Transcriptional regulation of Eomesodermin, and its roles in cell fate allocation during early mouse development.



Claire Samantha Simon

Lincoln College

Sir William Dunn School of Pathology

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity term 2015

Declaration

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

Acknowledgements

I would like to express my great appreciation to my supervisors Liz Robertson and Liz Bikoff for giving me the opportunity to work in their lab. I thank them for their sound advice, motivation, and expertise over the past four years and for helping me to develop as a research scientist.

I thank the many lab members past and present who have assisted me in my work. I am particularly grateful to Ita Costello for her mentorship, expert training and valuable guidance throughout my DPhil. Thank you to Pia Rantakari, Li Li and Xin Sun for their contributions to the enhancer deletion project. I also thank Ahmed Salman, Jane Rose and Tom Clague for their support in caring for mice. Additionally, I would like to thank those not directly involved in my projects: Arne Mould, Kim Cheng, Anna Senft, Andy Nelson, Salah Elias, Deb Bogani, Mohammed Ahmed, and Marc Morgan, for their useful discussions and for providing an enjoyable working environment.

I would also like to acknowledge my collaborators who have made this work possible. I give special thanks to Catherine Porcher and her post-doc Hedia Chagraoui for kindly hosting me at the Weatherall Institute of Molecular Medicine, Oxford. Thank you for your time, valuable advice and practical help on the haematopoiesis project. I thank Ben Davies and his team at the Transgenic Core Facility at the Wellcome Trust Centre for Human Genetics, Oxford, for the generation of transgenic mice. Thank you to Nigel Rust at the Flow Cytometry Facility at the Dunn School of Pathology, Oxford, for performing the FACS analysis. Thank you to Peter Koopman for providing the ER71 in situ probe. I also thank the mouse technicians in the Pathology Support Building at the Dunn School of Pathology, Oxford, for caring for the mice.

I dedicate this thesis to my Dad, Isy Simon, who passed away last year. His enthusiasm sparked my fascination with Science from an early age, which helped me get to where I am today. I thank my parents and sister for their love patience and support during my DPhil. To all of my friends in Lincoln College MCR thank you for making Oxford feel like home. I especially thank my wonderful boyfriend, Dan Brown, for being my rock throughout my DPhil and for his company during long study sessions. Your love, encouragement and reassurance have kept me going.

Transcriptional regulation of Eomesodermin, and its roles in cell fate allocation during early mouse development.

Claire Samantha Simon

Lincoln College
Sir William Dunn School of Pathology
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy
Trinity Term 2015

Abstract

Early embryonic development is governed by a tightly regulated series of cell fate decisions to form an entire organism. *Eomesodermin (Eomes*) is a T-box transcription factor dynamically expressed in the extra-embryonic ectoderm, anterior-visceral endoderm (AVE) and primitive streak during mouse gastrulation. Previous work from our lab shows that *Eomes* is essential within the AVE for establishment of the anterior-posterior (A-P) axis and within the primitive streak for epithelial-to-mesenchymal transition, and specification of cardiac mesoderm and definitive endoderm. However, the cis-acting regulatory elements that control the critical spatiotemporal expression domains of *Eomes* are not fully understood. Also, whether *Eomes* plays a functional role in specification of other mesodermal lineages is unknown. Here, I demonstrate an essential enhancer 8kb downstream from the TSS is necessary and sufficient for the correct expression of *Eomes* in the AVE and primitive streak. Mutants harboring a deletion of this enhancer exhibit A-P patterning defects and derivatives of the anterior primitive streak are compromised. The discovery of this essential enhancer provides greater understanding into how this critical gene is regulated during development. Moreover, this is the first characterisation of an *Eomes* enhancer that drives AVE expression. Additionally, I show that *Eomes* is required for the specification of haematopoietic mesoderm. Thus, implicating *Eomes* as a central player in the allocation of multiple cell fates within the primitive streak.

Table of contents

Chapter 1: Introduction 9
1.1 Preface 9
1.2 Post-implantation development 11
1.2.1 Anterior-Posterior axis specification 11
1.2.2 Gastrulation: the formation of the three germ layers 13
1.2.3 Cell fate allocation in the primitive streak 15
1.2.4 Nodal signalling pathway in early embryonic patterning 18
1.3 The T-box family of transcription factors 19
1.3.1 Brachyury expression and function 23
1.3.2 Eomes expression 24
1.3.3 Eomes functional requirements in the early embryo 25
1.3.4 Regulation of Eomes 26
1.4 Developmental enhancers 28
1.4.1 Eomes enhancers 30
1.5 Aims of thesis 31

Chapter 2: Materials and Methods 33
2.1 Materials 33
2.2 Molecular biology 35
2.2.1 Molecular cloning 35
2.2.2 Targeting vectors and transgenic constructs 36
2.2.3 DNA and RNA isolation from cells and tissues 38
2.2.4 Standard PCR 39
2.2.5 Quantitative real-time PCR 40
2.2.6 Riboprobe synthesis for in situ hybridisation 41
2.2.7 Southern blot 41
2.3 Tissue culture 43
2.3.1 mESC transfection 43
2.3.2 mESC maintenance 43
2.3.3 Culture of STOs and making feeder cells 44
2.3.4 Embryoid body differentiation of mESCs 44
2.3.5 Harvesting of embryoid bodies 45
2.3.6 Blast colony forming assay 46
2.3.7 Haematopoietic colony forming assay 47
2.4 Histology and cell biology 47
2.4.1 Whole-mount in situ hybridisation 47
2.4.2 β-galactosidase staining 49
2.4.3 Paraffin embedding 49
2.4.4 Immunofluorescence staining 49
2.4.5 Flow cytometry 50
2.5 Mouse strains and husbandry 50
2.6 Bioinformatics 51

Chapter 3: Targeted deletion of *Eomes* enhancers 52
3.1 Introduction 52
3.2 Identification of *Eomes* enhancers 53
3.3 Transgenic analysis of *Eomes* regulatory regions 60
3.3.1 The EH1 regulatory region has enhancer activity in the primitive streak 60
3.3.2 The EH3 regulatory region has enhancer activity in the primitive streak and anterior visceral endoderm 61
3.4 Targeted deletion of *Eomes* enhancer regions 65
3.4.1 Generation of a ΔEH3 targeted deletion 65
3.4.2 Gastrulation abnormalities in $Eomes^{\Delta EH3/N}$ embryos 69
3.4.3 Deletion of EH3 in the context of $Eomes^{GFP}$ reporter allele greatly diminishes expression 75
3.5 Discussion 82

3.5.1 The primitive streak enhancer EH1b is dispensable for proper regulation of Eomes 82
3.5.2 The anterior visceral endoderm and anterior primitive streak enhancer EH3 is required for correct embryonic patterning 83

Chapter 4: Generation of an *Eomes* driven improved Cre-recombinase mouse line 86
4.1 Introduction 86
4.2 Targeting of *Eomes* improved Cre allele 86
4.3 Fate mapping using the *Eomes* improved Cre allele 90
4.4 Discussion 94
4.4.1 EomesiCre progeny reside in the mesoderm component of the yolk sac blood islands 94
4.4.2 Eomes^{iCre} fate-map labels posterior mesoderm 95
4.4.3 Eomes^{iCre} mRNA cannot be detected in the ExE 96

Chapter 5: Investigating redundancy between *Eomes* and *Brachyury* during early mesoderm formation 98
5.1 Introduction 98
5.2 Comparison of *Eomes* and *Brachyury* T-box transcription factors 100
5.3 Dynamics of *Eomes* and *Brachyury* expression during early embryonic development and mesoderm specification 102
5.4 *Eomes* is required for haematopoietic progenitor specification 108
5.5 *Brachyury* cDNA expressed from the endogenous *Eomes* locus 116
5.6 Discussion 120
5.6.1 Eomes is expressed in the early streak and is essential for haematopoietic mesoderm formation 120
5.6.2 Eomes and Brachyury are co-expressed in the anterior region of the mid-streak 121
5.6.3 Eomes - Brachyury substitution experiment 122

Chapter 6: General Discussion 124
6.1 Aims, key results and future work 124
6.1.1 Identification and targeted deletion of Eomes enhancers 124
6.1.2 Generation of an Eomes driven improved-Cre recombinase mouse line 125
6.1.3 Investigating redundancy between Eomes and Brachyury during early mesoderm formation 126
6.2 Outstanding questions 128
6.3 Implications 131

References 132

List of figures

Figure 1.1: Fate decisions during early embryonic development. 10
Figure 1.2: Opposing signalling centres govern the formation of the embryonic anterior-posterior axis and position of the primitive streak. 12
Figure 1.3: The three germ layers are formed during gastrulation 14
Figure 1.4: Lineage allocation in the primitive streak. 16
Figure 1.5: The T-box transcription factors, *Eomes* and *Brachyury*, exhibit partially overlapping expression patterns during mouse gastrulation. 22
Figure 3.1: Identification of putative *Eomes* enhancers. 56
Figure 3.2: Predicted transcription factor binding sites located within mouse *Eomes* enhancers. 59
Figure 3.3: *Eomes* enhancer driven reporters recapitulate primitive streak and anterior visceral endoderm expression. 63
Figure 3.4: Generation of *Eomes*$^{\Delta EH3}$ mice, a mild hympomorph. 66
Figure 3.5: Class I *Eomes*$^{\Delta EH3/N}$ embryos display A-P axis patterning defects. 70
Figure 3.6: Class II *Eomes*$^{\Delta EH3/N}$ embryos have anterior primitive streak defects. 72
Figure 3.7: Loss of EH3 fails to perturb domains of *Eomes* expression in *Eomes* $^{\Delta EH3/\Delta EH3}$ embryos. 74
Figure 3.8: Deletion of EH3 in the context of an *Eomes*^{GFP} reporter allele. 76
Figure 3.9: Loss of the EH3 enhancer profoundly reduced the level of GFP reporter gene expression. 78
Figure 3.10: Generation of *Eomes*$^{\Delta EH1b}$ mice. 80
Figure 4.1: Generation of *Eomes*^{iCre} mice. 88
Figure 4.2: Assessing the fidelity of *iCre* expression from the *Eomes*^{iCre} reporter allele. 89
Figure 4.3: Fate mapping studies of *Eomes*^{iCre} at E7.5. 91
Figure 4.4: Fate mapping studies of *Eomes*^{iCre} at E8.5. 93
Figure 5.1: Sequence comparison of Eomes and Brachyury domain structure and protein coding information. 101
Figure 5.2: Partially overlapping mRNA expression domains of *Eomes* and *Brachyury* within the primitive streak 103
Figure 5.3: Unique and overlapping Eomes and Brachyury cell populations in the early embryo. 105
Figure 5.4: Kinetics of mesoderm markers during mESC differentiation. 107
Figure 5.5: *Eomes* is required for specification of cardiac and haematopoietic progenitors *in vitro* 110
Figure 5.6: *Eomes* is essential for haematopoietic differentiation *in vitro* 112
Figure 5.7: *Eomes* is required for specification of haematopoietic mesoderm *in vivo* 115
Figure 5.8: Strategy for replacing *Eomes* with *Brachyury* coding sequence 118

Abbreviations

A-P axis	Anterior-posterior axis
APS	Anterior primitive streak
AVE	Anterior visceral endoderm
BBR	Boehringer Blocking Reagent
bp	Basepairs
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
ChIP	Chromatin immunoprecipitation
Cre	Bacteriophage PI Cre recombinase
DE	Definitive endoderm
DEPC	Diethylpyrocarbonate
DIG	Digoxigenin
dNTP	Deoxynucleotide triphosphate
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
E	Embryonic day
EB	Embryoid body
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EGTA	Ethylene glycol tetraacetic acid
EMT	Epithelial-to-mesenchymal transition
Eomes	Eomesodermin
EPC	Ectoplacental cone
ESC	Embryonic stem cell
ExE	Extra-embryonic ectoderm
FACS	Fluorescence activated cell sorting
FIAU	Fialuridine
G418	Geneticin sulphate
H3K27Ac/Me	Histone 3, Lysine 27 acetylation / methylation
Hsp68	Heat shock protein 68 minimal promoter
kb	Kilobase
LacZ	E. Coli beta-galactosidase
MeOH	Methanol
mRNA	Messenger ribonucleic acid
MTG	Monothioglycerol
Neo	Neomycin phosphotransferase
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
P-D axis	Proximal-distal axis
PFA	Paraformaldehyde
PGK	Phosphoglycerate kinase promoter
Pol	Polymerase
PS	Primitive streak
qPCR	Quantitative polymerase chain reaction
RT	Reverse transcription
SDS	Sodium dodecyl sulphate
SNH	STO fibroblast Neomycin and Hygromycin resistant
SSC	Sodium chloride sodium citrate
T	Brachyury
TAE	Tris Acetate EDTA
TK	Thymidine kinase
TSS	Transcription start site
UV	Ultraviolet
VE	Visceral endoderm
WISH	Whole-mount in situ hybridisation

Chapter 1: Introduction

1.1 Preface

Mammalian embryonic development involves a tightly regulated series of lineage decisions to form an entire organism from a single cell, the totipotent zygote **(Figure 1.1)**. These in turn depend on a complex interaction of transcription factors, signalling molecules, position, timing and the history of a cell. The use of mouse embryonic stem cells (ESCs) and transgenic gene targeting technologies has greatly increased our understanding of cell fate allocation in the mouse. A gene can be removed, aberrantly expressed or replaced with reporters to understand its function. By modifying the mouse genome we hope to decipher the complicated gene regulatory networks that govern early embryonic development. Sophisticated ESC differentiation protocols allow comprehensive molecular analysis to further refine these models of cell specification.

