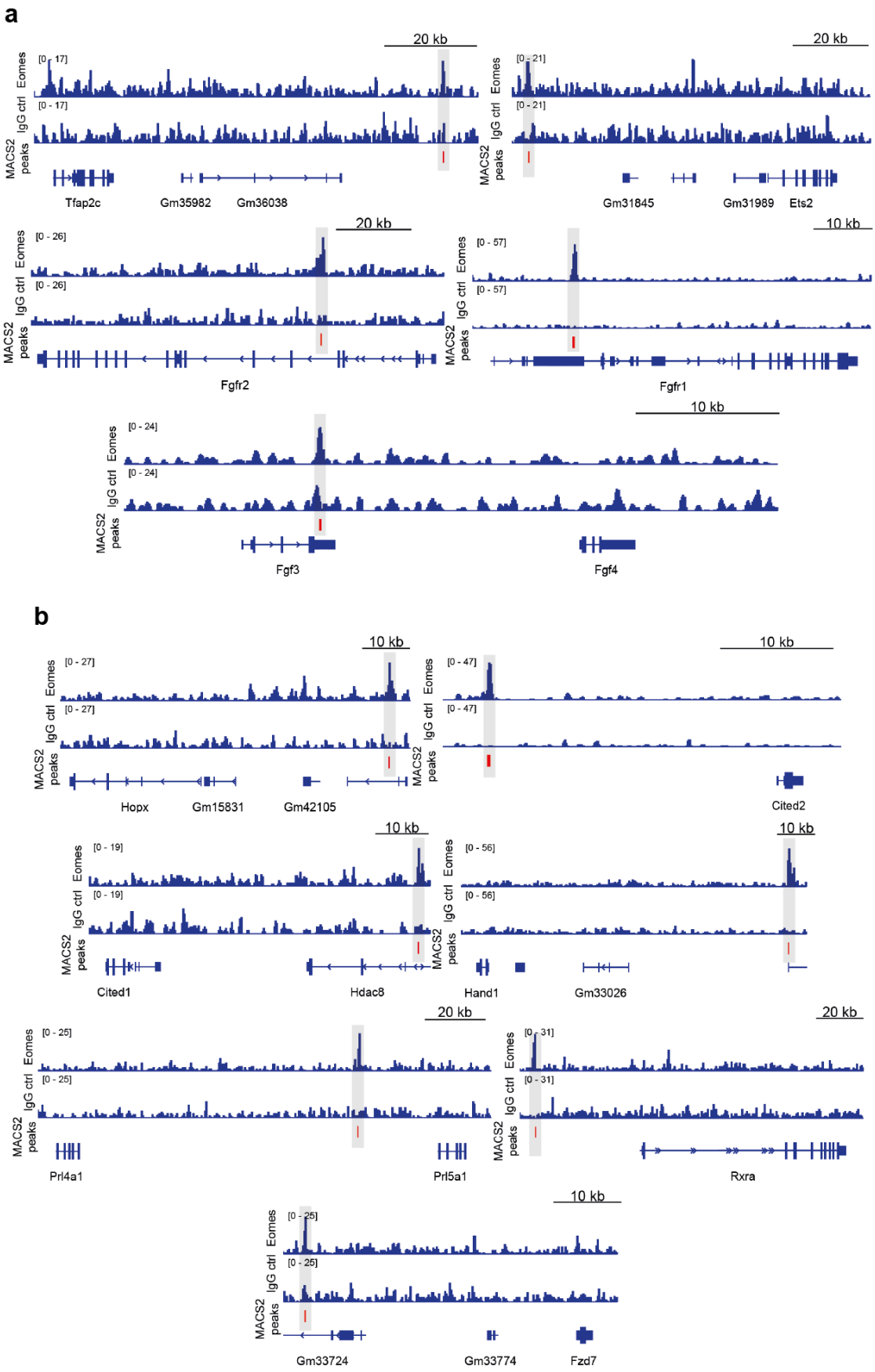


Figure 3.16 (continued on next page)

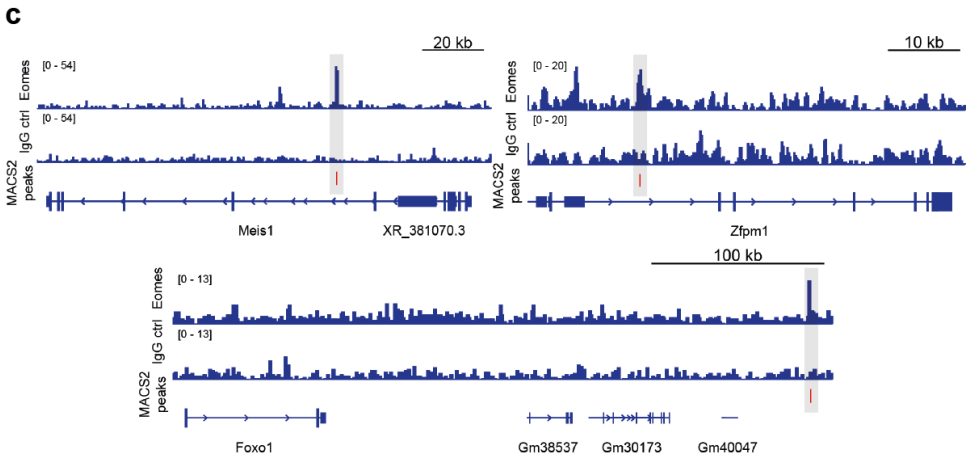


Figure 3.16: Eomes binds near genes associated with TSCs and TSC differentiation.
(a) TSC-expressed genes.
(b) Differentiating TSC-expressed genes. Hopx and Hand1 show biphasic expression, with higher expression in TSCs and cells differentiated for multiple days).
(c) Genes newly-identified as expressed in the TB lineage.

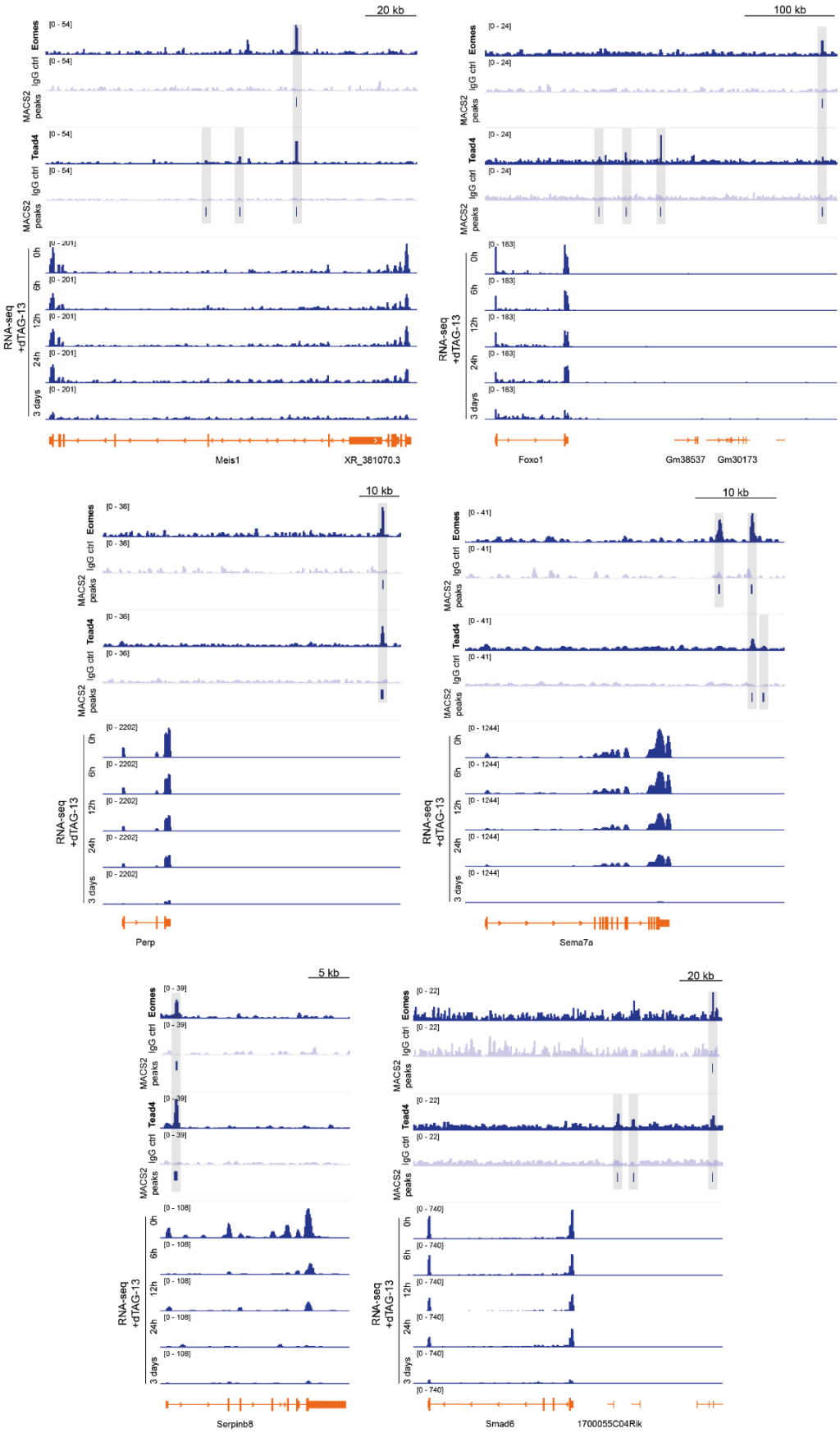


Figure 3.17: Eomes and Tead4 co-bind mostly to genes more highly expressed in untreated TSCs.

Eomes-bound genes up in TSCs	Eomes-bound genes up in Eomes depletion	Tead4-Eomes-co-bound genes up in TSCs	Tead4-Eomes-co-bound genes up in Eomes depletion
2700081O15Rik	3830417A13Rik	Acot12	3830417A13Rik
Acot12	Adamts8	Adgrl1	Adamts8
Acsl6	Asah2	Agap1	Asah2
Adcy2	Bcor	Apba2	Cd274
Adgrl1	Cd274	Arhgef26	D5Ertd579e
Adora2b	Csrnp1	Bambi	Dnah10
Agap1	D5Ertd579e	BC034090	Ehhadh
Ahcy	Dnah10	Bmp6	Eomes
Ankmy1	Ehhadh	Bmp8b	Ido2
Apba2	Eomes	Car2	Jade2
Arhgef26	Ggta1	Ccdc3	Oas1a
Arl4d	Ido2	Ccdc77	Pfkfb3
Arrdc4	Il11ra1	Chsy3	Pla2g7
Asph	Jade2	Cpne3	Rilpl2
Asphd2	Ldoc1	Crot	Trps1
B3gnt7	Lgals9	Cysltr1	Zbtb44
B4galnt3	Manba	Ddc	
Bambi	Megf11	Duox2	
BC034090	Oas1a	Dynlt1b	
Bmp6	Pard6b	Elmo1	
Bmp8b	Pax8	Emc10	
Car2	Pex5l	Entpd3	
Ccdc3	Pfkfb3	Eps8	
Ccdc77	Pla2g7	Erap1	
Chsy3	Psd4	Ezr	
Col27a1	Rilpl2	Fam124a	
Cpne3	Slc38a4	Fkbp5	
Crot	Tmsb10	Fmr1	
Ctbp2	Trps1	Foxo1	
Cwh43	Uckl1	Gas7	
Cysltr1	Zbtb44	Gbx2	
Ddah1		Gemin6	
Ddc		Gm1043	
Duox2		Gpr68	
Duoxa2		Gucd1	
Dynlt1b		Haus8	
Efnb3		Hs3st3b1	
Ehd3		Htra1	
Eif1b		Ivns1abp	
Elf5		Jakmip1	
Elmo1		Kcnn2	

Eomes-bound genes up in TSCs	Eomes-bound genes up in Eomes depletion	Tead4-Eomes-co-bound genes up in TSCs	Tead4-Eomes-co-bound genes up in Eomes depletion
Emc10		Kctd14	
Entpd3		Kdr	
Eps8		Kif27	
Erap1		Lig4	
Etv5		Meis1	
Ezr		Mfge8	
Fam124a		Mgat5b	
Fam43b		Myrip	
Fgf3		N4bp2	
Fgfr1		Nebl	
Fkbp5		Nlrp12	
Fmr1		Nrp2	
Foxo1		Perp	
Fzd7		Pgap1	
Gabbr2		Popdc3	
Gas7		Pou3f1	
Gbx2		Prss46	
Gemin6		Prss50	
Gm1043		Ptpn3	
Gnao1		Rai14	
Gpc1		Ranbp17	
Gpr68		Rassf10	
Gucd1		Rftn1	
Haus8		Ripk2	
Hk1		Ryr3	
Hlx		Scarf2	
Hs3st3b1		Sema7a	
Htra1		Serpinb8	
Hunk		Slc45a1	
Irx3		Slitrk4	
Itch		Smad6	
Ivns1abp		Smyd3	
Jakmip1		Soga3	
Kcnn2		Sord	
Kctd14		Srd5a1	
Kdr		Sulf2	
Khdrbs3		Tbx21	
Kif27		Tinagl1	
Lig4		Tmem114	
Lpo		Tmem40	
Meis1		Tmem64	

Eomes-bound genes up in TSCs	Eomes-bound genes up in Eomes depletion	Tead4-Eomes-co-bound genes up in TSCs	Tead4-Eomes-co-bound genes up in Eomes depletion
Mfge8		Trim36	
Mgat5b		Trim9	
Mn1		Trp73	
Myc		Tsku	
Myrip		Tspan9	
N4bp2		Uaca	
Nebl		Vldlr	
Nfatc1		Vstm2a	
Ninj2		Zfp820	
Nlrp12		Zfpm1	
Nrarp		Zic5	
Nrp2			
Pbx3			
Pdzrn4			
Perp			
Pgap1			
Pkdcc			
Plbd2			
Plxna4			
Popdc3			
Pou3f1			
Ppp1r14c			
Prss46			
Prss50			
Ptp4a3			
Ptpn3			
Rai14			
Ranbp17			
Rassf10			
Reln			
Rftn1			
Ripk2			
Ryr3			
Scarf2			
Sema7a			
Serpinb8			
Slc20a1			
Slc45a1			
Slitrk4			
Smad6			
Smyd3			

Eomes-bound genes up in TSCs	Eomes-bound genes up in Eomes depletion	Tead4-Eomes-co-bound genes up in TSCs	Tead4-Eomes-co-bound genes up in Eomes depletion
Soga3			
Sord			
Srd5a1			
Sulf2			
Tbx21			
Tinagl1			
Tmem114			
Tmem40			
Tmem64			
Trim36			
Trim9			
Trp73			
Tsku			
Tspan9			
Ttbk1			
Uaca			
Ube2ql1			
Vldlr			
Vps37c			
Vstm2a			
Zfp42			
Zfp820			
Zfpm1			
Zic5			
Zmat4			

Table 3.8: Differentially expressed genes from Eomes protein depletion timecourse experiment that are also Eomes-bound or Eomes/Tead4-co-bound (merged across timepoints).

CHAPTER 4: INVESTIGATING EOMES FUNCTIONS IN THE SPECIFICATION OF DE AND MIDLINE PROGENITORS

4.1 INTRODUCTION

Past experiments have established a critical role for Eomes in specification of numerous epiblast -derived progenitors during gastrulation. Eomes binds to and regulates expression of Runx1, a TF essential for conferring haemogenic competency in the ExM, and regulates accessibility of sites normally bound by the TF SCL (stem cell leukemia) to specify primitive erythrocytes (Harland et al., 2021). Via binding of Eomes to critical T-box enhancers of Mesp1/2, known master regulators of cardiac fate, Eomes is also essential in promoting cardiac mesoderm identity (Costello et al., 2011). By characterising epiblast-specific Eomes KO embryos, it has also been demonstrated that Eomes is essential for DE and midline (node/notochord) specification, as well as correct initiation of gastrulation. Epiblast-specific Eomes KO embryos fail to upregulate PS/DE marker Cer1, as well as DE markers Hhex and FoxA2. The node marker Shh and the morphological node itself are also absent. These embryos also fail to downregulate E-cadherin and nascent mesoderm cells are blocked in their ability to delaminate and ingress through the PS, resulting in a profound gastrulation block and inability to specify the three germ layers (Arnold et al., 2008).

Fate mapping experiments using $Eomes^{Cre}$ x $ROSA26^{LacZ}$ embryos, in which Eomes expression activates expression of the LacZ reporter, indicated that $Eomes^{+}$ cells selectively contribute to the cardiac progenitors of the myocardium and endocardium, as well as the DE and node and its derivative the notochord. Moreover, chimera experiments assessing the developmental potential of Eomes-null ES cells in WT embryos clearly show a cell-autonomous requirement for Eomes in the specification of these discrete progenitor populations, as Eomes-null cells were specifically excluded from the heart, gut tube and notochord of E9.5 embryos (Costello et al., 2011).

To date, however, little is known about the roles of Eomes in DE and midline specification. This Chapter aims to provide a molecular insight into how Eomes functions to specify these discrete lineages. Since it is not practical to study this in the normal gastrulating embryo I first developed and refined an *in vitro* protocol that accurately recapitulates the specification of DE and midline progenitors using embryoid body (EB) differentiation cultures. I then exploited the Eomes$^{deg/deg}$ ESC lines described in Chapter 3 to rapidly and efficiently remove endogenous Eomes protein. This confirmed that Eomes activity is essential for specification of DE progenitors *in vitro*. By exploiting the ability to ablate Eomes function at specific timepoints during EB differentiation in combination with bulk RNA- and ATAC-seq I have identified Eomes-dependent genes known to be essential in specification of the DE and node/notochord, and further characterised the time window in which Eomes depletion exhibits its effects. I have also identified potential chromatin sites that may be involved in the regulation of these genes.

In addition to the bulk experiments described below, this project will also benefit from the single-cell "multi-ome" (RNA- and ATAC-seq) experiments that we have carried out in collaboration with Bertie Göttgens' group at the Stem Cell Institute of the University of Cambridge. A preliminary analysis of the data acquired is presented here, but we have yet to carry out in-depth analysis of the effect of Eomes protein in individual cells/cell types. This will allow us to circumvent the noise inherent in bulk experiments and definitively correlate transcriptional and accessibility states within the individual cells/cell types specified in the EBs.

4.2 Optimisation of a DE progenitor EB differentiation protocol in mESCs

In normal embryogenesis, DE progenitors are specified during gastrulation in a portion of the cells traversing the anterior primitive streak (APS). This region of the PS is associated with high levels of Nodal signalling, along with WNT signalling. In contrast, the APS is a BMP-low environment, as the cells are situated distal to the BMP-expressing proximal epiblast. This unique signalling environment is thought to promote the DE GRN, while suppressing ectodermal and mesodermal gene expression. These observations in the normal embryo were used to devise an *in vitro* mESC differentiation protocol utilising cytokines and small molecules to mimic the APS signalling environment.

In order to study the molecular role of Eomes in the specification of DE *in vitro*, I optimised an EB-based DE differentiation model (Figure 4.1a) based on the assessment of cell surface receptors CXCR4 and cKit, which are known to combinatorially mark DE progenitors (Green et al., 2011; Li et al., 2011; Mfopou et al., 2014; Oss-Ronen et al., 2020; Pour et al., 2022). I also used the Eomes$^{deg/deg}$ mESCs (described in Chapter 3), which contain the endogenous Eomes reporter mCherry, to monitor endogenous Eomes expression. I initially cultured mESCs in 2i/LIF conditions (a defined medium) for 2-3 days in order to promote ground-state pluripotency. This was done to ensure a consistent starting cell population, circumvent potential variability caused by non-defined, serum-containing medium, and therefore guarantee a consistent differentiation result. I then generated EBs by removing LIF and cultured them for 2 days in serum-free medium. To promote a DE progenitor identity, on day 2.0 of differentiation I added a high concentration of ActA (50 ng/ml) to the medium as a Nodal agonist, and cultured the EBs for 3 days in graded concentrations of WNT agonist CHIR-99021 (“CHIR”) and BMP inhibitor LDN-193189 (“LDN”). To assess the performance of the protocol across genetic backgrounds, I trialled it on Eomes$^{deg/deg}$ cell lines generated on two different backgrounds, CCE and E14-Runx-Venus (Figure 4.1b,c).

In the case of both genetic backgrounds, intermediate concentrations (50-100 nM) of LDN resulted in the most efficient specification (highest percentage) $CXCR4^+/cKit^+$ cells. In all conditions but one (no CHIR, maximum LDN), the CCEs were less efficient in $CXCR4^+/cKit^+$ cell specification than the E14-RV cells, which may be due to genetic background differences resulting in different signalling and gene expression responses to the same external molecular stimuli (Figure 4.2a,b). On the CCE background, an intermediate concentration of CHIR (1μM) contributed towards the highest percentage of $CXCR4^+/cKit^+$ cells (Figure 4.2a), whereas this was variable in the E14-RV cells (either intermediate or high, depending on the LDN concentration) (Figure 4.2b).

The conditions I settled on were 50ng/ml ActA, 1μM CHIR, and 50nM LDN, as these produced the highest percentage of $CXCR4^+/cKit^+$ cells in the CCE $Eomes^{deg/deg}$ cell line (41.2%), and the second-highest in the E14-RV $Eomes^{deg/deg}$ cell line (60.7%). In the E14-RV line, the difference between the first- and second-most efficient cytokine/small molecule combinations was negligible (63.3 vs. 60.7%), and since the second-most efficient combination required half the concentration of both CHIR and LDN compared to the first, it would be more resource-efficient (Figure 4.1b,c).

Earlier experiments with 'condition grids' similar to the ones in Figure 4.1b,c are included in Chapter 4 Appendix Figure 4.19, showing similar results for the variety of concentrations used for CHIR and LDN.

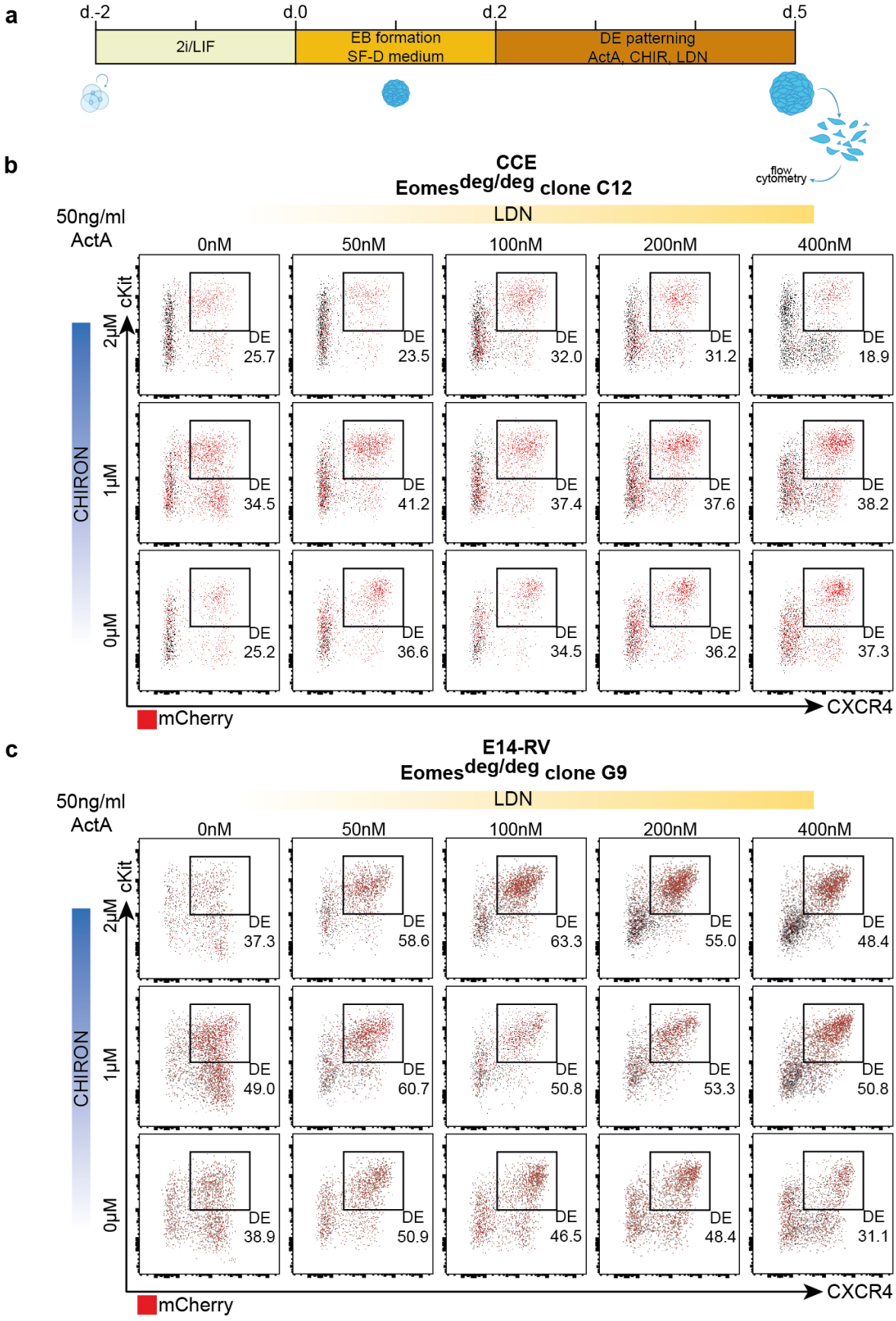


Figure 4.1 (legend on next page)

Figure 4.1: Optimisation of an EB-based DE differentiation protocol in mESCs.
(a) Cartoon summary of DE differentiation timeline.
(b,c) Flow cytometry plots of CXCR4 and cKit cell surface markers, as well as mCherry reporter expression in day 5.0 EBs cultured in various conditions. A high concentration of ActA, and intermediate concentrations of WNT agonist and BMP inhibitor small molecules result in higher efficiency of DE progenitor specification both in cells of CCE (b) and E14-RV (c) genetic backgrounds.

Interestingly, in the CCE background line, at the highest CHIR concentrations (highest exogenous WNT activation), $CXCR4^+/cKit^+$ cell specification efficiency peaks at a higher concentration of LDN than in the two lower CHIR concentration trendlines (Figure 4.2a). As BMP is known to activate Wnt3 signalling at the onset of gastrulation, thus promoting Nodal signalling, which is essential for DE specification (Ben-Haim et al., 2006), this might reflect a requirement for higher exogenous WNT signalling promotion in an environment with increased BMP inhibition. This is reflected in the E14-RV background line as well (Figure 4.2b), with the highest CHIR concentration culture conditions yielding a greater percentage of DE progenitors at a higher concentration of LDN (and thus, greater BMP signalling inhibition). However, even greater concentrations of LDN overall appear to reduce the efficiency of $CXCR4^+/cKit^+$ cell specification, suggesting that a baseline level of BMP is required, possibly due to its involvement in the WNT/Nodal axis activation.

As the $Eomes^{deg/deg}$ lines contain an mCherry reporter, I also assessed the presence of mCherry in the EB cell populations as a proxy for Eomes expression in all differentiation conditions (Figure 4.1b,c; Figure 4.3). In both genetic backgrounds, mCherry protein is enriched in the $CXCR4^+/cKit^+$ population relative to the total EB cell population. Consistent with the generally lower efficiency of $CXCR4^+/cKit^+$ specification in the CCE background, mCherry presence was also lower in CCE EB cells compared to E14-RV. However, between the $CXCR4^+/cKit^+$ populations of the two genetic backgrounds, there was no significant difference in the percentage of $mCherry^+$ cells. This suggests that, although the EBs from the two genetic backgrounds overall have significant differences in Eomes expression, the two double-positive populations are more similar in this respect.

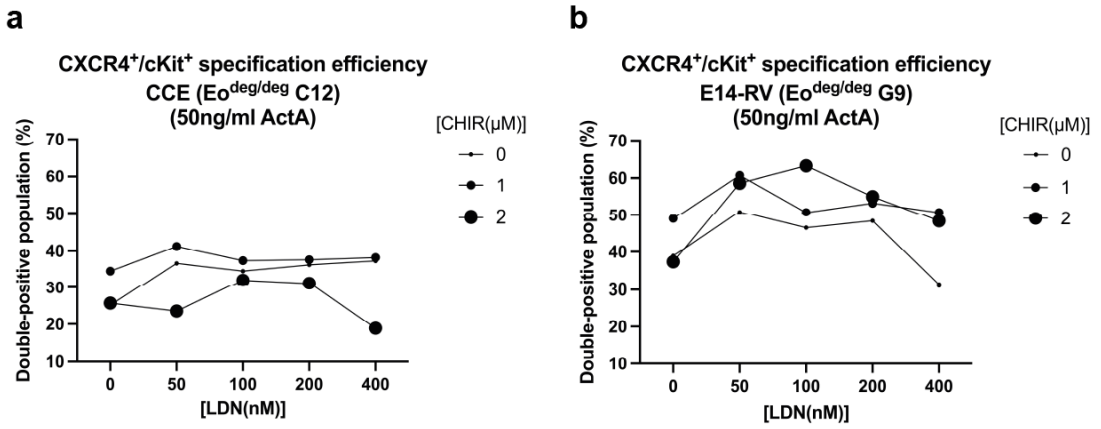


Figure 4.2: Variable effects of graded CHIR and LDN concentrations in CXCR4⁺/cKit⁺ cell specification efficiency.

(a) Scatterplot of CXCR4⁺/cKit⁺ specification efficiency in the CCE background as a function of CHIR and LDN concentration. An intermediate concentration of CHIR results in the highest percentage of CXCR4⁺/cKit⁺ population regardless of LDN concentration, whereas the highest CHIR concentration appears to have an inhibitory effect compared to no CHIR. Additionally, at the highest CHIR concentration, comparatively higher concentrations of LDN are required for greater efficiency.

(b) Scatterplot of CXCR4⁺/cKit⁺ specification efficiency in the E14-RV background as a function of CHIR and LDN concentration. Some WNT activation via CHIR addition to the differentiation medium is required to increase DE specification efficiency, but the effect of the highest CHIR concentration isn't consistently inhibitory as it is in the case of the CCE background. However, as in the CCE background, at the highest CHIR concentration, a higher LDN concentration is required to maximise DE progenitor specification relative to the intermediate and low CHIR concentrations.

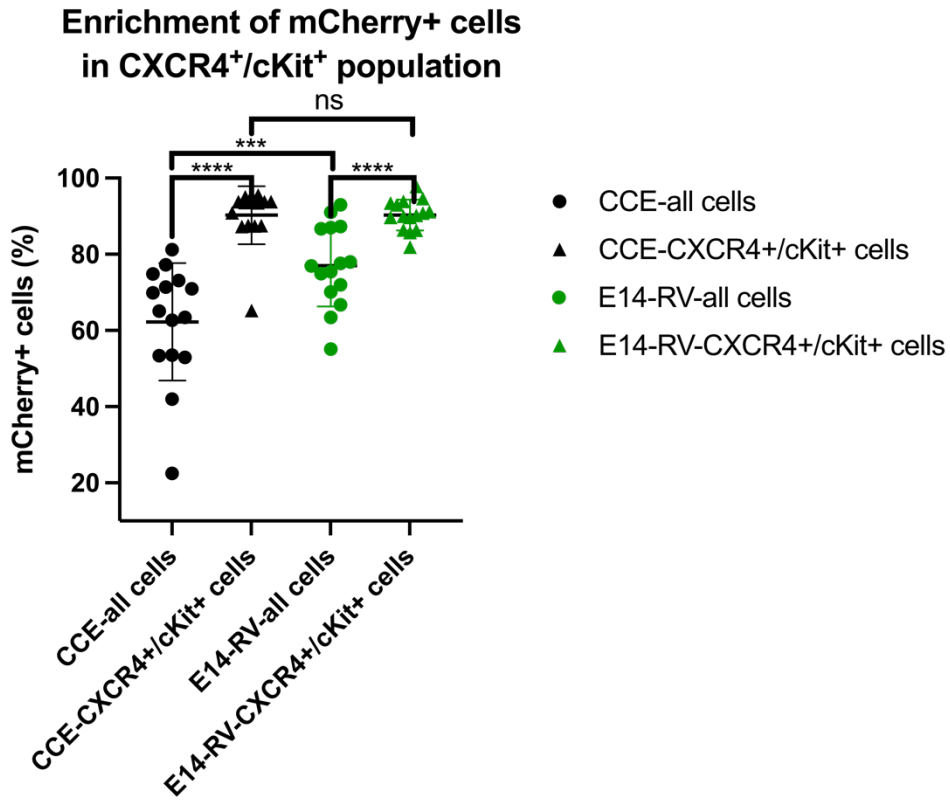


Figure 4.3: mCherry reporter is enriched in CXCR4⁺/cKit⁺ populations of EBs from both CCE and E14-RV genetic backgrounds.
Bar graph of percentage of mCherry⁺ cells in all cells of EBs from each genetic background, as well as only CXCR4⁺/cKit⁺ cells of EBs. This suggests that during the specification of this cell population, Eomes expression has, at one point, been higher in these cells relative to all cells in the EB. Paired *t* tests were carried out to determine whether differences in mCherry⁺ cells were significant (****, P value≤0.0001; ***, P≤0.001; ns, not significant). Error bars indicate standard deviation (SD). n=15 in each group.

4.3 CXCR4+/cKit+ cells in DE EB cultures express DE lineage markers

To further confirm that the day 5.0 CXCR4+/cKit+ cell population produced from the DE EB differentiation protocol have a DE identity, RNA-seq was carried out on a sorted CXCR4+/cKit+ population of both CCE and E14-RV genetic backgrounds, as well as the corresponding undifferentiated mESCs.

Only uniquely-mapping reads were assessed during the RNA-seq analysis. To assess library quality, the fraction of reads in genes and exons was assessed, with a higher proportion indicating a library with little contamination and little unspliced mRNA. Additionally, contamination from rRNA and reads aligning to mitochondrial DNA was quantified (Appendix Figure 4.20a). Finally, a QQ plot was generated to assess if there were any significant differences in probe quantitation between samples (Appendix Figure 4.20b).

Principal component analysis (PCA) demonstrated that the greatest variation between samples arose between the double-positive cells and the ESCs, with double-positive cells grouping on one end of PC1 (which accounts for 76% of inter-sample variance) and ESCs on the other (Figure 4.4a). This suggested that genetic background contributed comparatively little to inter-sample variance, with PC2 accounting for 14%. The remaining PCs contributed very little to variance (Figure 4.4b).

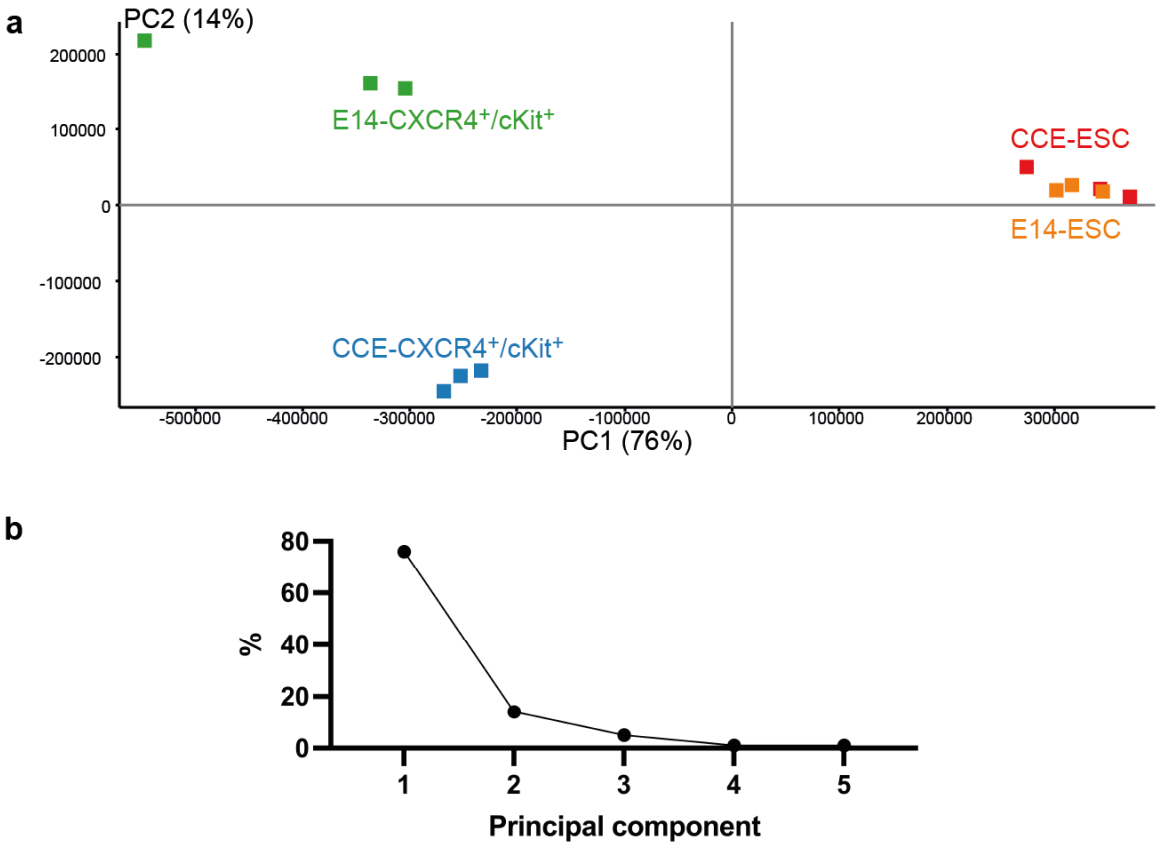


Figure 4.4: CXCR4+/cKit+ transcriptomes are distinct from those of undifferentiated ESCs.
(a) PCA of RNA-seq replicates. PC1 accounts for most of the difference between samples. (b) Scatterplot of the percentage contributions of each PC to inter-sample variance. Most of the transcriptional variance is accounted for by PC1.