The transcription factor Eomes is a key regulator of fate decisions during embryonic development. The aim of this thesis is to identify cis-acting regulatory sequences (enhancers) that govern the expression of this essential gene in different lineages and, in turn, the functional role *Eomes* plays in cell specification.

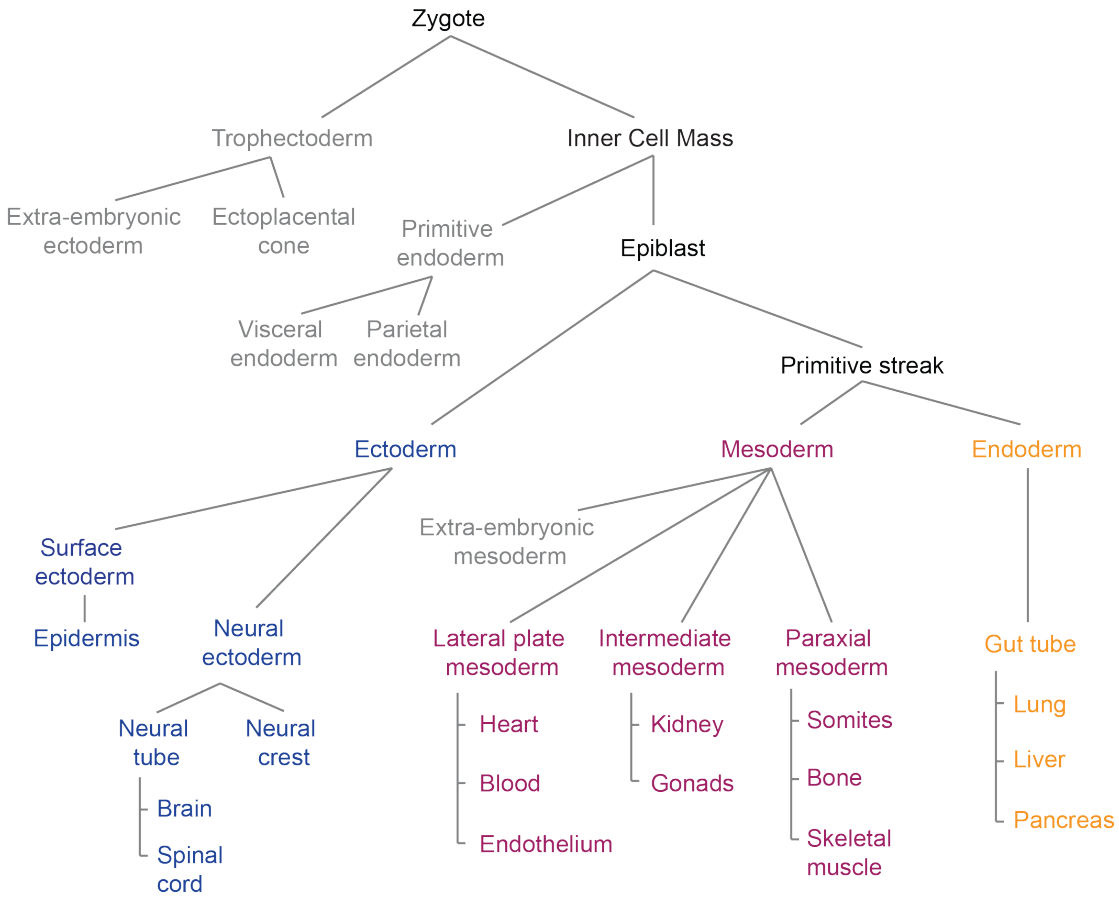


Figure 1.1: Fate decisions during early embryonic development.
Embryonic development involves a tightly regulated series of lineage decisions. The totipotent zygote differentiates to form extra-embryonic tissues (grey) and the embryo proper (black). The three germ layers that give rise to all somatic tissues are ectoderm (blue), mesoderm (red) and endoderm (yellow). Modified from LifeMap Sciences (discovery.lifemapsc.com).

1.2 Post-implantation development

1.2.1 Anterior-Posterior axis specification

Anterior-posterior (A-P) axis formation is initiated at embryonic day 5.5 (E5.5). The visceral endoderm is exposed to Nodal signalling from the epiblast (Brennan et al., 2001). The cells in the distal most region of the visceral endoderm furthest from inhibitory extra-embryonic ectoderm (ExE) derived BMP become specified as the distal visceral endoderm (DVE), and begin to express genes such as *Lefty1, Cer1* and *Hex* **(Figure 1.2)** (Takaoka and Hamada, 2012). Additional cells recruited to form the anterior visceral endoderm (AVE) express the same set of genes (Takaoka et al., 2011). At E6.5, the DVE together with the AVE migrates proximally, towards the embryonic/extra-embryonic junction and establishes the anterior side of the embryo (Beddington and Robertson, 1998; Srinivas et al., 2004; Takaoka and Hamada, 2012).

At E6.5, two opposing signalling centres set up the morphogen gradients that pattern the embryo (Arnold and Robertson, 2009; Beddington and Robertson, 1999; Tam and Loebel, 2007). BMP4 secreted by the ExE induces high levels of Wnt3 and Nodal within the epiblast. Conversely, the AVE secretes the agonists Lefty (Nodal inhibitor), Cer1 (Nodal and BMP inhibitor) and Dkk1 (Wnt inhibitor) that deplete the underlying epiblast of Nodal and Wnt activities, thus specifying anterior fate and the future head region. The proximal posterior epiblast exposed to the highest levels of Wnt, BMP and Nodal signalling gives rise to the primitive streak (Perea-Gomez et al., 2002).

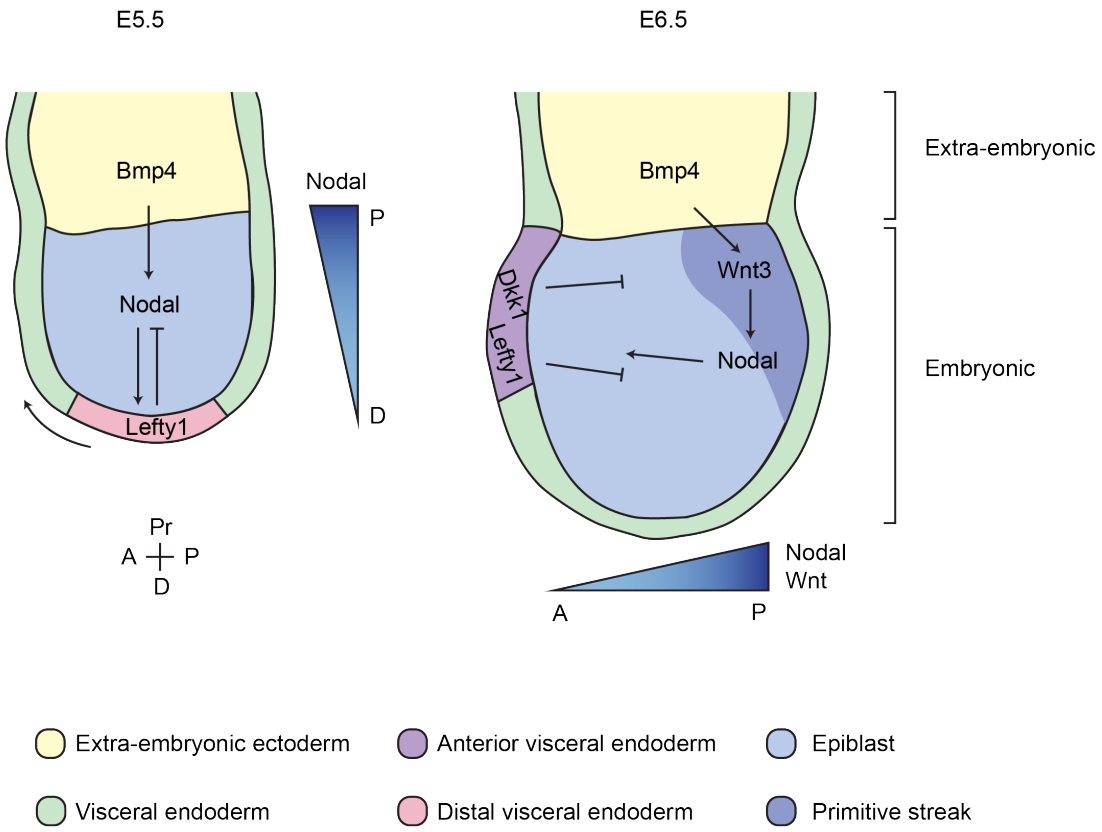


Figure 1.2: Opposing signalling centres govern the formation of the embryonic anterior-posterior axis and position of the primitive streak.
At E5.5 BMP4 signalling in the extra-embryonic ectoderm (ExE) induces high levels of Nodal signalling in the epiblast, which in turn specifies the distal visceral endoderm (DVE) in the overlying visceral endoderm (VE). BMP4 inhibits DVE formation in more proximal regions of visceral endoderm, thus DVE formation is restricted to the most distal tip. The DVE expresses the Nodal inhibitor *Lefty1* setting up a gradient of Nodal within the epiblast along the proximal-distal (Pr-D) axis. Additional cells are recruited that express the same set of markers as the DVE, the anterior visceral endoderm (AVE). The AVE/DVE migrate proximally, halting before the extra-embryonic boundary, thus determining the anterior side of the embryo. At E6.5 the migrated AVE secretes Wnt and Nodal inhibitors (including *Dkk1* and *Lefty1*), which set up an anterior-posterior (A-P) gradient of Nodal and Wnt activity within the epiblast. The primitive streak forms in the proximal-posterior region of the epiblast, with the highest Nodal and Wnt3 activity. Modified from Takaoka and Hamada (2012).

1.2.2 Gastrulation: the formation of the three germ layers

Gastrulation converts the single-layered pluripotent epiblast into the three germ layers; ectoderm, mesoderm and endoderm, the progenitors of all somatic cell types in the adult organism. The site where mesoderm and endoderm delaminate, undergoing an epithelial-to-mesenchymal transition (EMT), is the primitive streak (Tam and Loebel, 2007). The nascent mesoderm subsequently migrates to form a layer between the epiblast and visceral endoderm **(Figure 1.3)**. Definitive endoderm intercalates onto the surface of the embryo, displacing the overlying visceral endoderm (Kwon et al., 2008).

At the onset of gastrulation, the primitive streak is restricted to the proximal-posterior region of the embryo, adjacent to the ExE. As gastrulation proceeds, the embryo continues to grow, and the primitive streak elongates to reach the distal end of the embryo. Epiblast cells that did not traverse the streak are fated to become ectoderm and form the innermost layer.