The highest-contributing components to PC1 were assessed to determine their function in ESCs and DE cells (Figure 4.5a). As expected from the PCA, the highest contributors to PC1 clustered in two broad groups, one more highly expressed in the CXCR4$^+$/cKit$^+$ cells, and one in the ESCs. Gene ontology analysis of the two gene groups indicated that those more highly expressed in the CXCR4$^+$/cKit$^+$ cells are involved, among others, in processes such as (i) EMT and cell motility associated with gastrulation, (ii) the specification or morphogenesis of mesodermal and various definitive endodermal cell lineages, (iii) embryonic axis specification, and (iv) promotion of active chromatin marks. Conversely, the genes more highly expressed in ESCs were associated with processes to do with (i) gene silencing, (ii) maintenance of the stem cell state, and (iii) response to LIF. Interestingly, there is a noticeable difference in the expression level of the CXCR4$^+$/cKit$^+$-high genes between the double-positive cells of CCE and E14-RV background, with their expression stronger in the E14-RV double-positive cells.

Differential expression analysis was carried out between the CXCR4$^+$/cKit$^+$ cell and ESC transcriptomes. 8,787 genes were found to be significantly overexpressed in CXCR4$^+$/cKit$^+$ cells, and 9,342 were significantly overexpressed in the ESCs. I selected a few established markers of each cell type and found that the DE master regulators Foxa2 and Sox17 (Dufort et al., 1998; Kanai-Azuma et al., 2002), as well as various other DE markers, are significantly overexpressed in the CXCR4$^+$/cKit$^+$ cells compared to ESCs, as well as various other DE markers. In contrast, expression of the Yamanaka pluripotency factors (Takahashi & Yamanaka, 2006), as well as pluripotency markers Nanog and Esrrb (Chambers et al., 2003; Festuccia et al., 2012; Mitsui et al., 2003), was significantly reduced in the double-positive population relative to ESCs. Finally, in order to ensure that I could differentiate between DE and primitive endoderm, I assessed Sox7 expression (a primitive endoderm marker (Artus et al., 2011; Futaki et al., 2004; Murakami et al., 2004)), and found that it was reduced in expression compared to undifferentiated mESCs (Figure 4.5b).

Overall, these results indicate that the CXCR4$^+$/cKit$^+$ cell transcriptomes are markedly different from those of undifferentiated ESCs. The two groups of cells have significantly different expression of 18,129 genes. The genes accounting for the greatest difference between the two groups are involved either in peri-gastrulation events, such as embryonic axis specification, gastrulation-associated cell motility, or specification of mesodermal and endodermal lineages; or alternatively, in events associated with maintenance of pluripotency. Additionally, DE markers are significantly upregulated in the double-positive cells (and particularly so in those of the E14-RV genetic background), while pluripotency factors significantly downregulated. Collectively this analysis suggested that the double-positive cells correspond to a gastrulation-stage embryo, with GO terms for lineages arising at the APS (mesendoderm, cardiac mesoderm, definitive endoderm) strongly suggesting that the cells largely correspond to APS-derived lineages.

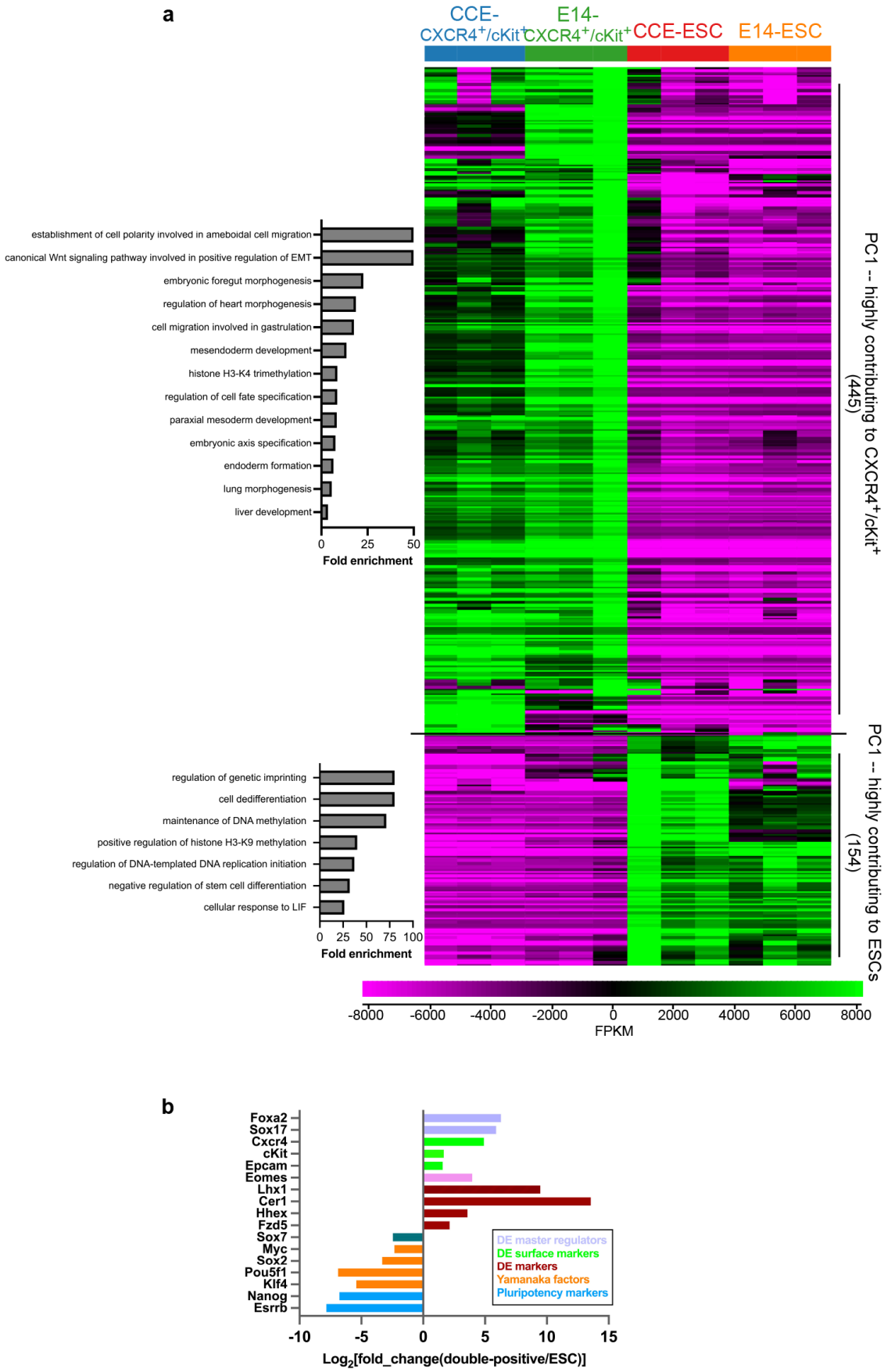


Figure 4.5 (legend on next page)

Figure 4.5: CXCR4+/cKit+ cells from day 5.0 DE EBs express DE and APS markers and downregulate pluripotency markers.
(a) The highest-contributing genes of PC1 are expressed highly either in the double-positive day 5.0 EB cells, or in the ESCs. Enriched GO terms for each of these two groups are shown beside them.
(b) Selected DE and pluripotency markers are significantly differentially expressed between day 5.0 EB CXCR4+/cKit+ cells and ESCs.

4.4 Eomes protein ablation in DE EB cultures results in specific loss of CXCR4+/cKit+ double-positive cell population

As described in Chapter 3, I demonstrated the functionality of the Eomes^{deg} allele by carrying out a degradation time course for Eomes protein by adding dTAG-13 to day 3.0 DE EBs and sampling them for WB analysis at selected timepoints. Eomes protein was undetectable via WB after 30 minutes of dTAG-13 presence and remained so after several hours of incubation (Figure 3.1d, 3.2b).

In order to determine the time window during which Eomes activity is functionally required in specifying the DE progenitor population, I carried out intracellular flow cytometry to characterise the dynamics of Eomes expression in the DE EB cultures using the E14-RV Eomes$^{deg/deg}$ mESCs (clone G9) (Figure 4.6). Eomes protein is initially detected around day 2.5, 12h after cytokine and small molecule addition to the differentiation medium. It then peaks at day 3.0 with approximately 80% of cells being Eomes+ and decreases considerably by day 4.0 with 18% Eomes+ cells, becoming virtually undetectable by day 4.5. This suggests a window of action during the period of maximal Eomes expression, between day 2.5-4.0.

Although I did not carry out any intracellular flow cytometry experiments for Eomes and mCherry at the same time, it is noteworthy that the peak percentage of Eomes+ cells (day 3.0, ~80%) (Figure 4.6) is very close to the percentage of mCherry+ cells in day 5.0 E14-RV EBs (Figure 4.3, E14-RV, all cells), suggesting mCherry is a reliable marker of past or present Eomes expression. It also highlights the stability of mCherry, as it is still

present in cells at a time (day 5.0) when almost all of them have already lost Eomes expression (Figure 4.6).

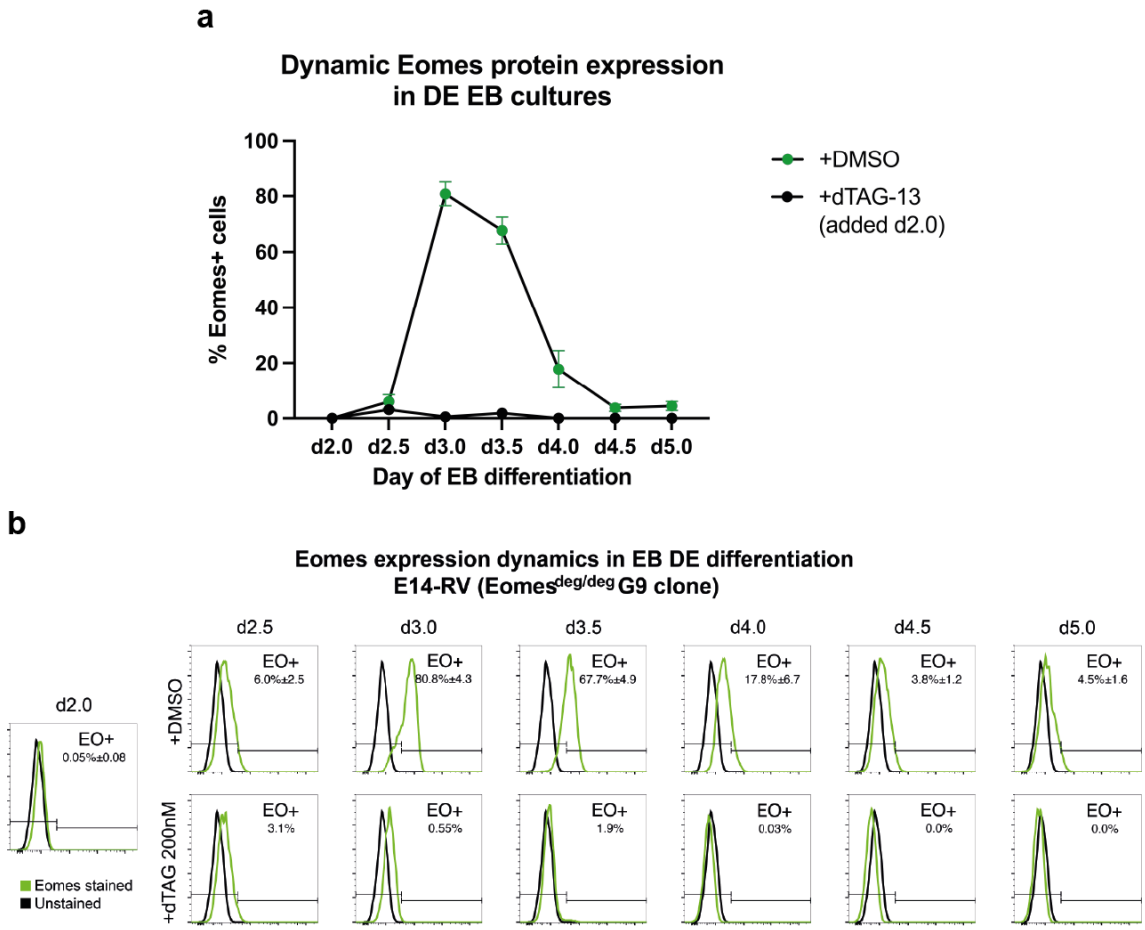


Figure 4.6: Dynamic Eomes expression in DE EB cultures.
(a) Scatterplot of Eomes⁺ cell percentage at different times during DE EB differentiation cultures. dTAG-13 addition on day 2.0 of DE differentiation almost completely ablates presence of Eomes protein in EBs.
(b) Histograms of Eomes expression in DMSO- and dTAG-13-treated DE EBs. Cells unstained for Eomes are used as a measure of comparison in each plot. N=3 in each DMSO-treated timepoint, N=1 in dTAG-treated groups. Error bars indicate SD, and percentages in flow panels indicate mean+/-SD.

To more precisely define the critical window of Eomes activity, I added dTAG-13 to the culture medium at different timepoints, followed by assessment of DE marker CXCR4, cKit, and EpCAM expression via flow cytometry on day 5.0 of differentiation (Figure 4.7a). Addition of dTAG-13 on or before day 2.5 resulted in complete loss of the CXCR4$^+$/cKit$^+$/EpCAM$^+$ population. [Although I did not initially use EpCAM (epithelial cell adhesion molecule) to assess DE specification, it is also highly expressed in CXCR4$^+$/cKit$^+$ DE progenitors (Green et al., 2011), and I used it from this point onwards as further confirmation of the DE identity of the cells.] The loss was partial (as measured by percentage of the total cell population) when dTAG-13 was added between day 3.0-3.5 of the differentiation. dTAG-13 addition from day 4.0 onwards did not have a significant effect on specification of the DE population. Although addition of dTAG-13 on day 4.5 did seem to reduce the DE population, I considered this an experimental artefact, as I observed a similar reduction in the DE population in the vehicle (DMSO) control, with no significant difference between the dTAG-13- and DMSO-treated EBs from the day 4.5 addition timepoint (Figure 4.7b).

By contrast, dTAG-13 addition at earlier timepoints (day 2.0-2.5) had a dramatic effect on the day 5.0 CXCR4 single-positive population (Figure 4.7c), which was greatly increased as a percentage of the total EB cell population. This effect was lost when dTAG-13 was added after that timepoint, in which case the CXCR4$^+$ population was actually lower than untreated day 5.0 EBs. Conversely, the cKit single-positive population decreased at early dTAG-13 addition timepoints (Figure 4.7d), but this effect was lost after day 2.5. These data suggest that cells Eomes-depleted EBs fail to upregulate cKit after day 2.5, thus explaining the increase in CXCR4 single-positive and decrease in cKit single-positive cells at Eomes depletion on or before day 2.5.

Finally, I assessed the presence of the cell surface marker EpCAM in the CXCR4$^+$/cKit$^+$ population of dTAG-13- and DMSO-treated EBs (Figure 4.7e). As the DE is an epithelial embryonic tissue, and EpCAM is involved in cell-cell adhesion in epithelia (Litvinov et al.,

1994), I used EpCAM expression as a proxy in assessing the quality of the DE identity of the double-positive CXCR4+/cKit+ cells. The CXCR4+/cKit+ population generated from untreated EBs is almost uniformly positive for EpCAM (>95%). DMSO-treated EBs showed no significant difference in EpCAM expression relative to untreated EBs. In the case of dTAG-13 treatment, addition on day 3.0 or later (i.e. the EBs that had time to specify a significant CXCR4+/cKit+ population) showed a small reduction in the percentage of EpCAM+ cells. However, EBs with dTAG-13 added before day 3.0 (i.e. EBs containing no significant population of CXCR4+/cKit+ cells) showed significantly fewer EpCAM+ cells. This suggests that in this last group, DE cells cannot be properly specified, and the few CXCR4+/cKit+ cells that are formed are qualitatively different from the DE cells specified in the other conditions.

Collectively these experiments demonstrate that addition of dTAG-13 at early timepoints results in specific loss of the CXCR4+/cKit+ putative DE population. DE specification appears to occur gradually over the course of approximately 36h, between day 2.5 and day 4.0 of DE EB differentiation, and is acutely dependent on Eomes expression on day 2.0-2.5 of differentiation.

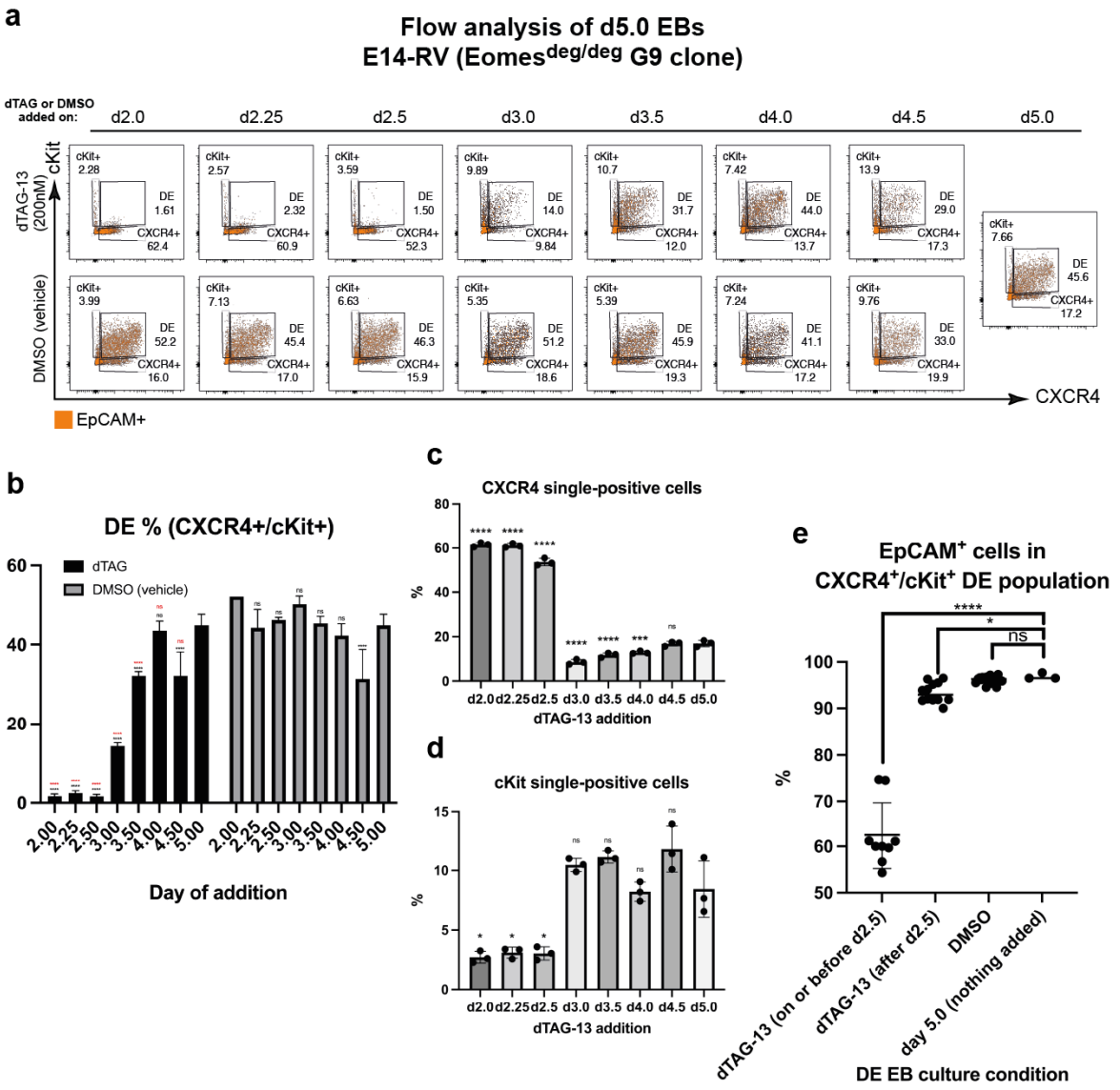


Figure 4.7: Time-dependent effect of Eomes protein degradation on DE population in EB cultures.

(a) Representative live flow cytometry plots for DE markers CXCR4, cKit, and EpCAM in day 5.0 DE EBs. EpCAM⁺ cells are coloured orange. The time of addition of dTAG-13 determines the magnitude of the impact of Eomes degradation on the CXCR4⁺/cKit⁺ population.

(b) Bar graph of CXCR4⁺/cKit⁺ cells in day 5.0 EBs treated with dTAG-13 or DMSO control at different timepoints during DE EB differentiation. Unpaired *t* tests were carried out to determine if the difference between groups was significant. Black significance notations refer to the comparison to day 5.0 within each of the dTAG and DMSO groups, whereas red significance notations refer to same-timepoint comparisons between the dTAG and DMSO groups. N=3 in all groups except for DMSO d2.0, where N=1.

(c,d) Bar graphs comparing CXCR4 (c) and cKit (d) single-positive populations on day 5.0 between DE EBs treated with dTAG-13 or DMSO control at different timepoints. N=3 in all groups.

(e) Bar graph of EpCAM⁺ cell percentage in CXCR4⁺/cKit⁺ cell population of DE EBs treated in various conditions. N=9, 12, 19, 3. In (c), (d), and (e), unpaired *t* tests were carried out to determine significant differences between groups. Error bars indicate SD. ****, P≤0.0001; ***, P≤0.001; *, P≤0.05; ns, not significant.

4.5 Eomes degradation results in dysregulation of endodermal, APS, and mesodermal markers

In order to exploit the potential offered by the Eomes-degron system, I designed an experimental plan to assess the direct effects of Eomes protein loss in differentiating cultures (Figure 4.8). Within the first few hours of Eomes loss, any changes observed are likely to be a direct consequence, rather than secondary effects from dysregulation of Eomes downstream targets.

Based on the findings above, I decided to assess the effect of Eomes protein loss on both the transcriptome (through RNA-seq) and genome-wide chromatin accessibility (through ATAC-seq) at timepoints shortly after dTAG-13 addition, namely at 3, 6, 12, and 24h. I chose day 2.5 of DE EB culture as the starting point, as my previous experiments identified this as the beginning of the Eomes "window of action," essential for specification of the DE population.

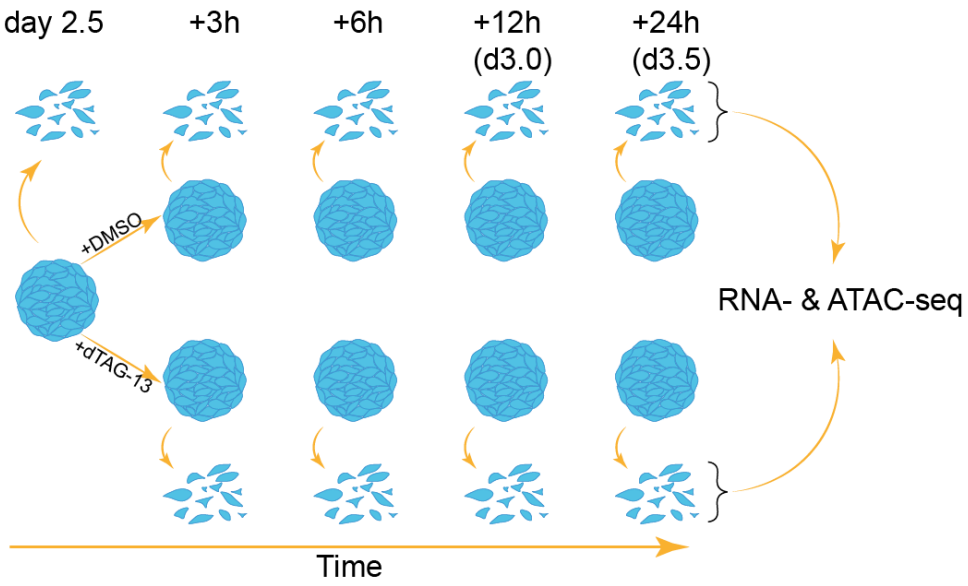
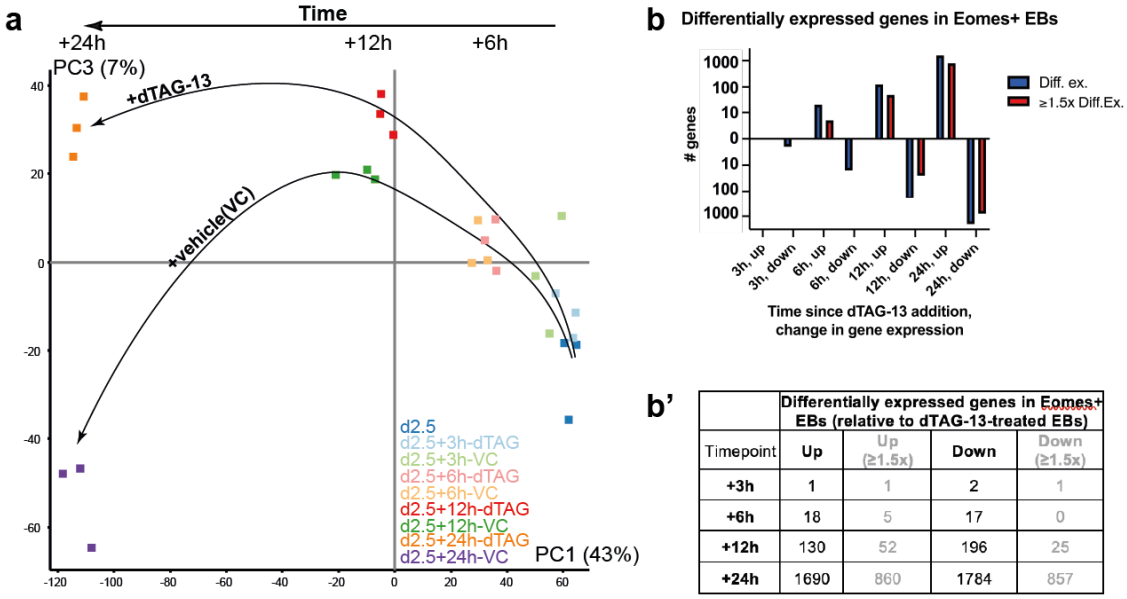


Figure 4.8: Experimental layout for RNA- and ATAC-seq experiments in DE EBs upon Eomes degradation.
Each condition was carried out in triplicate and samples were divided for RNA- and ATAC-seq.

Samples of dTAG-13- and control DMSO-treated EBs were collected and processed for RNA-seq. Similarly to previous RNA-seq experiments described in this thesis, uniquely-mapping reads were assessed only. To assess library quality, the fraction of reads in genes and exons was assessed, with a higher proportion indicating a library with little contamination and little unspliced mRNA. Additionally, contamination from rRNA and reads aligning to mitochondrial DNA was quantified (Appendix Figure 4.21a). A QQ plot was generated to assess if there were any significant differences in probe quantitation between samples (Appendix Figure 4.21b).

PCA of the RNA-seq samples (Figure 4.9a) indicated that there is a separation of the transcriptomes both temporally (PC1), with transcriptomes gradually changing over time, as well as based on presence or absence of Eomes protein (PC3). Separation of samples along PC3 became apparent at the +12h timepoint and was more pronounced by +24h.

The resulting data sets were assessed for significantly differentially expressed transcripts. I narrowed down the number of differentially expressed genes by setting a threshold of ≥1.5-fold differential expression (Figure 4.9b). Between timepoints, differentially expressed genes increase by about 10-fold (Figure 4.9b') and there is a roughly even distribution between genes upregulated and downregulated upon Eomes depletion, which is consistent with its known functions both as a gene activator and repressor.



b'

	Differentially expressed genes in Eomes+ EBs (relative to dTAG-13-treated EBs)			
Timepoint	Up	Up (≥1.5x)	Down	Down (≥1.5x)
+3h	1	1	2	1
+6h	18	5	17	0
+12h	130	52	196	25
+24h	1690	860	1784	857

Figure 4.9: Eomes depletion results in a distinct differentiation trajectory and gene expression signature in EBs.

(a) PCA of RNA-seq reads from an Eomes depletion time course. The differentiation timepoint (PC1) and the depletion (or not) of Eomes protein (PC3) are major sources of transcriptional variance in Eomes$^{deg/deg}$ EBs.

(b, b') Log_{10}-base bar graph and table of number of differentially expressed genes by timepoint. The number of differentially expressed genes increases over time from the moment of dTAG-13 addition, with an even distribution between upregulated and downregulated genes.

A closer examination of the differentially expressed genes at different timepoints reveals that at +3h, there are only three differentially expressed genes in the Eomes-depleted EBs (Figure 4.10a). Cebpb, a gene involved in placental labyrinth patterning in the ExE-derived chorion (Simmons et al., 2008), is significantly downregulated. Conversely, Nkx1-2 and Foxb1 are both upregulated in Eomes-depleted EBs, suggesting they may be normally repressed by Eomes. Nkx1-2 is expressed in epiblast cells fated to give rise to all three germ layers, but mostly neuromesodermal progenitor-derived ectoderm and mesoderm (Albors et al., 2018). Foxb1 is also mostly expressed in ectoderm and paraxial mesoderm (Zhao et al., 2007). These differences in RNA expression suggest a potential role for Eomes in repression of TFs that would normally initiate alternative ectodermal or mesodermal cell fate trajectories.

At +6h (Figure 4.10a'), several genes are significantly downregulated in Eomes-depleted EBs, including **(i)** APS markers Goosecoid (Gsc) and Mixl1 (Blum et al., 1992; Pearce & Evans, 1999), **(ii)** Wnt11, which is expressed in the node and notochord and required for appropriate axis elongation during gastrulation (Andre et al., 2015), **(iii)** Adamts9, a metalloprotein involved in the cardiac lineage (Benz et al., 2016; Enomoto et al., 2010; Kern et al., 2010) and **(iv)** Cxcl12, which encodes the CXCR4 receptor ligand (Bleul et al., 1996; Oberlin et al., 1996). In contrast, genes that are upregulated upon Eomes depletion include **(i)** Fn1 (fibronectin), an ECM component required for normal morphogenesis of the embryo (George et al., 1993; Georges-Labouesse et al., 1996), and **(ii)** Lefty1, a Nodal antagonist expressed in the node and notochord and which has a role in left-right patterning (Meno et al., 1998).

By the +12h timepoint (Figure 4.10a''), Eomes depletion prevents upregulation of an increasing number of APS derivative markers: **(i)** Foxa2, Cer1, and Gsc, DE markers, **(ii)** Mesp1, a master regulator of cardiogenesis (Kitajima et al., 2000; Saga, 1998), **(iii)** Lhx1, which has roles in the DE, node, and midline (Costello et al., 2015), and chordin (Chrd), a BMP antagonist expressed in the node and notochord (Anderson et al., 2002; Bachiller

et al., 2000; Yang & Klingensmith, 2006). Other significantly downregulated genes include the EMT regulator Snai1 (Ciruna & Rossant, 2001), and components of major embryonic signalling pathways including Bmp7 (Kim et al., 2019; Solloway & Robertson, 1999), which is involved, redundantly with Bmp5, in specification of various mesodermal lineages, including cardiac progenitors; and Cyp26a1, a retinoic acid catabolising enzyme expressed in the anterior midline cells (Ribes et al., 2007; Sakai et al., 2001).

By +24h following Eomes protein depletion (Figure 4.10a'''), additional DE markers that fail to become upregulated relative to the control DE EBs include Sox17, Hhex, Cldn6 (Anderson et al., 2008), Cxcr4, Foxa1, and node and notochord marker Foxj1 (Alten et al., 2012).

By contrast, genes that are significantly upregulated 24h after Eomes depletion include **(i)** Wnt3a, a signalling protein critical for paraxial mesoderm specification and appropriate posterior axial elongation (Cunningham et al., 2015; Dunty et al., 2008; Takada et al., 1994), **(ii)** Cdx2, which at this stage of development is also involved in appropriate A-P axis elongation and posterior paraxial mesoderm specification (Chawengsaksophak et al., 2004; Gaunt et al., 2005) **(iii)** Hoxa3, a TF involved in early organogenesis, including roles in tracheal epithelium, thymic, and parathyroid endoderm (Bertrand et al., 2011; Chojnowski et al., 2016; Gaufo et al., 2003), **(iv)** Hoxb2, also an organogenesis-stage TF involved in hindbrain patterning (Ferretti et al., 2000), and **(v)** Meis2, a homeobox gene involved in forebrain, neural crest cell, and somitic mesoderm differentiation (Cecconi et al., 1997; Fabik et al., 2020; Guibentif et al., 2021; Wen et al., 2017).

In summary, at the +24h timepoint, genes significantly differentially upregulated in Eomes^{+} EBs are associated with GO related mostly to endoderm and mesendoderm development as well as paraxial mesoderm. Notably, "embryonic foregut morphogenesis" is also one of the enriched terms (Figure 4.10b). By contrast, enriched terms in genes

significantly differentially upregulated in Eomes-depleted EBs are associated with ectodermal lineages, as well as paraxial mesoderm and interferon response (Figure 4.10b').

Collectively, the RNA-seq experiments indicate that Eomes performs a variety of functions throughout the first 24h of its activity in the differentiation protocol. Its earliest effect appears to be downregulation of TFs that promote ectodermal and certain mesodermal identities (paraxial). Soon after, it begins to promote expression of TF regulators and markers of cell types produced at the APS, including DE, cardiac mesoderm, axial mesoderm (which will give rise to the node and notochord), and anterior midline mesendoderm. Coincidently, it suppresses the expression of paraxial/presomitic/somitic mesoderm and (neuro)ectodermal regulators. Eomes also regulates the expression of signalling pathway components, such as BMPs and WNTs, and RA, as well as ECM components such as fibronectin and Adamts9. Eomes might also be involved in priming genes for later expression, or suppressing their expression until a later, appropriate developmental stage. A potential candidate for this category is the TF Hoxa3, which is only expressed in DE-derived tissues in early organogenesis.

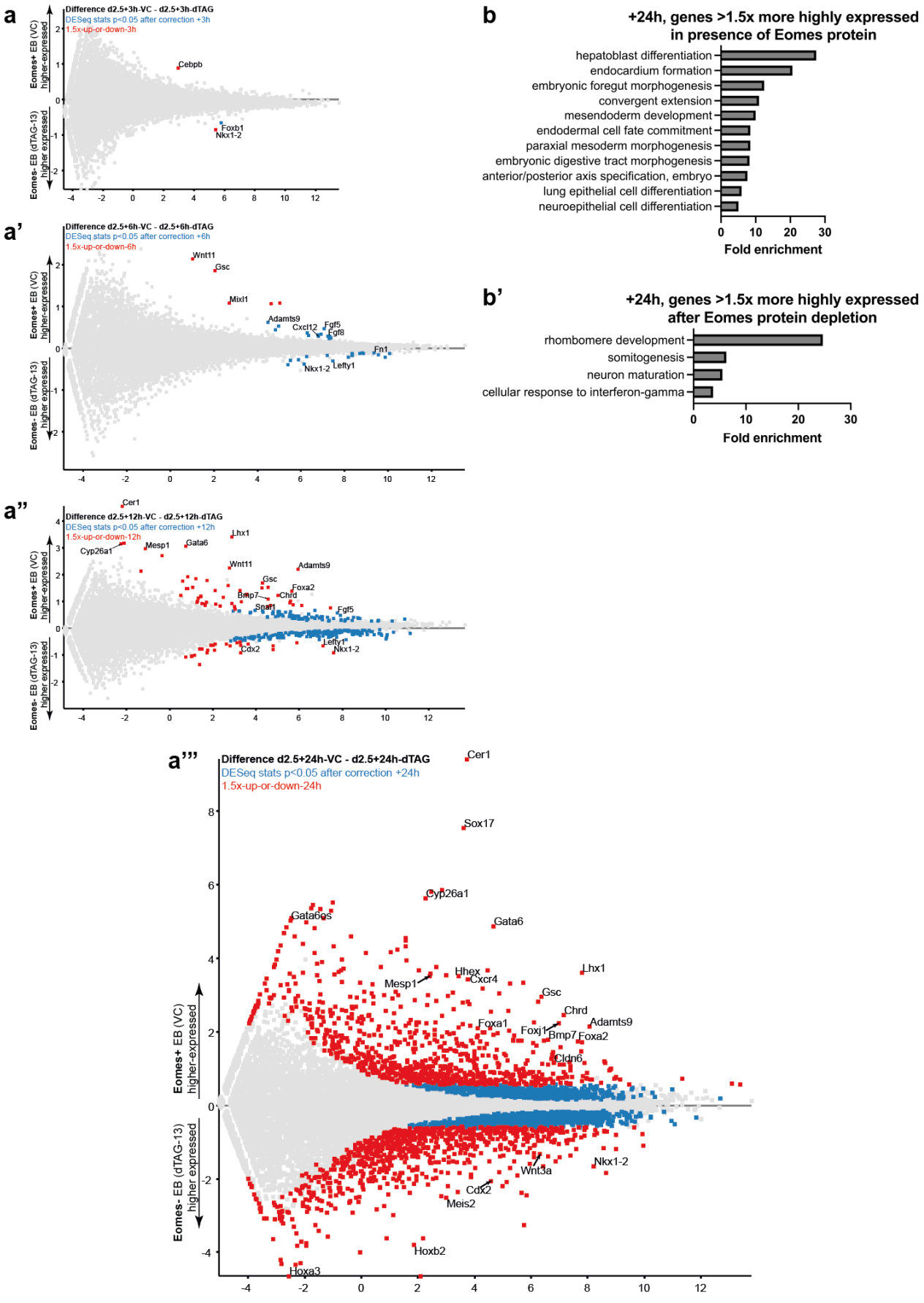


Figure 4.10 (legend on next page)

Figure 4.10: Eomes depletion results in rapid downregulation of PS, DE, and anterior mesodermal markers.
(a-a’’’) MA (log ratio/mean average) plots of differentially expressed genes at different times post-dTAG-13 addition, with selected genes highlighted.
(b, b’) Enriched gene ontology terms for differentially expressed genes upregulated in EBs where Eomes has not been degraded (b) and where it has been degraded (b’), 24h after addition of DMSO or dTAG-13 respectively.