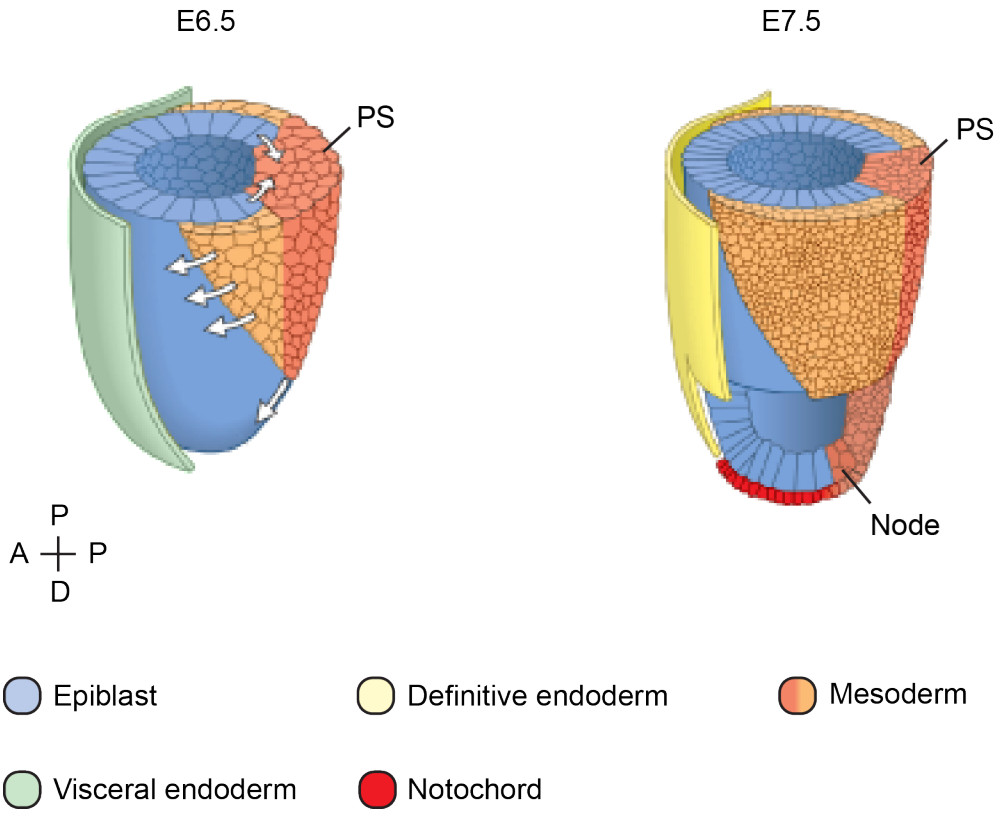


Figure 1.3: The three germ layers are formed during gastrulation
The three germ layers; definitive endoderm, mesoderm and ectoderm, are formed during gastrulation between E6.5 – E7.5. The primitive streak (PS) is the site through which epiblast cells undergo EMT and either intercalate onto the surface of the embryo (definitive endoderm) or migrate over the epiblast to form the middle layer (mesoderm). The inner layer (ectoderm) is formed of epiblast cells that do not traverse the streak. Arrows indicate the direction of cell migration. Modified from Wolpert et al. (2011).

1.2.3 Cell fate allocation in the primitive streak

The timing and position of a cell exiting the streak determines its fate **(Figure 1.4A)** (Lawson et al., 1991). The first cells to delaminate from the early posterior primitive streak give rise to extra-embryonic yolk sac mesoderm that in turn form endothelium and erythrocytes within the yolk sac blood islands (Kinder et al., 1999). Next at mid-streak stages, cardiac and cranial mesoderm migrates to the anterior side of the embryo. Then at late streak stages, the extra-embryonic, lateral plate, and paraxial mesoderm emerge from posterior, intermediate and anterior positions, respectively. Finally, the anterior-most primitive streak (APS: anterior primitive streak) cell population gives rise to axial mesoderm (prechordal plate, notochord and node) and the definitive endoderm **(Figure 1.4B)** (Arnold and Robertson, 2009; Kinder et al., 2001)

Morphogen gradients formed by the TGF-β family members BMP and Nodal pattern the primitive streak. Highest levels of BMP and Nodal are required for the allocation of posterior and anterior fates, respectively. Wnt and FGF signalling molecules are also essential for cell fate specification within the streak (Tam and Loebel, 2007).

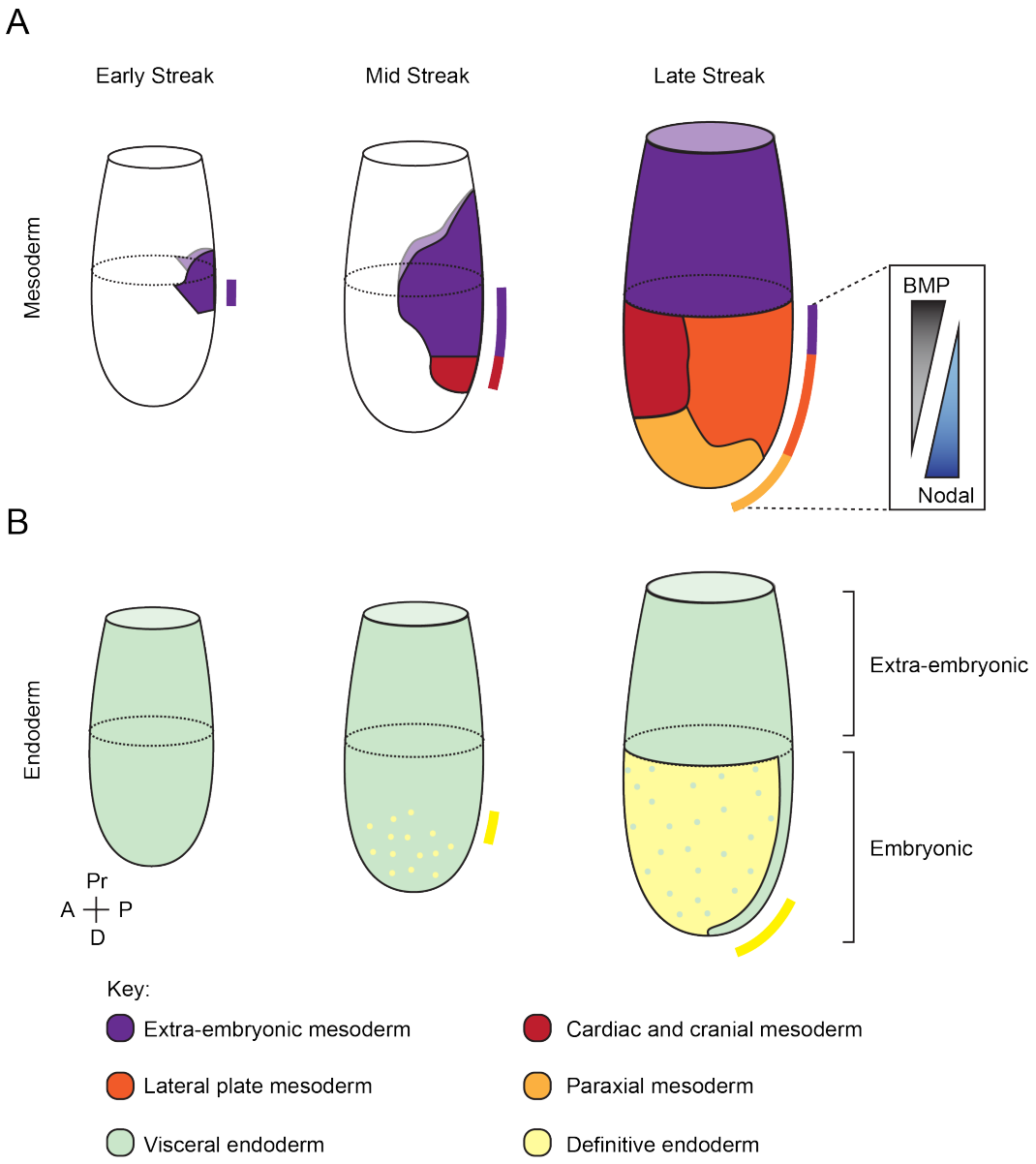


Figure 1.4: Lineage allocation in the primitive streak.
Representation of migrating mesoderm and intercalating definitive endoderm (DE) germ layers during gastrulation as they emerge from different positions within the primitive streak (coloured bar). (A) The first mesoderm cells to migrate through the early-streak are extra-embryonic mesoderm (purple) that are fated to become erythrocytes of the yolk sac blood islands. As gastrulation proceeds, the streak extends. The next population to exit the streak at more anterior level is the cardiac and cranial mesoderm. At late-streak stage, extra-embryonic, lateral plate and paraxial mesoderm populations are allocated at posterior, intermediate, and anterior regions, respectively. Levels of BMP and Nodal signalling pattern the streak. (B) definitive endoderm is specified in the anterior primitive streak, by high levels of Nodal signalling. The definitive endoderm migrates, intermingled with mesoderm, and intercalates onto the surface of the embryo, replacing the overlying visceral endoderm. Modified from Kinder et al. (2001) and Kwon et al. (2008).

1.2.3.1 Allocation of haematopoietic and cardiac lineages

Haematopoiesis, the generation of blood cells, occurs at different times and at distinct sites during development (Jagannathan-Bogdan and Zon, 2013). The first wave of haematopoiesis, so called primitive haematopoiesis, takes place in the yolk sac at approximately E8.0 (Baron, 2013). Extra-embryonic mesoderm emerging at early and mid-streak stages migrates to the yolk sac, and differentiates to form the blood islands (Kinder et al., 1999). The endothelial cells and erythrocytes within the blood islands are thought to arise from a bipotent haemangioblast. Consistent with this idea, *in vitro* differentiation assays demonstrate mESCs and isolated E7.5 embryonic cells differentiate into both haematopoietic and endothelial cell types (Choi et al., 1998; Huber et al., 2004). On the other hand, *in vivo* clonal-analysis and lineage tracing experiments indicate these populations do not share a common progenitor (Kinder et al., 1999; Padron-Barthe et al., 2014).

Cardiac mesoderm emerges slightly later than yolk sac mesoderm from the primitive streak (Kinder et al., 2001). Recent work in mESC differentiation system has demonstrated that the timing of *Mesp1* (a bHLH transcription factor) expression patterns the immature mesoderm in a context dependent fashion. Early expression of *Mesp1* specifies haematopoietic cell types, whereas later or longer expression specifies cardiac fates (Chan et al., 2013). One possibility is that *Mesp1* controls mesoderm fate allocation in the embryo by controlling the timing of EMT in the primitive streak via regulation of *E-cadherin* (Lindsley et al., 2008).

Within the cardiac lineage, *Mesp1* is a direct target of the T-box transcription factor Eomes (Costello et al., 2011; van den Ameele et al., 2012). However, another T-box transcription factor, Brachyury, also has the ability to induce *Mesp1* expression in the cardiac lineage (David et al., 2011). Which of these T-box family members controls

Mesp1 in the haematopoietic lineage has yet to be elucidated.

1.2.4 Nodal signalling pathway in early embryonic patterning

Nodal, a TGF-β family member, binds Type I and Type II serine threonine kinase receptors, which in turn phosphorylate the downstream effectors; Smad2 and Smad3. Phosphorylated Smad2/3 transcription factors, in association with the co-smad, Smad4, translocate to the nucleus and activate target genes containing the consensus sequence CAGAC (SBE: Smad Binding Element) (Wu and Hill, 2009). Smad specificity and function is dependent on interactions with transcriptional co-activators (Schmierer and Hill, 2007). For example, Smad2/3 associate with DNA-binding transcription factors such as FoxH1, MIX/BIX family members and Eomes during endoderm differentiation (Kim et al., 2011; Labbe et al., 1998; Pereira et al., 2011; Teo et al., 2011). Additionally, they also recruit cofactors, such as the chromatin modifiers p300 (histone acetylase) and Jmjd3 (histone demethylase) to activate target genes (Dahle et al., 2010; Feng et al., 1998; Gaarenstroom and Hill, 2014; Kartikasari et al., 2013).

Genetic experiments have shown that *Nodal* is essential for proximal-distal and anterior-posterior axis specification and mesendoderm formation (Brennan et al., 2001). Furthermore, dose-dependent Nodal signals pattern the mesoderm and endoderm along the streak, with the highest levels required to specify anterior primitive streak progenitors (Vincent et al., 2003). Nodal function in embryonic patterning is highly conserved amongst vertebrates (Shen, 2007).

Dose-dependent Nodal activities within the epiblast are tightly regulated through a combination of protease cleavages and feedback loops. The cleavage by ExE

proprotein convertases Pace4 and Furin releases the mature active ligand (Mesnard et al., 2011). Nodal in turn, increases Pace4 and Furin expression in the neighbouring ExE. This positive-feedback loop ensures highest levels of Nodal signalling are restricted to the posterior, proximal region of the embryo. Additionally, this regulatory interaction enhances BMP4 signalling from the ExE, and consequently induces *Wnt3* expression in the posterior epiblast. Subsequently, Wnt3 maintains *Nodal* expression in the primitive streak (Ben-Haim et al., 2006). Furthermore, Nodal itself via its FoxH1-dependent auto regulatory enhancer reinforces this signalling pathway (Norris et al., 2002). These various mechanisms are crucial for setting up graded Nodal signalling responsible for embryonic patterning.

1.3 The T-box family of transcription factors

The T-box family of transcription factors containing a common DNA binding domain, the T-box, which is conserved among metazoans (Minguillon and Logan, 2003). Besides the founding member, *Brachyury* (also known as *T*, for "short tail"), additional T-box genes were subsequently discovered based on sequence homology (Bollag et al., 1994). T-box transcription factors function as activators or repressors depending on context (Naiche et al., 2005).

The T-box binds a consensus DNA sequence termed the T-half site, although individual family members have different preferences for the number, spacing and orientation that they bind (Sinha et al., 2000). Specificity appears to reside in subtle differences within the T-box (Conlon et al., 2001), which allows preferences for certain target genes.

Interactions via N- and C- domains with other proteins and co-factors further increase

T-box binding specificity as well as influence gene activation. T-box transcription factors can form complexes with other transcription factors (e.g. homeodomain, GATA zinc finger, and LIM domain proteins) and recruit chromatin modifiers to activate target genes (Naiche et al., 2005). For example, Tbx5 and Gata4 interact during cardiac development, which when disrupted causes cardiac septal defects (Garg et al., 2003). Moreover, T-bet, Brachyury, Eomes and Tbx5 are all able to interact with the histone methyltransferase RbBp5 and de-methyltransferase Jmjd3 when co-transfected in EL4 (lymphoma) cells (Miller et al., 2008).

Furthermore, T-box factors have been shown to interact with components of the TGF-β signalling pathway to specify cell fates during gastrulation. For example, Eomes directly complexes with Smad2 to activate mesoderm genes in *Xenopus* (Picozzi et al., 2009). Through this mechanism T-box factors can specify different fates within the primitive streak in a context dependent fashion, determined by the levels of BMP (Smad1, mesoderm) or Nodal (Smad2, definitive endoderm) signalling (Bernardo et al., 2011; Costello et al., 2011; Faial et al., 2015; Teo et al., 2011).

T-box family members play essential roles in early embryogenesis. Mutational analysis has revealed that T-box family members are involved a variety of important developmental processes including, but not limited to: A-P (*Eomes*) and L-R axis specification (*Brachyury* and *Tbx6*), somitogenesis (*Brachyury*, *Tbx6* and *Tbx18*), cardiogenesis (*Eomes, Tbx1, Tbx2, Tbx3, Tbx5, Tbx18* and *Tbx20*), and vasculogenesis in the extra-embryonic yolk sac (*Tbx3*) and allantois (*Brachyury* and *Tbx4*) (Naiche et al., 2005; Papaioannou, 2014; Showell et al., 2004; Wardle and Papaioannou, 2008). During gastrulation in mouse, two T-box factors *Eomes* and *Brachyury* are expressed within the primitive streak at early and mid-streak stages and both regulate EMT, and specify mesoderm. *Eomes* also specifies endoderm in the

primitive streak **(Figure 1.5)**. *Tbx6* is expressed in the primitive streak at late-streak stages and specifies paraxial mesoderm.

Interestingly, ChIP-seq has revealed that different T-box family members often bind the same genomic loci during mesoderm and mesendoderm development. In *Xenopus*, the T-box transcription factors Eomes, Xbra and VegT act collaboratively to specify mesoderm. These T-Box TFs are recruited to the same genomic loci during gastrulation and individually they are able to directly activate a common set of mesoderm genes (Gentsch et al., 2013). In humans, ChIP-seq binding sites of BRACHYURY and EOMES largely overlap at both mesoderm and endoderm genes in differentiating human ESCs (Faial et al., 2015). In mouse, the similarity of T-box factor occupancy during mesoderm specification has not been assessed genome-wide. Although, T-half sites at the mouse *Mesp1/2* loci can be occupied by Eomes, Brachyury or Tbx6 (Costello et al., 2011; David et al., 2011; Yasuhiko et al., 2006).

Sequential activities of T-box transcription factors during mesoderm development implicate this family as key components of the gene regulatory networks that organise mesoderm migration, specification and patterning (Papaioannou, 2014). Indeed, this appears to be an evolutionary conserved role of T-box genes (Wardle and Papaioannou, 2008).

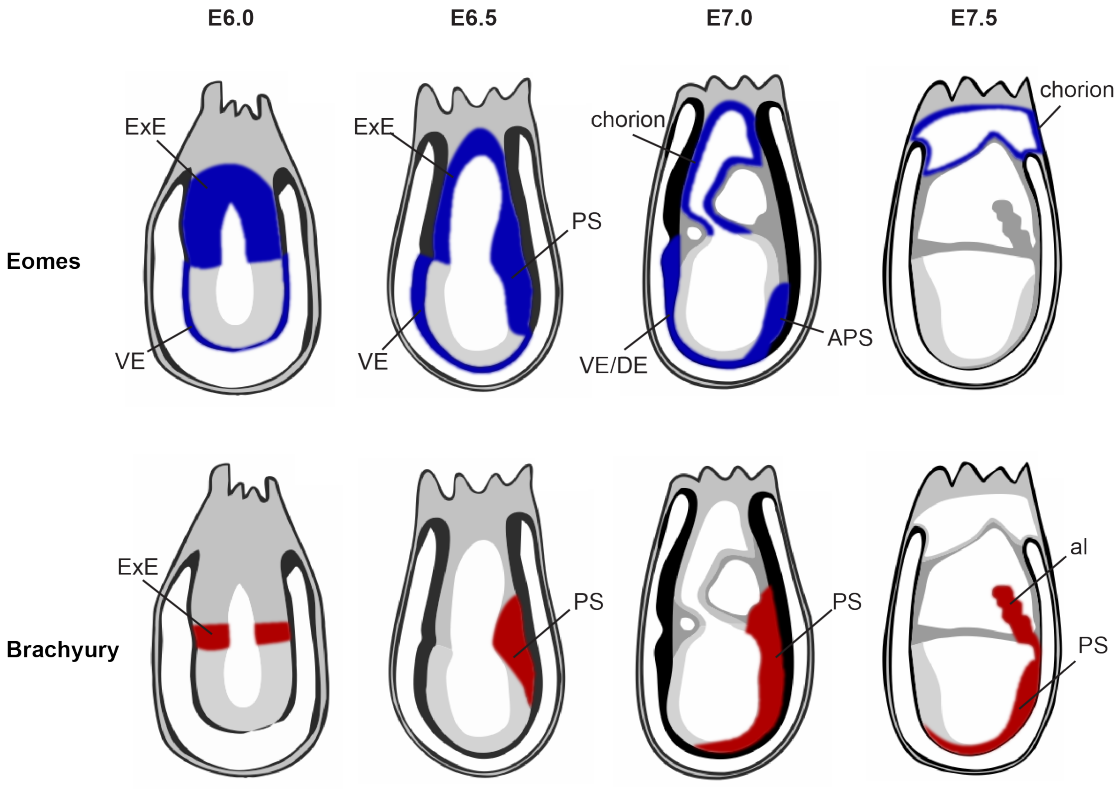


Figure 1.5: The T-box transcription factors, *Eomes* and *Brachyury*, exhibit partially overlapping expression patterns during mouse gastrulation.

Eomes is expressed in the extra-embryonic ectoderm (ExE) and chorion and the visceral endoderm (VE) during gastrulation. At E6.5 *Eomes* is expressed in the primitive streak (PS), becoming restricted to the anterior primitive streak (APS) and newly emerged definitive endoderm (DE) at E7.0. By E7.5 *Eomes* is no longer detected in the embryonic region, and expression only persists in the chorion. *Brachyury* is initially expressed in a radial ring at the extra-embryonic-embryonic boundary before overt streak formation at E6.0. Brachyury is expressed in the primitive streak throughout gastrulation, and in the allantoic bud (al).

1.3.1 Brachyury expression and function

The founding member of the T-box family *Brachyury,* is initially expressed radially at the extra-embryonic-embryonic boundary prior to overt streak formation. During gastrulation, *Brachyury* is strongly expressed in the primitive streak, as well as the node, notochord, and allantois **(Figure 1.5)** (Beddington et al., 1992; Inman and Downs, 2006b). Wnt3 followed by Wnt3a signals are required to activate and maintain *Brachyury* expression within the primitive streak. Expression is restricted to the primitive streak in part because inhibitory factors produced by the anterior visceral endoderm block the spread of Wnt signalling to the anterior epiblast (Rivera-Perez and Magnuson, 2005; Yamaguchi et al., 1999).

Heterozygous *Brachyury* mutants have truncated tails; giving the gene the name "short tail" abbreviated to T. Homozygous *T* mutants fail to develop a functional allantois and consequently lack proper vasculature and fail to form a chorioallantoic placenta. The embryos die at E10.5 due to malnutrition (Inman and Downs, 2006a). Additionally, the node, notochord and posterior mesoderm are defective. The severity of the mesodermal abnormalities varies depending on the specific mutation. The failure to maintain the node and notochord in *T* mutants causes patterning defects of the neural tube, posterior somites (which are absent or abnormal beyond the 7th somite pair) and L-R abnormalities (Conlon et al., 1995; King et al., 1998). T^{wis} and T^{c} mutants both produce a truncated protein with dominant negative activities, leading to severe defects that extend more anteriorly in comparison with the original *T* mutation (Beddington et al., 1992).

As for *Eomes* epiblast conditional mutants, which display a block in EMT due to the absence of *Mesp1* expression (Arnold et al., 2008; Costello et al., 2011; van den Ameele et al., 2012), in *T* mutants, the primitive streak becomes thickened because

cells lack the ability to migrate away from the streak (Wilson et al., 1995). A similar mechanism to promote EMT may also exist for *Brachyury* (David et al., 2011).

Brachyury target genes include genes involved in cardiac mesoderm formation (*Mesp1*), paraxial and somitic mesoderm (*Cdx2, Mesp2, Msgn1, Tbx6* and *Axin2*) and node and notochord function (*Cer1, Gsc, Mixl1* and *Lefty1*) (Faial et al., 2015; Lolas et al., 2014). Additionally, *Fgf8* is required for cell migration within the primitive streak, and is a direct target of Brachyury (Lolas et al., 2014; Sun et al., 1999).

Chimeric analyses demonstrate that *Brachyury* acts in a cell autonomous and dose dependent fashion controlling the emergence of nascent mesoderm cells from the primitive streak (Beddington et al., 1992; Rashbass et al., 1991; Wilson and Beddington, 1997). Absence of functional *Brachyury* protein causes the cells to remain in the streak. Whereas, increased levels of *Brachyury* expression result in a premature exit from the streak. Graded *Brachyury* expression seems to balance tissue expansion and maintenance of progenitor cell populations within the primitive streak (and later in the tailbud) to ensure complete axial elongation (Neijts et al., 2014; Wilson and Beddington, 1997).

1.3.2 Eomes expression

Eomes (also known as *T-brain 2: Tbr2*) expression has been characterised in several extra-embryonic and embryonic tissues during development. Firstly, in extra-embryonic lineages *Eomes* is expressed at the blastocyst stage and is restricted to the trophectoderm (TE), and persists within the extra-embryonic ectoderm (ExE) at post-implantation stages and in the chorion at late-streak stages (Ciruna and Rossant, 1999). *Eomes* is also strongly expressed in the visceral endoderm overlying the

epiblast E5.5 (Arnold et al., 2009; Ciruna and Rossant, 1999; Nowotschin et al., 2013).