4.6 Eomes protein loss results in changes to chromatin accessibility near APS/DE-related genes

In parallel to the RNA-seq experiments, I carried out a differential chromatin accessibility analysis between dTAG- and DMSO-treated EBs at each timepoint. I then analysed the identified regions using Stanford’s Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010; Tanigawa et al., 2022) in order to determine the nearest TSSs associated with each differentially accessible site and to assess enrichment of GO terms in the identified genes.

The first timepoint at which differentially accessible sites are detected is +12h. There are 7 differentially accessible sites out of a total of 50,816 total accessible sites identified. These are associated with 12 genes, two of which are also differentially expressed in my RNA-seq dataset. At the +24h timepoint, there are 157 differentially accessible sites (out of a total of 36,872), associated with 286 genes, of which 109 are also differentially expressed at that timepoint (Table 4.1). In Eomes+ EBs, a greater number of regions had increased accessibility, compared to the number of regions that became less accessible.

Timepoint	+12h (d3.0)		+24h (d3.5)	
Accessible sites (# peaks called)	50,816		36,872	
Differentially accessible sites	7 (0.01%)		157 (0.4%)	
More/less accessible when Eomes present	6	1	103	54
Number of associated genes	12		286	
Of those genes, also differentially expressed at same time	2		109	

Table 4.1: Summary statistics of differential accessibility analysis in Eomes degradation time course experiment carried out in DE EBs.

At the +12h timepoint, only two differentially expressed genes are associated with differentially accessible peaks. FoxA2 is significantly downregulated upon Eomes depletion, and it is also associated with a chromatin accessibility site approximately 37kb upstream of its TSS, which fails to become accessible upon Eomes depletion (Figure 4.11).

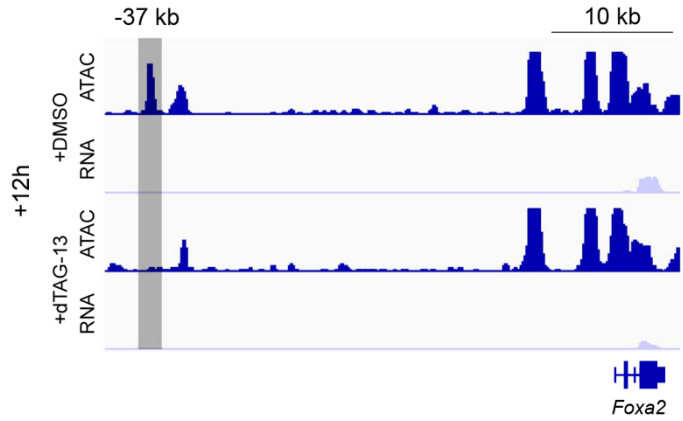


Figure 4.11: Foxa2 is associated with one of the earliest-arising differentially accessible sites between Eomes-depleted and Eomes-non-depleted DE EBs.
Genomic RNA- and ATAC-seq tracks of FoxA2 and its upstream chromatin region. The accessible chromatin peak at -37 kb from the Foxa2 TSS, highlighted in grey, is absent after 12h of dTAG-13 treatment. The reduction in Foxa2 transcripts (in lilac) is also noticeable.

At the +24h timepoint, the 103 genomic regions that are more accessible in the Eomes+ EBs are associated with 188 genes. Of these, 58 are more highly expressed in Eomes+ conditions compared to dTAG-treated samples. 16 were less highly expressed. By contrast, there are 100 genes associated with the 54 regions that become less accessible in the Eomes+ samples. Of those, 30 genes are significantly underexpressed in the Eomes+ samples, while 5 are significantly overexpressed. Overall though an increase in accessibility appears to correlate with an increase in transcript levels, and vice versa (Table 4.2).

		Overexpressed in Eomes+ EBs	**Overexpressed in Eomes- EBs**
Genes associated with 103 Eomes+ accessible sites	188	**58**	16
Genes associated with 54 Eomes+ inaccessible sites	100	5	**30**

Table 4.2: Positive correlation between chromatin accessibility and increase in expression level at the +24h dTAG-13 addition timepoint.
Note that the total number of associated genes is 286 (Table 4.1), and not 288, which is the sum of the genes from the two rows, as there is an overlap of 2 genes between the two groups (Aurkc and Usp29, neither of whose expression is significantly different).

The genes associated with Eomes+-accessible and -inaccessible peaks at +12h and +24h can be found in Chapter 4 Appendix Tables 4.3 and 4.4 respectively.

The majority of differentially accessible sites are 50-500 kb from the nearest TSS (54%), followed by 5-50 kb from the nearest TSS (35%), suggesting they largely represent distal enhancer regions (Figure 4.12a). Furthermore, GO analysis indicates that the peak-associated genes are enriched for terms related to abnormal development of DE derivatives and the notochord, as well as abnormal rostral-caudal (A-P) axis patterning (Figure 4.12b). Examination of the individual genes that contribute toward these GO terms reveals that most are upregulated in the presence of Eomes protein. Additionally, a portion of those upregulated genes are also found near sites that exhibit increased accessibility in the presence of Eomes protein (Figures 4.12c, 4.13a).

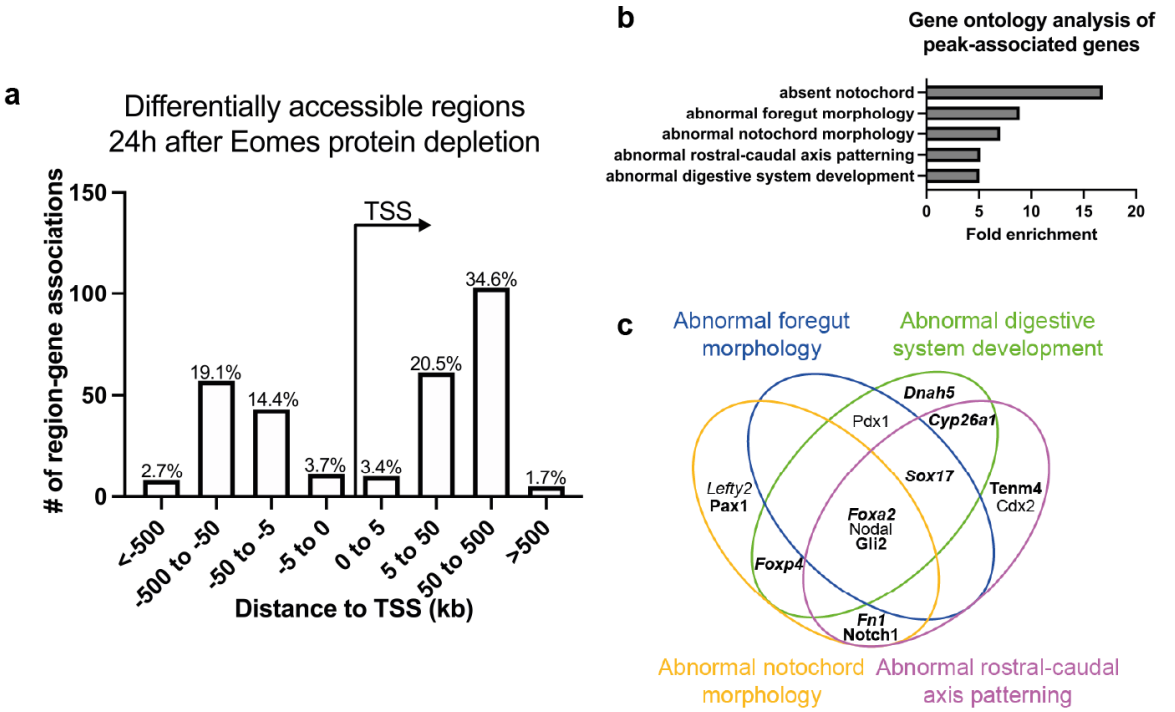


Figure 4.12: Differentially accessible genomic sites at +24h from dTAG-13 addition are found near genes related to APS-derived lineages, as well as rostral-caudal patterning.

(a) Bar graph of the distribution of accessible sites relative to their nearest TSS. The majority of differentially accessible sites are found distally from annotated genes' TSS, at a distance between 50 and 500 kb.

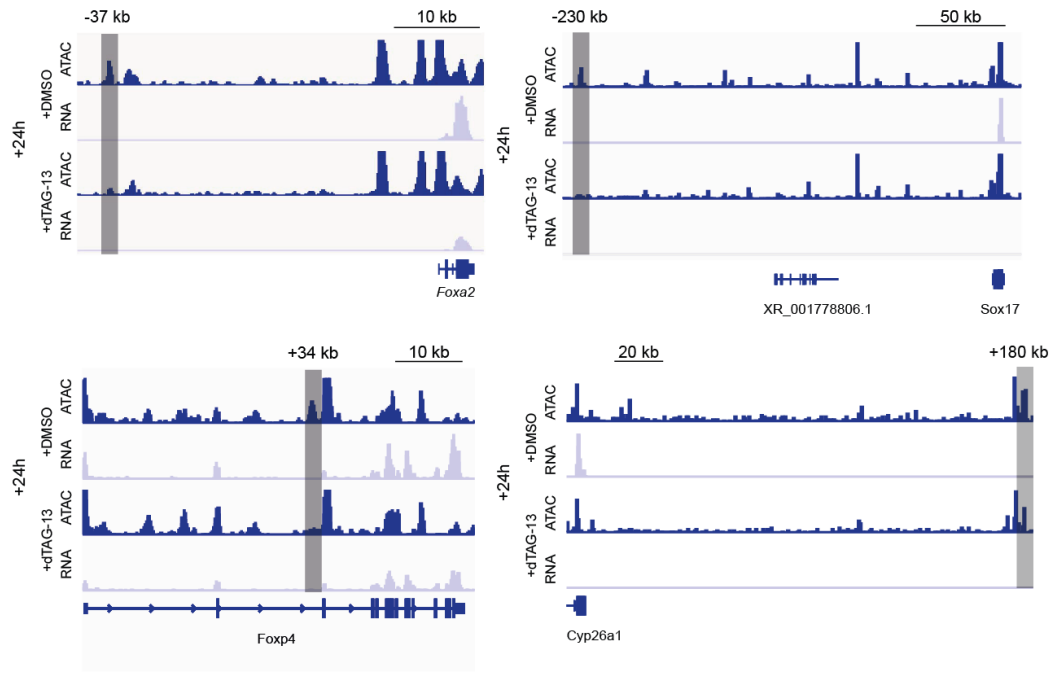
(b) Gene ontology analysis of genes associated with differentially accessible sites indicates that they are associated with terms related to APS derivative cell types (notochord, digestive system) and rostral-caudal patterning.

(c) Venn diagram of genes associated with selected gene ontology terms from (b). Terms in bold indicate genes significantly overexpressed in the presence of Eomes protein (see Figure 4.10a'''). Italicised terms indicate genes found near genomic sites that are more accessible when Eomes protein in present.

Analysis of enriched genomic sequence motifs at the sites of differential accessibility showed that loci with greater accessibility in the Eomes+ samples are enriched for the Eomes binding motif, as well as the binding motif for FoxA2, and several other Fox TFs such as FoxA1 (another DE marker). GATA factor binding motifs were also enriched, including Gata4 and Gata6, both of which are expressed in the DE, as well as the Lhx1 and Smad2/3 motifs (the latter being intracellular effectors of Nodal signalling). Finally, Nanog motifs were also enriched, suggesting Eomes may be involved in regulating pluripotency-associated genes (Figure 4.13b). In contrast, sites that become accessible upon Eomes protein depletion are enriched for the binding motifs of the Wnt signalling nuclear effectors TCF and LEF, suggesting an Eomes involvement in regulating the Wnt signalling response (Figure 4.13b').

Overall, the ATAC-seq data suggests that Eomes regulates accessibility of chromatin regions associated with genes important in DE and node development. Eomes may co-bind regions along with other TFs critical in DE specification and downstream regionalisation and development, such as FoxA2, FoxA1, and GATA factors; as well as Lhx1, which is critical for AME and node specification (Costello et al., 2011). Additionally, Eomes may be involved in the repression of pluripotency genes via association with Nanog-regulated genes, which has previously been demonstrated in human DE differentiation (Teo et al., 2011). Finally, Eomes may be involved in modulating the response to extracellular signalling through its involvement in regulation of the target loci of intracellular effectors such as Smad2/3 and TCF/LEF.

a



b

Motif enrichment of Eomes+ accessible sites			
Consensus motif	**Transcription factor**	**P-value**	**% sequences with motif**
	FoxA1	1e-20	42.72
	FoxA2	1e-14	32.04
	Gata6	1e-5	22.33
	Tbox:Smad, Smad2/3	1e-4	9.71
	Gata4	1e-4	22.33
	Eomes	1e-3	36.89
	Nanog	1e-2	56.31
	Lhx1	1e-2	18.45

b'

Motif enrichment of Eomes- accessible sites			
Consensus motif	**Transcription factor**	**P-value**	**% sequences with motif**
	Tcfl2	1e-6	12.96
	Tcf3	1e-5	16.67
	Lef1	1e-3	22.22
	Tcf7	1e-2	11.11
	Tcf12	1e-2	22.22

Figure 4.13: Differentially accessible genomic sites at +24h are enriched for binding motifs of DE-specific TFs and intracellular signalling effectors.
(a) Genomic tracks of selected loci exhibiting both increased expression levels and accessibility in the presence of Eomes protein. The grey shading in each panel indicates the site of differential accessibility.
(b, b') HOMER (Hypergeometric Optimization of Motif EnRichment) analysis reveals enriched sequence motifs at differentially accessible genomic regions. HOMER analysis of regions more accessible in Eomes presence is in b, whereas HOMER analysis of regions less accessible in Eomes presence are in b'.

4.7 The timing of Eomes depletion affects anterior-posterior gene expression within the CXCR4+/cKit+ cell population

Because GO term analysis indicated that Eomes promotes the upregulation of genes associated with embryonic foregut morphogenesis (see Figure 4.10b), I investigated whether Eomes may play a role in influencing the A-P identity of DE cells. In order to do this, I varied the duration of Eomes protein presence in the EB cultures by removing it at different timepoints of the 5-day differentiation protocol, and isolating the CXCR4+/cKit+ population by FACS (see Figure 4.7a). I then used RT-qPCR to assess the expression of Sox2 and Cdx2, as they are markers of the anterior and posterior gut tube respectively at early organogenesis stages (E9.5) (Sherwood et al., 2009).

It should be noted that due to the varying duration of Eomes protein expression in the EB cultures, the percentage of the CXCR4+/cKit+ cells that are generated differs, as demonstrated in Figure 4.7a.

If duration of Eomes activity influences A-P identity, then I would expect to see a reverse correlation between levels of Sox2 and Cdx2 expression in the CXCR4+/cKit+ cells from each distinct Eomes depletion timepoint. Since terms associated with foregut morphogenesis were upregulated in Eomes+ EBs, I expected to see greater Sox2 levels in cells where Eomes had been depleted at later stages, and conversely higher Cdx2 levels in those cells where Eomes protein was depleted at earlier timepoints. However, this did not appear to be the case, as both Sox2 and Cdx2 levels were highest in the double-positive cell population in which Eomes was depleted at the earliest timepoint (day 3.0) (Figure 4.14). This may indicate that loss of Eomes protein may have a more generalised effect on transcription, rather than specifically influencing the A-P identity of DE cells. This is also in agreement with previous findings suggesting that DE cells are not initially specified with a particular A-P identity, as DE cells grafted to other parts of the gut tube begin to express markers associated with the host location's identity (Wells & Melton,

2000). However, more markers need to be assayed, or RNA-seq carried out, in order to determine with greater certainty whether there are any transcriptional signs of an A-P identity due to the variable duration of Eomes protein presence in the DE EB culture.

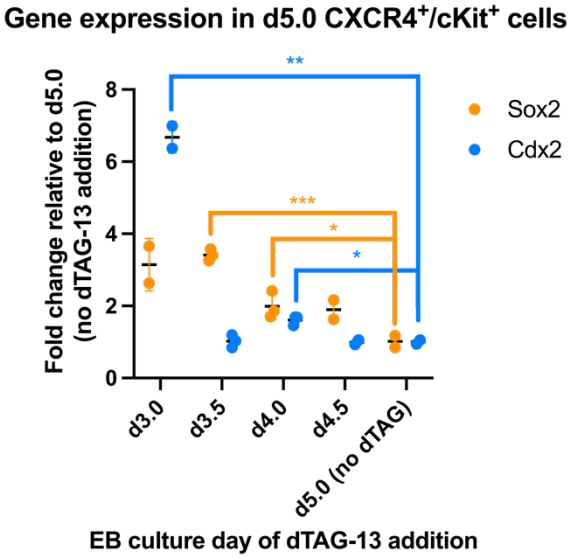


Figure 4.14: Depletion of Eomes protein at different timepoints of the DE EB differentiation protocol affects expression of A-P DE markers Sox2 and Cdx2.
Bar graph of fold-change of Sox2 and Cdx2 expression in CXCR4⁺/cKit⁺ cells from day 5.0 DE EB upon Eomes depletion at different timepoints, relative to double-positive cells from Eomes-non-depleted day 5.0 EBs. Eomes was depleted at different timepoints by dTAG-13 addition, and CXCR4⁺/cKit⁺ cells were isolated by FACS on day 5.0. However, this effect does not result in an inverse correlation between Sox2 and Cdx2 levels, as I would expect if the cells adopted a specific A-P identity based on the duration of their Eomes protein expression prior to degradation. Each experimental condition was done in triplicate, except when Eomes was depleted on d3.0, d4.5, or d5.0, which were carried out in duplicate. Unpaired *t* tests were carried out to determine whether differences in gene expression were significant (***, P≤0.001; **, P≤0.01, *, P≤0.05). Error bars indicate standard deviation (SD).

4.8 Deletion of Sox17- and Foxa2-associated differentially accessible regions does not significantly impact CXCR4+/cKit+ cell specification

Depletion of Eomes in DE EB cultures results in rapid and significant reduction of DE marker expression, including DE master regulators Foxa2 and Sox17. Specifically, FoxA2 mRNA expression levels are significantly reduced 12h after dTAG-13 addition, whereas Sox17 transcripts are significantly reduced after 24h (see Figure 4.10a'', a'''). Additionally, each of these two genes is associated with a chromosome region which exhibits reduced accessibility at the same timepoints (Figures 4.11, 4.13a).

Previous published findings have associated these regions with cis-regulation of FoxA2 and Sox17 expression. The FoxA2-associated differentially accessible region I identified in my ATAC-seq experiment corresponds to an enhancer region upstream of the FoxA2 TSS (the –37 kb enhancer) identified in Wang et al. (2017). The authors reported that the –37 kb enhancer (similarly to two other putative enhancers they identified at –50 and –53 kb) is bound by ß-catenin in an ActA-dependent manner (Figure 4.15a,c). More specifically, EBs formed from mESCs followed by ActA addition for two hours on day 3 exhibited Tcf3, ß-catenin, and Smad2/3 binding (Figure 4.15b). In contrast, ß-catenin binding to the –37 kb enhancer was significantly reduced when a Nodal/activin receptor kinase inhibitor was added to the EB medium, ß-catenin binding to the locus was reduced (Figure 4.15a). If Eomes regulates accessibility to this site, it may be participating in a WNT/ActA/BMP signalling axis regulating FoxA2 expression.

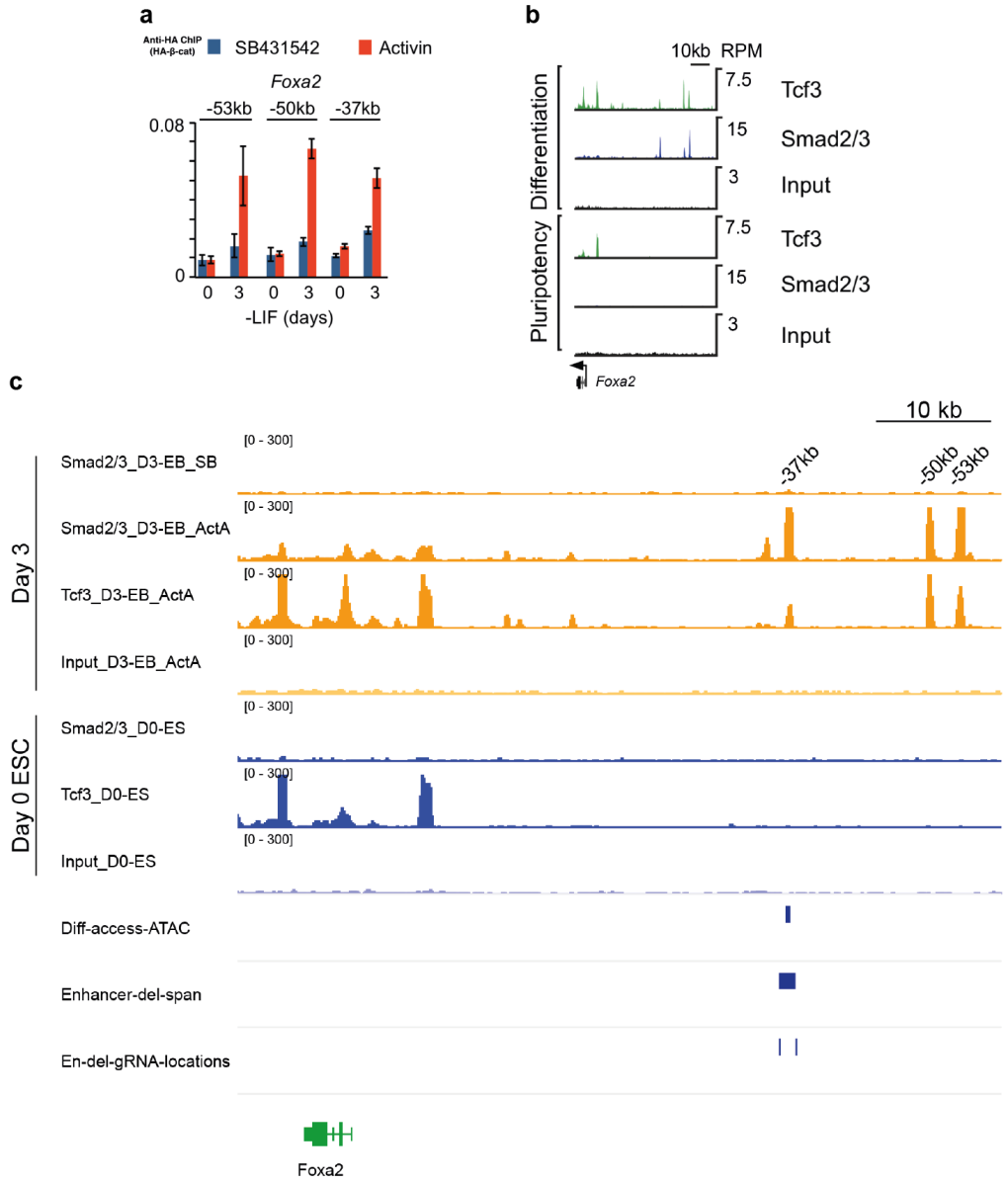


Figure 4.15: FoxA2-associated region identified as differentially accessible in ATAC-seq experiments is Tcf3/β-catenin/Smad2/3-bound in DE differentiation.
(a) ChIP-qPCR demonstrates that enhancers upstream of FoxA2 gene are bound by β-catenin in an ActA-dependent manner during mesendodermal differentiation, showing reduced binding upon administration of small molecule SB-431542 (inhibitor of ALK receptors, through which TGF-β superfamily members, including ActA, signal), and an increase upon ActA treatment [figure reproduced from Wang et al. (2017)].
(b) Genomic data tracks showing Tcf3 (intracellular WNT signalling response element) and Smad2/3 (intracellular effector of ActA) both binding the FoxA2 -37, -50, and -53 elements upon ActA EB treatment [figure reproduced from Wang et al. (2017)].
(c) Smad2/3 and Tcf3 binding of FoxA2 -37, -50, and -53 elements only occurs in ActA-treated day 3 EBs, and not in SB-431542-treated EBs or ESCs [data from Wang et al. (2017)]. Also indicated are the differentially-accessible region identified in my ATAC-seq experiments, the location of crRNAs ("gRNAs") designed for the enhancer deletion, and the resulting deleted region, spanning the length of the differentially-accessible region.

Meanwhile, the Sox17 differentially accessible region is homologous to a region in the human genome which becomes differentially methylated upon DE induction, acquiring both H3K27me3 and H3K4me3 marks, remodelling the chromatin to a poised state (Tsankov et al., 2015) (Figure 4.16a). In addition, the region becomes bound by EOMES, SOX17, and FOXA2 upon DE induction (Figure 4.16b). The two homologous regions between mouse and human both have a distance of approximately 240 kb from the Sox17 gene. Eomes may therefore be involved in increasing accessibility to the region and promoting its differential methylation, causing it to adopt a poised state.

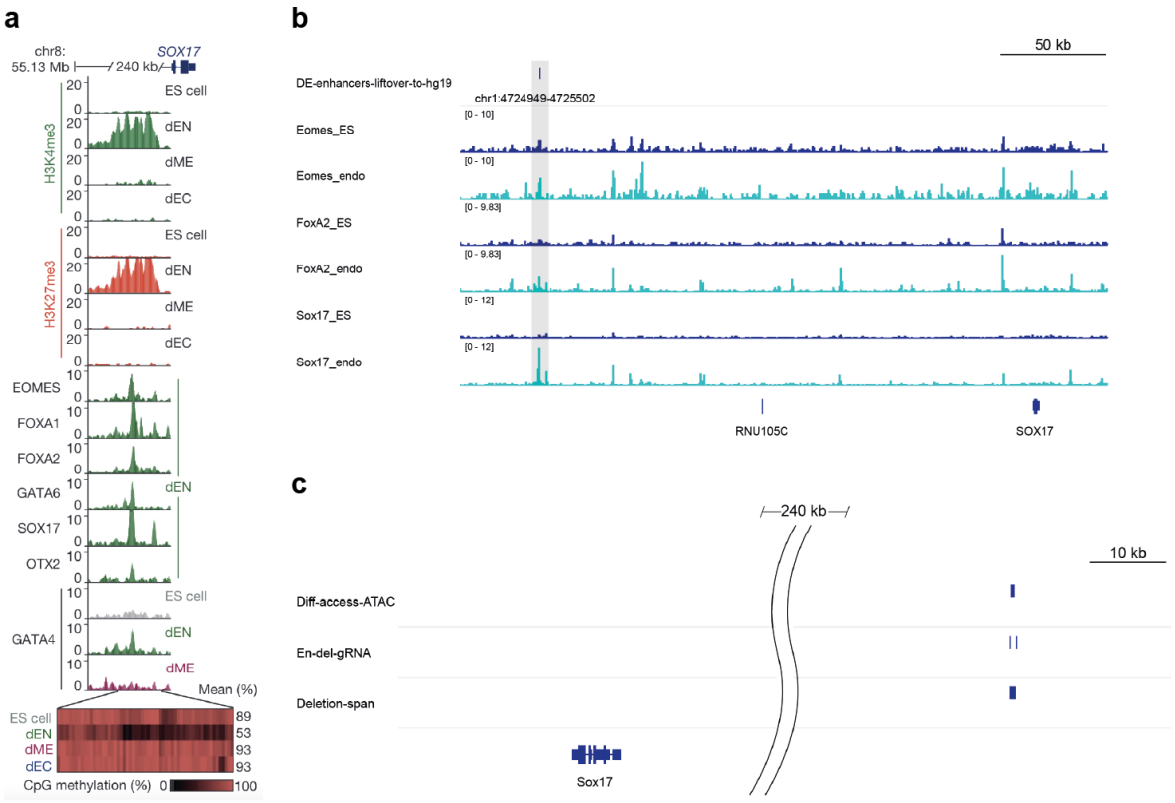


Figure 4.16: Sox17-associated differentially accessible region is homologous to a human region that is differentially methylated specifically in DE.
(a) Genomic tracks for a region in the human genome homologous to the SOX17-associated differentially-accessible region identified in my ATAC-seq experiments. This region exhibits differential methylation exclusively during endodermal differentiation, binding of multiple DE-associated TFs, as well as CpG island demethylation exclusively in DE [figure reproduced from Tsankov et al. (2015)].
(b) Genomic tracks showing FOXA2, SOX17, and EOMES binding in DE from Tsankov et al. (2015), including the region identified by genomic liftover as homologous between human and the differentially-accessible Sox17-associated region identified in my ATAC-seq experiments.
(c) Genomic tracks indicating the differentially-accessible region near Sox17, gRNA locations, and the resulting deleted region after Cas9-mediated targeting.

In order to assess the significance of these two genomic regions in specification of the CXCR4⁺/cKit⁺ population of the DE EBs, I targeted them for deletion via a Cas9-mediated strategy using two crRNAs for each as outlined in Figure 4.15c, Figure 4.16c. I deleted them individually or both (sequentially) in E14-RV mESCs, isolated homozygously deleted clones, and carried out a DE EB differentiation. However, by the time of the submission of this thesis I was only able to carry out the differentiation in single enhancer deletion (FoxA2 or Sox17) clones.

There was no significant difference between the percentage of CXCR4+/cKit+ cells specified in the parental cell line and the Sox17 enhancer deletion clones. In contrast, there was a significant increase in the percentage of double-positive cells in the FoxA2 enhancer deletion cells (Figure 4.17a,b). However, the cells I used as a control were the parental E14-RV line, and not mock-transfected, low-density plated, and subcloned E14-RV cells. Therefore, the significant difference in CXCR4+/cKit+ cell specification may be due to genetic drift from subcloning the enhancer-deleted cells in order to isolate homozygous deletion clones, rather than due to a genuine effect from the putative enhancer deletion.

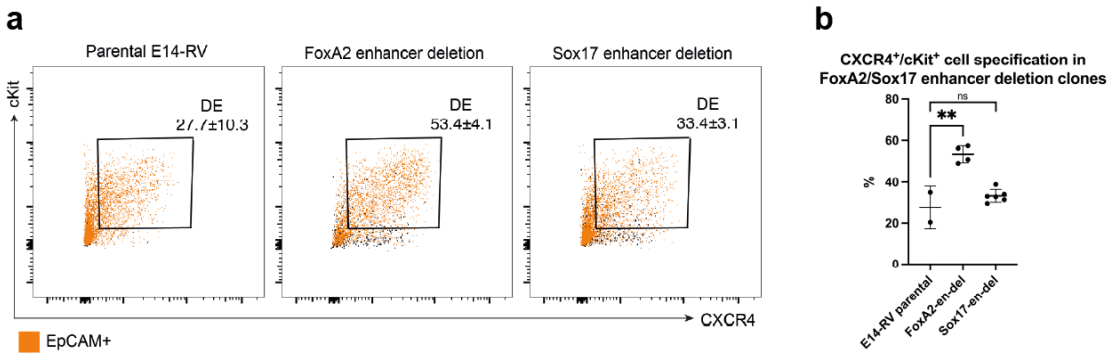


Figure 4.17: Deletion of FoxA2- and Sox17-associated differentially accessible regions does not result in reduction of the percentage of CXCR4⁺/cKit⁺ cells in DE EB cultures.
(a) Flow cytometry plots for E14-RV parental, FoxA2 and Sox17 enhancer deletion clones. EpCAM+ cells are shown in orange. (n=2 replicates for the E14-RV parental clone; n=4 Foxa2 enhancer deletion clones; n=6 Sox17 enhancer deletion clones)
(b) Bar graph showing percentages of individual clones. Student's t-tests were carried out to determine if there is a significant difference between samples. The error bars correspond to SD. **, P≤0.01; ns: not significant.

4.9 Single-cell RNA- and ATAC-sequencing reveal specific loss of DE and notochord cells in Eomes-depleted DE EB cultures

In order to investigate in greater depth the effects of Eomes protein depletion on cells in my DE EB culture system, I have initiated a collaboration with Bart Theeuwes and Mai-Linh Ton, PhD students in Bertie Göttgens' group at the Stem Cell Institute of the University of Cambridge. Using the 10x Genomics platform, we carried out single-cell RNA- and ATAC-seq (thus acquiring single-cell "multi-ome" data) on day 2.5, 3.0, and 4.5 EBs, as shown in Figure 4.18a. The resulting data were mapped onto the mouse gastrulation atlas produced by Pijuan-Sala et al. (2019).

On day 2.5, most cells in the EBs correspond to the epiblast and PS. In control (DMSO)-treated EBs, cells then follow two major trajectories: one appearing to pass through an APS identity and acquiring both notochord and DE/gut fates; and one appearing to pass through a nascent mesoderm identity and acquiring a paraxial/somitic mesoderm fate. In contrast, cells recovered from dTAG-13-treated ("knock-down", KD) EBs almost entirely fail to specify APS, DE/gut, and notochord (Figure 4.18b). This is also apparent in the proportional breakdown of EB cell types in Figure 4.18c. Here, it becomes apparent that on day 3.0, the APS population has failed to become specified, while by day 4.5, there is a complete absence of DE/gut (light and dark pink) and notochord (dark blue) cells, while there is a proportional increase in the number of paraxial (ice blue) and somitic (navy blue) mesoderm. These findings correspond with what is known about the essential role of Eomes in DE specification (and its derivative, the gut tube) (Arnold et al., 2008). It should be noted that the mouse gastrulation atlas represented in the panels of Figure 4.18b in grey corresponds only to the RNA modality, and therefore the data summarised there and also in Figure 4.18c do not incorporate the scATAC-seq data. The ATAC-seq data is represented in Figure 4.18d, which highlights the chromatin accessibility differences found between the cells of Eomes-depleted and non-depleted EBs.

As only a preliminary analysis has been carried out on the DE multiome data at this point, it will be particularly interesting to further analyse these datasets in order to determine how both DE and notochord lineages appear to be promoted by presence of Eomes protein. This is especially interesting considering that the ATAC profiles of these cells appear to be more similar (as shown in the second panel of the top row of Figure 4.18d, where blue notochord cells are interspersed among pink DE cells) compared to their transcriptomes, which appear more distinct (as seen in the second panel of the second row of Figure 4.18d). Additionally, it will be interesting to investigate what alternative GRNs are activated upon Eomes depletion to promote the paraxial/somitic mesodermal fates overrepresented in Eomes-depleted EBs. In summary, these data, perhaps in combination with more time-resolved data from intermediate differentiation timepoints, will help to shed light on the numerous diverse functions carried out by Eomes during gastrulation of the mouse embryo.

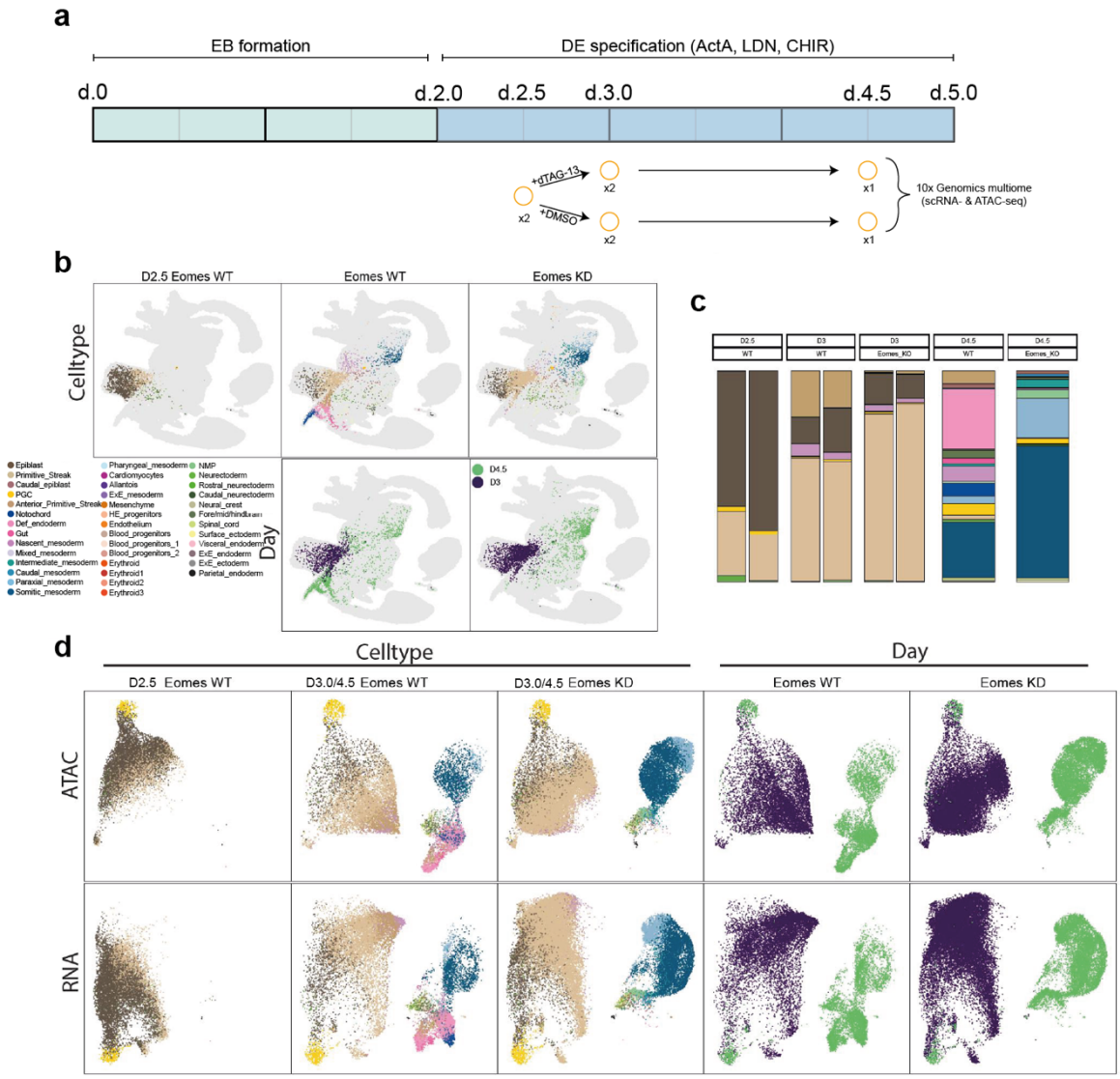


Figure 4.18: Single-cell multiome (RNA- & ATAC-seq) on cells from DE EBs shows that Eomes depletion results in loss of anterior primitive streak, DE, and notochord cells.
(a) Cartoon summary of DE EB differentiation protocol, including timepoint and number of samples included in the multiome experiment.
(b) Mapping of scRNA-seq modality to the gastrulation atlas, coloured by mapped cell type (top) and collection day (bottom).
(c) Mapped cell type proportions for each replicate.
(d) UMAPs for the ATAC (top) and RNA (bottom) modalities colours by cell type and collection day.

4.10 Summary of findings

4.10.1: Optimisation of an EB differentiation protocol promoting APS/DE identity

I optimised an EB culture protocol by varying the cytokines and small molecules of the differentiation medium, aiming to best mimic the environment encountered by epiblast cells ingressing through the APS. This environment is high in Nodal signals and has active Wnt signalling, but low BMP signalling due to its distance from the proximal epiblast. Accordingly, I varied the concentrations of ActA, CHIR, and LDN in the medium, and assessed the generation of $CXCR4^+/cKit^+(/EpCAM^+)$ cells. These are established markers used to assess DE specification. Although genetic background contributed to variability in differentiation efficiency, I identified a set of conditions that gave rise to almost 50% double-positive cells.

RNA-seq of the double-positive population further indicated that their transcriptomes are vastly different from those of undifferentiated ESCs, with pluripotency markers significantly downregulated in the double-positive cells. The double-positive cells exhibit much higher expression of known APS and DE markers, such as FoxA2, Sox17, Mixl1, Hhex, Cer1, and Lhx1. This indicates that the double-positive population represents a population of APS cells and APS derivatives, such as the DE.

4.10.2: Eomes is essential for specification of the $CXCR4^+/cKit^+$ population

I took advantage of the $Eomes^{deg/deg}$ ESC lines that I previously generated in order to investigate the requirement for Eomes in my EB differentiation protocol. dTAG-13 addition to the EB differentiation medium at various timepoints indicated that there is a time-sensitive requirement for Eomes, with addition on or before day 2.5 of differentiation completely ablating the double-positive population, and with addition on day 4.0 or later resulting in no effect on the population. Addition at intermediate timepoints resulted in part of the population (proportionally) being specified.

I investigated markers of A-P DE identity Sox2 and Cdx2 in double-positive populations where Eomes was depleted at these intermediate timepoints, to assess if Eomes presence in the EB culture for variable amounts of time impacted the A-P identity of the specified double-positive cells. I was not able to observe an inverse relationship between Sox2 and Cdx2, however, which suggested, as is currently believed, that there was no distinct A-P identity in the different double-positive populations, and that A-P identity is acquired at a later stage through inductive signals from neighbouring tissues.

4.10.3: Eomes depletion results in dysregulation of lineage-specific markers

I carried out an Eomes depletion time course experiment in my EB culture system, followed by RNA- and ATAC-seq. This enabled me to assess transcriptional and chromatin accessibility changes at very early stages after Eomes depletion, allowing me to identify direct effects of Eomes loss.

Many markers of APS, node/notochord, and DE identity are downregulated, including Gsc, Hhex, Cer1, Lhx1, Foxj1, Mixl1, cardiac marker Mesp1, and DE master regulators FoxA2 and Sox17. Interestingly, some of the earliest-affected genes are also markers of mesodermal and ectodermal lineages, which are, however, upregulated upon Eomes depletion. These results agree with what is currently known about Eomes and its functions in the APS lineages. Eomes may be suppressing expression of markers from other lineages (ectoderm, paraxial mesoderm), an effect which has been reported in the past (Tosic et al., 2019), thereby further safeguarding the specification of APS/DE progenitors.

4.10.4: Eomes depletion results in chromatin accessibility changes

At 12h after Eomes depletion, the first changes in chromatin accessibility are detected. Differentially accessible sites are found near genes enriched for GO terms associated with notochord and gut tube defects, as well as A-P patterning. Most differentially accessible sites display increased accessibility in Eomes+ EBs, with fewer becoming accessible upon Eomes depletion. This suggests a role of Eomes in increasing accessibility to enhancer

regions regulating gene expression. Interestingly, the differentially accessible sites include regions near FoxA2 and Sox17 which have previously been reported to be bound by Nodal and Wnt intracellular effectors to regulate expression of both genes. This suggests Eomes may be synergising with these signalling pathways to promote expression of these genes.

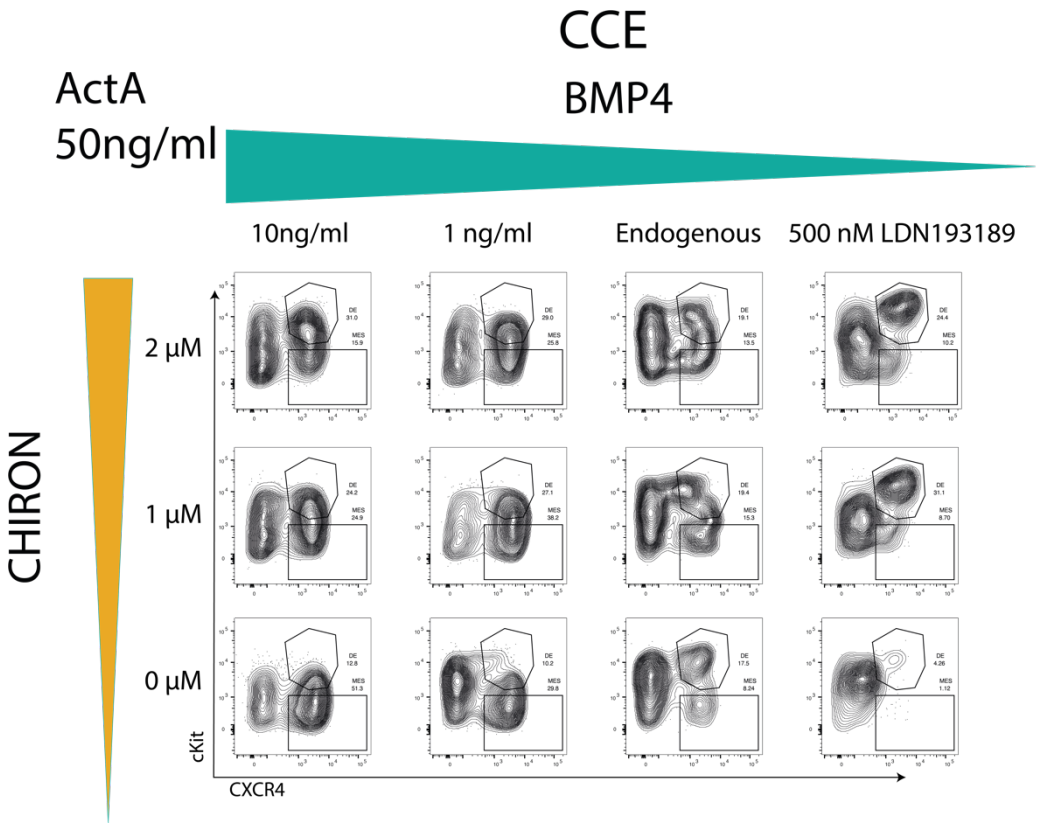
However, Cas9-mediated deletion of those differentially accessible sites did not result in defective EB differentiation (at least in terms of $CXCR4^+/cKit^+$ specification). There may be other regions, not identified in my differential accessibility assay, that are acted upon by Eomes to regulate FoxA2 and Sox17 expression.

4.10.5: Single-cell multi-ome experiment in EBs may aid in parsing Eomes effects in APS lineages

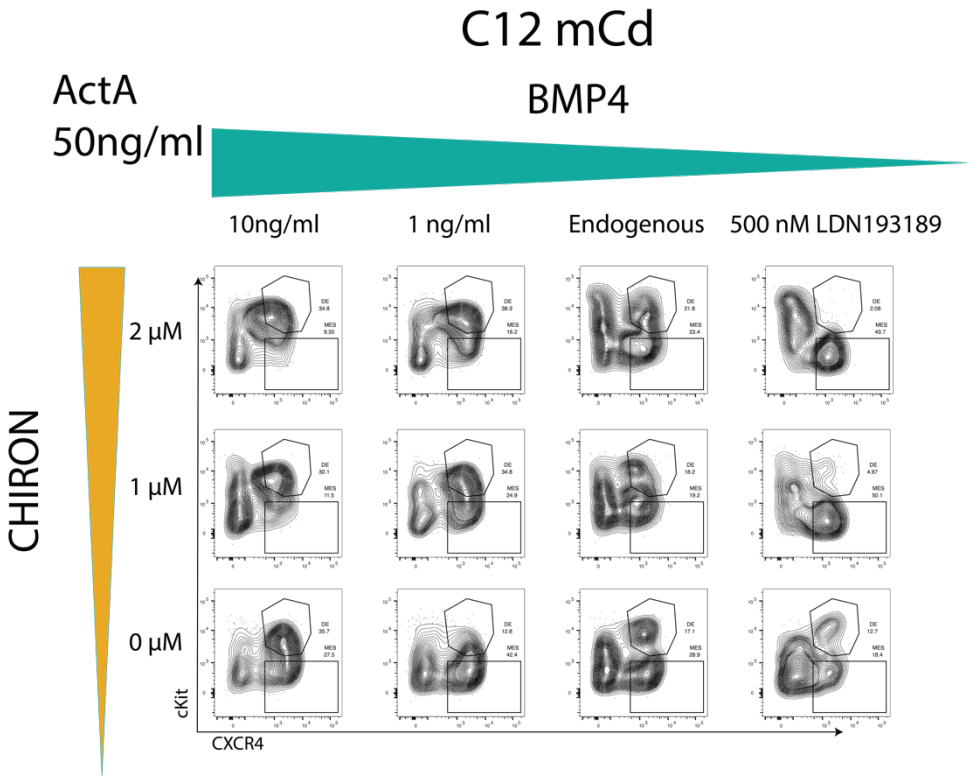
In collaboration with the Göttgens group at the University of Cambridge, we are carrying out the analysis of a single-cell multi-ome experiment. Eomes+ and Eomes-depleted EBs were subjected to single-cell RNA- and ATAC-seq, and the RNA-seq results were mapped onto the previously published mouse gastrulation atlas (Pijuan-Sala et al., 2019). This has allowed us to identify the precise cell type composition of the EBs in Eomes+ and Eomes-depleted cells, and confirmed that the APS and the lineages originating from it are specifically lost in Eomes-depleted EBs, while paraxial/somitic mesoderm is significantly enriched. Further analysis of the data generated, and integration with the single-cell ATAC-seq data, will allow us to correlate chromatin states with transcriptional states within individual cells/cell types, and to further parse the activities of Eomes within individual Eomes-dependent cell types.

4.11 Appendix

a



b



c

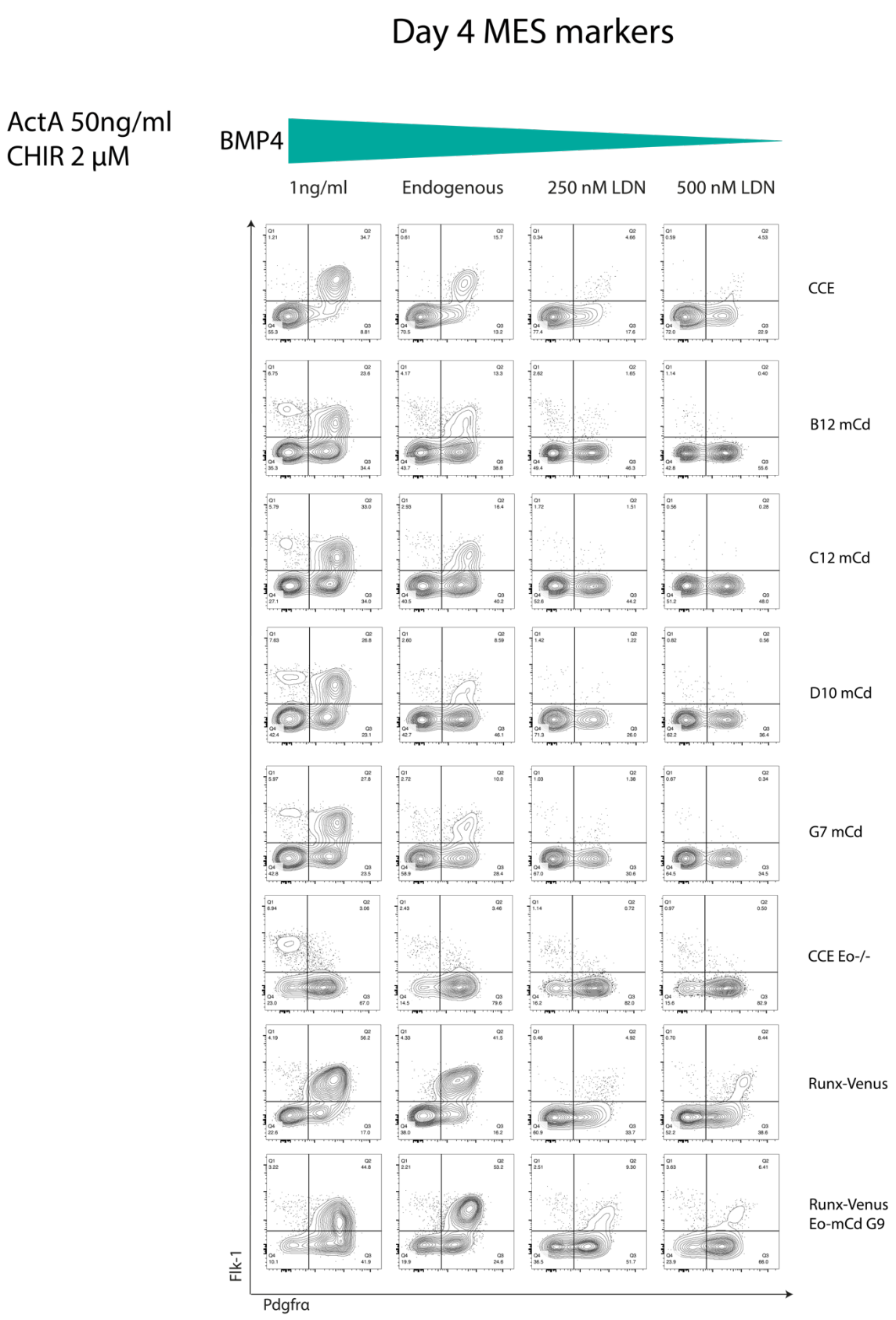


Figure 4.19 (legend on next page)

Figure 4.19: Previous flow cytometry experiments carried out to optimise conditions for APS/DE EB differentiation.
(a-b) (a) WT CCE mESCs and (b) Eomes$^{deg/deg}$ clone C12 (on CCE background) day 5.0 EBs assessed for CXCR4 and cKit expression. Cells were cultured on days 2.0-5.0 in high ActA, with variable CHIR and BMP/LDN concentrations. It is noteworthy that in the parental cells, LDN treatment (BMP inhibition) results in the most distinct double-positive population, whereas this effect is not seen in the Eomes$^{deg/deg}$ C12 clone. This experiment, however, did not sufficiently assess finer gradations in LDN concentration (going from 0 nM to 500 nM), which was addressed in further experiments.
(c) Day 5.0 EBs (WT and Eomes$^{deg/deg}$, as well as Eomes-null) assessed for CXCR4 and cKit expression. Cells were cultured in high ActA and a fixed CHIR concentration, with variable BMP/LDN concentration. As in (a-b), it is noteworthy that the CCE parental line, as well as the E14-RV parental and Eomes$^{deg/deg}$ lines (G9) develop the highest percentage of double-positive cells in LDN-containing medium, whereas CCE-derived Eomes$^{deg/deg}$ lines (B12, C12, D10, G7) do so in BMP- and LDN-free medium (with endogenous levels of BMP signalling). Eomes-null cells also completely fail to specify a double-positive population, indicating an essential requirement for Eomes in the specification of this population. This experiment also did not sufficiently assess the outcome of finer gradations in LDN concentration, which was explored more fully in Figure 4.1 of this thesis.

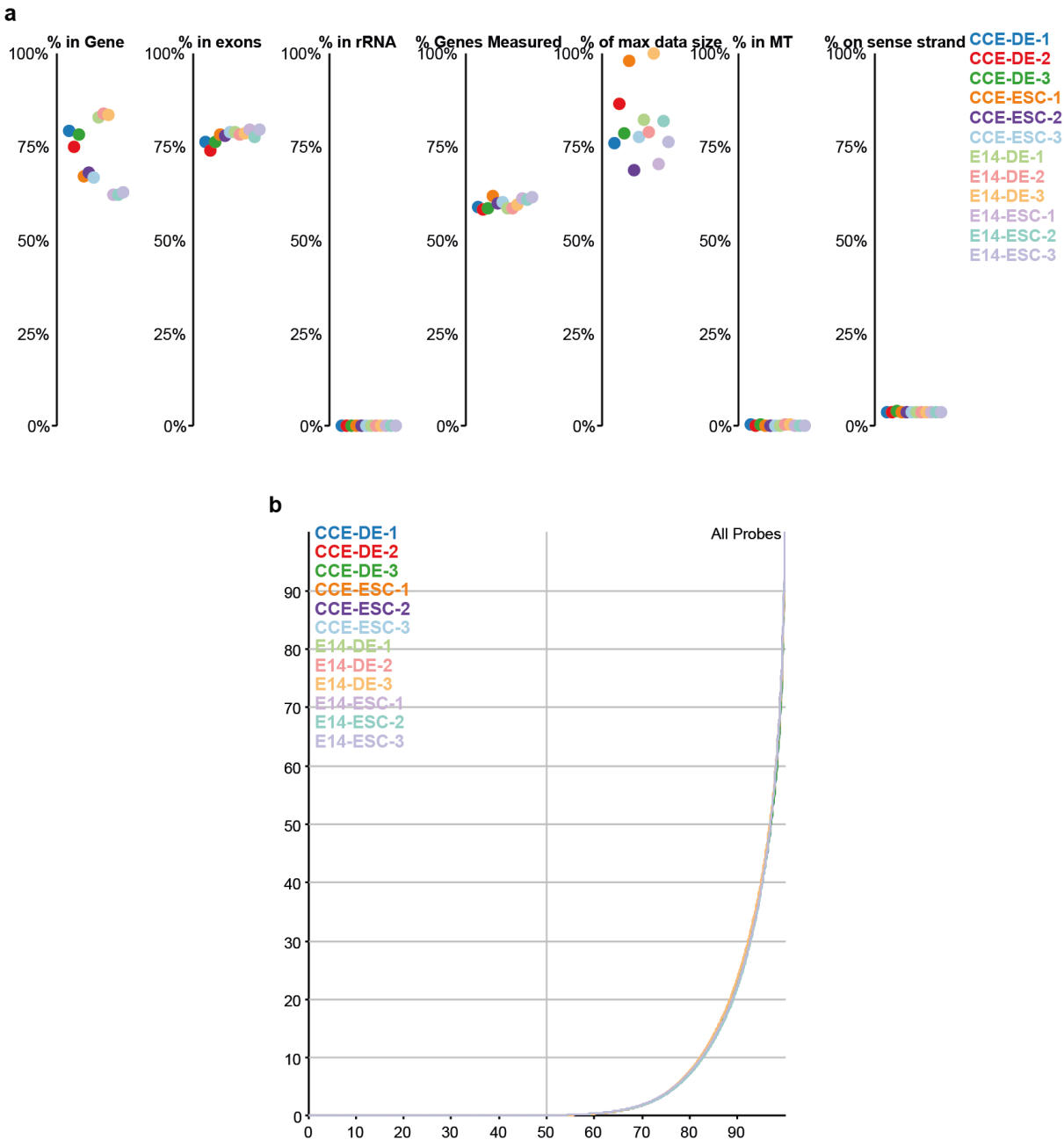


Figure 4.20: Quality control for RNA-seq of undifferentiated mESCs and day 5.0 CXCR4⁺/cKit⁺ EB cells.

(a) Quality control of RNA-seq samples. Statistics shown for each sample.

(b) QQ plot for RNA-seq samples. For each sample, the sum of quantitations for each probe is calculated. Then, probes are ordered from lowest to highest quantitation. The x axis shows the percentile position of probes against the percentile quantitation until that point. Closely-aligning traces indicate that the data between samples is well-normalised.

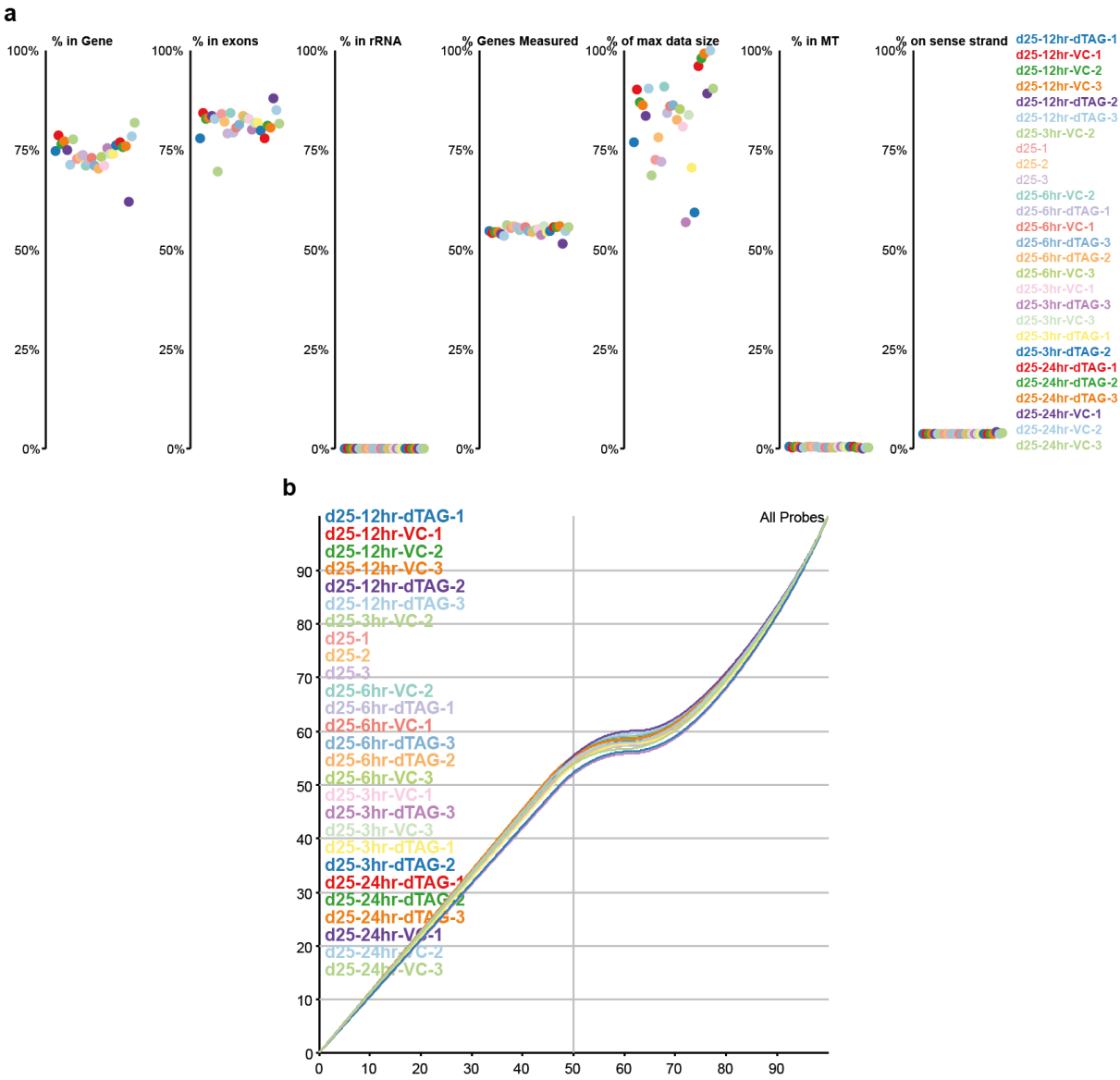


Figure 4.21: Quality control for Eomes$^{deg/deg}$ G9 clone EB differentiation timecourse RNA-seq.
(a) Quality control of RNA-seq samples. Statistics shown for each sample.
(b) QQ plot for RNA-seq samples. For each sample, the sum of quantitations for each probe is calculated. Then, probes are ordered from lowest to highest quantitation. The x axis shows the percentile position of probes against the percentile quantitation until that point. Closely-aligning traces indicate that the data between samples is well-normalised.

Genes associated with differentially accessible peaks at the +12h timepoint
Depdc1b
Foxa2
Lrrc55
Nars2
Olfm4
Olfr987
Pcdh8
Pde4d
Snn
Sstr4
Tenm4
Ttll7

Table 4.3: List of genes associated with Eomes+/- -differentially accessible genomic sites at +12h of dTAG/VC treatment.

Genes associated with differentially accessible peaks at the +24h timepoint							
1190002N15Rik	C77080	Egfl7	Gpc5	Lmx1b	Nfe2l3	Scd2	Ttc7b
4930502E18Rik	Calm1	Egln3	Grik1	Lnx1	Nfib	Scd4	Tub
5530401A14Rik	Camk2d	Eif3f	Grik3	Lrig3	Nipal2	Scyl1	Tubb1
9130230L23Rik	Car4	Eif4ebp2	Grsf1	Lrrc1	Nipal4	Sema3f	Ube2v2
9630041A04Rik	Ccdc188	Enah	Gse1	Lrrc4	Nodal	Sema6d	Usp29
A230072I06Rik	Ccnd2	Eno1	Gsg1l	Ltn1	Notch1	Serbp1	Usp32
Abr	Cd5	Eogt	H2afz	Ly86	Npas3	Sergef	Vapb
Ackr3	Cd6	Epb41l3	H3f3b	Lypd6	Nrde2	Serpina10	Vps37c
Actn1	Cdh19	Ephb1	Habp2	Lypd6b	Nsd3	Sgpp2	Wbscr17
Adam19	Cdh26	Epp13	Hes6	Maf	Olfr1505	Sh3bp4	Wnt10a
Adamts7	Cdh4	Eps8l3	Hist1h1d	Mapk10	Olfr338	Slc16a7	Wwox
Adgrb1	Cdk5r2	Evx1	Hist1h4f	March10	Otulin	Slc22a6	Xrcc5
Adipor2	Cdx2	F2	Hk1	March4	Parp11	Slc24a5	Yars
Agap1	Chic2	Fam105a	Hkdc1	Marco	Pax1	Slc31a2	Zbp1
Agpat3	Clip4	Fam19a2	Hnrnpa2b1	Mcph1	Pck1	Slc9a9	Zbtb14
Alk	Clvs1	Fam19a4	Hoxa13	Mdfi	Pdx1	Snd1	Zbtb43
Alpk2	Cnot6	Fam35a	Hunk	Mex3b	Per2	Socs2	Zcchc6
Angpt2	Col12a1	Fam92b	Id2	Mis18a	Pgd	Sox14	Zdhhc21
Ank2	Cops8	Farsb	Igfbp1	Mkrn1	Phox2b	Sox17	Zfp36l1
Ankmy1	Cox7a2	Fbxl14	Igfbp3	Mlip	Pid1	Sphkap	Zfp667
Ankrd16	Cryba4	Fkbp15	Il17rd	Mn1	Plpp5	Srsf6	
Arhgap24	Csf1	Flt1	Il20rb	Mob1b	Pomp	Sstr4	
Arhgef3	Csf3r	Fmod	Inhbb	Mrc2	Pop1	Stx16	
Armc9	Ctdsp2	Fn1	Irx3	Mreg	Ppp4r4	Sumo1	
Asic2	Cxcr4	Fnbp1l	Itga8	Mrfap1	Ptgs1	Syt15	
Asph	Cxx1b	Foxa2	Jag1	Mroh5	Ptprn2	Tbc1d2b	
Atp23	Cyp26a1	Foxp4	Jakmip1	Mrpl15	Ptprt	Tcf7l2	
Aurkc	Cypt3	Frmd8	Kcnc1	Mrpl42	Pxdn	Tcp10a	
Auts2	Cytl1	Fto	Kcng1	Msx1	Rassf8	Tenm3	
B3gnt7	D430042O09Rik	Gadd45a	Kcnk12	Mvb12b	Rere	Tenm4	
Bach1	Dctd	Galk1	Kctd12b	Myo10	Rin3	Terf1	
Basp1	Ddit4l	Gas1	Kidins220	Myof	Rps6ka2	Tex29	
Bcar3	Ddx1	Gfpt2	Kif1b	N6amt1	Rreb1	Thsd7b	
Bhlhe41	Dennd2a	Gli2	Klhl29	Nars2	Rrp1	Timm22	
Bloc1s4	Dnah5	Gm1043	Lbr	Ncbp3	Rtn4r	Tle4	
Brd2	Dner	Gm973	Lefty2	Ncl	Rufy3	Tmc3	
Btg2	Dsel	Gnat1	Lgmn	Nedd4l	Rwdd2b	Tpo	
C1ql2	Efcab1	Gpc1	Limch1	Nfatc2	Sbspon	Trio	

Table 4.4: List of genes associated with Eomes+/- -differentially accessible genomic sites at +24h of dTAG/VC treatment.
In orange are highlighted those genes near Eomes+ accessible sites.

CHAPTER 5: FINAL DISCUSSION

5.1: CHALLENGES, POTENTIAL SOLUTIONS, AND FUTURE WORK IN MOUSE DEGRON ALLELE APPLICATION

5.1.1: Discussion

We generated the Eomes^{deg} mouse line in the hope that it would be valuable as a tool with which to more finely delete Eomes, for example at different developmental stages, rather than through a total epiblast-null after a Sox2.Cre deletion. Harland et al. (2021) demonstrated that Eomes is required for a brief window of time to prime the chromatin landscape in order to enable the specification of haemogenic endothelium, but not for later development or differentiation. This could be demonstrated with a time-sensitive degron system, but not the full epiblast-null Cre-based system.

As described in Chapter 2, although the Eomes$^{deg/deg}$ mice are viable and fertile, we were only successful in fully depleting Eomes-degron protein when treating embryos with dTAG *ex vivo*. When carrying out *in vivo* dTAG injections, we found that Eomes was inconsistently and variably degraded.

These results from *in vivo* dTAG treatments gave us some uncertainty as to the consistency of the technical aspects of the procedure (such as the efficiency of IP injections). Eomes$^{deg/+}$ x Eomes$^{deg/deg}$ crosses enabled us to compare degron-homozygous embryos that theoretically should lose all Eomes protein upon dTAG addition, to litter-matched heterozygous embryos that do not (which was confirmed by genotyping of each embryo). This enabled us to determine that Eomes depletion does occur in Eomes$^{deg/deg}$ embryos, sometimes to completion, but very often only occurs partially. This variation occurs within litter-matched embryos as well, suggesting that the success or efficiency of each injection is not the cause of this effect.

Due to the expression of Eomes in multiple tissues at the onset of gastrulation and the variable response to *in utero* dTAG injections, we suspect that our more inconsistent results

are due to a gradient of sensitivity to degradation due to differing expression dynamics of Eomes between different tissues of the early gastrula. During gastrulation, Eomes is upregulated in cells of the PS in a very dynamic fashion, with expression beginning, peaking, and ending within approximately 36h. Therefore, the administered dTAG dose is likely insufficient to fully deplete Eomes protein, or at best to maintain full depletion for long. In contrast, Eomes in the ExE lineage is depleted somewhat more effectively, and even more so in the VE. This suggests that expression levels are closer to steady-state in those lineages (consistent with what is known about Eomes expression in the TE and VE), making Eomes depletion more efficient. This also explains the considerably more rapid and robust recovery of Eomes protein in the PS after *in vitro* washout experiments.

This may also play into the apparent lack of a phenotype in Eomes$^{deg/deg}$ embryos subjected to *in utero* dTAG treatment followed by later-stage harvesting (24-48h later). dTAG may be metabolised or removed too soon after administration for Eomes depletion to persist at length, or Eomes might persist or recover despite dTAG presence.

In contrast with our model, Abuhashem et al. (2022b) demonstrated consistent and full degradation of a degron-tagged Nelfb allele, and protein recovery only occurred after several hours of the *in vivo* injection. Nelfb is a housekeeping gene required for transcription, a ubiquitous function, and thus its levels are likely not to fluctuate dramatically, resulting in more efficient depletion after *in vivo* dTAG injections (similarly to Eomes in the VE).

5.1.2: Precautions when employing degron systems

These experiments demonstrated that while the dTAG system was fully functional in *ex vivo*, challenges still presented themselves in our attempts to degrade Eomes in post-implantation embryos *in utero*. Therefore, it is important to consider those limitations to the implementation of a degron-based system that may not be theoretically obvious. Firstly, it is important to consider the timing of dTAG injection relative to embryonic implantation

and establishment of a connection between the maternal and fetal circulatory systems (although we are not certain as to the extent that this may have caused the lack of Eomes degradation at E5.5). Additionally, it is also critical to consider the dynamics of expression of the gene/protein of interest, as the system may be more successfully applied to housekeeping or other genes with a low or steady level of expression. Although Abuhashem et al. (2022b) were successful in fully degrading Nelfb *in utero* at E6.5, we faced challenges with inconsistent Eomes degradation at the same developmental stage. Therefore, researchers interested in implementing this system should consider rigorous validation of the (persistent) depletion of their protein of interest, and/or alternative methods such as *ex vivo* culture, if possible.

5.1.3: Future experiments

As discussed in Chapter 2, we were unable to identify any developmental phenotypic defects in embryos 24 or 48 h after *in vivo* dTAG injections. It may be that Eomes protein levels recover rapidly and robustly enough to prevent defects from developing. However, it may be useful to investigate later timepoints, such as E9.5 or E10.5, to determine if there are any defects or developmental delays that only become apparent at mid-gestation (such as anterior midline defects). It is already known that $Eomes^{+/-}$ mice are viable, fertile, and phenotypically normal. *In vitro* differentiation protocols suggest that an 80% reduction in Eomes expression levels are required to observe a phenotype (Simon et al, 2017). If the developmental situation *in utero* matches the one *in vitro*, and Eomes levels need to be reduced by over 80% in order for defects to become apparent, then it may be that Eomes protein recovery after dTAG administration is rapid enough to surpass the 20% threshold before defects become apparent. Alternatively, the protein loss may simply result in relative developmental delay, which would indicate a high degree of developmental plasticity.