In the embryo proper, *Eomes* is initially expressed in the proximal-posterior region of the epiblast, prior to overt streak formation. *Eomes* is transiently expressed in the early primitive streak **(Figure 1.5)**, and at late-streak stages becomes restricted to the anterior primitive streak, nascent mesoderm and definitive endoderm. Slightly later, expression becomes undetectable (Arnold et al., 2008; Ciruna and Rossant, 1999; Russ et al., 2000). Fate-mapping studies reveal that descendants of *Eomes* expressing cells reside in the cardiac mesoderm, vasculature, head mesenchyme, and definitive endoderm (Costello et al., 2011).

In addition, *Eomes* transcripts are present in the retina, forebrain, limb and olfactory lobes at E14.5 (Mao et al., 2008; Russ et al., 2000). In the adult, *Eomes* is expressed in NK cells and CD8+ T cells of the adaptive immune system and intermediate progenitor cells (IPC) in the hippocampus (Gordon et al., 2012; Hodge et al., 2008; Pearce et al., 2003).

1.3.3 Eomes functional requirements in the early embryo

Eomes null mice die at around the time of implantation due to defective trophectoderm development. However, heterozygotes are viable (Russ et al., 2000). Conditional ablation of *Eomes* expression in the epiblast results in a thickening of the primitive streak, due to a block in epithelial-to-mesenchymal transition (EMT). *Eomes* promotes EMT by activation of *Mesp1/2* and down regulating E-cadherin (Arnold et al., 2008; Costello et al., 2011). These mutant embryos lack a node, cardiac mesoderm and definitive endoderm. The loss of definitive endoderm and cardiac mesoderm

reflects essential cell-autonomous requirements during specification of these cell lineages (Arnold et al., 2008; Costello et al., 2011; Russ et al., 2000; Teo et al., 2011; van den Ameele et al., 2012).

In the context of low levels of Nodal signalling within the early streak, Eomes directly activates *Mesp1/2* to specify cardiac fates (Costello et al., 2011; van den Ameele et al., 2012). However, at late-streak stages, in the context of high levels of Nodal signalling within the anterior primitive streak, Eomes specifies definitive endoderm and inhibits mesoderm fates. Eomes in association with Smad2 directly activates definitive endoderm genes (e.g. *Mixl1, Foxa2, Sox17* and *Gsc*) and further potentiates *Eomes* expression (Izumi et al., 2007; Teo et al., 2011).

Conditional ablation of *Eomes* in the visceral endoderm results in constriction at the extra-embryonic and embryonic boundary and the failure of A-P axis and primitive streak formation, reflecting *Eomes'* requirement for correct AVE specification and migration, in part due to activation of *Lhx1* (Nowotschin et al., 2013).

1.3.4 Regulation of Eomes

Pluripotency transcription factors exist in precarious balance with each other, inhibiting differentiation to particular lineages whilst specifying others (Loh and Lim, 2011). In the epiblast, mESCs and hESCs, pluripotency factors Oct4, Sox2 and Nanog transcriptionally repress *Eomes*, hence preventing overt trophoblast, mesoderm and endoderm differentiation. (Teo et al., 2011; Yuan et al., 2009). Paradoxically, in hESCs it has been shown that pluripotency factors NANOG, OCT4 and TBX3 initiate *Eomes* expression at the onset of differentiation (Beyer et al., 2013; Kartikasari et al., 2013; Teo et al., 2011).

Nodal is required for *Eomes* expression within the primitive streak and in the ExE at gastrulation (Brennan et al., 2001; Guzman-Ayala et al., 2004). Compound *Eomes* and *Nodal* mouse mutants have a more severe phenotype than single heterozygous mice (which are phenotypically normal), demonstrating that they interact genetically during gastrulation (Arnold et al., 2008). In *Xenopus*, the Nodal agonist Activin strongly induces *Eomes* expression via an Activin response element (ARE) in its promoter (Ryan et al., 1996). Nodal signalling also activates *Eomes* expression in human definitive endoderm cultures (Brown et al., 2011).

Further evidence that *Eomes* is regulated by Nodal signalling in multiple lineages comes from analysis of *Smad2* mouse mutants. In chimeric embryos *Smad2* null mESCs contribute to all mesodermal lineages, but are unable to colonise the definitive endoderm (Tremblay et al., 2000). Consistent with this, in mouse definitive endoderm cultures, *Smad2* null cells have severely reduced levels of *Eomes* expression (Costello et al., 2011). Interestingly, in human definitive endoderm cultures Smad2 but not Smad3 mediates Nodal dependent *Eomes* expression (Brown et al., 2011).

Smad2 null E5.5 embryos lack *Eomes* expression in the embryonic visceral endoderm. Similarly, when *Smad2* null mESCs are converted by Gata6 overexpression to extra-embryonic endoderm (XEN) like cells, *Eomes* induction is significantly decreased relative to wild type (Nowotschin et al., 2013). Thus, Nodal signalling is essential for *Eomes* expression in the visceral endoderm.

Not only is *Eomes* a downstream effector of Nodal signalling, but Eomes also acts co-operatively binding with Smad2 to regulate key gene expression, thus forming a coherent feed-forward loop to ensure robust commitment of cell fate. In *Xenopus*, a

similar gene regulatory network has been described involving *Nodal* and *Eomes* for the specification of paraxial mesoderm (Fukuda et al., 2010).

1.4 Developmental enhancers

Developmental enhancers direct tissue-specific gene expression during embryogenesis. These cis-acting regulatory elements tend to be several hundred base pairs in length and contain multiple transcription factor binding sites (Lagha et al., 2012) . Enhancers function as information hubs; integrating signalling cues and cell-type specific transcription factors to produce different transcriptional outputs (Buecker and Wysocka, 2012). Transcription factor complexes bound to enhancers recruit a range of co-activators such as histone modifiers (e.g. histone acetylase p300), chromatin remodelling factors (e.g. BAF complex) or proteins that act as a bridge to the basal transcriptional machinery (e.g. Mediator) (Fong et al., 2012; Ong and Corces, 2011).

Over short distances enhancers can directly recruit activators to initiate transcription due to their close proximity to the promoter. However, over longer distances, enhancer-promoter interactions depend on additional factors such as CTCF and Cohesin that mediate DNA looping (Fong et al., 2012), to bring together the enhancer and promoter regions, looping out the intervening DNA (Bulger and Groudine, 2011; Kulaeva et al., 2012). Recent evidence showing physical interactions between both promoter and enhancers has been provided by chromatin conformation capture techniques such as 3C (de Wit and de Laat, 2012).

The status of an enhancer; active, poised or inactive, is often tissue specific. Active enhancers driving gene expression are bound by PolII, cofactors, and tissue specific

transcription factors and are typically depleted of nucleosomes, and therefore particularly sensitive to DNA nucleases. Furthermore, active histone marks H3K4me1 and H3K27ac and the presence of short non-coding RNA, termed enhancer RNA (eRNA), can further identify active enhancers (Buecker and Wysocka, 2012). Inactive enhancers silenced by transcriptional repressors are devoid of the active H3K27ac mark and instead are characterised by the polycomb mediated H3K27me3 repressive mark.

Poised enhancers are bound by transcription factors and co-activators, but are not associated with PolII. They often display the histone mark H3K4me1 and DNaseI hypersensitivity, but are distinguished from active enhancers by H3K27me3 instead of H3K27ac (Creyghton et al., 2010). Poised enhancers are particularly important during development, when cell-type specific genes are expressed rapidly in response to extrinsic signalling cues to initiate differentiation (Lagha et al., 2012). Enhancer "priming" may also play a role in the conversion of an inactive to a poised state. In this case, a transcription factor binds repressed heterochromatic chromatin to open up the DNA and promote binding of additional factors. For example, Foxa1 acts as a pioneer factor at an *Afp* enhancer during ESC differentiation, displacing histones to provide access to Smad2/4 for subsequent activation of *Afp* expression (Taube et al., 2010).

Enhancer sequences are sufficient to drive expression of a reporter gene from a minimal promoter in the context of transgenic mice or alternatively in transfection assays using luciferase reporter constructs (Yeom et al., 1996). Targeted enhancer deletions definitively demonstrates the ability to direct correct spatiotemporally restricted expression in vivo (Norris and Robertson, 1999).

1.4.1 Eomes enhancers

Mouse transgenic studies analysing highly conserved regions have characterised cis-acting regulatory elements; a 2.8kb limb and retinal enhancer active at E14.5 located 3.5kb upstream of the transcriptional start site (TSS) (Mao et al., 2008), a forebrain enhancer at E10.5 located 160kb upstream of the TSS (Pernaute et al., 2010b) and a primitive streak enhancer 23kb upstream of the TSS (Wolf et al., 2012) are sufficient for driving tissue specific *Eomes* expression. However, functional consequences resulting from loss of these enhancers remain unknown. Moreover candidate enhancers that regulate the tissue specific *Eomes* expression in the visceral endoderm or extra-embryonic ectoderm have yet to be characterised.

In humans, much insight into the molecular mechanisms governing *EOMES* enhancer-promoter activation has been achieved by utilising ChIP-seq during mesendoderm differentiation of hESCs. In hESCs, *EOMES* is not expressed, the promoter is bivalent (H3K4me3, H3K27me3), and a widely studied enhancer 6.2kb enhancer upstream of the EOMES TSS is in a poised state (H3K27me3, H3K4me1) (Rada-Iglesias et al., 2011). In pluripotent cells, NANOG, OCT4, SOX2, SMAD2, TEAD4 and YAP bind this enhancer region (Beyer et al., 2013; Teo et al., 2011). YAP prevents POLII elongation of at the *EOMES* promoter (Estaras et al., 2015). Additional inhibitory complexes (e.g. PRC2 or NuRD) are probably also present to repress transcription.

Upon mesendoderm differentiation, YAP, TEAD4 and SOX2 binding are lost, but SMAD2, OCT4 and NANOG binding persists. Additional recruitment of FOXH1, JMJD3 and β-CATENIN promotes enhancer-promoter interactions (Funa et al., 2015; Kartikasari et al., 2013; Kim et al., 2011). Recruitment of the histone demethylase JMJD3 closely correlates with loss of the H3K27me3 repressive mark at the promoter

region. SMAD2 promotes POLII mediated transcriptional elongation, initiating expression (Estaras et al., 2015). Slightly later, EOMES itself in a complex with SMAD2 binds to the enhancer to reinforce expression in an auto-regulatory positive feedback loop (Teo et al., 2011).

The corresponding human enhancer region can drive expression of luciferase reporter constructs in cell culture and transgenic assays in the shield region of *Zebrafish* (analogous to the primitive streak) (Estaras et al., 2015; Rada-Iglesias et al., 2011; Teo et al., 2011). An analogous enhancer region 7kb upstream of the mouse *Eomes* TSS has also been characterised and shown by 3C to interact with the *Eomes* promoter during definitive endoderm differentiation (Kartikasari et al., 2013). Additionally, Jmjd3 and Tbx3 bind to this region upon mESC differentiation. The enhancer-promoter interactions are Tbx3 and Jmjd3 dependent. However, this enhancer was not tested for its ability to drive reporter expression.

1.5 Aims of thesis

Developmental genes often have multiple cis-regulatory elements that may act together, or redundantly, to ensure robust and correct temporally and spatially restricted gene expression patterns (Lagha et al., 2012). *Eomes* lies within a 300kb gene desert, which potentially contains numerous regulatory regions. In this study, my research has exploited knowledge of *Eomes* enhancer activation discovered in human mesendoderm differentiation to identify additional regulatory elements that act during mouse gastrulation. The tissue specific activity of the enhancers was examined by transgenic analysis in mouse embryos. Targeted deletions of these enhancer regions have definitively tested whether these regions are necessary for correct *Eomes* expression during mouse development.

Eomes is one of the earliest markers expressed within the primitive streak and is known to be essential for EMT. However, a possible functional role in additional mesodermal lineages (e.g. haematopoietic) has yet to be investigated. I aimed to generate an *Eomes* improved Cre recombinase allele to investigate contributions made by *Eomes* expressing cells in additional mesodermal cells types and for conditional ablation experiments within the trophectoderm, visceral endoderm and primitive streak tissues. Additionally, I utilised *Eomes* conditional mutant embryos and null mESCs to discover the functional role in haematopoietic mesoderm fate allocation.