Although dTAG small molecule delivery *in vivo*/*in utero* has only been reported via IP injection (Abuhashem et al., 2022b; Nabet et al., 2018, 2020), it should be possible to assess alternative delivery methods for their efficiency, such as intravenous delivery. However, as

IP injection generally permits delivery of a greater volume (Turner et al., 2011), this likely makes it the most convenient small molecule delivery method in this case.

Several methods have been developed for *ex vivo* post-implantation embryonic culture, of varying costliness and permitting culture for various lengths of time and up to various developmental stages (Aguilera-Castrejon et al., 2021; Rivera-Pérez et al., 2010). The Robertson lab is currently setting up an *ex vivo* longer-term embryo 'roller culture system' (Rivera-Pérez et al., 2010) that can be exploited to carry out post-implantation dTAG experiments while maintaining a steady dTAG concentration, as maternal metabolism, blood circulation, and systemic clearance of the small molecule need not be a concern. Here, the embryo will be harvested and cultured *in vitro*, allowing development to be monitored. This will also allow precise addition of dTAG at particular embryonic stages, something that is not possible *in utero*, where the precise timing of oocyte fertilisation and implantation is unknown. Hence, dTAG can be added at pre-streak, early-streak, or late-streak stages, in order to determine the effects on DE, the node, haematoendothelial and cardiac mesoderm, and axial mesendoderm development. Combining imaging and scRNA-seq methods will hopefully allow us to tease apart the precise roles of Eomes in the development of each lineage. Long-term impact of Eomes loss can also be assessed, as the embryos can now be cultured for multiple days. As *ex vivo* culture is inherently an artificial environment, careful attention should be paid to ensure experimental reproducibility and validation of experimental findings.

5.2: Eomes in the TE lineage

5.2.1: Discussion

The current prevailing theory of the role played by Eomes in TSCs is one of regulation of core TSC GRN components. However, the degron-based experiments detailed in Chapter 3 suggest a potentially more expanded and complex set of functions that previously

appreciated. Utilising the degron-tagged TSC lines, I was able to dissect the effects of Eomes protein loss at high temporal resolution. Eomes appears to be regulating numerous genes related to the cytoskeleton, the ECM, and cell-cell contacts. In addition to cell shape and environment, it may be modulating the signalling environment, allowing the sustained development of the TE lineage.

If microenvironment modulation is indeed a key function of Eomes in the TE lineage, this could go some way in explaining the peri-implantation lethal phenotype of Eomes-null embryos. Indeed, it could be that Eomes enables successful uterine implantation through modulation of the blastocyst's surroundings.

As discussed in Chapter 3, several of the genes identified have better-documented roles in the human TE/placental lineage compared to the mouse, thus suggesting several new potential targets for investigation in mouse TE/placental development.

5.2.2: Limitations of current experiments

The set of experiments described in Chapter 3 has a few notable drawbacks. Firstly, the second, smaller Eomes peptide "fragment" identified via WB remains a mystery, as we were not able to determine which portion(s) of the Eomes peptide sequence it represents. We also don't know what impact this peptide had, if any, on the stem cell state and behaviour more generally, of these TSC lines. However, the mouse line from which the TSCs were derived are viable and fertile, so if this fragment is present *in vivo* then it appears not to have any epistatic or dominant-negative functions. Secondly, the cell-density effect on the transcriptome was non-negligible, and although I attempted to compensate for it in my analysis, this was a less-than-perfect solution. After Eomes depletion, TSCs continue to divide for several days before becoming fully differentiated and senescent, making changes in cell density inevitable. Finally, the clear transcriptional differences between the $Eomes^{deg/deg}$ clones may have been an additional source of noise, although upon Eomes protein depletion, trends common in all three TSC lines did emerge.

It is also worth highlighting that TSCs are an *in vitro* artefact. Caution should therefore be taken when extrapolating *in vitro* findings to a hypothesised *in vivo* equivalent. As it is uncertain whether TSCs are representative of the TE lineage at pre- or post-implantation stages (or both), it is not entirely clear which developmental stage(s) these findings are most representative of. ExE- and EPC-specific equivalent *in vitro* "cell states" have not been isolated (as these are likely mixed pools of stem cell and transient progenitor populations that are difficult to replicate *ex vivo*). It could be that findings about Eomes in TSCs are representative of Eomes' function in the pre-implantation TE, and its roles may be distinct in the proximal TSCs of the post-implantation ExE. Therefore, it may be beneficial for the investigation of Eomes functions in the TE lineage to make use of the Eomes^{deg} mouse line. Timed Eomes depletion may help identify stage-specific Eomes functions in the TE lineage, thus clarifying the current findings.

5.2.3: Future experiments

It will be useful in future analysis to compare the transcriptomes of untreated Eomes$^{deg/deg}$ TSCs to published TSC transcriptomes to determine with greater certainty if they do indeed represent bona fide TSCs. It may also be informative to carry out IF staining for established TSC markers, in order to determine their expression levels relative to WT TSCs, as well as karyotyping them to confirm that they are free from gross chromosomal abnormalities. Finally, assessing the epigenetic state of established TSC-expressed loci, such as Elf5 promoter demethylation (Ng et al., 2008), will provide further evidence of the clones' identity.

As the RNA-seq data showed great variability between clones, further analysis will be useful to determine if this is an analysis artefact, or whether it represents genuine inter-clone transcriptional variability.

Furthermore, it would be informative to validate some of the putative Eomes-regulated genes revealed by the differential expression analysis of the RNA-seq data by assessing the perturbation of these genes (via *in situ* hybridisation or IF) in the TE compartment of dTAG-

treated Eomes$^{deg/deg}$ blastocysts (discussed in Chapter 2). This would provide convincing validation in the early embryo of data collected in embryo-derived cell lines.

Eomes performs several context-dependent activities within the different lineages it is required in, which suggests lineage-specific protein partnerships may be critical in these activities. We are therefore exploiting a recently-developed technique, RIME (rapid immunoprecipitation mass spectrometry of endogenous protein) (Mohammed et al., 2016), in order to identify TSC-specific Eomes protein interactors, in collaboration with the company Active Motif. RIME is not dependent on tagging the protein of interest and has previously been successfully used to determine Eomes protein interactors in memory T cells (Istaces et al., 2019). Briefly, RIME involved light fixation of TSCs with formaldehyde, snap-freezing, protein preparation and IP of Eomes (with a rabbit polyclonal antibody) followed by mass spectrometry of the cross-linked protein complexes. The control for this experiment is IP with a species-matched total IgG antibody. Alternative controls we might have considered was carrying out the IP with the anti-Eomes antibody in dTAG-treated and thus Eomes-depleted TSCs. Once putative interactors of Eomes in TSCs have been identified, we intend to confirm their interaction via co-IP and WB. We further intend to carry out chromatin occupancy assays for the confirmed interactors to determine regions of chromatin binding overlap, and assessment of the mode of action of these interactors by examining the proximity of Eomes and interactor binding motifs. We may thus further refine our understanding of Eomes TSC-specific regulatory functions.

Further assessment of changes in chromatin organisation (conformation capture experiments) and changes chromatin modifications triggered by Eomes depletion may further contribute to our understanding of Eomes in the TE lineage and its roles in maintaining TSC “stemness” as well as promoting embryo-uterine interactions. These experiments will enable us to identify if Eomes is actively responsible for maintaining the chromatin landscape in the “steady stem cell state” of TSCs. Finally, assessment of chromatin-chromatin interactions via Hi-C will enable us to determine definitively the

genomic regions that Eomes-bound loci interact with and regulate, rather than just relying on the "nearest gene" rule.

In carrying out further studies on Eomes' role in the TSC lineage (and others), it would be very useful to assess its expression in other species during embryonic development. This will enable identification of potential conserved functions across species and could be of interest in the identification of Eomes-centred GRNs.

In addition to examining TSCs, the *in vitro* equivalent of the TE, it may also prove beneficial to carry out experiments with Eomes$^{deg/deg}$ blastocysts. Specifically, an *ex vivo* blastocyst implantation assay (Bedzhov et al., 2014; Ichikawa et al., 2022), combined with Eomes dTAG-mediated depletion at different stages, will permit the assessment of Eomes requirement in the TE lineage at high temporal resolution.

5.3: Eomes roles in the APS and its daughter lineages

5.3.1: Discussion

In Chapter 4, I describe the optimisation an EB-based differentiation protocol employing cytokines and small molecules aimed at reproducing the conditions found in the APS, with the aim of specifying the maximal possible percentage of cells toward a DE identity. Assessment of established cell surface markers by flow cytometry, as well as RNA-seq, gave me confidence that this protocol robustly generates a population of cells exhibiting early DE identity. Unsurprisingly, cells of different genetic backgrounds behaved differently under the same conditions, indicating that there may be persistent epigenetic differences between ESC lines impacting cellular response to extracellular signalling.

I was then able to demonstrate, through use of homozygously-tagged Eomes^{deg} ESC clones, that depletion of Eomes protein is sufficient to result in loss of the putative DE progenitor CXCR4^{+}/cKit^{+} population. This effect is timing-dependent and suggests that there is a

window of approximately 36h during which DE progenitors appear to be continuously and steadily specified. This roughly corresponds to the duration of Eomes protein expression in the DE EB cultures (as discussed in Chapter 4).

Although the DE population is specified steadily over the course of approximately 1.5 days, cells specified at different timepoints do not yet appear to correspond to clear DE subpopulations along the A-P axis, as determined by A-P marker RT-qPCR analysis. This corresponds to what is already known about DE specification, i.e. that regionalisation occurs as a distinct step after specification (Sherwood et al., 2009).

I then investigated the earliest-affected genes upon Eomes depletion during DE differentiation, in terms of RNA expression and chromatin accessibility. It may be an intuitive conclusion that earlier-affected genes are more likely to be directly regulated by Eomes. However, certain genes may only appear differentially expressed at later timepoints because their expression only begins (and can thus only be affected) at later stages of the differentiation. Therefore, a more convincing way to establish direct regulation by Eomes of a specific gene of interest (GOI) could be by assessing chromatin occupancy of Eomes (discussed below, in 5.3.3, Future experiments). If a region proximal to the GOI is identified as Eomes-occupied, that region may then be deleted to attempt to establish a direct causal link between Eomes and regulation of that gene.

Several known endodermal and mesodermal markers are some of the earliest-affected genes, including DE master regulators FoxA2 and Sox17, whose expression is reduced, but which also are associated with differentially accessible chromatin regions known from previous studies to be targets of intracellular effectors of cell signalling. This suggested a possible mechanism by which Eomes binds to these regions, promoting expression of these master regulators of DE identity. Deletion of these regions, however, did not result in loss of the $CXCR4^+/cKit^+$ population. In fact, in the case of the FoxA2-associated region, there was an increase in the double-positive population. This suggests that, if these regions do indeed play a role in expression of FoxA2 and Sox17, this role is likely non-exclusive and

redundant with other regulatory mechanisms. Indeed, in the case of FoxA2, two other possible candidates for regulatory elements are readily found in the –50 and –53 enhancers identified in Wang et al. (2017). These two enhancers, although they did not appear differentially accessible in my ATAC-seq experiments, are bound by ß-catenin and Tcf3 in an ActA-dependent manner similarly to the –37 enhancer that I deleted. Another non-mutually-exclusive explanation is that the regions I selected for deletion did not include the full extent of the relevant enhancer region, thereby still permitting TF binding and enhancer activity.

Investigation of the differentially accessible sites in the Eomes depletion time course experiments suggests further possible modes for Eomes function in APS lineages. Eomes+ accessible sites are enriched, perhaps unexpectedly, for the Nanog consensus sequence. Nanog prevents precocious loss of pluripotency through various mechanisms. Among others, it can suppress the expression of differentiation genes through sustaining H3K27me3 repressive histone marks (Heurtier et al., 2019). At later developmental stages, Eomes may be activating or enabling access to some of those previously Nanog-repressed genes. As Nanog is still expressed in the posterior side of the epiblast during gastrulation (Osorno et al., 2012), this suggests a potential competitive interaction with Eomes.

Conversely, sites accessible upon Eomes depletion are enriched for multiple members of the Tcf and Lef families. These are intracellular effectors of Wnt signalling, and the fact that their consensus motifs are enriched in Eomes-depletion-accessible sites suggests that, when Eomes protein is present, it may function to restrict accessibility of such sites. Wnt is broadly known in vertebrate development as a posteriorizing agent, responsible for suppressing anterior/ectodermal identity and promoting posterior/meso- and endodermal identity (Yamaguchi, 2001). However, in order for the different lineages emerging through the PS to acquire their distinct identities, it may be necessary to differentially modulate Wnt signalling. For example, in the case of APS derivatives, Eomes may function to restrict access to certain Wnt target genes with Tcf/Lef binding sites. Further investigation on the

effect of Eomes on signalling pathway components may shed light on how it functions to specify a particular cell lineage. It has been demonstrated in previous research that Eomes is sufficient to replace Wnt signalling during activin-mediated mesodermal and endodermal differentiation in a human ESC model (Yoney et al., 2022), which strongly suggests an overlap or interplay between Eomes' and Wnt's downstream effects. Further investigation into this, for example by investigating the effect of Eomes depletion on Wnt/ß-catenin target genes, regions or Eomes/ß-catenin co-localisation, or Eomes overexpression in ß-catenin-null cells, may further elucidate the complex signalling pathways essential in specification of the three germ layers.

Finally, it is worth noting that one of the earliest effects of Eomes protein depletion in my EB system is the upregulation of genes known to be expressed in paraxial mesoderm. Foxb1 marks nascent paraxial mesoderm (Zhao et al., 2007). Consistent with this, the single-cell multiome data reveal that Eomes-depleted EBs have a much higher proportion of paraxial and somitic mesoderm cells compared to Eomes+ EBs, suggesting that part of the role of Eomes in the APS is to suppress paraxial mesodermal identity. Interestingly, Luke Harland, who also carried out his DPhil research in the Robertson lab, makes a similar comment in his thesis, based ChIP-seq data suggesting that Eomes binds near genes associated with absent somite formation and failure of somite differentiation (Harland, 2020). Notably, Eomes-null cells have been shown in chimera experiments to contribute to paraxial mesoderm (Costello et al., 2011). It may therefore be interesting to investigate how Eomes may be actively suppressing certain cell identities in the PS, while concomitantly promoting others.

5.3.2: Single-cell multi-ome

Thanks to the multi-ome experiment in combination with the gastrulation atlas (Pijuan-Sala et al., 2019), we were able to identify the precise cell type composition of the EBs that I generated with my differentiation protocol. This will enable a much more detailed parsing of the molecular events resulting in DE specification. We determined that, consistent with

what is known about Eomes expression and function from fate mapping and chimera experiments, its ablation results in loss of the APS lineage as well as its daughter lineages, the node/notochord and the DE.

It is also interesting to note the apparent similarity of the chromatin accessibility profiles of the node/notochord and DE cells. Investigating how this results in distinct transcriptomes and vastly different developmental roles in the mouse embryo may yield useful insights in the specification of these lineages.

Further analysis of the multiome data, currently being carried out in collaboration with the Göttgens group, may also contribute to the identification of differentially accessible or expressed loci that were potentially 'diluted out' from my bulk experiments. This may explain the lack of a loss-of-function phenotype in my enhancer deletion EB differentiations. Other potential Eomes-regulated chromosomal regions controlling FoxA2 and Sox17 expression may not have been identified as differentially accessible due to this dilutive effect.

In addition to the above, the multi-ome experiment will permit integration of the RNA and ATAC data, thereby offering the opportunity to correlate transcriptional state with chromatin accessibility at single-cell resolution, as discussed in Chapter 4. It will also permit the reconstruction of the GRNs regulating specification of the cell types generated in Eomes+ versus Eomes-depleted EBs, thus elucidating the higher-order control of cell identity acquisition during gastrulation.

Further investigation of the data will enable us to examine cell type-specific Eomes functions in isolation, as the bulk experiments that I have carried out offer an "average" view across all cells and cell types. Overall, this dataset may be able to shed more light on how Eomes functions to specify this variety of closely-appositioned cell types that have vastly differing functions in the emerging embryo.

5.3.3: Future experiments

By adding dTAG-13 to the DE EB cultures at ever earlier timepoints, I was able to determine that day 2.5-4.0 is the critical window for $CXCR4^+/cKit^+$ cell specification. It would be interesting to carry out the inverse experiment, i.e. to add dTAG-13 on day 2.5 and wash it out at ever later timepoints, in order to establish if these EB cultures retain the capacity for $CXCR4^+/cKit^+$ cell specification upon Eomes protein re-establishment, and if so, for how long after dTAG-13 addition and Eomes protein depletion. This would be an interesting way to investigate whether Eomes-depleted EBs (or at least some cells within them) remain at all developmentally "paused," or, rather, whether they rapidly and uniformly assume an alternative differentiation trajectory, thereby losing the capacity to specify APS/DE progenitors.

Despite the lack of a profound phenotype, I am currently working on generating double enhancer deletion ESC lines (of the two differentially accessible sites identified in my experiments) in order to investigate their phenotype during EB differentiation. The method I used to test clones for deletion involved carrying out PCR at the junctions of the targeted region, as well as internal to the region (all of which, in deletion clones, I expected to return a negative PCR result), as well as a PCR specific to the repair template I used for the targeting (which I expected to appear positive in deletion clones). However, this strategy lacks the capacity to detect potential larger-scale (multi-kb) deletions on one allele. To circumvent this issue, it may be useful to employ ddPCR (droplet digital PCR), which, in addition to amplifying, is additionally capable of quantifying target DNA regions (Taylor et al., 2017).

Although the putative Foxa2/Sox17 enhancer regions I deleted were determined by the differentially accessible peaks that were revealed during the ATAC-seq data analysis, a more thorough way to determine the boundaries of the putative enhancer sites would have been to investigate the conservation of accessible sites between species, as this may be an indicator of a conserved, and therefore critical, enhancer region. This holds for the putative

Foxa2/Sox17 enhancers I investigated, as well as the suspected 10kb super-enhancer region upstream of Eomes. ATAC data from bona fide mouse embryo cells (and in particular, scATAC data from recently-published early development atlases) may give an indication as to whether a region appears to possess (super) enhancer characteristics, and in particular if it does so across different cell types (e.g. between TSCs and APS progenitors, as examined in this thesis).

Additionally, it will be useful to assess other potential factors contributing to the regulation of DE-critical loci, such as chromatin modifications. I am therefore aiming to carry out chromatin occupancy assays for active and inactive chromatin marks, in order to establish if, and how, Eomes presence/absence regulates them. Most critically, though, chromatin occupancy assays for Eomes protein itself will be key in formulating a more comprehensive model for Eomes function within the DE lineage.

5.4: General principles of Eomes regulatory functions between TSC and APS lineages

Although I did not carry out in-depth comparisons between the functions of Eomes in the TSC and DE lineage in the context of this thesis, some broad principles emerge from the two individual datasets. Firstly, based on the degradation timecourse RNA-seq in the two lineages, Eomes appears to be broadly acting as a gene activator in both, as many more genes become downregulated upon Eomes protein depletion compared to the genes that are upregulated at the same time. This is aligned with what is known about Eomes as a gene-activating TF.

Nevertheless, the similarities appear to end there, as the genes regulated by Eomes are vastly different in the two lineages. It is clear, therefore, that Eomes acts context-dependently, with its activities likely modulated by cell type-specific co-factors [as suggested by the enrichment of binding motifs for such putative co-factors in the differential

accessibility dataset of the APS/DE EB differentiation (Foxa2, Gata6, Lhx1), and the Eomes CUT&RUN dataset in TSCs (Cdx2, AP-2γ, Gata3, Elf5)]. An exception to the above is the rare occasion when there is an apparent 'context crossover,' where baseline control of a gene by Eomes is altered even in contexts where that gene is not normally expressed/required – for instance, as discussed in this thesis, Cebpb, a gene expressed in the TB lineage, is downregulated upon Eomes depletion in the APS/DE EB differentiation.

CHAPTER 6: MATERIALS & METHODS

6.1 METHODS

6.1.1: Stem cell maintenance and differentiation

mESC lines (CCE, CCE $Eo^{deg/deg}$, E14, E14 $Eo^{deg/deg}$, E14-RV, E14-RV $Eo^{deg/deg}$, Foxa2 Sox17 enhancer deletion clones) were maintained in feeder-free, serum + LIF conditions as previously described (Costello et al., 2011). Cells were dissociated in 0.25% Trypsin (Gibco, 25200-056) and passaged every 2 days. 72 hours prior to EB differentiation, cells were plated at a density of $0.6x10^6$/ml, and 48 hours prior to differentiation, cells were washed in PBS (Gibco, 14190-094) and the medium replaced with serum-free ESC media containing 50% Neurobasal medium (Gibco, cat. no. 21103049), 50% DMEM/F12 (Gibco, cat. no. 11320033), and supplemented with 0.5x N2 (Gibco, cat. no. 17502048) and 0.5x B27 (Gibco, cat. no. 17504044), 1% Pen/Strep (Gibco, 15140-122), 1% L-glutamine (Gibco, 25030-024), 0.05% BSA (Gibco, 15260037), 1μM PD-0325091 (University of Dundee, MRC-PPU unit), 3μM CHIR99021 (University of Dundee, MRC-PPU unit), and 1000U/ml LIF (Millipore, ESG1107). On d.0 cells were washed with PBS, dissociated with TrypLE (Gibco, 12605-010), and neutralised in 10% FBS (Gibco, 10500-064), 90% IMDM (Gibco, 12440053). Cells were seeded at a density of $5x10^4$/ml in SF-D medium (Nostro et al. 2008) on 6cm non-TC-treated dishes (Thermo Fisher, 150340) and cultured on an orbital shaker at 70 rpm for 24h without growth factors in order to induce EB formation. On d.2.0, EBs were expanded into 10 cm dishes (Thermo Fisher, 101VR20) coated in 5% PolyHEMA (Sigma, P3932-10G), in SF-D medium supplemented with growth factors and small molecules. For DE specification, SF-D was supplemented with 50ng/ml Activin A (R&D Systems, 338-AC), 1mM CHIR 99021, and 50 nM LDN-193189 (Stemgent, 04-0074). For haematoendothelial specification, SF-D was supplemented with 5ng/ml recombinant human (rh) VEGF (R&D Systems, 293-VE), 10ng/ml rhBMP4 (R&D Systems, 314-BP), and 5ng/ml Activin A.

TSCs were maintained in serum-free, defined medium originally described in Kubaczka et al. (2014). DMEM/F12 (Thermo Fisher, 21331-020) was supplemented with L-glutamine

(Thermo Fisher, 25030-24), Pen/Strep, L-ascorbic acid (SigmaAldrich, A8960-5G), sodium bicarbonate (SigmaAldrich, S5761-500G), sodium selenite (SigmaAldrich, S5261-10G), and insulin (SigmaAldrich, I5523-50MG). For differentiation medium, medium was supplemented with transferrin (SigmaAldrich, T8158-100MG). For TSC maintenance medium (Tx4+), medium was additionally supplemented with TGFβ1 (Peprotech, 100-21 1µg), FGF4 (SigmaAldrich, F8424-25UG), and heparin (SigmaAldrich, H3149). TSCs were cultured on Matrigel-coated (Corning, 354230) tissue culture plates per Kubaczka et al. (2014). Cells were passaged when they reached 80-90% confluence using 0.05% Trypsin (Gibco, 25300-054).

For protein degradation in degron-tagged clones, small molecule dTAG-13 (Bio-Techne, 6605) was resuspended in DMSO per manufacturer's instructions and then diluted to a final concentration of 200nM in culture medium. DMSO at the same dilution was used as a vehicle control. For dTAG washout experiments, EBs were washed thrice in IMDM and resuspended in dTAG-free medium.

6.1.2: Generation of Eomes-degron-P2A-mCherry repair template

An IDT Megamer® was ordered containing the sequence for the Eomes-degron-P2A-mCherry KI allele, flanked by homology sequences for the 5' (the 3' end of the Eomes exon 6 CDS) and 3' (3' UTR) homology arms. The Megamer was amplified with primers containing restriction digest sites EcoRV and Nb.BsmI on the 5' end of the repair template (5'-EcoRV-Nb.BsmI-3'), and NotI on the 3' end of the repair template (all enzymes from eNew England Biolabs). The PCR product and a pBlueScript KS- plasmid were then digested with EcoRV and NotI, and the PCR fragment was ligated into the plasmid with NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, E2621). The plasmid was amplified in NEB Stable cells (New England Biolabs, C3040H) following manufacturer's instructions. The plasmid was then digested with Nb.BsmI and NotI, denatured using Denaturing Gel-loading Buffer (Biodynamics Laboratory, DS612), and run out on an agarose gel, from which the single-stranded fragment was cut out and extracted using the ZymocleanTM Gel RNA recovery kit

(Zymo Research, R1011). The fragment was then freeze-dried and resuspended in IDT duplex buffer, Nanodropped and used as a ssODN repair template (described below) at 0.5pmol.

6.1.3: Cas9-mediated mESC modification

Custom synthetic crRNA and ssODNs (Integrated DNA Technologies, USA) were resuspended in IDT duplex buffer. Sequences are provided in Table 6.2. crRNA sequences were designed using IDT's Custom Alt-R® CRISPR-Cas9 guide RNA design tool (https://eu.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE). Feeder-free and feeder-dependent cell medium was changed to Pen/Strep-free on the morning of the electroporation. Cas9 ribonucleoprotein (RNP) complexes were assembled immediately prior to electroporation following manufacturer's protocol (Alt-R CRISPR-Cas9 System: Delivery of ribonucleoprotein complexes into Jurkat T cells using the Neon® Transfection System). Briefly, 4.4µl of 100µM tracrRNA, 4.4µl of 100µM crRNA, and 1.2µl if IDT Duplex Buffer were mixed, heated at 95ºC for 5 min using a thermocycler and then cooled to RT. The RNP complex was formed by diluting Alt-R® S.p. HiFi Cas9 Nuclease V3 in Buffer R to 36µM and mixing it 1:1 with the crRNA:tracrRNA duplex. mESCs were electroporated using the Neon transfection system (V=1600V, pulse width=10ms, number of pulses=3) with the appropriate ssODN (50pmol) and RNP. If two or three crRNAs were used, then RNPs were produced separately and then mixed 1:1 or 1:1:1 prior to electroporation.

For generation of the $Eo^{deg/deg}$ clones, a custom crRNA targeting the UAG translational stop in exon 6 of Eomes (Harland et al., 2021) was used. A ssDNA fragment was designed to insert a degron tag, a P2A self-cleaving peptide, and an mCherry fluorescent protein, all in-frame with the Eomes exon 6. An RNP was generated as described above. $2*10^5$ cells per electroporation in 10µl buffer R were electroporated with 1µl RNP and 2µl ssODN as described above. Low-density plating was performed after 2-3 days and colonies picked after 7-10 days and screened by PCR. Homozygous clones were selected and the full-length insert was amplified by PCR and Sanger sequenced.

For generation of Eomes, Foxa2 and Sox17 enhancer deletion clones, the same Cas9-mediated DNA modification strategy as above was used, with different gRNAs and restriction site-containing ssODN repair templates. The deletion repair site was amplified by PCR and homology directed repair (via the ssODN template) was confirmed by restriction digestion. Absence of genomic reintegration of the deleted enhancer sequences was confirmed by PCRs of targets internal to the deletion region.

6.1.4: Mouse strains and genotyping

All animal procedures were carried out in accordance with Home Office (UK) regulations and approved by the Oxford Local Animal Welfare and Ethical Committee. Two independent feeder-dependent $Eomes^{deg/deg}$ mESC clones were used to generate germline chimeras. Genotyping primers for the offspring are listed in Table 6.3. The strain was maintained on a C57Bl/6 background.

6.1.5: Immunostaining

EBs, embryos, and cells on glass coverslips were fixed in 4% PFA for 15 min at RT or 1% PFA at 4ºC overnight. Samples were washed with 0.1% Triton-X in PBS, permeabilised with 0.5% Triton-X in PBS for 15 min, washes with 0.1% Triton-X, blocked in 5% donkey serum, 0.2% BSA, 0.1% Triton-X, incubated in primary antibody overnight at 4ºC in blocking solution, washed with 0.1% Triton-X, incubated with fluorophore-conjugated secondary antibody at RT for 2 h in blocking solution, washed with 0.1% Triton-X and counterstained with DAPI. Fluorescent images were captured on an Olympus FV1000. Antibodies are listed in Table 6.1.

6.1.6: Western blots

Cells or EBs were lysed in RIPA buffer (50mM Tris pH8.0, 150mM NaCl, 1% Igepal, 0.5% sodium deoxycholate, 0.1% SDS). Protein was quantified by Bradford assay with the DC™ Protein Assay Kit I (Bio-Rad, cat. no. 5000111) on a Jenway Genova DNA Life Science Analyser. Samples were denatured at 98ºC for 10 min in Laemmli sample buffer (BioRad, 1610747)

with 10% β-mercaptoethanol, run at 90V on a Mini-Protean® PAGE gel (BioRad, 4561094) in a Mini Trans-Blot® Cell, and transferred onto PVDF membrane for 75 minutes at 90V. The membrane was rinsed with dH_2O, optionally Ponceau stained, and washed in TBST (0.1%) for 10 minutes. Membrane was blocked with EveryBlot blocking buffer (BioRad, 12010020), or 5% non-fat milk in TBST for 1h at RT and the primary antibody was added and incubated on a shaking platform at 4ºC overnight. Membrane was washed with TBST and secondary antibody was added in blocking buffer. Membrane was washed TBST and ECL prime (SigmaAldrich, RPN2236) was added to the membrane per manufacturer's instructions. The membrane was exposed to X-ray film and developed in a dark room. When necessary, the membrane was stripped with stripping buffer (2.9g glycine, 20ml SDS in 2l H_2O, pH2.2) and blocked, washed, and exposed to antibody as described above.

6.1.7: Flow cytometry and FACS

EBs were dissociated in TrypLE and neutralised in FACS buffer (PBS supplemented with 1% Pen/Strep, 2% FCS). Fluorophore-conjugated antibodies were diluted to appropriate concentration in flow buffer, cells were resuspended in the solution and stained on ice for 30 min. Cells were washed in FACS buffer twice, stained with 1:5000 DAPI in FACS buffer (BD Biosciences, 564907) for 15 minutes on ice. Cells were again washed and resuspended in FACS buffer and analysed on a BD Fortessa X20. Fluorescence-activated cell sorting (FACS) was carried out on a BD FACSAria Fusion.

For intracellular flow cytometry, cells were dissociated and stained with a viability dye (ThermoFisher, L34955) and fixed using Fixation/Permeabilization Concentrate and Diluent (cat. nos. 00-5123-43, 00-5223-56) for 40 minutes on ice, protected from light. Cells were washed in FACS buffer and permeabilised using Permeabilization Buffer (eBioscience, cat. no. 00-8333-56). Antibody was added to the permeabilization buffer and cells were incubated at room temperature for 30 minutes, protected from light. Cells were washed and resuspended in FACS buffer.

Flow cytometry analysis was carried out on FlowJo™ v10.8.1 Software (BD Life Sciences). Antibodies used are listed in Table 6.1.

6.1.8: RNA isolation and qPCR

RNA was isolated from cells using QuickRNA Microprep (Zymo Research, cat. no. R1051) according to manufacturer's instructions. cDNA was generated with SuperScriptIII and oligo-dT primers (Invitrogen, 18080-051) and qPCR was performed with the QuantiTect SybrGreen PCR kit (Qiagen, 204143) on the Quant Studio 3 (Thermo Fisher). Data analysis was carried out using the ΔΔCt method. Antibodies used are listed in Table 6.3.

6.1.9: RNA-seq

450 ng RNA per sample were provided to the Wellcome Trust Centre for Human Genetics for ribodepletion, library preparation, and sequencing.

6.1.10: RNA-seq analysis

Paired-end reads were aligned to the mm10 mouse genome using STAR aligner with default settings (Dobin et al., 2013). Bam files were imported to SeqMonk 1.48.0 (RRID:SCR_001913) (http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/). Differential expression was assessed in SeqMonk using DESeq2 (Love et al., 2014) using default settings ($p < 0.05$, multiple testing correction) and the significantly differentially expressed genes were filtered by intensity difference to include those most differentially expressed among genes with similar expression levels, thus accounting for greater expression variability in lowly-expressed genes. PCA plots and hierarchical clustering was carried out in SeqMonk.

6.1.11: ATAC-seq

Tagmentation and indexing of single-cell suspensions ($6x10^4$ cells) was carried out as described in Buenrostro et al. (2013), using the cells from triplicate experiments also used to carry out RNA-seq. To control for Tn5 transposase sequence bias, 100 ng gDNA of the cells was also tagmented and indexed with custom Nextera index primers taken from table S1 in

Buenrostro et al. (2013). Libraries were purified by upper and lower size-selection with Agencourt AMPure XP beads (Beckman Coulter, A63880) (0.6:1 --> 1.5:1 beads:sample ratio). Library size and concentration was determined with the Agilent High Sensitivity D1000 ScreenTape system. Samples were multiplexed and sequenced on a 300-cycle paired-end NextSeq500 Illumina platform.

6.1.12: ATAC-seq analysis

Paired-end reads were aligned to the mm10 mouse genome using Bowtie2 (Langmead & Salzberg, 2012), sorted and mitochondrial reads discarded using Samtools (Li et al., 2009), duplicates removed, and biological replicates downsampled to contain the same number of reads using Picard (http://broadinstitute.github.io/picard/). Bigwig files were generated using deepTools2 (Ramírez et al., 2016). Peaks were called using MACS2 (Zhang et al., 2008) with the tagmented gDNA sample as a control. MACS2-called peaks with p-value $<10^3$ were used in downstream analyses. Loci of significant differential accessibility were identified through DiffBind (Stark & Brown, 2011) on the Galaxy Project's public server at usegalaxy.org (Afgan et al., 2016) using bam files and a bed file of all identified peaks in each sample. GREAT analysis was carried out to identify peak-gene associations (McLean et al., 2010). Enrichment of transcription factor binding motifs in the differentially accessible peaks was carried out on HOMER (Heinz et al., 2010). Genomic snapshots were visualised in IGV (Robinson et al., 2011) from previously-generated bigwig files.

6.1.13: CUT&RUN

Cleavage Under Targets & Release Using Nuclease (CUT&RUN) and library prep was carried out as described in Skene et al. (2018) with the following differences or specific options: 500,000 cells were used per sample, samples were bound to antibody at 4ºC overnight, the targeted chromatin release was carried out for 30 minutes instead of 10, and the alternative DNA extraction method for quantitative recovery of <80bp fragments was carried out. DNA fragment quantitation was carried out on a Qubit 4 Fluorometer with the dsDNA HS assay kit (ThermoFisher, Q32851) following manufacturer's instructions, and library prep was

carried out with an Illumina NEBNext Ultra II DNA library prep kit (New England Biolabs, E7645S). Library yield was again quantified with the same method as above and size distribution of the libraries was determined with the Agilent High Sensitivity D1000 ScreenTape system. Samples were multiplexed and sequenced on a 150-cycle paired-end NextSeq500 Illumina platform.

6.1.14: CUT&RUN analysis

Paired-end reads were aligned to the mm10 mouse and sacCer3 yeast genomes using Bowtie2, and sorted using Samtools. Successfully aligned reads against the yeast genome were quantified with Samtools flagstat, and bam files were downsampled to normalise based on the yeast spike-in DNA. Peaks were called using MACS2 (Zhang et al., 2008) against IgG CUT&RUN controls on Galaxy Project's public server at usegalaxy.org (Afghan et al., 2016). Peak-gene associations and motif enrichment analyses were carried out using GREAT and HOMER tools as described above in ATAC-seq analysis section.

6.1.15: 10X multiome

EBs were dissociated and neutralised as described above in the Flow Cytometry section. Nuclear isolation was carried out according to the 10X protocol for cell lines and PBMCs (https://assets.ctfassets.net/an68im79xiti/6t5iwATCRaHB4VWOJm2Vgc/bdfd23cdc1d0a321487c8b231a448103/CG000365_DemonstratedProtocol_NucleiIsolation_ATAC_GEX_Sequencing_RevB.pdf). Nuclei were counted manually in trypan blue and were diluted such that a maximum of 16,000 were taken forward for 10x Multiome library preparation. Library preparation and sequencing were carried out at the CRUK-CI Genomics Core.

6.2: MATERIALS

6.2.1: Antibodies

Use	Product	Source	Catalogue number	Dilution
Immunofluorescence	Rabbit polyclonal α-mouse Eomes	Abcam	ab23345	1:200
	Rat monoclonal α-mouse Eomes	Invitrogen	14-4875-82	1:500
	Goat polyclonal α-Foxa2	R&D Systems	AF2400	1:100
	Donkey anti-rat IgG	Invitrogen	A21208	1:400
	Donkey anti-goat IgG	Invitrogen	A11058	1:400
	Donkey anti-mouse IgG	Invitrogen	A21202	1:400
	Vectashield antifade mounting medium with DAPI	Vector Laboratories	H-1200	N/A
Flow cytometry	AF488 α-mouse Eomes	Invitrogen	53-4875-82	1:400
	BV605 α-mouse CD184 (CXCR4)	BioLegend	146519	1:100
	BV711 α-mouse CD177 (cKit)	BioLegend	105835	1:160
	APC α-mouse CD326 (EpCAM)	BioLegend	118214	1:80
	APC α-mouse Flk-1	Thermo Fisher	17-5821-81	1:40
	PeCy7 α-mouse PdgfRa	Thermo Fisher	25-1401-82	1:160
	DAPI	BD Pharmigen	564907	1:5000
	Viability dye	Invitrogen	L34955	N/A (per manufacturer instructions)
Western blot	Rabbit polyclonal α-mouse Eomes	Abcam	ab23345	1:2000
	Rat monoclonal α-mouse Eomes	Thermo Fisher	14-4875-82	1:10,000
	Rabbit polyclonal α-ß-tubulin	Cell Signaling	2146	1:2000
	Donkey α-rabbit IgG, HRP conjugated	Amersham	NA934V	1:2000
	Goat α-rat IgG, HRP conjugated	Cell Signaling	7077	1:2000
CUT&RUN	Rabbit polyclonal α-mouse Eomes	Abcam	ab23345	1:50
	Mouse monoclonal α-human Tead4	Abcam	ab58310	1:100
	Rabbit monoclonal α-H3K27me3	Cell Signaling	9733	1:100
	Rabbit IgG	Santa Cruz	sc-2027	1:100
	Mouse IgG	Santa Cruz	sc-2025	1:100
	Guinea pig α-rabbit IgG	Antibodies-Online	ABIN101961	1:100
	Mouse IgG2a	Antibodies-Online	ABIN2749088	1:100
	Rabbit α-goat IgG	Abcam	ab6697	1:100
	Mouse IgG,κ	BD Biosciences	554121	1:100

Table 6.1: List of antibodies.

6.2.2: Cas9 targeting reagents

Product Type	Name	Sequence	Application
*: Phosphorothioated bases			
crRNA (sequence without PAM)	gRNA#1	AAGGTTAAAATAATGCTCTAGGG	Generation of Eomes-degron ESC lines
	Foxa2_en_up_1	GGTTCCCAACGAGACTCTGT	Generation of Foxa2 enhancer deletion ESC lines
	Foxa2_en_down_1	GTGCCAACATCGATTCCAGC	
	Sox17_en_up_1	TATTACCACGGTTCTAGTAC	Generation of Sox17 enhancer deletion ESC lines
	Sox17_en_down_1	GGGAGCATTCTGCTATACAT	
	Eomes_en_up_1	CAGGAGTCGCAACTCACCAC	Generation of Eomes enhancer deletion ESC lines
	Eomes_en_down_1	CAAGAGCCCCCACAAGGTTA	
	Eomes_en_mid	AAAAGAAGGTTATCAGCACG	
tracrRNA	Alt-R® CRISPR-Cas9 tracrRNA	N/A (invariable) product # 1072532	All Cas9-based genome editing
Cas9 protein	Alt-R® S.p. HiFi Cas9 Nuclease V3	N/A product # 1081061	All Cas9-based genome editing
Ultramer™ ssDNA Oligo	Foxa2_1_1_EcoRI	T*G*GAGAGAGCAGGGCTTAGAGAAAAGATGCATTTTTCTTTTCAATCTAATTGCTTTGTTCTATTAAGGTAGGTTTCTCCAGAATTCGGAATCGATGTTGGCACGCTACCATCAAACACAGCATGGATGCATTTGGCAGGGGCTGGGGTGTGG*G*T	Repair template for generation of Foxa2 enhancer deletion ESC lines
	Foxa2_1_1_SpeI	T*G*GAGAGAGCAGGGCTTAGAGAAAAGATGCATTTTTCTTTTCAATCTAATTGCTTTGTTCTATTAAGGTAGGTTTCTCCAACTAGTGGAATCGATGTTGGCACGCTACCATCAAACACAGCATGGATGCATTTGGCAGGGGCTGGGGTGTGGG*T*A	
	Sox17_1_1_EcoRI	A*G*ACTTCATAGATTCCCCCGTTGCACCCAGAATTTCTCTGAACTTAAGTCAGCTCAGCAAAAATGGAAAGTGAATGTCAGGGCAAAACCTGAATTCCATAGGGCATGAATCCTAGATATATTTATGTTTAATTTATTTTAATTACATTTATA*T*G	Repair template for generation of Sox17 enhancer deletion ESC lines
	Sox17_1_1_SpeI	A*G*ACTTCATAGATTCCCCCGTTGCACCCAGAATTTCTCTGAACTTAAGTCAGCTCAGCAAAAATGGAAAGTGAATGTCAGGGCAAAACCTACTAGTCATAGGGCATGAATCCTAGATATATTTATGTTTAATTTATTTTAATTACATTTATAT*G*T	
	Eomes_1_1_SphI	A*A*GGTTTGGGGGTGCCTTTGAGGGCTAGGATGGTGTTGCTTTTCAAGGTGACATCGTTGCTGGGATTAGAAACCCCGCCTGCATGCTTAAGGAGCACAGCTCCGCGGTACAATGTTTGCCCAGCGTGCCTGAGGCCCTGGGTCTAAGGCCTCA*G*C	Repair template for generation of Eomes enhancer deletion ESC lines
	Eomes_1_1_SpeI	A*A*GGTTTGGGGGTGCCTTTGAGGGCTAGGATGGTGTTGCTTTTCAAGGTGACATCGTTGCTGGGATTAGAAACCCCGCCTACTAGTTTAAGGAGCACAGCTCCGCGGTACAATGTTTGCCCAGCGTGCCTGAGGCCCTGGGTCTAAGGCCTCA*G*C	

Table 6.2 (continued on next page)

Repair template for generation of Eomes-degron-mCherry reporter KI ESC lines as excised from pBS backbone (ssDNA)
CATTCGAATTCTACTTCCAGTACCTTGCTCCCATATGGTATTAAGTCCTTGCCCCTGCAGACATCC
CATGCCCTGGGGTATTACCCTGACCCGACCTTCCCTGCTATGGCAGGGTGGGGAGGCCGTGGCG
CTTATCAGAGGAAGATGGCAGCTGGACTACCATGGACATCCAGAATGAGCCCACCTGTCTTCCCA
GAAGATCAGCTTGCCAAGGAAAAAGTTAAAGAAGAGATTAGTTCCTCCTGGATAGAGACTCCCCC
CTCCATCAAGTCTCTAGACTCCAGCGACTCCGGGGTGTACAACAGCGCTTGCAAGAGAAAGCGCC
TGTCTCCCAGCACCCCCAGCAATGGAAACTCGCCCCCCATAAAGTGTGAGGACATTAACACTGAA
GAGTACAGTAAAGACACCTCCAAAGGCATGGGGGCTTATTATGCTTTTTACACAAGTCCCAGTGG
ATCCGAGACTCCTGGTACATCAGAGTCCGCCACACCCGAAAGTGGAGTGCAGGTGGAAACCATCT
CCCCAGGAGATGGACGAACTTTTCCAAAACGTGGACAAACTTGTGTAGTGCACTACACCGGGATG
CTTGAAGATGGAAAGAAAGTTGATTCCTCCCGGGACAGAAACAAGCCCTTTAAGTTTATGCTAGG
CAAGCAGGAGGTGATCCGAGGCTGGGAAGAAGGGGTTGCCCAGATGAGTGTGGGTCAGAGAGC
CAAACTGACTATATCTCCAGATTATGCCTATGGTGCCACTGGGCACCCAGGCATCATCCCACCACA
TGCCACTCTCGTCTTCGATGTGGAGCTTCTAAAACTGGAAGGAAGCGGAGCTACTAACTTCAGCC
TGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTATGGTGAGCAAGGGCGAGGAGG
ATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGC
CACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGAACTCAAACAGCTAAAC
TTAAAGTTACTAAAGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTAC
GGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGA
GGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGA
CTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCG
ACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGA
GGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGA
CGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCA
ACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACTATTGTAGAACAATATGAACGAGCTG
AAGGTCGACATTCCACCGGCGGCATGGACGAGCTGTACAAGTAGAGCATTATTTTAACCTTTAAC
CTTTCCAGATGGACCTGGTGGTGTTTTGTTGTGTTCTTTGCCTAGGTTACCAAAAAGACTTTTGC
CTTCCACCTTTGATGTATCCTGTTTTGTGCAATTCTCTGAGAGAAGGCGCCAAAGCTTTGCCACG
GCCGCAGGTAACTAAACTGAAGCAGACCTAGCATTGGTTTTAAAGTGACTGGAGGACGCAGGAG
TGTTTTTAAAACTGAAAAGGTCGTTCAAGGTGCTGGATTGATTCATTTATGGGAAACGAGAAATG
TTCAGAAAAAGCAGGCTATGAAGAACGAGTGCCCGGTGCTATTAAATGAATTTCAAAGACAAGAA
AACATGACAGAAAAAAACAAAACAAAACAAACTTTGGCTA

Table 6.2: List of reagents for Cas9-mediated genome editing.
All supplied by Integrated DNA Technologies except for the degron KI ssDNA repair template, which was generated as described in 5.1.2.

6.2.3: Genotyping primers

Name	Sequence	Pair with	Product size	Use
AB_38	gtgtacaacagcgcttgcaa	AB_41	585 bp WT, 1728 mCherry degron allele	$Eo^{deg/deg}$ clone genotyping
AB_40	cccacaacgaggactacact	AB_41	515 bp mCherry degron allele	$Eo^{deg/deg}$ clone and ABDE mouse genotyping
AB_41	tctgggctgctgtttgctta		-	
AB_41_F2	TCTAGACTCCAGCGACTCCG	AB_41	607 bp WT alelle, 1750 bp mCherry allele	ABDE mouse genotyping
eo-en-f1	GGGCTAGGATGGTGTTGCTT		-	
eo-en-r1	CCTCACGTTGCCTACTGTGT	eo-en-f1	299 bp in WT	Eomes enhancer deletion genotyping
eo-en-del-r1	ATAGCGACCTCAATACGGCG	eo-en-f1	255 bp in deletion	
foxa2-en-f2	CTAGAACCTCGTGTGGCCTTC	-	442 bp in WT	Foxa2 enhancer deletion genotyping
foxa2-en-r2	AGGCCTGCACTCTTCTTAGTG			
foxa2-en-f1	TTCCAGCTAAGGAACAGCGA	-	349 bp in deletion	
foxa2-en-del-r2	GGCCCACAAACAACATCACC			
sox17-en-f1	CCTGGCTGCCATCTAGCATT	-	336 bp in WT	Sox17 enhancer deletion genotyping
sox17-en-r1	TGGTGAGGTGGCTCAAAGTC			
sox17-en-f2	TCAGCTCTGGTGCTTTTCCT	-	441 bp in deletion	
sox17-en-del-r3	GGTGAGGTGGCTCAAAGTCA			
eo-junc-f1	TGGTGGAGCTCACTTGGAAC	-	284 bp	For detection of products internal to Eomes enhancer deletion
eo-junc-r1	ATAGCGACCTCAATACGGCG			
eo-en1-f2	AGTTGGGTCCAGGCGAATC	-	324 bp	
eo-en1-r2	CTGCGAGAGCTCTAGGGACC			
eo-en2-f1	CACACTCTCCAGGTACGGGT	-	140 bp	
eo-en2-r1	TGGCCCATAAAGTTCCACTCC			
eo-int-f1	GCGGCCAGAAGTGGTTACTA	-	129 bp	
eo-int-r1	TCAAATGGGTCTCTCCGTGC			
eo-int-f2	ATGCCTCACAAACTGCCTCA	-	427 bp	
eo-int-r2	CAAGCCGAGGGAAGAGAGTC			
foxa2-junc-f2	ATAGCATCCTTGGAGAAGGTCG	-	322 bp	For detection of products internal to Foxa2 enhancer deletion
foxa2-junc-r2	GAATTGCCAGGTCCCTCAAGA			
foxa2-int-f1	CCCTGTGATTTGGCCTGGAT	-	264 bp	
foxa2-int-r1	CCTTCCAGTCTGGCACCATT			
foxa2-int-f2	AATGGTGCCAGACTGGAAGG	-	159 bp	
foxa2-int-r2	ATCGCTGCCTGCTAGTTACC			
sox17-junc-f1	GTCAGTCAGCTCTGGTGCTT	-	394 bp	For detection of products internal to Sox17 enhancer deletion
sox17-junc-r1	GCGTCCCTAGCACTTACAGG			
sox17-int-f1	CCTGTAAGTGCTAGGGACGC	-	169 bp	
sox17-int-r1	ATTACGGCTGATAGAGCGGC			
sox17-int-f2	CGTGTTGTCAGCCGCTCTAT	-	238 bp	
sox17-int-r2	ACGCAGGTCCATTCATGTCA			
sox2_f5	agg aaa ggg ttc ttg ctg gg	-		For RT-qPCR of Cdx2
sox2_r5	gac cac gaa aac ggt ctt gc			
cdx2_f2	cca gct cac ttt tcc tcc tg	-		For RT-qPCR of Sox2
cdx2_r2	cag cag tcc cta gga agc c			

Table 6.3: Primers used for genotyping and RT-qPCR.

REFERENCES

Aberle, H., Bauer, A., Stappert, J., Kispert, A., & Kemler, R. (1997). ß-catenin is a target for the ubiquitin–proteasome pathway. *The EMBO Journal*, *16*(13), 3797–3804. https://doi.org/10.1093/emboj/16.13.3797

Abuhashem, A., Chivu, A. G., Zhao, Y., Rice, E. J., Siepel, A., Danko, C. G., & Hadjantonakis, A.-K. (2022). RNA Pol II pausing facilitates phased pluripotency transitions by buffering transcription. *Genes & Development*, *36*(13–14), 770–789. https://doi.org/10.1101/gad.349565.122

Abuhashem, A., Lee, A. S., Joyner, A. L., & Hadjantonakis, A.-K. (2022). Rapid and efficient degradation of endogenous proteins in vivo identifies stage-specific roles of RNA Pol II pausing in mammalian development. *Developmental Cell*, *57*(8), 1068-1080.e6. https://doi.org/10.1016/j.devcel.2022.03.013

Adachi, K., Nikaido, I., Ohta, H., Ohtsuka, S., Ura, H., Kadota, M., Wakayama, T., Ueda, H. R., & Niwa, H. (2013). Context-Dependent Wiring of Sox2 Regulatory Networks for Self-Renewal of Embryonic and Trophoblast Stem Cells. *Molecular Cell*, *52*(3), 380–392. https://doi.org/10.1016/j.molcel.2013.09.002

Adamson, S. L., Lu, Y., Whiteley, K. J., Holmyard, D., Hemberger, M., Pfarrer, C., & Cross, J. C. (2002). Interactions between Trophoblast Cells and the Maternal and Fetal Circulation in the Mouse Placenta. *Developmental Biology*, *250*(2), 358–373. https://doi.org/10.1006/dbio.2002.0773

Adjaye, J., Huntriss, J., Herwig, R., BenKahla, A., Brink, T. C., Wierling, C., Hultschig, C., Groth, D., Yaspo, M., Picton, H. M., Gosden, R. G., & Lehrach, H. (2005). Primary Differentiation in the Human Blastocyst: Comparative Molecular Portraits of Inner Cell Mass and Trophectoderm Cells. *STEM CELLS*, *23*(10), 1514–1525. https://doi.org/10.1634/stemcells.2005-0113

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., & Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, *44*(W1), W3–W10. https://doi.org/10.1093/nar/gkw343

Aguilera-Castrejon, A., Oldak, B., Shani, T., Ghanem, N., Itzkovich, C., Slomovich, S., Tarazi, S., Bayerl, J., Chugaeva, V., Ayyash, M., Ashouokhi, S., Sheban, D., Livnat, N., Lasman, L., Viukov, S., Zerbib, M., Addadi, Y., Rais, Y., Cheng, S., … Hanna, J. H. (2021). Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature*, *593*(7857), 119–124. https://doi.org/10.1038/s41586-021-03416-3

Albors, A. R., Halley, P. A., & Storey, K. G. (2018). Lineage tracing axial progenitors using Nkx1-2CreERT2 mice defines their trunk and tail contributions. *Development*. https://doi.org/10.1242/dev.164319

Alexander, J., & Stainier, D. Y. R. (1999). A molecular pathway leading to endoderm formation in zebrafish. *Current Biology*, *9*(20), 1147–1157. https://doi.org/10.1016/S0960-9822(00)80016-0

Alten, L., Schuster-Gossler, K., Beckers, A., Groos, S., Ulmer, B., Hegermann, J., Ochs, M., & Gossler, A. (2012). Differential regulation of node formation, nodal ciliogenesis and cilia positioning by *Noto* and *Foxj1*. *Development*, *139*(7), 1276–1284. https://doi.org/10.1242/dev.072728

Álvarez-Aznar, A., Martínez-Corral, I., Daubel, N., Betsholtz, C., Mäkinen, T., & Gaengel, K. (2020). Tamoxifen-independent recombination of reporter genes limits lineage tracing and mosaic analysis using CreERT2 lines. *Transgenic Research*, *29*(1), 53–68. https://doi.org/10.1007/s11248-019-00177-8

Anderson, R. M., Lawrence, A. R., Stottmann, R. W., Bachiller, D., & Klingensmith, J. (2002). Chordin and noggin promote organizing centers of forebrain development in the mouse. *Development*, *129*(21), 4975–4987. https://doi.org/10.1242/dev.129.21.4975

Anderson, W. J., Zhou, Q., Alcalde, V., Kaneko, O. F., Blank, L. J., Sherwood, R. I., Guseh, J. S., Rajagopal, J., & Melton, D. A. (2008). Genetic targeting of the endoderm with claudin-6CreER. *Developmental Dynamics*, *237*(2), 504–512. https://doi.org/10.1002/dvdy.21437

Andre, P., Song, H., Kim, W., Kispert, A., & Yang, Y. (2015). *Wnt5a* and *Wnt11* regulate mammalian anterior-posterior axis elongation. *Development*. https://doi.org/10.1242/dev.119065

Ang, S. L., Wierda, A., Wong, D., Stevens, K. A., Cascio, S., Rossant, J., & Zaret, K. S. (1993). The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. *Development*, *119*(4), 1301–1315. https://doi.org/10.1242/dev.119.4.1301

Anzalone, A. v., Koblan, L. W., & Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nature Biotechnology*, *38*(7), 824–844. https://doi.org/10.1038/s41587-020-0561-9

Aragón, E., Wang, Q., Zou, Y., Morgani, S. M., Ruiz, L., Kaczmarska, Z., Su, J., Torner, C., Tian, L., Hu, J., Shu, W., Agrawal, S., Gomes, T., Márquez, J. A., Hadjantonakis, A.-K., Macias, M. J., & Massagué, J. (2019). Structural basis for distinct roles of SMAD2 and SMAD3 in FOXH1 pioneer-directed TGF-β signaling. *Genes & Development*, *33*(21–22), 1506–1524. https://doi.org/10.1101/gad.330837.119

Argelaguet, R., Clark, S. J., Mohammed, H., Stapel, L. C., Krueger, C., Kapourani, C.-A., Imaz-Rosshandler, I., Lohoff, T., Xiang, Y., Hanna, C. W., Smallwood, S., Ibarra-Soria, X., Buettner, F., Sanguinetti, G., Xie, W., Krueger, F., Göttgens, B., Rugg-Gunn, P. J., Kelsey, G., … Reik, W. (2019). Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature*, *576*(7787), 487–491. https://doi.org/10.1038/s41586-019-1825-8

Argelaguet, R., Lohoff, T., Li, J. G., Nakhuda, A., Drage, D., Krueger, F., Velten, L., Clark, S. J., & Reik, W. (2022). Decoding gene regulation in the mouse embryo using single-cell multi-omits. *bioRxiv*, 2022.06.15.496239. https://doi.org/10.1101/2022.06.15.496239

Arkell, R. M., Fossat, N., & Tam, P. P. (2013). Wnt signalling in mouse gastrulation and anterior development: new players in the pathway and signal output. *Current Opinion in Genetics & Development*, *23*(4), 454–460. https://doi.org/10.1016/j.gde.2013.03.001

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J. K., & Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling

in pregastrulation mammalian development. *Proceedings of the National Academy of Sciences*, *95*(9), 5082–5087. https://doi.org/10.1073/pnas.95.9.5082

Arman, E., Haffner-Krausz, R., Gorivodsky, M., & Lonai, P. (1999). Fgfr2 is required for limb outgrowth and lung-branching morphogenesis. *Proceedings of the National Academy of Sciences*, *96*(21), 11895–11899. https://doi.org/10.1073/pnas.96.21.11895

Arnold, S. J., Hofmann, U. K., Bikoff, E. K., & Robertson, E. J. (2008). Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development*, *135*(3), 501–511. https://doi.org/10.1242/dev.014357

Arnold, S. J., Sugnaseelan, J., Groszer, M., Srinivas, S., & Robertson, E. J. (2009). Generation and analysis of a mouse line harboring GFP in the Eomes/Tbr2 locus. *Genesis*, *47*(11), 775–781. https://doi.org/10.1002/dvg.20562

Artus, J., Piliszek, A., & Hadjantonakis, A.-K. (2011). The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. *Developmental Biology*, *350*(2), 393–404. https://doi.org/10.1016/j.ydbio.2010.12.007

Bachiller, D., Klingensmith, J., Kemp, C., Belo, J. A., Anderson, R. M., May, S. R., McMahon, J. A., McMahon, A. P., Harland, R. M., Rossant, J., & de Robertis, E. M. (2000). The organizer factors Chordin and Noggin are required for mouse forebrain development. *Nature*, *403*(6770), 658–661. https://doi.org/10.1038/35001072

Bahar Halpern, K., Vana, T., & Walker, M. D. (2014). Paradoxical Role of DNA Methylation in Activation of FoxA2 Gene Expression during Endoderm Development. *Journal of Biological Chemistry*, *289*(34), 23882–23892. https://doi.org/10.1074/jbc.M114.573469

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C. J., Mclauchlan, H., Klevernic, I., Arthur, J. S. C., Alessi, D. R., & Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. *Biochemical Journal*, *408*(3), 297–315. https://doi.org/10.1042/BJ20070797

Bao, H., Liu, D., Xu, Y., Sun, Y., Mu, C., Yu, Y., Wang, C., Han, Q., Liu, S., Cai, H., Liu, F., Kong, S., Deng, W., Cao, B., Wang, H., Wang, Q., & Lu, J. (2020). Hyperactivated Wnt-ß-catenin signaling in the absence of sFRP1 and sFRP5 disrupts trophoblast differentiation through repression of Ascl2. *BMC Biology*, *18*(1), 151. https://doi.org/10.1186/s12915-020-00883-4

Bao, J., Ma, H.-Y., Schuster, A., Lin, Y.-M., & Yan, W. (2013). Incomplete cre-mediated excision leads to phenotypic differences between *Stra8-iCre; Mov10l1* $^{lox/lox}$ and *Stra8-iCre; Mov10l1* $^{lox/\Delta}$ mice. *Genesis*, *51*(7), 481–490. https://doi.org/10.1002/dvg.22389

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., & Surani, M. A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature*, *461*(7268), 1292–1295. https://doi.org/10.1038/nature08534

Bardot, E. S., & Hadjantonakis, A. K. (2020). Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. *Mechanisms of Development*, *163*. https://doi.org/10.1016/j.mod.2020.103617

Basilicata, M. F., Frank, M., Solter, D., Brabletz, T., & Stemmler, M. P. (2016). Inappropriate cadherin switching in the mouse epiblast compromises proper signaling between the

epiblast and the extraembryonic ectoderm during gastrulation. *Scientific Reports*, *6*(1), 26562. https://doi.org/10.1038/srep26562

Beddington, R. S. P., & Robertson, E. J. (1998). Anterior patterning in mouse. *Trends in Genetics*, *14*(7), 277–284. https://doi.org/10.1016/S0168-9525(98)01499-1

Bedzhov, I., Leung, C. Y., Bialecka, M., & Zernicka-Goetz, M. (2014). In vitro culture of mouse blastocysts beyond the implantation stages. *Nature Protocols*, *9*(12), 2732–2739. https://doi.org/10.1038/nprot.2014.186

Benchetrit, H., Herman, S., van Wietmarschen, N., Wu, T., Makedonski, K., Maoz, N., Yom Tov, N., Stave, D., Lasry, R., Zayat, V., Xiao, A., Lansdorp, P. M., Sebban, S., & Buganim, Y. (2015). Extensive Nuclear Reprogramming Underlies Lineage Conversion into Functional Trophoblast Stem-like Cells. *Cell Stem Cell*, *17*(5), 543–556. https://doi.org/10.1016/j.stem.2015.08.006

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E. J., & Constam, D. B. (2006). The Nodal Precursor Acting via Activin Receptors Induces Mesoderm by Maintaining a Source of Its Convertases and BMP4. *Developmental Cell*, *11*(3), 313–323. https://doi.org/10.1016/j.devcel.2006.07.005

Benz, B. A., Nandadasa, S., Takeuchi, M., Grady, R. C., Takeuchi, H., LoPilato, R. K., Kakuda, S., Somerville, R. P. T., Apte, S. S., Haltiwanger, R. S., & Holdener, B. C. (2016). Genetic and biochemical evidence that gastrulation defects in Pofut2 mutants result from defects in ADAMTS9 secretion. *Developmental Biology*, *416*(1), 111–122. https://doi.org/10.1016/j.ydbio.2016.05.038

Berneau, S. C., Ruane, P. T., Brison, D. R., Kimber, S. J., Westwood, M., & Aplin, J. D. (2019). Investigating the role of CD44 and hyaluronate in embryo-epithelial interaction using an in vitro model. *MHR: Basic Science of Reproductive Medicine*, *25*(5), 265–273. https://doi.org/10.1093/molehr/gaz011

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., Capecchi, M., & Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. *Developmental Biology*, *353*(2), 266–274. https://doi.org/10.1016/j.ydbio.2011.02.029

Betts, D. H., & Tobias, I. C. (2015). Canine Pluripotent Stem Cells: Are They Ready for Clinical Applications? *Frontiers in Veterinary Science*, *2*. https://doi.org/10.3389/fvets.2015.00041

Bibikova, M., Golic, M., Golic, K. G., & Carroll, D. (2002). Targeted Chromosomal Cleavage and Mutagenesis in Drosophila Using Zinc-Finger Nucleases. *Genetics*, *161*(3), 1169–1175. https://doi.org/10.1093/genetics/161.3.1169

Bleul, C. C., Farzan, M., Choe, H., Parolin, C., Clark-Lewis, I., Sodroski, J., & Springer, T. A. (1996). The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature*, *382*(6594), 829–833. https://doi.org/10.1038/382829a0

Blum, M., Gaunt, S. J., Cho, K. W. Y., Steinbeisser, H., Blumberg, B., Bittner, D., & de Robertis, E. M. (1992). Gastrulation in the mouse: The role of the homeobox gene goosecoid. *Cell*, *69*(7), 1097–1106. https://doi.org/10.1016/0092-8674(92)90632-M

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., & Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*, *326*(5959), 1509–1512. https://doi.org/10.1126/science.1178811

Bogutz, A. B., Oh-McGinnis, R., Jacob, K. J., Ho-Lau, R., Gu, T., Gertsenstein, M., Nagy, A., & Lefebvre, L. (2018). Transcription factor ASCL2 is required for development of the glycogen trophoblast cell lineage. *PLOS Genetics*, *14*(8), e1007587. https://doi.org/10.1371/journal.pgen.1007587

Boheler, K. R., Czyz, J., Tweedie, D., Yang, H.-T., Anisimov, S. v., & Wobus, A. M. (2002). Differentiation of Pluripotent Embryonic Stem Cells Into Cardiomyocytes. *Circulation Research*, *91*(3), 189–201. https://doi.org/10.1161/01.RES.0000027865.61704.32

Brickman, J. M., & Serup, P. (2017). Properties of embryoid bodies. *WIREs Developmental Biology*, *6*(2). https://doi.org/10.1002/wdev.259

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*(8), 1176–1185. https://doi.org/10.1002/stem.666

Brunst, K. J., Tignor, N., Just, A., Liu, Z., Lin, X., Hacker, M. R., Bosquet Enlow, M., Wright, R. O., Wang, P., Baccarelli, A. A., & Wright, R. J. (2018). Cumulative lifetime maternal stress and epigenome-wide placental DNA methylation in the PRISM cohort. *Epigenetics*, *13*(6), 665–681. https://doi.org/10.1080/15592294.2018.1497387

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*, *10*(12), 1213–1218. https://doi.org/10.1038/nmeth.2688

Cambuli, F., Murray, A., Dean, W., Dudzinska, D., Krueger, F., Andrews, S., Senner, C. E., Cook, S. J., & Hemberger, M. (2014). Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast. *Nature Communications*, *5*(1), 5538. https://doi.org/10.1038/ncomms6538

Caussinus, E., Kanca, O., & Affolter, M. (2012). Fluorescent fusion protein knockout mediated by anti-GFP nanobody. *Nature Structural & Molecular Biology*, *19*(1), 117–121. https://doi.org/10.1038/nsmb.2180

Cecconi, F., Proetzel, G., Alvarez-Bolado, G., Jay, D., & Gruss, P. (1997). Expression of Meis2, a Knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. *Developmental Dynamics*, *210*(2), 184–190. https://doi.org/10.1002/(SICI)1097-0177(199710)210:2<184::AID-AJA10>3.0.CO;2-E

Cernilogar, F. M., Hasenöder, S., Wang, Z., Scheibner, K., Burtscher, I., Sterr, M., Smialowski, P., Groh, S., Evenroed, I. M., Gilfillan, G. D., Lickert, H., & Schotta, G. (2019). Pre-marked chromatin and transcription factor co-binding shape the pioneering activity of Foxa2. *Nucleic Acids Research*, *47*(17), 9069–9086. https://doi.org/10.1093/nar/gkz627

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*(5), 643–655. https://doi.org/10.1016/S0092-8674(03)00392-1

Chan, M. M., Smith, Z. D., Grosswendt, S., Kretzmer, H., Norman, T. M., Adamson, B., Jost, M., Quinn, J. J., Yang, D., Jones, M. G., Khodaverdian, A., Yosef, N., Meissner, A., & Weissman, J. S. (2019). Molecular recording of mammalian embryogenesis. *Nature*, *570*(7759), 77–82. https://doi.org/10.1038/s41586-019-1184-5

Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J., & Beck, F. (2004). Cdx2 is essential for axial elongation in mouse development. *Proceedings of the National Academy of Sciences*, *101*(20), 7641–7645. https://doi.org/10.1073/pnas.0401654101

Chen, L., Yabuuchi, A., Eminli, S., Takeuchi, A., Lu, C.-W., Hochedlinger, K., & Daley, G. Q. (2009). Cross-regulation of the Nanog and Cdx2 promoters. *Cell Research*, *19*(9), 1052–1061. https://doi.org/10.1038/cr.2009.79

Chen, Y.-H., & Yu, J. (2012). Ectopic expression of Fgf3 leads to aberrant lineage segregation in the mouse parthenote preimplantation embryos. *Developmental Dynamics*, *241*(11), 1651–1664. https://doi.org/10.1002/dvdy.23851

Cheng, S., Pei, Y., He, L., Peng, G., Reinius, B., Tam, P. P. L., Jing, N., & Deng, Q. (2019). Single-Cell RNA-Seq Reveals Cellular Heterogeneity of Pluripotency Transition and X Chromosome Dynamics during Early Mouse Development. *Cell Reports*, *26*(10), 2593-2607.e3. https://doi.org/10.1016/j.celrep.2019.02.031

Cho, L. T. Y., Wamaitha, S. E., Tsai, I. J., Artus, J., Sherwood, R. I., Pedersen, R. A., Hadjantonakis, A.-K., & Niakan, K. K. (2012). Conversion from mouse embryonic to extra-embryonic endoderm stem cells reveals distinct differentiation capacities of pluripotent stem cell states. *Development*, *139*(16), 2866–2877. https://doi.org/10.1242/dev.078519

Choi, K.-H., Lee, D.-K., Kim, S. W., Woo, S.-H., Kim, D.-Y., & Lee, C.-K. (2019). Chemically Defined Media Can Maintain Pig Pluripotency Network In Vitro. *Stem Cell Reports*, *13*(1), 221–234. https://doi.org/10.1016/j.stemcr.2019.05.028

Chojnowski, J. L., Trau, H. A., Masuda, K., & Manley, N. R. (2016). Temporal and spatial requirements for Hoxa3 in mouse embryonic development. *Developmental Biology*, *415*(1), 33–45. https://doi.org/10.1016/j.ydbio.2016.05.010

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J., & Voytas, D. F. (2010). Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*, *186*(2), 757–761. https://doi.org/10.1534/genetics.110.120717

Chrysanthou, S., Senner, C. E., Woods, L., Fineberg, E., Okkenhaug, H., Burge, S., Perez-Garcia, V., & Hemberger, M. (2018). A Critical Role of TET1/2 Proteins in Cell-Cycle Progression of Trophoblast Stem Cells. *Stem Cell Reports*, *10*(4), 1355–1368. https://doi.org/10.1016/j.stemcr.2018.02.014

Ciemerych, M. A., & Sicinski, P. (2005). Cell cycle in mouse development. *Oncogene*, *24*(17), 2877–2898. https://doi.org/10.1038/sj.onc.1208608

Ciruna, B. G., & Rossant, J. (1999). Expression of the T-box gene Eomesodermin during early mouse development. *Mechanisms of Development*, *81*(1–2), 199–203. https://doi.org/10.1016/S0925-4773(98)00243-3

Ciruna, B., & Rossant, J. (2001). FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. *Developmental Cell*, *1*(1), 37–49. https://doi.org/10.1016/S1534-5807(01)00017-X

Clackson, T., Yang, W., Rozamus, L. W., Hatada, M., Amara, J. F., Rollins, C. T., Stevenson, L. F., Magari, S. R., Wood, S. A., Courage, N. L., Lu, X., Cerasoli, F., Gilman, M., & Holt, D. A. (1998). Redesigning an FKBP–ligand interface to generate chemical dimerizers with novel specificity. *Proceedings of the National Academy of Sciences*, *95*(18), 10437–10442. https://doi.org/10.1073/pnas.95.18.10437

Clark, S. J., Argelaguet, R., Lohoff, T., Krueger, F., Drage, D., Göttgens, B., Marioni, J. C., Nichols, J., & Reik, W. (2022). Single-cell multi-omics profiling links dynamic DNA methylation to cell fate decisions during mouse early organogenesis. *Genome Biology*, *23*(1), 202. https://doi.org/10.1186/s13059-022-02762-3

Clements, D., & Woodland, H. R. (2000). Changes in embryonic cell fate produced by expression of an endodermal transcription factor, Xsox17. *Mechanisms of Development*, *99*(1–2), 65–70. https://doi.org/10.1016/S0925-4773(00)00476-7

Cole, R. J., & Paul, J. (1965). Properties of Cultured Preimplantation Mouse and Rabbit embryos, and Cell Strains Derived from Them. In G. E. W. Wolstenholme & M. O'Connor (Eds.), *Ciba Foundation Symposium - Preimplantation Stages of Pregnancy* (pp. 82–122). Little, Brown, and Co. https://doi.org/10.1002/9780470719435.ch5

Collignon, J., Varlet, I., & Robertson, E. J. (1996). Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature*, *381*(6578), 155–158. https://doi.org/10.1038/381155a0

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., & Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development*, *120*(7), 1919–1928. https://doi.org/10.1242/dev.120.7.1919

Costantini, F., & Lacy, E. (1981). Introduction of a rabbit ß-globin gene into the mouse germ line. *Nature*, *294*(5836), 92–94. https://doi.org/10.1038/294092a0

Costello, I., Nowotschin, S., Sun, X., Mould, A. W., Hadjantonakis, A.-K., Bikoff, E. K., & Robertson, E. J. (2015). Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. *Genes & Development*, *29*(20), 2108–2122. https://doi.org/10.1101/gad.268979.115

Costello, I., Pimeisl, I.-M., Dräger, S., Bikoff, E. K., Robertson, E. J., & Arnold, S. J. (2011). The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. *Nature Cell Biology*, *13*(9), 1084–1091. https://doi.org/10.1038/ncb2304

Coucouvanis, E., & Martin, G. R. (1995). Signals for death and survival: A two-step mechanism for cavitation in the vertebrate embryo. *Cell*, *83*(2), 279–287. https://doi.org/10.1016/0092-8674(95)90169-8