Eomes and *Brachyury* are T-box transcription factors that exhibit partially overlapping expression within the primitive streak. Whether they act collaboratively or redundantly with each other during gastrulation in mouse remains an unresolved issue. I performed gene-swapping experiments to directly test functional redundancy of these T-box factors.

Overall, I aimed to expand our knowledge of the cis-regulatory elements governing *Eomes* expression during mouse gastrulation and the functional roles played by *Eomes* in mesoderm fate allocation within the streak. Greater understanding of lineage decisions in the embryo can be used to improve *in vitro* differentiation protocols, which will be of great use to model diseases and design effective cell therapies.

Chapter 2: Materials and Methods

2.1 Materials

Crude lysis buffer	50mM KCl, 10mM Tris-HCl pH 8.3, 1.5mM $MgCl_2$, 0.45% NP40, 0.45% Tween-20, 0.5mg/ml Proteinase K
DEPC-PBT	1x PBS, 0.01% DEPC, 0.01% Tween-20
MAB buffer	100mM Maleic acid, 150mM NaCl, 0.1% Tween-20
Modified Church and Gilbert buffer	500mM sodium phosphate, pH 7.2, 1mM EDTA, 250mM NaCl, 5% SDS, 1% BSA, 100µg/ml denatured fish sperm DNA
Southern blot DNA lysis buffer	100mM Tris, pH7.5, 50mM Na_2EDTA, 0.5% SDS, 1mg/ml Proteinase K
Southern blot denaturation solution	1.5M NaCl, 500mM NaOH
Southern blot depurination solution	120mM HCl
Southern blot low stringency buffer	2x SSC, 1% SDS
Southern blot medium stringency buffer	0.5x SSC, 0.1% SDS
Southern blot high stringency buffer	0.3x SSC, 0.1% SDS

WISH alkaline phosphatase (AP) buffer	100mM NaCl, 100mM Tris-HCl pH 9.5, 50mM $MgCl_2$, 0.1% Tween-20
WISH hybridisation buffer	50% Formamide, 5x SSC, 0.1% Tween-20, 0.1% SDS, 50µg/ml Heparin in 4x SSC, 50µg/ml yeast tRNA, 60mM Citric acid
WISH wash buffer I	50% Formamide, 5x SSC, 60mM Citric acid, 5% SDS
WISH wash buffer II	50% Formamide, 2x SSC, 24 mM Citric acid, 1% SDS, 0.01% Tween-20
X-gal buffer	1x PBS, 5mM EGTA, 2mM $MgCl_2.6H_2O$, 0.04% NP40, 0.01% Deoxychoic acid
X-gal fixative solution	1% Formaldehyde, 0.2% Glutaraldehyde in X-gal buffer
X-gal stain solution	5mM K_3Fe, 5mM K_4Fe, 0.5 mg/ml X-gal in X-gal buffer

2.2 Molecular biology

2.2.1 Molecular cloning

1 – 1.5 µl of plasmid DNA was transformed into DH5α chemically competent E. Coli (18265-017; Invitrogen) according to the manufacturer's instructions. Bacteria were gown in LB medium (L7275; Sigma) with 50µg/ml ampicillin (A0166-5G; Sigma). Purification of plasmid DNA was carried out with QIAprep Spin Miniprep Kit (27106; Qiagen) or HiSpeed Plasmid Maxi prep kit (12662; Qiagen) according to the manufacturer's protocol. DNA concentration was quantified using a Nanodrop ND 1000 spectrophotometer (Thermo Scientific).

DNA from PCR and restriction endonuclease digestions of plasmids were cut out from 0.8% - 1% agarose gel after electrophoresis and purified using the High Pure PCR Product Purification kit (11732676001; Roche). Cut DNA inserts were ligated into a linearised plasmid vector using the Fast-Link DNA Ligation kit (LK6201H; Epicentre). 50-200ng of vector DNA was added to an excess of insert DNA, in a molar ratio of 3:1 insert:vector. The DNA mixture was then incubated with 1.5µl 10x Fast-Link Ligation Buffer, 1.5µl 10mM ATP, 1µl Fast-Link DNA ligase and made up to a total volume of 15µl with H_2O for 30 mins at room temperate. This was followed by 15 min incubation at 70°C to heat inactivate the DNA ligase. PCR fragments were ligated into pCR-XL-TOPO using the TOPO-XL PCR Cloning kit (45-0008; Invitrogen) according to the manufacturer's protocol.

All restriction endonucleases were obtained from New England Biolabs. Oligonucleotides for PCR reactions and linker regions were obtained from Life Technologies.

2.2.2 Targeting vectors and transgenic constructs

2.2.2.1 EH3 targeting vector

The EH3 targeting vector was generated by recombineering by Pia Rantakari. The targeting vector includes a 5' homology region of a 3.6 kb fragment of the *Eomes* locus that includes Exon 6 and a 3' homology arm of 5.1 kb fragment of the *Eomes* locus. Primers used to recombine EH3 with a FLP recognition targeted (FRT) flanked PGK.Neomycin selection cassette were 5'-GGC TGG GGT TGG GGA AGG AGT GTT TGC CCT GGA GAT GCA AGA TTG TGC TCG GAT CCA ATT AAC CCT CAC TAA AGG GC-3' and 5'-GGT CCC AGA AGT TTG GAG GAC GGG AAA GAC TGT CCA CAG CTC AGG TAT ATC GAA GTT ATA AGC TTG AAG TTC CTA TAC TTT C -3', therefore deleting 656bp EH3. Outside of the 3' homology arm of this construct was PGK.DTA (diphtheria toxin A) cassette for negative selection.

2.2.2.2 EH1b targeting vector

The EH1b targeting vector was cloned into a pBSIIKS(-) plasmid (Stragene) and includes a 5' homology region 5.8kb SpeI-EcoRV fragment and a 3' homology region 5kb KpnI-EcoRI fragment of the *Eomes* locus, thereby deleting 2kb of EH1b. The homology arms surround a FRT flanked PGK.Neomycin positive selection cassette integrated in place of the enhancer. A PGK.DTA cassette for negative selection was inserted outside of the 5' arm homology region.

2.2.2.3 $Eomes^{iCre}$ targeting vector

The targeting vector for the generation of the $Eomes^{iCre}$ targeted allele comprises a 8.25kb HpaI–HpaI fragment of the *Eomes* locus (Arnold et al., 2008). An iCre β -globin polyA cassette (Shimshek et al., 2002) and a LoxP-flanked PGK.Neomycin positive resistance cassette were integrated between the SphI site at the translational

start site and an EagI site, thereby deleting 500bp of the 5'coding region of Exon 1. The 3' homology region of the targeting construct was flanked with a pMCI.TK (thymidine kinase) negative selection cassette.

2.2.2.4 *Eomes*^{T} targeting vector

Brachyury (*T*) cDNA was cloned from RNA isolated from day 3 embryoid bodies (EBs). RNA was reverse transcribed to cDNA using Superscript III First Strand Synthesis System (18080-051; Life Technologies). The targeting vector for the generation of the *Eomes*^{T} targeted allele comprises a 8.25kb HpaI–HpaI fragment of the *Eomes* locus (Arnold et al., 2008). *T* cDNA with a SV40 poly A tail and a LoxP-flanked PGK.Neomycin positive resistance cassette were integrated between the SphI site at the translational start site and an EagI site, thereby deleting 500bp of the 5'coding region of Exon 1. The 3' homology region of the targeting construct was flanked with a pMCI.TK negative selection cassette.

2.2.2.5 Linearisation of targeting vectors

Targeting vectors were linearised by digesting 100µg of plasmid DNA with 500U of the appropriate restriction endonuclease: EH1b with XhoI, EH3 with ApaLI, *Eomes*^{iCre} and *Eomes*^{T} with NotI and digested for 2 hours at 37°C, followed by phenol:chloroform extraction and ethanol precipitation of the DNA. Precipitated DNA was resuspended in Tris buffer, to a concentration of 1µg/µl.

2.2.2.6 EH1 and EH3 transgenic constructs

EH1 (4.6 kb HincII – KpnI fragment of *Eomes* locus) or EH3 (696bp amplified by PCR with F -GCCCTGGAGATGCAAGATTG and R-CAGCTCAGGTATATCTTCTGGC primers) sequences were cloned into a

pBSIIKS(-) plasmid (Stragene) upstream of a hsp68 promoter, LacZ cassette and SV40 polyA signal (Sasaki and Hogan, 1996). The plasmid was then linearised with NotI and purified by the transgenics core (Wellcome Trust Centre for Human Genetics, Oxford, UK) before microinjection into 2-cell stage blastocyst. The following day, the transgenic embryos were transferred into E0.5 pseudopregnant females. Embryos were collected at E6.5-E7.5 and analysed by X-gal staining. Alternatively, adult positive transgenic mice were identified by PCR and used to found stable transgenic mouse lines.

2.2.3 DNA and RNA isolation from cells and tissues

RNA was isolated from ESCs and embryoid bodies (EBs) using Qiashredder homogoniser (79654; Qiagen), RNeasy mini kit (74104; Qiagen) and RNase-Free DNase Set (79254; Qiagen) according to the manufacturer's protocol.

DNA was extracted from putatively targeted ESC clones in 96-well plates as described (Behringer et al., 2013). Briefly, 50µl of Southern Blot DNA lysis buffer was added per well and incubated 60°C in a humid chamber overnight. The next day, DNA was precipitated by the addition of 100µl of 75mM NaCl in 98.5% ethanol per well for 30 mins at room temperature. The precipitated genomic DNA was then washed 3 times with 70% ethanol and dried overnight. Genomic DNA was then used immediately for Southern blot analysis, or stored at -20˚C until required.

To isolate DNA from ESCs grown on 6cm dishes, 500µl of Southern blot DNA lysis buffer was added per plate, incubated briefly at room temperature, before being transferred to a 1.5ml microcentrifuge tube and incubated at 55˚C overnight. The next day, DNA was extracted by phenol:chloroform and isopropanol precipitation and then

dissolved in 10mM Tris·Cl, pH 8.5 to a final concentration of 1μg/μl.

For PCR genotyping, ear notches were taken from weaned mice. For embryonic stages, yolk sac (E8.5 to E9.5) or ectoplacental cone (E6.5 to E7.5) was removed. Crude extraction of DNA was performed by the addition of 100μl (ear notches) or 30μl (yolk sac and ectoplacental cone) crude lysis buffer and incubation overnight at 55˚C. The following day, samples were heat inactivated by incubating at 95˚C for 15mins before genotyping by PCR.

2.2.4 Standard PCR

PCR genotyping was carried out using PCR ReddyMix PCR Mastermix (AB-0575; Thermo Scientific) and 1μl of each primer (0.5μM each final concentration) in a reaction volume of 15μl. 1μl crude DNA from tissues, or 0.2μl phenol:chloroform purified cell DNA was used as a template. PCR cloning was carried out using a high fidelity Platinum *pfx* DNA polymerase (11708; Invitrogen) according to the manufacturer's instructions. Reactions were performed on a T100 Thermal Cycler (Biorad) and cycling conditions were optimised for each primer pair. Primers ued in the study are provided in **Table 2.1**. *Eomes*^{CA}, *Eomes*^{N}, *Eomes*^{GFP}, Flk1-LacZ and Sox2-Cre mice were genotyped according to published protocols (Arnold et al., 2008; Arnold et al., 2009; Ema et al., 2006; Hayashi et al., 2002).

Eomes$^{\Delta EH3}$	F: TCGTTGAGTGGTGAGCAGGGAG R1: TTTGGAGGACGGGAAAGACTG R2: AGCGAGGACATCCACGGAAAAC Annealing temperature: 55˚C. Products: Wild type 369bp, ΔEH3 264bp.
Eomes$^{\Delta EH1b}$	F1: GGCTATTGCCTCCATACAGC F2: TTGCGTTTGTTGGGTTTTGG R: CCATCACTGGGAGAGTAGGC Annealing temperature: 56˚C. Products: Wild type 427bp, ΔEH1b 712bp.
Eomes^{T}	F1: TGTTGGTAAACAGCCCACAA F2: TGAGCTCCGAGCGCTACTAC R: GGTGGAATTTGAGCCATAGG Annealing temperature: 56˚C. Products: Wild type 429bp, *Eomes*^{T} 338bp.
Eomes^{iCre}	F: AAGGAAAGGGGCACCTACAATC R1: AGCACTCTCCAGCGAGTAGAAGTG R2: TTTCCTGTTGTTCAGCTTGC Annealing temperature: 55˚C. Products: Wild type 302bp, *Eomes*^{iCre} 440bp.
Transgenics	F: TTACCAGGCCGAAGCAGCGTTGTTG R: GCGGCAGTAAGGCGGTCGGGATAGT Annealing temperature: 59˚C. Product: LacZ 300bp.