Cunningham, T. J., Kumar, S., Yamaguchi, T. P., & Duester, G. (2015). *Wnt8a* and *Wnt3a* cooperate in the axial stem cell niche to promote mammalian body axis extension. *Developmental Dynamics*, *244*(6), 797–807. https://doi.org/10.1002/dvdy.24275

de Clercq, K., Pérez-García, V., van Bree, R., Pollastro, F., Peeraer, K., Voets, T., & Vriens, J. (2021). Mapping the expression of transient receptor potential channels across murine placental development. *Cellular and Molecular Life Sciences*, *78*(11), 4993–5014. https://doi.org/10.1007/s00018-021-03837-3

Dear, T. N., & Boehm, T. (1999). Diverse mRNA expression patterns of the mouse calpain genes Capn5, Capn6 and Capn11 during development. *Mechanisms of Development*, *89*(1–2), 201–209. https://doi.org/10.1016/S0925-4773(99)00214-2

Dewar, A. D. (1957). BODY WEIGHT CHANGES IN THE MOUSE DURING THE OESTROUS CYCLE AND PSEUDOPREGNANCY. *Journal of Endocrinology*, *15*(2), 230–233. https://doi.org/10.1677/joe.0.0150230

Dietrich, B., Haider, S., Meinhardt, G., Pollheimer, J., & Knöfler, M. (2022). WNT and NOTCH signaling in human trophoblast development and differentiation. *Cellular and Molecular Life Sciences*, *79*(6), 292. https://doi.org/10.1007/s00018-022-04285-3

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dong, Y.-L., Reddy, D. M., Green, K. E., Chauhan, M. S., Wang, H.-Q., Nagamani, M., Hankins, G. D. V., & Yallampalli, C. (2007). Calcitonin Gene-Related Peptide (CALCA) Is a Proangiogenic Growth Factor in the Human Placental Development1. *Biology of Reproduction*, *76*(5), 892–899. https://doi.org/10.1095/biolreprod.106.059089

Donnison, M., Beaton, A., Davey, H. W., Broadhurst, R., L'Huillier, P., & Pfeffer, P. L. (2005). Loss of the extraembryonic ectoderm in *Elf5* mutants leads to defects in embryonic patterning. *Development*, *132*(10), 2299–2308. https://doi.org/10.1242/dev.01819

Donnison, M., Broadhurst, R., & Pfeffer, P. L. (2015). Elf5 and Ets2 maintain the mouse extraembryonic ectoderm in a dosage dependent synergistic manner. *Developmental Biology*, *397*(1), 77–88. https://doi.org/10.1016/j.ydbio.2014.10.011

Du, Y., Xie, W., Zhang, F., & Liu, C. (2019). *Chimeric Mouse Generation by ES Cell Blastocyst Microinjection and Uterine Transfer* (pp. 99–114). https://doi.org/10.1007/978-1-4939-8831-0_6

Dufort, D., Schwartz, L., Harpal, K., & Rossant, J. (1998). The transcription factor HNF3beta is required in visceral endoderm for normal primitive streak morphogenesis. *Development*, *125*(16), 3015–3025. https://doi.org/10.1242/dev.125.16.3015

Duncan, W. C., Shaw, J. L. v., Burgess, S., McDonald, S. E., Critchley, H. O. D., & Horne, A. W. (2011). Ectopic Pregnancy as a Model to Identify Endometrial Genes and Signaling Pathways Important in Decidualization and Regulated by Local Trophoblast. *PLoS ONE*, *6*(8), e23595. https://doi.org/10.1371/journal.pone.0023595

Dunn, N. R., Koonce, C. H., Anderson, D. C., Islam, A., Bikoff, E. K., & Robertson, E. J. (2005). Mice exclusively expressing the short isoform of Smad2 develop normally and are viable and fertile. *Genes & Development*, *19*(1), 152–163. https://doi.org/10.1101/gad.1243205

Dunty, W. C., Biris, K. K., Chalamalasetty, R. B., Taketo, M. M., Lewandoski, M., & Yamaguchi, T. P. (2008). Wnt3a/ß-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. *Development*, *135*(1), 85–94. https://doi.org/10.1242/dev.009266

Engert, S., Burtscher, I., Liao, W. P., Dulev, S., Schotta, G., & Lickert, H. (2013). Wnt/ß-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. *Development*, *140*(15), 3128–3138. https://doi.org/10.1242/dev.088765

Enomoto, H., Nelson, C. M., Somerville, R. P. T., Mielke, K., Dixon, L. J., Powell, K., & Apte, S. S. (2010). Cooperation of two ADAMTS metalloproteases in closure of the mouse palate identifies a requirement for versican proteolysis in regulating palatal mesenchyme proliferation. *Development*, *137*(23), 4029–4038. https://doi.org/10.1242/dev.050591

Epinat, J.-C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., Patin, A., Zanghellini, A., Pâques, F., & Lacroix, E. (2003). A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Research*, *31*(11), 2952–2962. https://doi.org/10.1093/nar/gkg375

Erlebacher, A., Price, K. A., & Glimcher, L. H. (2004). Maintenance of mouse trophoblast stem cell proliferation by TGF-ß/activin. *Developmental Biology*, *275*(1), 158–169. https://doi.org/10.1016/j.ydbio.2004.07.032

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, *292*(5819), 154–156. https://doi.org/10.1038/292154a0

Fabik, J., Kovacova, K., Kozmik, Z., & Machon, O. (2020). Neural crest cells require Meis2 for patterning the mandibular arch via the Sonic hedgehog pathway. *Biology Open*, *9*(6). https://doi.org/10.1242/bio.052043

Fan, X., Rai, A., Kambham, N., Sung, J. F., Singh, N., Petitt, M., Dhal, S., Agrawal, R., Sutton, R. E., Druzin, M. L., Gambhir, S. S., Ambati, B. K., Cross, J. C., & Nayak, N. R. (2014). Endometrial VEGF induces placental sFLT1 and leads to pregnancy complications. *Journal of Clinical Investigation*, *124*(11), 4941–4952. https://doi.org/10.1172/JCI76864

Feldman, B., Poueymirou, W., Papaioannou, V. E., DeChiara, T. M., & Goldfarb, M. (1995). Requirement of FGF-4 for Postimplantation Mouse Development. *Science*, *267*(5195), 246–249. https://doi.org/10.1126/science.7809630

Felsher, D. W., & Bishop, J. M. (1999). Transient excess of MYC activity can elicit genomic instability and tumorigenesis. *Proceedings of the National Academy of Sciences*, *96*(7), 3940–3944. https://doi.org/10.1073/pnas.96.7.3940

Fendereski, M., Neupane, B., Nazneen, F., Bai, F., & Guo, Y.-L. (2022). Mouse Trophoblast Cells Can Provide IFN-Based Antiviral Protection to Embryonic Stem Cells via Paracrine Signaling. *The Journal of Immunology*, *208*(12), 2761–2770. https://doi.org/10.4049/jimmunol.2100679

Ferretti, E., Marshall, H., Popperl, H., Maconochie, M., Krumlauf, R., & Blasi, F. (2000). Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. *Development*, *127*(1), 155–166. https://doi.org/10.1242/dev.127.1.155

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S. R., & Chambers, I. (2012). Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells. *Cell Stem Cell*, *11*(4), 477–490. https://doi.org/10.1016/j.stem.2012.08.002

Fisher, J. B., Pulakanti, K., Rao, S., & Duncan, S. A. (2017). GATA6 is essential for endoderm formation from human pluripotent stem cells. *Biology Open*. https://doi.org/10.1242/bio.026120

Fu, S., Fei, Q., Jiang, H., Chuai, S., Shi, S., Xiong, W., Jiang, L., Lu, C., Atadja, P., Li, E., & Shou, J. (2011). Involvement of Histone Acetylation of Sox17 and Foxa2 Promoters during Mouse Definitive Endoderm Differentiation Revealed by MicroRNA Profiling. *PLoS ONE*, *6*(11), e27965. https://doi.org/10.1371/journal.pone.0027965

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J., & Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes & Development*, *16*(7), 784–789. https://doi.org/10.1101/gad.968802

Funa, N. S., Schachter, K. A., Lerdrup, M., Ekberg, J., Hess, K., Dietrich, N., Honoré, C., Hansen, K., & Semb, H. (2015). β-Catenin Regulates Primitive Streak Induction through Collaborative Interactions with SMAD2/SMAD3 and OCT4. *Cell Stem Cell*, *16*(6), 639–652. https://doi.org/10.1016/j.stem.2015.03.008

Futaki, S., Hayashi, Y., Emoto, T., Weber, C. N., & Sekiguchi, K. (2004). Sox7 Plays Crucial Roles in Parietal Endoderm Differentiation in F9 Embryonal Carcinoma Cells through Regulating Gata-4 and Gata-6 Expression. *Molecular and Cellular Biology*, *24*(23), 10492–10503. https://doi.org/10.1128/MCB.24.23.10492-10503.2004

Gadue, P., Huber, T. L., Paddison, P. J., & Keller, G. M. (2006). Wnt and TGF-β signaling are required for the induction of an *in vitro* model of primitive streak formation using embryonic stem cells. *Proceedings of the National Academy of Sciences*, *103*(45), 16806–16811. https://doi.org/10.1073/pnas.0603916103

Gardner, R. L., & Papaioannou, V. E. (1975). Differentiation in the trophectoderm and inner cell mass. In M. Balls & A. E. Wild (Eds.), *The Early Development of Mammals, 2nd Symposium of the British Society for Developmental Biology* (pp. 107–132). Cambridge University Press.

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, *109*(39). https://doi.org/10.1073/pnas.1208507109

Gaufo, G. O., Thomas, K. R., & Capecchi, M. R. (2003). Hox3 genes coordinate mechanisms of genetic suppression and activation in the generation of branchial and somatic motoneurons. *Development*, *130*(21), 5191–5201. https://doi.org/10.1242/dev.00730

Gaunt, S. J., Drage, D., & Trubshaw, R. C. (2005). cdx4/lacZ and cdx2/lacZ protein gradients formed by decay during gastrulation in the mouse. *The International Journal of Developmental Biology*, *49*(8), 901–908. https://doi.org/10.1387/ijdb.052021sg

Genbacev, O., Larocque, N., Ona, K., Prakobphol, A., Garrido-Gomez, T., Kapidzic, M., Bárcena, A., Gormley, M., & Fisher, S. J. (2016). Integrin α4-positive human trophoblast progenitors: functional characterization and transcriptional regulation. *Human Reproduction*, *31*(6), 1300–1314. https://doi.org/10.1093/humrep/dew077

George, E. L., Georges-Labouesse, E. N., Patel-King, R. S., Rayburn, H., & Hynes, R. O. (1993). Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development*, *119*(4), 1079–1091. https://doi.org/10.1242/dev.119.4.1079

Georges-Labouesse, E. N., George, E. L., Rayburn, H., & Hynes, R. O. (1996). Mesodermal development in mouse embryos mutant for fibronectin. *Developmental Dynamics*, *207*(2), 145–156. https://doi.org/10.1002/(SICI)1097-0177(199610)207:2<145::AID-AJA3>3.0.CO;2-H

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., & Ruddle, F. H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proceedings of the National Academy of Sciences*, *77*(12), 7380–7384. https://doi.org/10.1073/pnas.77.12.7380

Green, M. D., Chen, A., Nostro, M.-C., d'Souza, S. L., Schaniel, C., Lemischka, I. R., Gouon-Evans, V., Keller, G., & Snoeck, H.-W. (2011). Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nature Biotechnology*, *29*(3), 267–272. https://doi.org/10.1038/nbt.1788

Grosswendt, S., Kretzmer, H., Smith, Z. D., Kumar, A. S., Hetzel, S., Wittler, L., Klages, S., Timmermann, B., Mukherji, S., & Meissner, A. (2020). Epigenetic regulator function through mouse gastrulation. *Nature*, *584*(7819), 102–108. https://doi.org/10.1038/s41586-020-2552-x

Guibentif, C., Griffiths, J. A., Imaz-Rosshandler, I., Ghazanfar, S., Nichols, J., Wilson, V., Göttgens, B., & Marioni, J. C. (2021). Diverse Routes toward Early Somites in the Mouse Embryo. *Developmental Cell*, *56*(1), 141-153.e6. https://doi.org/10.1016/j.devcel.2020.11.013

Guo, F., Li, L., Li, J., Wu, X., Hu, B., Zhu, P., Wen, L., & Tang, F. (2017). Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Research*, *27*(8), 967–988. https://doi.org/10.1038/cr.2017.82

Hancock, S. N., Agulnik, S. I., Silver, L. M., & Papaioannou, V. E. (1999). Mapping and expression analysis of the mouse ortholog of Xenopus Eomesodermin. *Mechanisms of Development*, *81*(1–2), 205–208. https://doi.org/10.1016/S0925-4773(98)00244-5

Haque, A., Engel, J., Teichmann, S. A., & Lönnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine*, *9*(1), 75. https://doi.org/10.1186/s13073-017-0467-4

Harland, L. T. G. (2020). *Eomesodermin governs the hemogenic competence of murine mesodermal progenitors*. University of Oxford.

Harland, L. T. G., Simon, C. S., Senft, A. D., Costello, I., Greder, L., Imaz-Rosshandler, I., Göttgens, B., Marioni, J. C., Bikoff, E. K., Porcher, C., de Bruijn, M. F. T. R., & Robertson, E. J. (2021). The T-box transcription factor Eomesodermin governs haemogenic competence of yolk sac mesodermal progenitors. *Nature Cell Biology*, *23*(1), 61–74. https://doi.org/10.1038/s41556-020-00611-8

Hart, A. H., Hartley, L., Sourris, K., Stadler, E. S., Li, R., Stanley, E. G., Tam, P. P. L., Elefanty, A. G., & Robb, L. (2002). Mixl1 is required for axial mesendoderm morphogenesis and patterning in the murine embryo. *Development*, *129*(15), 3597–3608. https://doi.org/10.1242/dev.129.15.3597

Hasegawa, A., MOCHIDA, K., OGONUKI, N., HIROSE, M., TOMISHIMA, T., INOUE, K., & OGURA, A. (2017). Efficient and scheduled production of pseudopregnant female mice for embryo transfer by estrous cycle synchronization. *Journal of Reproduction and Development*, *63*(6), 539–545. https://doi.org/10.1262/jrd.2017-068

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., & Saitou, M. (2011). Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. *Cell*, *146*(4), 519–532. https://doi.org/10.1016/j.cell.2011.06.052

Hayashi, S., Lewis, P., Pevny, L., & McMahon, A. P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Mechanisms of Development*, *119*, S97–S101. https://doi.org/10.1016/S0925-4773(03)00099-6

He, P., Williams, B. A., Trout, D., Marinov, G. K., Amrhein, H., Berghella, L., Goh, S.-T., Plajzer-Frick, I., Afzal, V., Pennacchio, L. A., Dickel, D. E., Visel, A., Ren, B., Hardison, R. C., Zhang, Y., & Wold, B. J. (2020). The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature*, *583*(7818), 760–767. https://doi.org/10.1038/s41586-020-2536-x

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., & Glass, C. K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell*, *38*(4), 576–589. https://doi.org/10.1016/j.molcel.2010.05.004

Hemberger, M., Hanna, C. W., & Dean, W. (2020). Mechanisms of early placental development in mouse and humans. *Nature Reviews Genetics*, *21*(1), 27–43. https://doi.org/10.1038/s41576-019-0169-4

Hemberger, M., Nozaki, T., Masutani, M., & Cross, J. C. (2003). Differential expression of angiogenic and vasodilatory factors by invasive trophoblast giant cells depending on depth of invasion. *Developmental Dynamics*, *227*(2), 185–191. https://doi.org/10.1002/dvdy.10291

Hemberger, M., Udayashankar, R., Tesar, P., Moore, H., & Burton, G. J. (2010). ELF5-enforced transcriptional networks define an epigenetically regulated trophoblast stem cell compartment in the human placenta. *Human Molecular Genetics*, *19*(12), 2456–2467. https://doi.org/10.1093/hmg/ddq128

Heurtier, V., Owens, N., Gonzalez, I., Mueller, F., Proux, C., Mornico, D., Clerc, P., Dubois, A., & Navarro, P. (2019). The molecular logic of Nanog-induced self-renewal in mouse embryonic stem cells. *Nature Communications*, *10*(1), 1109. https://doi.org/10.1038/s41467-019-09041-z

Hogan, B. L. M., Cooper, A. R., & Kurkinen, M. (1980). Incorporation into Reichert's membrane of laminin-like extracellular proteins synthesized by parietal endoderm cells of the mouse embryo. *Developmental Biology*, *80*(2), 289–300. https://doi.org/10.1016/0012-1606(80)90405-4

Holland, A. J., Fachinetti, D., Han, J. S., & Cleveland, D. W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proceedings of the National Academy of Sciences*, *109*(49). https://doi.org/10.1073/pnas.1216880109

Home, P., Saha, B., Ray, S., Dutta, D., Gunewardena, S., Yoo, B., Pal, A., Vivian, J. L., Larson, M., Petroff, M., Gallagher, P. G., Schulz, V. P., White, K. L., Golos, T. G., Behr, B., & Paul, S. (2012). Altered subcellular localization of transcription factor TEAD4 regulates first mammalian cell lineage commitment. *Proceedings of the National Academy of Sciences*, *109*(19), 7362–7367. https://doi.org/10.1073/pnas.1201595109

Huang, Y., Osorno, R., Tsakiridis, A., & Wilson, V. (2012). In Vivo Differentiation Potential of Epiblast Stem Cells Revealed by Chimeric Embryo Formation. *Cell Reports*, *2*(6), 1571–1578. https://doi.org/10.1016/j.celrep.2012.10.022

Hudson, C., Clements, D., Friday, R. v, Stott, D., & Woodland, H. R. (1997). Xsox17α and -ß Mediate Endoderm Formation in Xenopus. *Cell*, *91*(3), 397–405. https://doi.org/10.1016/S0092-8674(00)80423-7

Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C., & Birchmeier, W. (2000). Requirement for ß-Catenin in Anterior-Posterior Axis Formation in Mice. *Journal of Cell Biology*, *148*(3), 567–578. https://doi.org/10.1083/jcb.148.3.567

Ibarra-Soria, X., Jawaid, W., Pijuan-Sala, B., Ladopoulos, V., Scialdone, A., Jörg, D. J., Tyser, R. C. V., Calero-Nieto, F. J., Mulas, C., Nichols, J., Vallier, L., Srinivas, S., Simons, B. D., Göttgens, B., & Marioni, J. C. (2018). Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation. *Nature Cell Biology*, *20*(2), 127–134. https://doi.org/10.1038/s41556-017-0013-z

Ichikawa, T., Zhang, H. T., Panavaite, L., Erzberger, A., Fabrèges, D., Snajder, R., Wolny, A., Korotkevich, E., Tsuchida-Straeten, N., Hufnagel, L., Kreshuk, A., & Hiiragi, T. (2022). An ex vivo system to study cellular dynamics underlying mouse peri-implantation development. *Developmental Cell*, *57*(3), 373-386.e9. https://doi.org/10.1016/j.devcel.2021.12.023

Istaces, N., Splittgerber, M., Lima Silva, V., Nguyen, M., Thomas, S., Le, A., Achouri, Y., Calonne, E., Defrance, M., Fuks, F., Goriely, S., & Azouz, A. (2019). EOMES interacts with RUNX3 and BRG1 to promote innate memory cell formation through epigenetic reprogramming. *Nature Communications*, *10*(1), 3306. https://doi.org/10.1038/s41467-019-11233-6

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., & Benvenisty, N. (2000). Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Molecular Medicine (Cambridge, Mass.)*, *6*(2), 88–95.

Izumi, N., Era, T., Akimaru, H., Yasunaga, M., & Nishikawa, S. (2007). Dissecting the Molecular Hierarchy for Mesendoderm Differentiation Through a Combination of Embryonic Stem Cell Culture and RNA Interference. *STEM CELLS*, *25*(7), 1664–1674. https://doi.org/10.1634/stemcells.2006-0681

Jaber, M., Radwan, A., Loyfer, N., Abdeen, M., Sebban, S., Khatib, A., Yassen, H., Kolb, T., Zapatka, M., Makedonski, K., Ernst, A., Kaplan, T., & Buganim, Y. (2022). Comparative parallel multi-omics analysis during the induction of pluripotent and trophectoderm states. *Nature Communications*, *13*(1), 3475. https://doi.org/10.1038/s41467-022-31131-8

Jiang, S., Huang, Z., Li, Y., Yu, C., Yu, H., Ke, Y., Jiang, L., & Liu, J. (2023). Single-cell chromatin accessibility and transcriptome atlas of mouse embryos. *Cell Reports*, *42*(3), 112210. https://doi.org/10.1016/j.celrep.2023.112210

Kalkan, T., & Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1657), 20130540. https://doi.org/10.1098/rstb.2013.0540

Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P. L., & Hayashi, Y. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development*, *129*(10), 2367–2379. https://doi.org/10.1242/dev.129.10.2367

Kern, C. B., Wessels, A., McGarity, J., Dixon, L. J., Alston, E., Argraves, W. S., Geeting, D., Nelson, C. M., Menick, D. R., & Apte, S. S. (2010). Reduced versican cleavage due to Adamts9 haploinsufficiency is associated with cardiac and aortic anomalies. *Matrix Biology*, *29*(4), 304–316. https://doi.org/10.1016/j.matbio.2010.01.005

Khalil, A. M. (2020). The genome editing revolution: review. *Journal of Genetic Engineering and Biotechnology*, *18*(1), 68. https://doi.org/10.1186/s43141-020-00078-y

Kidder, B. L., & Palmer, S. (2010). Examination of transcriptional networks reveals an important role for TCFAP2C, SMARCA4, and EOMES in trophoblast stem cell maintenance. *Genome Research*. https://doi.org/10.1101/gr.101469.109

Kim, H.-S., Neugebauer, J., McKnite, A., Tilak, A., & Christian, J. L. (2019). BMP7 functions predominantly as a heterodimer with BMP2 or BMP4 during mammalian embryogenesis. *ELife*, *8*. https://doi.org/10.7554/eLife.48872

Kim, S. W., Yoon, S.-J., Chuong, E., Oyolu, C., Wills, A. E., Gupta, R., & Baker, J. (2011). Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. *Developmental Biology*, *357*(2), 492–504. https://doi.org/10.1016/j.ydbio.2011.06.009

Kim, Y. S., Fan, R., Kremer, L., Kuempel-Rink, N., Mildner, K., Zeuschner, D., Hekking, L., Stehling, M., & Bedzhov, I. (2021). Deciphering epiblast lumenogenesis reveals proamniotic cavity control of embryo growth and patterning. *Science Advances*, *7*(11). https://doi.org/10.1126/sciadv.abe1640

Kimura-Yoshida, C., Tian, E., Nakano, H., Amazaki, S., Shimokawa, K., Rossant, J., Aizawa, S., & Matsuo, I. (2007). Crucial roles of Foxa2 in mouse anterior–posterior axis polarization via regulation of anterior visceral endoderm-specific genes. *Proceedings of the National Academy of Sciences*, *104*(14), 5919–5924. https://doi.org/10.1073/pnas.0607779104

Kitajima, S., Takagi, A., Inoue, T., & Saga, Y. (2000). MesP1 and MesP2 are essential for the development of cardiac mesoderm. *Development*, *127*(15), 3215–3226. https://doi.org/10.1242/dev.127.15.3215

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. *Cellular and Molecular Life Sciences*, *76*(18), 3479–3496. https://doi.org/10.1007/s00018-019-03104-6

Knöfler, M., & Pollheimer, J. (2013). Human placental trophoblast invasion and differentiation: a particular focus on Wnt signaling. *Frontiers in Genetics*, *4*. https://doi.org/10.3389/fgene.2013.00190

Kojima, Y., Kaufman-Francis, K., Studdert, J. B., Steiner, K. A., Power, M. D., Loebel, D. A. F., Jones, V., Hor, A., de Alencastro, G., Logan, G. J., Teber, E. T., Tam, O. H., Stutz, M. D., Alexander, I. E., Pickett, H. A., & Tam, P. P. L. (2014). The Transcriptional and Functional Properties of Mouse Epiblast Stem Cells Resemble the Anterior Primitive Streak. *Cell Stem Cell*, *14*(1), 107–120. https://doi.org/10.1016/j.stem.2013.09.014

Krendl, C., Shaposhnikov, D., Rishko, V., Ori, C., Ziegenhain, C., Sass, S., Simon, L., Müller, N. S., Straub, T., Brooks, K. E., Chavez, S. L., Enard, W., Theis, F. J., & Drukker, M. (2017). GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency. *Proceedings of the National Academy of Sciences*, *114*(45). https://doi.org/10.1073/pnas.1708341114

Krueger, C., Peat, J. R., Eckersley-Maslin, M. A., Hore, T. A., Mohammed, H., Andrews, S. R., Dean, W., & Reik, W. (2017). *A non-catalytic role of TET3 promotes open chromatin and enhances global transcription*.

Kuales, G., Weiss, M., Sedelmeier, O., Pfeifer, D., & Arnold, S. J. (2015). A Resource for the Transcriptional Signature of Bona Fide Trophoblast Stem Cells and Analysis of Their Embryonic Persistence. *Stem Cells International*, *2015*, 1–13. https://doi.org/10.1155/2015/218518

Kuang, H., Chen, Q., Fan, X., Zhang, Y., Zhang, L., Peng, H., Cao, Y., & Duan, E. (2009a). CXCL14 inhibits trophoblast outgrowth via a paracrine/autocrine manner during early pregnancy in mice. *Journal of Cellular Physiology*, *221*(2), 448–457. https://doi.org/10.1002/jcp.21877

Kuang, H., Chen, Q., Zhang, Y., Zhang, L., Peng, H., Ning, L., Cao, Y., & Duan, E. (2009b). The Cytokine Gene CXCL14 Restricts Human Trophoblast Cell Invasion by Suppressing Gelatinase Activity. *Endocrinology*, *150*(12), 5596–5605. https://doi.org/10.1210/en.2009-0570

Kubaczka, C., Senner, C., Araúzo-Bravo, M. J., Sharma, N., Kuckenberg, P., Becker, A., Zimmer, A., Brüstle, O., Peitz, M., Hemberger, M., & Schorle, H. (2014). Derivation and Maintenance of Murine Trophoblast Stem Cells under Defined Conditions. *Stem Cell Reports*, *2*(2), 232–242. https://doi.org/10.1016/j.stemcr.2013.12.013

Kubaczka, C., Senner, C. E., Cierlitza, M., Araúzo-Bravo, M. J., Kuckenberg, P., Peitz, M., Hemberger, M., & Schorle, H. (2015). Direct Induction of Trophoblast Stem Cells from Murine Fibroblasts. *Cell Stem Cell*. https://doi.org/10.1016/j.stem.2015.08.005

Kunath, T., Arnaud, D., Uy, G. D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R. L., Avner, P., & Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development*, *132*(7), 1649–1661. https://doi.org/10.1242/dev.01715

Kunath, T., Saba-El-Leil, M. K., Almousailleakh, 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*(16), 2895–2902. https://doi.org/10.1242/dev.02880

Kuo, C. T., Morrisey, E. E., Anandappa, R., Sigrist, K., Lu, M. M., Parmacek, M. S., Soudais, C., & Leiden, J. M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes & Development*, *11*(8), 1048–1060. https://doi.org/10.1101/gad.11.8.1048

Kurowski, A., Molotkov, A., & Soriano, P. (2019). FGFR1 regulates trophectoderm development and facilitates blastocyst implantation. *Developmental Biology*, *446*(1), 94–101. https://doi.org/10.1016/j.ydbio.2018.12.008

Kwon, J.-Y., Aldo, P., You, Y., Ding, J., Racicot, K., Dong, X., Murphy, J., Glukshtad, G., Silasi, M., Peng, J., Wen, L., Abrahams, V. M., Romero, R., & Mor, G. (2018). Relevance of placental type I interferon beta regulation for pregnancy success. *Cellular & Molecular Immunology*, *15*(12), 1010–1026. https://doi.org/10.1038/s41423-018-0050-y

Kwon, G. S., Viotti, M., & Hadjantonakis, A.-K. (2008). The Endoderm of the Mouse Embryo Arises by Dynamic Widespread Intercalation of Embryonic and Extraembryonic Lineages. *Developmental Cell*, *15*(4), 509–520. https://doi.org/10.1016/j.devcel.2008.07.017

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Latos, P. A., Goncalves, A., Oxley, D., Mohammed, H., Turro, E., & Hemberger, M. (2015a). Fgf and Esrrb integrate epigenetic and transcriptional networks that regulate self-renewal of trophoblast stem cells. *Nature Communications*. https://doi.org/10.1038/ncomms8776

Latos, P. A., Sienerth, A. R., Murray, A., Senner, C. E., Muto, M., Ikawa, M., Oxley, D., Burge, S., Cox, B. J., & Hemberger, M. (2015b). Elf5-centered transcription factor hub controls trophoblast stem cell self-renewal and differentiation through stoichiometry sensitive shifts in target gene networks. *Genes and Development*. https://doi.org/10.1101/gad.268821.115

Lawson, K. A., Dunn, N. R., Roelen, B. A. J., Zeinstra, L. M., Davis, A. M., Wright, C. V. E., Korving, J. P. W. F. M., & Hogan, B. L. M. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes & Development*, *13*(4), 424–436. https://doi.org/10.1101/gad.13.4.424

Lawson, K. A., & Pedersen, R. A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development*, *101*(3), 627–652. https://doi.org/10.1242/dev.101.3.627

Lee, B. K., Jang, Y., Kim, M., LeBlanc, L., Rhee, C., Lee, J., Beck, S., Shen, W., & Kim, J. (2019). Super-enhancer-guided mapping of regulatory networks controlling mouse trophoblast stem cells. *Nature Communications*. https://doi.org/10.1038/s41467-019-12720-6

Leonhard, W. N., Roelfsema, J. H., Lantinga-van Leeuwen, I. S., Breuning, M. H., & Peters, D. J. (2008). Quantification of Cre-mediated recombination by a novel strategy reveals a stable extra-chromosomal deletion-circle in mice. *BMC Biotechnology*, *8*(1), 18. https://doi.org/10.1186/1472-6750-8-18

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S., Göttgens, B., & Blanpain, C. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science*, *359*(6380), 1177–1181. https://doi.org/10.1126/science.aao4174

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 (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, F., He, Z., Li, Y., Liu, P., Chen, F., Wang, M., Zhu, H., Ding, X., Wangensteen, K. J., Hu, Y., & Wang, X. (2011). Combined activin A/LiCl/Noggin treatment improves production of mouse embryonic stem cell-derived definitive endoderm cells. *Journal of Cellular Biochemistry*, *112*(4), 1022–1034. https://doi.org/10.1002/jcb.22962

Liaw, A., Cunha, G. R., Shen, J., Cao, M., Liu, G., Sinclair, A., & Baskin, L. (2018). Development of the human bladder and ureterovesical junction. *Differentiation*, *103*, 66–73. https://doi.org/10.1016/j.diff.2018.08.004

Lim, S. M., Pereira, L., Wong, M. S., Hirst, C. E., van Vranken, B. E., Pick, M., Trounson, A., Elefanty, A. G., & Stanley, E. G. (2009). Enforced Expression of Mixl1 During Mouse ES Cell Differentiation Suppresses Hematopoietic Mesoderm and Promotes Endoderm Formation. *Stem Cells*, *27*(2), 363–374. https://doi.org/10.1634/stemcells.2008-1008

Lim, R., Tran, H. T., Liong, S., Barker, G., & Lappas, M. (2016). The Transcription Factor Interferon Regulatory Factor-1 (IRF1) Plays a Key Role in the Terminal Effector Pathways of Human Preterm Labor1. *Biology of Reproduction*, *94*(2). https://doi.org/10.1095/biolreprod.115.134726

Lin, J., Khan, M., Zapiec, B., & Mombaerts, P. (2016). Efficient derivation of extraembryonic endoderm stem cell lines from mouse postimplantation embryos. *Scientific Reports*, *6*(1), 39457. https://doi.org/10.1038/srep39457

Lin, Y., Xu, L., Jin, H., Zhong, Y., Di, J., & Lin, Q. (2009). CXCL12 enhances exogenous CD4+CD25+ T cell migration and prevents embryo loss in non-obese diabetic mice. *Fertility and Sterility*, *91*(6), 2687–2696. https://doi.org/10.1016/j.fertnstert.2008.01.109

Litvinov, S. v, Velders, M. P., Bakker, H. A., Fleuren, G. J., & Warnaar, S. O. (1994). Ep-CAM: a human epithelial antigen is a homophilic cell-cell adhesion molecule. *Journal of Cell Biology*, *125*(2), 437–446. https://doi.org/10.1083/jcb.125.2.437

Liu, L., Liu, X., Ren, X., Tian, Y., Chen, Z., Xu, X., Du, Y., Jiang, C., Fang, Y., Liu, Z., Fan, B., Zhang, Q., Jin, G., Yang, X., & Zhang, X. (2016). Smad2 and Smad3 have differential sensitivity in relaying TGFβ signaling and inversely regulate early lineage specification. *Scientific Reports*, *6*(1), 21602. https://doi.org/10.1038/srep21602

Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R., & Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nature Genetics*, *22*(4), 361–365. https://doi.org/10.1038/11932

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lucas, E. S., Vrljicak, P., Muter, J., Diniz-da-Costa, M. M., Brighton, P. J., Kong, C.-S., Lipecki, J., Fishwick, K. J., Odendaal, J., Ewington, L. J., Quenby, S., Ott, S., & Brosens, J. J. (2020). Recurrent pregnancy loss is associated with a pro-senescent decidual response during the peri-implantation window. *Communications Biology*, *3*(1), 37. https://doi.org/10.1038/s42003-020-0763-1

Luo, N., Cheng, W., Zhou, Y., Gu, B., Zhao, Z., & Zhao, Y. (2021). Screening Candidate Genes Regulating Placental Development from Trophoblast Transcriptome at Early Pregnancy in Dazu Black Goats (Capra hircus). *Animals*, *11*(7), 2132. https://doi.org/10.3390/ani11072132

Marsh, B., & Blelloch, R. (2020). Single nuclei RNA-seq of mouse placental labyrinth development. *ELife*, *9*. https://doi.org/10.7554/eLife.60266

Marsh, B., Zhou, Y., Kapidzic, M., Fisher, S., & Blelloch, R. (2022). Regionally distinct trophoblast regulate barrier function and invasion in the human placenta. *ELife*, *11*. https://doi.org/10.7554/eLife.78829

Martin, G. R., & Evans, M. J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proceedings of the National Academy of Sciences*, *72*(4), 1441–1445. https://doi.org/10.1073/pnas.72.4.1441

Martyn, I., Siggia, E. D., & Brivanlou, A. H. (2019). Mapping cell migrations and fates in a gastruloid model to the human primitive streak. *Development*. https://doi.org/10.1242/dev.179564

Matoba, S., Nakamuta, S., Miura, K., Hirose, M., Shiura, H., Kohda, T., Nakamuta, N., & Ogura, A. (2019). Paternal knockout of Slc38a4/SNAT4 causes placental hypoplasia associated with intrauterine growth restriction in mice. *Proceedings of the National Academy of Sciences*, *116*(42), 21047–21053. https://doi.org/10.1073/pnas.1907884116

McDole, K., Guignard, L., Amat, F., Berger, A., Malandain, G., Royer, L. A., Turaga, S. C., Branson, K., & Keller, P. J. (2018). In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell*, *175*(3), 859-876.e33. https://doi.org/10.1016/j.cell.2018.09.031

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., & Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*(5), 495–501. https://doi.org/10.1038/nbt.1630

Meno, C., Shimono, A., Saijoh, Y., Yashiro, K., Mochida, K., Ohishi, S., Noji, S., Kondoh, H., & Hamada, H. (1998). lefty-1 Is Required for Left-Right Determination as a Regulator of lefty-2 and nodal. *Cell*, *94*(3), 287–297. https://doi.org/10.1016/S0092-8674(00)81472-5

Merrill, B. J. (2012). Wnt Pathway Regulation of Embryonic Stem Cell Self-Renewal. *Cold Spring Harbor Perspectives in Biology*, *4*(9), a007971–a007971. https://doi.org/10.1101/cshperspect.a007971

Mesnard, D., Guzman-Ayala, M., & Constam, D. B. (2006). Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. *Development*, *133*(13), 2497–2505. https://doi.org/10.1242/dev.02413

Mfopou, J. K., Geeraerts, M., Dejene, R., van Langenhoven, S., Aberkane, A., van Grunsven, L. A., & Bouwens, L. (2014). Efficient definitive endoderm induction from mouse embryonic stem cell adherent cultures: A rapid screening model for differentiation studies. *Stem Cell Research*, *12*(1), 166–177. https://doi.org/10.1016/j.scr.2013.10.004

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., & Yamanaka, S. (2003). The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. *Cell*, *113*(5), 631–642. https://doi.org/10.1016/S0092-8674(03)00393-3