Table 2.1: PCR genotyping primers

2.2.5 Quantitative real-time PCR

RNA was reverse transcribed to cDNA using Superscript III First Strand Synthesis System (18080-051; Life Technologies). Quantitative RT-PCR was then carried out on cDNA samples in triplicate using SYBR-green kit (204143; Qiagen) on a Rotagene cycler (Qiagen) with the primers listed in **Table 2.2**. Relative gene expression was normalised to *Gapdh* and calculated as $2^{\Delta\Delta Ct}$.

ER71_F	CAGAGTCCAGCATTCACCAC
ER71_R (Lee et al., 2008)	AGGAATTGCCACAGCTGAAT
Gapdh_F	CAATGACCCCTTCATTGACC
Gapdh_R (Matsumoto et al., 2006)	GATCTCGCTCCTGGAAGATG
Eomes_F	TGTTTTCGTGGAAGTGGTTCTGGC
Eomes_R (Costello et al., 2011)	AGGTCTGAGTCTTGGAAGGTTCATTC
Mesp1_F	ACCCATCGTTCCAGTACGC
Mesp1_R	AGCATGTCGCTGCTGAAGA
Brachyury_F	GCTGTGACTGCCTACCAGCAGAATG
Brachyury_R (Ginis et al., 2004)	GAGAGAGAGCGAGCCTCCAAAC
EomesT_F	CTTTATAACCAGGCCGCGG
EomesT_R	CTCGGTACTGCAGGCTCTT

Table 2.2: qRT-PCR primers

2.2.6 Riboprobe synthesis for in situ hybridisation

To synthesis in situ probes, 10μg of plasmid DNA was digested for 3 hours at 37˚C with the relevant restriction endonuclease. Linearised DNA was purified by phenol:chloroform extraction and ethanol precipitation. 1μl of DNA precipitAid (A6587; AppliChem) was added to help visualise the DNA pellet, which was then washed in 70% ethanol, air dried, and resuspended in 15μl RNase-free H_2O. DIG labelling of probes was carried out using DIG RNA labelling mix (11277073910; Roche) with T7 RNA polymerase (10881767001; Roche) or T3 RNA polymerase (11031163001; Roche) according to the manufacturer's instructions in a total volume of 40μl, for 2 hours at 37˚C. DNA was then removed by incubation with 1μl DNase (M610A; Promega) for 20mins at 37˚C. RNA probe was purified over mini Quick Spin RNA Columns (11814427001; Roche) before being stored at -80˚C.

2.2.7 Southern blot

Southern blot of putatively targeted ESC clones was carried out as described in standard protocols (Behringer et al., 2013). Endonuclease digestion of genomic DNA

was carried out by adding 50µl restriction enzyme mix: 1x recommended NEB Buffer (NEB), 1x BSA (NEB), 100µg/ml RNaseA (Qiagen), 1mM spermidine and 100 units of enzyme (NEB), per well and incubating overnight at 37˚C in a humid chamber, with shaking.

The next day, samples were run on a 0.8% agarose gel to separate DNA fragments. Following electrophoresis, the gel was incubated in depurination solution for 10 mins, and then denaturation solution for 30 mins. The DNA was transferred onto Hybond XL membrane (RPN303S; GE Healthcare) overnight by capillary action.

The following day, DNA was UV cross-linked using Stratagene UV-crosslinker. Before hybridisation, membranes were incubated with modified Church and Gilbert buffer for 2 hours at 65˚C. Probes were radioactively labelled using a Random Primers DNA Labelling System (18187013; Invitrogen). 50ng of probe DNA was dissolved in H_2O to a final volume of 23µl, heated to 95˚C for 5 min and immediately cooled on ice. To the denatured probe the following was added; 2µl each of dATP, dGTP and dTTP solutions, 1 µl Random Primers Buffer Mixture, 5µl (approximately 50 µCi) [α-32P]dCTP, 3000 Ci/mmol, 10 µCi/µl and 1 µl Klenow Fragment. The synthesis reaction was incubated for 2 hours at room temperature, before being denatured by heating to 95˚C for 5 min, then cooled on ice.

Radioactively labelled probe was hybridised to membranes in fresh modified Church and Gilbert buffer overnight at 65˚C. To wash the blots, the membranes were incubated at 65˚C in low, medium and high stringency buffers for 30, 15 and 10 mins, respectively. After washing, the radioactively hybridised blots were exposed to x-ray film overnight at -80˚C to visualise bands.

EH3 targeting in wild type CCEs was screened using BamHI digest with a 5'external probe (8kb wild type, 5kb targeted), by Li Li. EH3 targeting in heterozygous *Eomes*^{GFP} cells was screened using SpeI digest with a 5' external probe (25kb or 22kb wild type or ΔEH3 locus, 11kb GFP, 7kb GFP ΔEH3). EH1b targeting was screened using 5' external probe on KpnI digested DNA (15kb wild type, 10kb targeted). *Eomes*^{iCre} and *Eomes*^{T} were screened with a 5' external probe on EcoRV digested DNA (15 kb wild type, 10 kb *Eomes*^{iCre} and 11kb *Eomes*^{T}). Correctly targeted *Eomes*^{iCre} and *Eomes*^{T} clones were transfected with pMCI.Cre and the resulting subclones screened using an internal probe on EcoRV-digested DNA to detect the 6kb excised allele.

2.3 Tissue culture

2.3.1 mESC transfection

1x10^7 mESCs (CCE cell line 129sv) were transfected by electroporation of 20μg of linearised construct DNA at 200V, 950uF, infinite resistance, 4mm cuvette in a single pulse. 2.5x10^6 cells from the electroporation were plated per 10cm feeder plate. Transfected mESCs were grown for 10 days. G418 selection medium was added on day 2 to positively select for clones with Neomycin resistance. If the construct contained a TK cassette, FIAU was added from day 4 to negatively select clones containing a random integration of the construct. Transfected colonies were picked on day 10 in a 96-well format, to be screened by Southern blot.

2.3.2 mESC maintenance

Feeder-dependent mESCs were cultured in ES medium; DMEM (11960-085; Invitrogen) supplemented with 15% fetal calf serum (26140-079; Gibco), 1% Penicillin/Streptomycin (15140-163; Invitrogen), 0.1mM 2-mercaptoethanol (M-

7522; Sigma), 1% Glutamine (25030-13; Invitrogen) on a feeder layer of inactivated SNH fibroblasts (STO, Neomycin, Hygromycin resistant).

Feeder-independent ESCs were cultured in feeder-free ES medium; DMEM (11960-085; Invitrogen) with 15% fetal calf serum (26140-079; Gibco), 1% Penicillin/Streptomycin (15140-163; Invitrogen), 0.1mM 2-mercaptoethanol (M-7522; Sigma), 1% Glutamine (25030-13; Invitrogen), 1% MEM Non-essential amino acids (11140; GIBCO), 1mM Sodium Pyruvate (S8636; Sigma), 1000U/ml LIF (ESG1107; ESGRO) on gelatinised (Sigma) plates.

2.3.3 Culture of STOs and making feeder cells

SNH fibroblast were maintained in STO medium; DMEM (11960-085; Invitrogen) supplemented with 10% bovine calf serum (SH0072.03; Hyclone), 1% Glutamine (25030-13; Invitrogen) and 1% Penicillin/Streptomycin (15140-163; Invitrogen). SNH fibroblasts were mitotically inactivated by treatment with 20µg/ml Mitomycin C (M0503; Sigma) in STO medium for 3 hours at 37˚C. The feeder cells were then washed 3 times with PBS, trypsinised for 3 min at 37˚C and resuspended in STO medium, centrifuged and resuspended in either STO medium or ES medium at a density of $3x10^5$ cells/ml. To generate a monolayer of feeder cells, the inactivated SNH cells were plated onto gelatinised plates at $5x10^4$ cells/cm^2.

2.3.4 Embryoid body differentiation of mESCs

For general mesoderm differentiation, mESCs were harvested by incubation with trypsin for 5 min at 37˚C. Cells were resuspended at $5x10^3$ cells/ml in feeder free ES medium without LIF and cultured in suspension to form EBs until required. This relates to experiments shown in Figure 3.9C,D and Figure 5.8E.

For mesoderm differentiation that enriches for haematopoietic mesoderm, mESC were cultured according to published protocols, with slight modifications (Keller et al., 2002). This relates to experiments shown in Figure 5.4, Figure 5.5, and Figure 5.6. Briefly, mESCs were cultured in IMDM (21980-032, Invitrogen), 15% fetal calf serum (26140-079; Gibco), 1% Penicillin/Streptomycin (15140-163; Invitrogen), 1% Glutamine (25030-13; Invitrogen), 1000U/ml LIF (ESG1107; ESGRO), 1.5 x 10^{-4} M MTG (M6145; Sigma) for 24 – 48hours before EB formation. ESCs were harvested by incubation with trypsin for 5 min at 37˚C. Cells were washed in IMDM 5% FCS and then cells were counted and resuspended at 1x10^4 cells/ml in EB differentiation medium: IMDM (21980-032, Invitrogen) supplemented with 15% fetal calf serum (26140-079; Gibco), 1% Glutamine (25030-13; Invitrogen), 1% Penicillin/Streptomycin (15140-163; Invitrogen), 300µg/ml transferrin (10652202001; Roche), 50µg/ml ascorbic acid (A4544; Sigma) and 4.68x10^{-4} MTG (M6145; Sigma). Using a multichannel pipette, droplets of 10µl (100 cells per droplet) were plated on the lid of 10cm dishes, with PBS in the dish to prevent dehydration. After 48hours of EB formation in the hanging drops, the EBs were pooled together in 10ml fresh EB differentiation medium and cultured in 10cm dishes until required. Alternatively to hanging drop formation, 5x10^4 cells were plated in suspension in the EB differentiation medium and culture in 10cm dishes until required.

2.3.5 Harvesting of embryoid bodies

Day 2.5 – 6 EBs were allowed to gravity settle in falcon tubes, and briefly washed in PBS. For RNA analysis, whole EBs were snap frozen in liquid nitrogen and stored at -80˚C before RNA extraction. For flow analysis or cell differentiation assays, EBs were dissociated by the addition of 3-5ml 0.25% trypsin and incubation at 37˚C for 3

mins. The reaction was quenched with 1ml fetal calf serum (26140-079; Gibco), before being passed twice through a P1000 pipette to generate a single cell suspension.

2.3.6 Blast colony forming assay

Wild type and *Eomes* null ESCs were grown as EBs under haematopoietic mesoderm conditions for 3.5 days in suspension culture at $1x10^5$ cells per plate. EBs were harvested and dissociated into single cells and passed through a filter. $1x10^7$ cells were incubated in 0.8µg Rat anti-Flk1 PE (Avas12a1, eBioscience) diluted in PBS +2% FCS for 30 mins on ice. Afterwards, cells were washed in PBE (PBS + 0.5% BSA, 2mM EDTA). Cells were resuspended in 60µl of PBE and 40µl of anti-PE microbeads (130-048-801; Miltenyi) was added and incubated on ice for 20min. Magnetically labelled Flk1+ cells were washed once in PBE and resuspended in 500µl PBE. Flk1+ cells were purified on MS MACS columns (130-042-201; Miltenyi) according to manufacturer's instructions. Briefly, $1x10^7$ magnetically labelled cells in 500µl PBE were passed through a pre-washed MS column by gravity flow. MS columns containing labelled cells were washed with 3x500µl PBE. 1ml PBE was used to collect the magnetically labelled cells on the MS column. Negative fractions were passed over a second column to recover more cells if required. Purity of fractions was checked by flow analysis.