Mittnenzweig, M., Mayshar, Y., Cheng, S., Ben-Yair, R., Hadas, R., Rais, Y., Chomsky, E., Reines, N., Uzonyi, A., Lumerman, L., Lifshitz, A., Mukamel, Z., Orenbuch, A.-H., Tanay, A., & Stelzer, Y. (2021). A single-embryo, single-cell time-resolved model for mouse gastrulation. *Cell*, *184*(11), 2825-2842.e22. https://doi.org/10.1016/j.cell.2021.04.004

Miura, S., & Mishina, Y. (2007). The DVE changes distal epiblast fate from definitive endoderm to neurectoderm by antagonizing nodal signaling. *Developmental Dynamics*, *236*(6), 1602–1610. https://doi.org/10.1002/dvdy.21166

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. *Nature Protocols*, *11*(2), 316–326. https://doi.org/10.1038/nprot.2016.020

Molkentin, J. D., Lin, Q., Duncan, S. A., & Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes & Development*, *11*(8), 1061–1072. https://doi.org/10.1101/gad.11.8.1061

Morawska, M., & Ulrich, H. D. (2013). An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast*, *30*(9), 341–351. https://doi.org/10.1002/yea.2967

Morrisey, E. E., Tang, Z., Sigrist, K., Lu, M. M., Jiang, F., Ip, H. S., & Parmacek, M. S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes & Development*, *12*(22), 3579–3590. https://doi.org/10.1101/gad.12.22.3579

Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., & Polakis, P. (1995). Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proceedings of the National Academy of Sciences*, *92*(7), 3046–3050. https://doi.org/10.1073/pnas.92.7.3046

Murakami, A., Shen, H., Ishida, S., & Dickson, C. (2004). SOX7 and GATA-4 Are Competitive Activators of Fgf-3 Transcription. *Journal of Biological Chemistry*, *279*(27), 28564–28573. https://doi.org/10.1074/jbc.M313814200

Murohashi, M., Nakamura, T., Tanaka, S., Ichise, T., Yoshida, N., Yamamoto, T., Shibuya, M., Schlessinger, J., & Gotoh, N. (2010). An FGF4-FRS2α-Cdx2 Axis in Trophoblast Stem Cells Induces Bmp4 to Regulate Proper Growth of Early Mouse Embryos . *Stem Cells*, *28*(1), 113–121. https://doi.org/10.1002/stem.247

Murray, A., Sienerth, A. R., & Hemberger, M. (2016). Plet1 is an epigenetically regulated cell surface protein that provides essential cues to direct trophoblast stem cell differentiation. *Scientific Reports*, *6*(1), 25112. https://doi.org/10.1038/srep25112

Murray, J. T., Campbell, D. G., Morrice, N., Auld, G. C., Shpiro, N., Marquez, R., Peggie, M., Bain, J., Bloomberg, G. B., Grahammer, F., Lang, F., Wulff, P., Kuhl, D., & Cohen, P. (2004). Exploitation of KESTREL to identify NDRG family members as physiological substrates for SGK1 and GSK3. *Biochemical Journal*, *384*(3), 477–488. https://doi.org/10.1042/BJ20041057

Nabet, B., Ferguson, F. M., Seong, B. K. A., Kuljanin, M., Leggett, A. L., Mohardt, M. L., Robichaud, A., Conway, A. S., Buckley, D. L., Mancias, J. D., Bradner, J. E., Stegmaier, K., & Gray, N. S. (2020). Rapid and direct control of target protein levels with VHL-

recruiting dTAG molecules. *Nature Communications*, *11*(1), 4687. https://doi.org/10.1038/s41467-020-18377-w

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. *Nature Chemical Biology*, *14*(5), 431–441. https://doi.org/10.1038/s41589-018-0021-8

Nagy, B., Berkes, E., Rigó, B., Bán, Z., Papp, Z., & Hupuczi, P. (2008). Under-Expression of CD24 in Pre-Eclamptic Placental Tissues Determined by Quantitative Real-Time RT-PCR. *Fetal Diagnosis and Therapy*, *23*(4), 263–266. https://doi.org/10.1159/000123612

Ng, R. K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., & Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nature Cell Biology*, *10*(11), 1280–1290. https://doi.org/10.1038/ncb1786

Niakan, K. K., Ji, H., Maehr, R., Vokes, S. A., Rodolfa, K. T., Sherwood, R. I., Yamaki, M., Dimos, J. T., Chen, A. E., Melton, D. A., McMahon, A. P., & Eggan, K. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes & Development*, *24*(3), 312–326. https://doi.org/10.1101/gad.1833510

Niakan, K. K., Schrode, N., Cho, L. T. Y., & Hadjantonakis, A.-K. (2013). Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. *Nature Protocols*, *8*(6), 1028–1041. https://doi.org/10.1038/nprot.2013.049

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. *Nature Methods*, *6*(12), 917–922. https://doi.org/10.1038/nmeth.1401

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R. O., Ogonuki, N., Makita, R., Kurihara, H., Morin-Kensicki, E. M., Nojima, H., Rossant, J., Nakao, K., Niwa, H., & Sasaki, H. (2009). The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. *Developmental Cell*, *16*(3), 398–410. https://doi.org/10.1016/j.devcel.2009.02.003

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., & Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mechanisms of Development*, *125*(3–4), 270–283. https://doi.org/10.1016/j.mod.2007.11.002

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., & Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. *Cell*, *123*(5), 917–929. https://doi.org/10.1016/j.cell.2005.08.040

Nowotschin, S., Costello, I., Piliszek, A., Kwon, G. S., Mao, C. an, Klein, W. H., Robertson, E. J., & Hadjantonakis, A. K. (2013). The T-box transcription factor eomesodermin is essential for AVE induction in the mouse embryo. *Genes and Development*, *27*(9). https://doi.org/10.1101/gad.215152.113

Nowotschin, S., Setty, M., Kuo, Y. Y., Liu, V., Garg, V., Sharma, R., Simon, C. S., Saiz, N., Gardner, R., Boutet, S. C., Church, D. M., Hoodless, P. A., Hadjantonakis, A. K., &

Pe'er, D. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature*, *569*(7756). https://doi.org/10.1038/s41586-019-1127-1

Oberlin, E., Amara, A., Bachelerie, F., Bessia, C., Virelizier, J.-L., Arenzana-Seisdedos, F., Schwartz, O., Heard, J.-M., Clark-Lewis, I., Legler, D. F., Loetscher, M., Baggiolini, M., & Moser, B. (1996). The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature*, *382*(6594), 833–835. https://doi.org/10.1038/382833a0

Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P. J., Chambers, I., & Wilson, V. (2012). The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. *Development*, *139*(13), 2288–2298. https://doi.org/10.1242/dev.078071

Oss-Ronen, L., Redden, R. A., & Lelkes, P. I. (2020). Enhanced Induction of Definitive Endoderm Differentiation of Mouse Embryonic Stem Cells in Simulated Microgravity. *Stem Cells and Development*, *29*(19), 1275–1284. https://doi.org/10.1089/scd.2020.0097

Ozawa, M., Yang, Q.-E., & Ealy, A. D. (2013). The expression of fibroblast growth factor receptors during early bovine conceptus development and pharmacological analysis of their actions on trophoblast growth in vitro. *REPRODUCTION*, *145*(2), 191–201. https://doi.org/10.1530/REP-12-0220

Paca, A., Séguin, C. A., Clements, M., Ryczko, M., Rossant, J., Rodriguez, T. A., & Kunath, T. (2012). BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. *Developmental Biology*, *361*(1), 90–102. https://doi.org/10.1016/j.ydbio.2011.10.013

Papaioannou, V. E. (2014). The T-box gene family: emerging roles in development, stem cells and cancer. *Development*, *141*(20), 3819–3833. https://doi.org/10.1242/dev.104471

Pearce, J. J. H., & Evans, M. J. (1999). Mml, a mouse Mix-like gene expressed in the primitive streak. *Mechanisms of Development*, *87*(1–2), 189–192. https://doi.org/10.1016/S0925-4773(99)00135-5

Pearton, D. J., Smith, C. S., Redgate, E., van Leeuwen, J., Donnison, M., & Pfeffer, P. L. (2014). Elf5 counteracts precocious trophoblast differentiation by maintaining Sox2 and 3 and inhibiting Hand1 expression. *Developmental Biology*, *392*(2), 344–357. https://doi.org/10.1016/j.ydbio.2014.05.012

Pijuan-Sala, B., Griffiths, J. A., Guibentif, C., Hiscock, T. W., Jawaid, W., Calero-Nieto, F. J., Mulas, C., Ibarra-Soria, X., Tyser, R. C. v., Ho, D. L. L., Reik, W., Srinivas, S., Simons, B. D., Nichols, J., Marioni, J. C., & Göttgens, B. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*, *566*(7745), 490–495. https://doi.org/10.1038/s41586-019-0933-9

Porteus, M. H., & Baltimore, D. (2003). Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science*, *300*(5620), 763–763. https://doi.org/10.1126/science.1078395

Posfai, E., Petropoulos, S., de Barros, F. R. O., Schell, J. P., Jurisica, I., Sandberg, R., Lanner, F., & Rossant, J. (2017). Position- and Hippo signaling-dependent plasticity during lineage segregation in the early mouse embryo. *ELife*, *6*. https://doi.org/10.7554/eLife.22906

Pour, M., Kumar, A. S., Farag, N., Bolondi, A., Kretzmer, H., Walther, M., Wittler, L., Meissner, A., & Nachman, I. (2022). Emergence and patterning dynamics of mouse-definitive endoderm. *IScience*, *25*(1), 103556. https://doi.org/10.1016/j.isci.2021.103556

Prozzillo, Y., Fattorini, G., Santopietro, M. V., Suglia, L., Ruggiero, A., Ferreri, D., & Messina, G. (2020). Targeted Protein Degradation Tools: Overview and Future Perspectives. *Biology*, *9*(12), 421. https://doi.org/10.3390/biology9120421

Pulit-Penaloza, J. A., Scherbik, S. v., & Brinton, M. A. (2012). Activation of Oas1a gene expression by type I IFN requires both STAT1 and STAT2 while only STAT2 is required for Oas1b activation. *Virology*, *425*(2), 71–81. https://doi.org/10.1016/j.virol.2011.11.025

Rajaraman, G., Murthi, P., Brennecke, S. P., & Kalionis, B. (2010). Homeobox Gene HLX Is a Regulator of HGF/c-met-Mediated Migration of Human Trophoblast-Derived Cell Lines1. *Biology of Reproduction*, *83*(4), 676–683. https://doi.org/10.1095/biolreprod.109.078634

Rakoczy, J., Padmanabhan, N., Krzak, A. M., Kieckbusch, J., Cindrova-Davies, T., & Watson, E. D. (2017). Dynamic expression of TET1, TET2, and TET3 dioxygenases in mouse and human placentas throughout gestation. *Placenta*, *59*, 46–56. https://doi.org/10.1016/j.placenta.2017.09.008

Ralston, A., Cox, B. J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P., Draper, J. S., & Rossant, J. (2010). Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development*, *137*(3), 395–403. https://doi.org/10.1242/dev.038828

Ralston, A., & Rossant, J. (2006). How Signaling Promotes Stem Cell Survival: Trophoblast Stem Cells and Shp2. *Developmental Cell*, *10*(3), 275–276. https://doi.org/10.1016/j.devcel.2006.02.007

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 Research*, *44*(W1), W160–W165. https://doi.org/10.1093/nar/gkw257

Rampon, C., Prandini, M.-H., Bouillot, S., Pointu, H., Tillet, E., Frank, R., Vernet, M., & Huber, P. (2005). Protocadherin 12 (VE-cadherin 2) is expressed in endothelial, trophoblast, and mesangial cells. *Experimental Cell Research*, *302*(1), 48–60. https://doi.org/10.1016/j.yexcr.2004.08.024

Renaud, S. J., Chakraborty, D., Mason, C. W., Rumi, M. A. K., Vivian, J. L., & Soares, M. J. (2015). OVO-like 1 regulates progenitor cell fate in human trophoblast development. *Proceedings of the National Academy of Sciences*, *112*(45). https://doi.org/10.1073/pnas.1507397112

Ribes, V., Fraulob, V., Petkovich, M., & Dollé, P. (2007). The oxidizing enzyme CYP26a1 tightly regulates the availability of retinoic acid in the gastrulating mouse embryo to ensure proper head development and vasculogenesis. *Developmental Dynamics*, *236*(3), 644–653. https://doi.org/10.1002/dvdy.21057

Rivera-Pérez, J. A., Jones, V., & Tam, P. P. L. (2010). *Culture of Whole Mouse Embryos at Early Postimplantation to Organogenesis Stages* (pp. 185–203). https://doi.org/10.1016/S0076-6879(10)76011-0

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Rodriguez, T. A., Casey, E. S., Harland, R. M., Smith, J. C., & Beddington, R. S. P. (2001). Distinct Enhancer Elements Control Hex Expression during Gastrulation and Early Organogenesis. *Developmental Biology*, *234*(2), 304–316. https://doi.org/10.1006/dbio.2001.0265

Rodriguez, T. A., Srinivas, S., Clements, M. P., Smith, J. C., & Beddington, R. S. P. (2005). Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm. *Development*, *132*(11), 2513–2520. https://doi.org/10.1242/dev.01847

Rostovskaya, M., Stirparo, G. G., & Smith, A. (2019). Capacitation of human naïve pluripotent stem cells for multi-lineage differentiation. *Development*, *146*(7). https://doi.org/10.1242/dev.172916

Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, *14*(12), 8096–8106. https://doi.org/10.1128/MCB.14.12.8096

Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S., & Polakis, P. (1996). Binding of GSK3β to the APC-β-Catenin Complex and Regulation of Complex Assembly. *Science*, *272*(5264), 1023–1026. https://doi.org/10.1126/science.272.5264.1023

Rugg-Gunn, P. J., Cox, B. J., Lanner, F., Sharma, P., Ignatchenko, V., McDonald, A. C. H., Garner, J., Gramolini, A. O., Rossant, J., & Kislinger, T. (2012). Cell-Surface Proteomics Identifies Lineage-Specific Markers of Embryo-Derived Stem Cells. *Developmental Cell*, *22*(4), 887–901. https://doi.org/10.1016/j.devcel.2012.01.005

Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A. J. R., Carlton, M. B. L., Pearce, J. J., Barton, S. C., Azim Surani, M., Ryan, K., Nehls, M. C., Wilsons, V., & Evans, M. J. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature*. https://doi.org/10.1038/35003601

Saga, Y. (1998). Genetic rescue of segmentation defect in MesP2-deficient mice by MesP1 gene replacement. *Mechanisms of Development*, *75*(1–2), 53–66. https://doi.org/10.1016/S0925-4773(98)00077-X

Saiz, N., & Plusa, B. (2013). Early cell fate decisions in the mouse embryo. *REPRODUCTION*, *145*(3), R65–R80. https://doi.org/10.1530/REP-12-0381

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J., & Hamada, H. (2001). The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. *Genes & Development*, *15*(2), 213–225. https://doi.org/10.1101/gad.851501

Sasaki, H., & Hogan, B. L. (1993). Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development*, *118*(1), 47–59. https://doi.org/10.1242/dev.118.1.47

Sauer, B. (1998). Inducible Gene Targeting in Mice Using the Cre/loxSystem. *Methods*, *14*(4), 381–392. https://doi.org/10.1006/meth.1998.0593

Scheibner, K., Schirge, S., Burtscher, I., Büttner, M., Sterr, M., Yang, D., Böttcher, A., Ansarullah, Irmler, M., Beckers, J., Cernilogar, F. M., Schotta, G., Theis, F. J., & Lickert, H. (2021). Epithelial cell plasticity drives endoderm formation during gastrulation. *Nature Cell Biology*, *23*(7), 692–703. https://doi.org/10.1038/s41556-021-00694-x

Schoenfelder, S., Mifsud, B., Senner, C. E., Todd, C. D., Chrysanthou, S., Darbo, E., Hemberger, M., & Branco, M. R. (2018). Divergent wiring of repressive and active chromatin interactions between mouse embryonic and trophoblast lineages. *Nature Communications*, *9*(1), 4189. https://doi.org/10.1038/s41467-018-06666-4

Schrode, N., Saiz, N., di Talia, S., & Hadjantonakis, A.-K. (2014). GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. *Developmental Cell*, *29*(4), 454–467. https://doi.org/10.1016/j.devcel.2014.04.011

Schrode, N., Xenopoulos, P., Piliszek, A., Frankenberg, S., Plusa, B., & Hadjantonakis, A.-K. (2013). Anatomy of a blastocyst: Cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. *Genesis*, *51*(4), 219–233. https://doi.org/10.1002/dvg.22368

Sferruzzi-Perri, A. N., Macpherson, A. M., Roberts, C. T., & Robertson, S. A. (2009). Csf2 Null Mutation Alters Placental Gene Expression and Trophoblast Glycogen Cell and Giant Cell Abundance in Mice1. *Biology of Reproduction*, *81*(1), 207–221. https://doi.org/10.1095/biolreprod.108.073312

Shahjalal, H. Md., Abdal Dayem, A., Lim, K. M., Jeon, T., & Cho, S.-G. (2018). Generation of pancreatic ß cells for treatment of diabetes: advances and challenges. *Stem Cell Research & Therapy*, *9*(1), 355. https://doi.org/10.1186/s13287-018-1099-3

Shimosato, D., Shiki, M., & Niwa, H. (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA Factors acquire the character of XEN cells. *BMC Developmental Biology*, *7*(1), 80. https://doi.org/10.1186/1471-213X-7-80

Sherwood, R. I., Chen, T.-Y. A., & Melton, D. A. (2009). Transcriptional dynamics of endodermal organ formation. *Developmental Dynamics*, *238*(1), 29–42. https://doi.org/10.1002/dvdy.21810

Simmons, D. G., Natale, D. R. C., Begay, V., Hughes, M., Leutz, A., & Cross, J. C. (2008). Early patterning of the chorion leads to the trilaminar trophoblast cell structure in the placental labyrinth. *Development*, *135*(12), 2083–2091. https://doi.org/10.1242/dev.020099

Simon, C. S., Downes, D. J., Gosden, M. E., Telenius, J., Higgs, D. R., Hughes, J. R., Costello, I., Bikoff, E. K., & Robertson, E. J. (2017). Functional characterisation of cis-regulatory elements governing dynamic Eomes expression in the early mouse embryo. *Development*. https://doi.org/10.1242/dev.147322

Sinner, D., Rankin, S., Lee, M., & Zorn, A. M. (2004). Sox17 and ß-catenin cooperate to regulate the transcription of endodermal genes. *Development*, *131*(13), 3069–3080. https://doi.org/10.1242/dev.01176

Skene, P. J., Henikoff, J. G., & Henikoff, S. (2018). Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nature Protocols*, *13*(5), 1006–1019. https://doi.org/10.1038/nprot.2018.015

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., & Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, *336*(6200), 688–690. https://doi.org/10.1038/336688a0

Solloway, M. J., & Robertson, E. J. (1999). Early embryonic lethality in Bmp5;Bmp7 double mutant mice suggests functional redundancy within the 60A subgroup. *Development*, *126*(8), 1753–1768. https://doi.org/10.1242/dev.126.8.1753

Srinivas, S., Rodriguez, T., Clements, M., Smith, J. C., & Beddington, R. S. P. (2004). Active cell migration drives the unilateral movements of the anterior visceral endoderm. *Development*, *131*(5), 1157–1164. https://doi.org/10.1242/dev.01005

Srivastava, V., & Ow, D. W. (2003). Rare instances of Cre-mediated deletion product maintained in transgenic wheat. *Plant Molecular Biology*, *52*(3), 661–668. https://doi.org/10.1023/A:1024839617490

Stark, R., & Brown, G. (2011). *DiffBind: differential binding analysis of ChIP-Seq peak data*. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nature Reviews Genetics*, *20*(11), 631–656. https://doi.org/10.1038/s41576-019-0150-2

Stavridis, M. P., Lunn, J. S., Collins, B. J., & Storey, K. G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development*, *134*(16), 2889–2894. https://doi.org/10.1242/dev.02858

Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., & Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development*. https://doi.org/10.1242/dev.01801

Sutherland, A. E., Calarco, P. G., & Damsky, C. H. (1988). Expression and function of cell surface extracellular matrix receptors in mouse blastocyst attachment and outgrowth. *Journal of Cell Biology*, *106*(4), 1331–1348. https://doi.org/10.1083/jcb.106.4.1331

Suzuki, S., Nozawa, Y., Tsukamoto, S., Kaneko, T., Imai, H., & Minami, N. (2015). Histone methyltransferase Smyd3 regulates early embryonic lineage commitment in mice. *REPRODUCTION*, *150*(1), 21–30. https://doi.org/10.1530/REP-15-0019

Takada, S., Stark, K. L., Shea, M. J., Vassileva, G., McMahon, J. A., & McMahon, A. P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes & Development*, *8*(2), 174–189. https://doi.org/10.1101/gad.8.2.174

Takahashi, S., Kobayashi, S., & Hiratani, I. (2018). Epigenetic differences between naïve and primed pluripotent stem cells. *Cellular and Molecular Life Sciences*, *75*(7), 1191–1203. https://doi.org/10.1007/s00018-017-2703-x

Takahashi, H., Takizawa, T., Matsubara, S., Ohkuchi, A., Kuwata, T., Usui, R., Matsumoto, H., Sato, Y., Fujiwara, H., Okamoto, A., Suzuki, M., & Takizawa, T. (2014). Extravillous trophoblast cell invasion is promoted by the CD44–hyaluronic acid interaction. *Placenta*, *35*(3), 163–170. https://doi.org/10.1016/j.placenta.2013.12.009

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Takaoka, K., & Hamada, H. (2012). Cell fate decisions and axis determination in the early mouse embryo. *Development*, *139*(1), 3–14. https://doi.org/10.1242/dev.060095

Takaoka, K., Nishimura, H., & Hamada, H. (2017). Both Nodal signalling and stochasticity select for prospective distal visceral endoderm in mouse embryos. *Nature Communications*, *8*(1), 1492. https://doi.org/10.1038/s41467-017-01625-x

Tam, P. P. L., & Beddington, R. S. P. (1992). Establishment and Organization of Germ Layers in the Gastrulating Mouse Embryo. In *Ciba Found Symp*. (Vol. 165, pp. 27–41). https://doi.org/10.1002/9780470514221.ch3

Tam, P. P. L., & Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nature Reviews Genetics*, *8*(5), 368–381. https://doi.org/10.1038/nrg2084

Tam, P. P., Parameswaran, M., Kinder, S. J., & Weinberger, R. P. (1997). The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. *Development*, *124*(9), 1631–1642. https://doi.org/10.1242/dev.124.9.1631

Tanaka, S. (2006). Derivation and Culture of Mouse Trophoblast Stem Cells In Vitro. In *Embryonic Stem Cell Protocols* (pp. 35–44). Humana Press. https://doi.org/10.1385/1-59745-037-5:35

Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A., & Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science*. https://doi.org/10.1126/science.282.5396.2072

Tanigawa, Y., Dyer, E. S., & Bejerano, G. (2022). Which TF is functionally important in your open chromatin data? *PLoS Computational Biology*, *18*(8), e1010378. https://doi.org/10.1371/journal.pcbi.1010378

Taylor, S. C., Laperriere, G., & Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Scientific Reports*, *7*(1), 2409. https://doi.org/10.1038/s41598-017-02217-x

Teo, A. K. K., Arnold, S. J., Trotter, M. W. B., Brown, S., Ang, L. T., Chng, Z., Robertson, E. J., Dunn, N. R., & Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes & Development*, *25*(3), 238–250. https://doi.org/10.1101/gad.607311

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*(7150), 196–199. https://doi.org/10.1038/nature05972

Theiler, K. (1989). *The House Mouse*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-88418-4

Thomas, P., & Beddington, R. (1996). Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. *Current Biology*, *6*(11), 1487–1496. https://doi.org/10.1016/S0960-9822(96)00753-1

Tiyaboonchai, A., Cardenas-Diaz, F. L., Ying, L., Maguire, J. A., Sim, X., Jobaliya, C., Gagne, A. L., Kishore, S., Stanescu, D. E., Hughes, N., de Leon, D. D., French, D. L., & Gadue, P. (2017). GATA6 Plays an Important Role in the Induction of Human Definitive

Endoderm, Development of the Pancreas, and Functionality of Pancreatic β Cells. *Stem Cell Reports*, *8*(3), 589–604. https://doi.org/10.1016/j.stemcr.2016.12.026

Todd, N., McNally, R., Alqudah, A., Jerotic, D., Suvakov, S., Obradovic, D., Hoch, D., Hombrebueno, J. R., Campos, G. L., Watson, C. J., Gojnic-Dugalic, M., Simic, T. P., Krasnodembskaya, A., Desoye, G., Eastwood, K.-A., Hunter, A. J., Holmes, V. A., McCance, D. R., Young, I. S., … McClements, L. (2021). Role of A Novel Angiogenesis FKBPL-CD44 Pathway in Preeclampsia Risk Stratification and Mesenchymal Stem Cell Treatment. *The Journal of Clinical Endocrinology & Metabolism*, *106*(1), 26–41. https://doi.org/10.1210/clinem/dgaa403

Tonami, K., Kurihara, Y., Aburatani, H., Uchijima, Y., Asano, T., & Kurihara, H. (2007). Calpain 6 Is Involved in Microtubule Stabilization and Cytoskeletal Organization. *Molecular and Cellular Biology*, *27*(7), 2548–2561. https://doi.org/10.1128/MCB.00992-06

Tosic, J., Kim, G.-J., Pavlovic, M., Schröder, C. M., Mersiowsky, S.-L., Barg, M., Hofherr, A., Probst, S., Köttgen, M., Hein, L., & Arnold, S. J. (2019). Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nature Cell Biology*, *21*(12), 1518–1531. https://doi.org/10.1038/s41556-019-0423-1

Tremblay, K. D., Hoodless, P. A., Bikoff, E. K., & Robertson, E. J. (2000). Formation of the definitive endoderm in mouse is a Smad2-dependent process. *Development*, *127*(14), 3079–3090. https://doi.org/10.1242/dev.127.14.3079

Tsankov, A. M., Gu, H., Akopian, V., Ziller, M. J., Donaghey, J., Amit, I., Gnirke, A., & Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. *Nature*, *518*(7539), 344–349. https://doi.org/10.1038/nature14233

Turlo, K. A., Gallaher, S. D., Vora, R., Laski, F. A., & Iruela-Arispe, M. L. (2010). When Cre-Mediated Recombination in Mice Does Not Result in Protein Loss. *Genetics*, *186*(3), 959–967. https://doi.org/10.1534/genetics.110.121608

Turner, P. V, Brabb, T., Pekow, C., & Vasbinder, M. A. (2011). Administration of substances to laboratory animals: routes of administration and factors to consider. *Journal of the American Association for Laboratory Animal Science : JAALAS*, *50*(5), 600–613.

Ullah, R., Naz, A., Akram, H. S., Ullah, Z., Tariq, M., Mithani, A., & Faisal, A. (2020). Transcriptomic analysis reveals differential gene expression, alternative splicing, and novel exons during mouse trophoblast stem cell differentiation. *Stem Cell Research & Therapy*, *11*(1), 342. https://doi.org/10.1186/s13287-020-01848-8

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics*, *11*(9), 636–646. https://doi.org/10.1038/nrg2842

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. (2009). Early Cell Fate Decisions of Human Embryonic Stem Cells and Mouse Epiblast Stem Cells Are Controlled by the Same Signalling Pathways. *PLoS ONE*, *4*(6), e6082. https://doi.org/10.1371/journal.pone.0006082

Vandereyken, K., Sifrim, A., Thienpont, B., & Voet, T. (2023). Methods and applications for single-cell and spatial multi-omics. *Nature Reviews Genetics*. https://doi.org/10.1038/s41576-023-00580-2

Vincent, S. D., Dunn, N. R., Hayashi, S., Norris, D. P., & Robertson, E. J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes & Development*, *17*(13), 1646–1662. https://doi.org/10.1101/gad.1100503

Wamaitha, S. E., del Valle, I., Cho, L. T. Y., Wei, Y., Fogarty, N. M. E., Blakeley, P., Sherwood, R. I., Ji, H., & Niakan, K. K. (2015). Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes & Development*, *29*(12), 1239–1255. https://doi.org/10.1101/gad.257071.114

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013). One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell*, *153*(4), 910–918. https://doi.org/10.1016/j.cell.2013.04.025

Wang, Q., Zou, Y., Nowotschin, S., Kim, S. Y., Li, Q. v., Soh, C.-L., Su, J., Zhang, C., Shu, W., Xi, Q., Huangfu, D., Hadjantonakis, A.-K., & Massagué, J. (2017). The p53 Family Coordinates Wnt and Nodal Inputs in Mesendodermal Differentiation of Embryonic Stem Cells. *Cell Stem Cell*, *20*(1), 70–86. https://doi.org/10.1016/j.stem.2016.10.002

Wang, Y., Yuan, P., Yan, Z., Yang, M., Huo, Y., Nie, Y., Zhu, X., Qiao, J., & Yan, L. (2021). Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. *Nature Communications*, *12*(1), 1247. https://doi.org/10.1038/s41467-021-21409-8

Watson, E. D., & Cross, J. C. (2005). Development of Structures and Transport Functions in the Mouse Placenta. *Physiology*, *20*(3), 180–193. https://doi.org/10.1152/physiol.00001.2005

Weinstein, D. C., Ruiz i Altaba, A., Chen, W. S., Hoodless, P., Prezioso, V. R., Jessell, T. M., & Darnell, J. E. (1994). The winged-helix transcription factor HNF-3β is required for notochord development in the mouse embryo. *Cell*, *78*(4), 575–588. https://doi.org/10.1016/0092-8674(94)90523-1

Wells, J. M., & Melton, D. A. (1999). Vertebrate Endoderm Development. *Annual Review of Cell and Developmental Biology*, *15*(1), 393–410. https://doi.org/10.1146/annurev.cellbio.15.1.393

Wells, J. M., & Melton, D. A. (2000). Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development*, *127*(8), 1563–1572. https://doi.org/10.1242/dev.127.8.1563

Wen, F., Tynan, J. A., Cecena, G., Williams, R., Múnera, J., Mavrothalassitis, G., & Oshima, R. G. (2007). Ets2 is required for trophoblast stem cell self-renewal. *Developmental Biology*, *312*(1), 284–299. https://doi.org/10.1016/j.ydbio.2007.09.024

Wen, J., Zeng, Y., Fang, Z., Gu, J., Ge, L., Tang, F., Qu, Z., Hu, J., Cui, Y., Zhang, K., Wang, J., Li, S., Sun, Y., & Jin, Y. (2017). Single-cell analysis reveals lineage segregation in early post-implantation mouse embryos. *Journal of Biological Chemistry*, *292*(23), 9840–9854. https://doi.org/10.1074/jbc.M117.780585

Wu, G., Li, R., Tong, C., He, M., Qi, Z., Chen, H., Deng, T., Liu, H., & Qi, H. (2019). Non-invasive prenatal testing reveals copy number variations related to pregnancy complications. *Molecular Cytogenetics*, *12*(1), 38. https://doi.org/10.1186/s13039-019-0451-3

Xu, W., Wen, Y., Liang, Y., Xu, Q., Wang, X., Jin, W., & Chen, X. (2021). A plate-based single-cell ATAC-seq workflow for fast and robust profiling of chromatin accessibility. *Nature Protocols*, *16*(8), 4084–4107. https://doi.org/10.1038/s41596-021-00583-5

Yagi, R., Kohn, M. J., Karavanova, I., Kaneko, K. J., Vullhorst, D., DePamphilis, M. L., & Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development*, *134*(21), 3827–3836. https://doi.org/10.1242/dev.010223

Yamaguchi, T. P. (2001). Heads or tails: Wnts and anterior–posterior patterning. *Current Biology*, *11*(17), R713–R724. https://doi.org/10.1016/S0960-9822(01)00417-1

Yang, J., & Jiang, W. (2020). The Role of SMAD2/3 in Human Embryonic Stem Cells. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.00653

Yang, Y.-P., & Klingensmith, J. (2006). Roles of organizer factors and BMP antagonism in mammalian forebrain establishment. *Developmental Biology*, *296*(2), 458–475. https://doi.org/10.1016/j.ydbio.2006.06.014

Yang, J., Lu, P., Li, M., Yan, C., Zhang, T., & Jiang, W. (2020). GATA6-AS1 Regulates GATA6 Expression to Modulate Human Endoderm Differentiation. *Stem Cell Reports*, *15*(3), 694–705. https://doi.org/10.1016/j.stemcr.2020.07.014

Yi, Y., Zhu, H., Klausen, C., Chang, H.-M., Inkster, A. M., Terry, J., & Leung, P. C. K. (2021). Dysregulated BMP2 in the Placenta May Contribute to Early-Onset Preeclampsia by Regulating Human Trophoblast Expression of Extracellular Matrix and Adhesion Molecules. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.768669

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*(7194), 519–523. https://doi.org/10.1038/nature06968

Yoney, A., Bai, L., Brivanlou, A. H., & Siggia, E. D. (2022). Mechanisms underlying WNT-mediated priming of human embryonic stem cells. *Development*, *149*(20). https://doi.org/10.1242/dev.200335

Yoon, Y., Huang, T., Tortelote, G. G., Wakamiya, M., Hadjantonakis, A.-K., Behringer, R. R., & Rivera-Pérez, J. A. (2015). Extra-embryonic Wnt3 regulates the establishment of the primitive streak in mice. *Developmental Biology*, *403*(1), 80–88. https://doi.org/10.1016/j.ydbio.2015.04.008

You, J., Wang, W., Chang, H.-M., Yi, Y., Zhao, H., Zhu, H., Sun, Y., Tang, M., Wang, C., Sang, Y., Feng, G., Cheng, S., Leung, P. C. K., & Zhu, Y.-M. (2021). The BMP2 Signaling Axis Promotes Invasive Differentiation of Human Trophoblasts. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.607332

Zadora, J., Singh, M., Herse, F., Przybyl, L., Haase, N., Golic, M., Yung, H. W., Huppertz, B., Cartwright, J. E., Whitley, G., Johnsen, G. M., Levi, G., Isbruch, A., Schulz, H., Luft, F. C., Müller, D. N., Staff, A. C., Hurst, L. D., Dechend, R., & Izsvák, Z. (2017). Disturbed Placental Imprinting in Preeclampsia Leads to Altered Expression of DLX5, a Human-Specific Early Trophoblast Marker. *Circulation*, *136*(19), 1824–1839. https://doi.org/10.1161/CIRCULATIONAHA.117.028110

Zang, J., Yan, M., Zhang, Y., Peng, W., Zuo, J., Zhou, H., Gao, G., Li, M., Chu, Y., & Ye, Y. (2022). MiR-326 inhibits trophoblast growth, migration, and invasion by targeting PAX8 via Hippo pathway. *Reproductive Biology and Endocrinology*, *20*(1), 38. https://doi.org/10.1186/s12958-022-00909-2

Zhang, H., Fraser, S. T., Papazoglu, C., Hoatlin, M. E., & Baron, M. H. (2009). Transcriptional Activation by the Mixl1 Homeodomain Protein in Differentiating Mouse Embryonic Stem Cells. *Stem Cells*, *27*(12), 2884–2895. https://doi.org/10.1002/stem.203

Zhang, H. T., & Hiiragi, T. (2018). Symmetry Breaking in the Mammalian Embryo. *Annual Review of Cell and Developmental Biology*, *34*(1), 405–426. https://doi.org/10.1146/annurev-cellbio-100617-062616

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 Biology*, *9*(9), R137. https://doi.org/10.1186/gb-2008-9-9-r137

Zhao, T., Zhou, X., Szabó, N., Leitges, M., & Alvarez-Bolado, G. (2007). Foxb1-driven Cre expression in somites and the neuroepithelium of diencephalon, brainstem, and spinal cord. *Genesis*, *45*(12), 781–787. https://doi.org/10.1002/dvg.20356

Zhou, W., Menkhorst, E., & Dimitriadis, E. (2022). Characterization of chloride intracellular channel 4 in the regulation of human trophoblast function. *Placenta*, *119*, 24–30. https://doi.org/10.1016/j.placenta.2022.01.012

Zhou, Y., Damsky, C. H., Chiu, K., Roberts, J. M., & Fisher, S. J. (1993). Preeclampsia is associated with abnormal expression of adhesion molecules by invasive cytotrophoblasts. *Journal of Clinical Investigation*, *91*(3), 950–960. https://doi.org/10.1172/JCI116316

Zhu, K., Liu, Y., Fan, C., Zhang, M., Cao, H., He, X., Li, N., Chu, D., Li, F., Zou, M., Hua, J., Wang, H., Wang, Y., Fan, G., & Zhang, S. (2020). Etv5 safeguards trophoblast stem cells differentiation from mouse EPSCs by regulating fibroblast growth factor receptor 2. *Molecular Biology Reports*, *47*(12), 9259–9269. https://doi.org/10.1007/s11033-020-05969-4

Zimmer, A. M., Pan, Y. K., Chandrapalan, T., Kwong, R. W. M., & Perry, S. F. (2019). Loss-of-function approaches in comparative physiology: is there a future for knockdown experiments in the era of genome editing? *Journal of Experimental Biology*, *222*(7). https://doi.org/10.1242/jeb.175737