Blast colony assays were performed in triplicate with purified Flk1+ cell fractions. $5x10^4$ cells per 35mm plate were cultured in IMDM (21980-032, Invitrogen) supplemented with 15% fetal calf serum (26140-079; Gibco), 1% Glutamine (25030-13; Invitrogen), 1% Penicillin/Streptomycin (15140-163; Invitrogen), 300µg/ml transferrin (10652202001; Roche), 25µg/ml ascorbic acid (A4544; Sigma), $4x10^{-4}$

MTG (M6145; Sigma), 5ng/ml VEGF (450-32; Peprotech), 5ng/ml IL6 (216-16; Peprotech), 100ng/ml SCF (250-03; Peprotech), 25% D4T conditioned medium and 1% methylcellulose. Cells were cultured for 3-4 days and blast colonies scored on the basis of morphology.

2.3.7 Haematopoietic colony forming assay

Day 6 EBs grown under haematopoietic mesoderm conditions were dissociated to single cells and 2 x 10^4 cells were plated onto a 3.5 cm dish in Methylcellulose-Based Medium with Recombinant Cytokines and EPO for Mouse Cells (MethoCult GF M3434; Stem Cell Technologies), in triplicate. At day 10 of differentiation colonies were scored based on morphology as primitive erythroid (EryP), definitive erythroid (EryD), Granulocyte/Macrophage (GM) or Mixed.

2.4 Histology and cell biology

2.4.1 Whole-mount in situ hybridisation

E6.5 to E9.5 embryos were fixed in 4% PFA overnight at 4°C and dehydrated through a MeOH/PBT series and stored at -20°C until required. Whole-mount in situ hybridisation (WISH) was performed according to published protocols (Behringer et al., 2013). Briefly, embryos were rehydrated through a MeOH/PBT series, bleached for 1 hour in 6% H_2O_2 (6% H_2O_2, in PBT) and washed with PBT. Subsequently, embryos were permeablised for the relevant time period (E6.5: 4 min, E7.5: 5 min, E8.5: 6 min, E9.5: 7 min) in proteinase K solution (10mg/ml PK in PBT) and the reaction was quenched with Glycine (2mg/ml in PBT). Thereafter, the embryos were rinsed in PBT and post-fixed in 4%PFA, 0.2% Glutaraldehyde/PBT for 20 mins and then washed again in PBT. Embryos were incubated in a mixture with equal volumes of PBT and WISH hybridisation solution for 10 mins before being placed in

WISH hybridisation solution and incubated at 70˚C for 2 hours. For hybridisation, embryos were moved into fresh pre-warmed WISH hybridisation solution containing 50-100ng of DIG labelled probe and incubated overnight at 70˚C. Published probes used were as follows; *Eomes*, *Shh*, *Otx2*, *AFP*, *Foxa2*, *T* and *ER71* (Arnold et al., 2008; DiTacchio et al., 2012). A new iCre probe was generated from a sub-clone of the *Eomes*^{iCre} targeting vector.

The following day, embryos were washed in Wash Buffer I 3 x 30 min at 70˚C, followed by 3 washes in Wash Buffer II for 30 mins each at 65˚C. Embryos were then washed at room temperature in MAB buffer before blocking with 10% heat inactivated sheep serum (S22; Millipore), 2% Boehringer Blocking Reagent (BBR; 11096176001; Roche) in MAB buffer for 90 mins. Blocking solution was then replaced with pre-adsorbed anti-DIG AP- conjugated antibody (Anti-Digoxigenin-AP Fab Fragments antibody; 11093274910; Roche) in 1% heat inactivated sheep serum, 2% BBR/MAB buffer and incubated at 4˚C overnight. Embryos were washed with 1% levamisole (L9756; Sigma) in MAB buffer approximately 10 times throughout the day, and then incubated overnight in 1% levamisole in MAB at 4˚C.

Embryos were washed in AP buffer 3 x 10 mins, before incubation with BM Purple AP Substrate (1442074001; Roche) at room temperature, in dark. Intensity of staining was monitored periodically time maximise signal and minimise background. The reaction was quenched with PBS + 2mMEDTA. Stained embryos were post-fixed in 4% PFA, 0.1% Glutaraldehyde in PBS, then washed in PBS and photographed in PBS or in 80% glycerol.

2.4.2 β-galactosidase staining

LacZ activity from $Rosa^{26R}$ allele, transgenic enhancer reporters or Flk1LacZ was visualised by whole-mount X-gal staining performed according to standard protocols (Behringer et al., 2013). Embryos were fixed in X-gal buffer for 30min at 4°C, washed in X-gal buffer and stained in X-gal stain solution overnight at 37°C. Embryos were post fixed in 4% PFA, before being photographed in whole-mount or processed for histological analysis.

2.4.3 Paraffin embedding

Embryo samples were dehydrated through an ethanol series (25%, 50%, 70%, 90%, 100%), cleared in 2 washes of histoclear (HS-200; National Diagnostic). Embryos were then transferred through a 1:1 histoclear/paraffin wax and 2 changes of paraffin wax at 60°C before embedding in fresh paraffin using a Leica EG1160 embedding station. Sections were cut to 0.8 μm thickness using a Leica RM2125 RT rotary microtome and counterstained with NuclearFastRed (H-3403; Vector labs) according to standard protocols.

2.4.4 Immunofluorescence staining

E6.5-E7.5 embryos or day 4 EBs were fixed in 4% PFA in PBS at room temperature for 30 min. Samples were then washed in 0.1% Triton-X in PBS briefly 3 times, followed by permeabilisation in 0.5% Triton-X in PBS for 15 min. Samples were washed in 0.1% Triton-X in PBS and then blocked in 0.1% Triton-X, 0.2% BSA, 5% donkey serum in PBS 2hours at room temperature. Following this, samples were incubated with primary antibodies in block solution overnight at 4°C.

The next day, samples were washed in 0.1% Triton-X in PBS briefly 3 times and then

for 15 min, before incubation with secondary antibodies or Phalloidin stain in block solution at room temperature in the dark for 2hours. Samples were washed thoroughly in 0.1% Triton-X in PBS briefly 5 times, and then 3 times for 15 mins. The final wash contained 1:5,000 DAPI stain. Finally, samples were mounted in Vectashield mounting media containing DAPI (H-1200; Vector Laboratories) on chamber slides (155409, LabTek) and left overnight at 4˚C. Primary antibodies used were goat anti-Brachyury (N-19; Santa Cruz) at a 1:100 dilution, rabbit anti-TBR2/Eomes (Abcam) at a 1:100 dilution and rabbit anti-GFP AlexaFluor 488 (Invitrogen) at a 1:200 dilution. Phalloidin AlexaFluor 633 (Molecular Probes) was used at 1:000 dilution. Antibody staining was visualised by using the following secondary antibodies; donkey anti-rabbit IgG AlexaFluor 488 (Molecular Probes) at a dilution of 1:200, and donkey anti-goat IgG AlexaFluor 594 (Molecular Probes) at a dilution of 1:200. Images were acquired using an Olympus FV1000 inverted confocal microscope.

2.4.5 Flow cytometry

Day 4 EBs were incubated in 0.25% trypsin for 5 mins at 37˚C and passed through a 20-guage needle to dissociate them into single cells. For cell staining $1x10^6$ cells were incubated with 0.5µg rat anti-Pdgfrα APC (APA5, eBioscience) and 0.5µg rat anti-Flk1 PE (Avas12a1, eBioscience) diluted in PBS +2% FCS + 2% P/S + 0.1% NaN_3 for 30 mins on ice. Afterwards, cells were washed with PBS +2% FCS + 2% P/S + 0.1% NaN_3, before being resuspended in a volume of 0.5ml. FACS analysis was performed using a BD FACSCalibur 4 (BD Biosciences) and data was analysed using FlowJo.

2.5 Mouse strains and husbandry

Successfully targeted CCE (129/sv) mESCs were injected into C57Bl/6 blastocysts to

generate chimeric mice. High percentage male chimeras were bred to C57Bl/6 females and the progeny were typed by multiplex PCR at standard conditions. Mice were maintained on a mixed genetic background 129/Sv and C57Bl/6.

2.6 Bioinformatics

ChIP-seq data was downloaded from GEO (www.ncbi.nlm.nih.gov/geo/), with the following accession codes; GSE26097, GSE19461 and GSE29422 (Brown et al., 2011; Kim et al., 2011; Teo et al., 2011). ChIP-seq data for SMAD2/3, SMAD4, FOXH1 and EOMES binding in human ESCs or human definitive endoderm were aligned to the *EOMES* locus on the UCSC Genome browser Human Mar. 2006 (NCBI36/hg18) Assembly (http://genome.ucsc.edu/).

ChIP-seq and DNaseI Hypersensitivity data in mouse ESCs was obtained from the ENCODE project (http://genome.ucsc.edu/ENCODE) and visualised on the UCSC Genome browser Mouse July 2007 (NCBI37/mm9) Assembly (http://genome.ucsc.edu/). The tracks and UCSC accession codes used are as follows: H3K4me1 (wgEncodeEM001681), H3K27me3 (wgEncodeEM002709), H3K27ac (wgEncodeEM002497), polII (wgEncodeEM001704), p300 (wgEncodeEM001705) and DNase-seq (wgEncodeEM001728).

Eomes enhancer sequences were analysed for consensus sequences using JASPAR (http://jaspar.genereg.net/) and STAMP (http://www.benoslab.pitt.edu/stamp/) tools with a cut-off score of 80 or an E-value $<8x10^{-5}$, respectively.

Clustal Omega alignment of the mouse Eomes protein (707 amino acid; UniProt ID: O54839) and mouse Brachyury protein (436 amino acid; UniProt ID: P20293) was performed on the UniProt website (www.uniprot.org/align).

Chapter 3: Targeted deletion of *Eomes* enhancers

3.1 Introduction

Eomes is an essential transcription factor controlling cell fate decisions during gastrulation (Arnold et al., 2008; Costello et al., 2011). Therefore, the timing of expression is tightly controlled in different cellular and spatial contexts. *Eomes* is expressed in dynamic spatiotemporal domains in the post implantation embryo. Before overt streak formation, *Eomes* is expressed in the visceral endoderm, and extra-embryonic ectoderm (ExE). At early-streak stages, *Eomes* is additionally expressed in the most proximal posterior region of the epiblast, where the primitive streak forms. As gastrulation proceeds, *Eomes* continues to be expressed in the primitive streak, nascent mesoderm, anterior visceral endoderm (AVE) and ExE. By late-streak stages, *Eomes* is down regulated in the posterior streak and expression remains only in the most anterior region of the streak and newly emerged definitive endoderm and the chorion (Arnold et al., 2009; Ciruna and Rossant, 1999; Russ et al., 2000).

Previous studies have described *Eomes* enhancers that promote tissue-specific expression in the mouse embryonic retina, limb, and primitive streak (Mao et al., 2008; Pernaute et al., 2010a; Wolf et al., 2012). These enhancers were identified solely by sequence conservation between species. However, using conservation as the sole method to identify enhancers is risky as many highly conserved noncoding elements seem to be non-functional (Ahituv et al., 2007).

Genome-wide high throughput ChIP-seq for transcription factors and histone modifications has provided new information about tissue-specific enhancers (Ong and Corces, 2011; Rada-Iglesias et al., 2011). Recent studies in human ESCs (hESCs) and

differentiated definitive endoderm (hDE) cultures has provided considerable insight into molecular mechanisms governing *Eomes* expression (Beyer et al., 2013; Estaras et al., 2015; Funa et al., 2015; Kim et al., 2011; Teo et al., 2011). It remains unclear whether these regulatory events also control *Eomes* expression in the mouse embryo.

The objective of this study is to identify mouse cis-regulatory regions that control *Eomes* expression during gastrulation by using a combined approach utilising sequence conservation and published ChIP-seq datasets. I aim to test whether these enhancers are sufficient to drive reporter expression and if they are truly required for proper expression of *Eomes* during early embryonic development. This will be the first study to generate enhancer deletions at the *Eomes* locus to validate their function. With this, I hope to build a complete picture of how the *Eomes* locus is regulated and how several enhancers may interact together to provide the dynamic spatiotemporal expression domains of *Eomes* in the early embryo.

3.2 Identification of *Eomes* enhancers

Eomes is regulated by Nodal signalling, and together they act co-operatively to induce differentiation of definitive endoderm and maintain *Eomes* expression in an auto-regulatory manner (Teo et al., 2011). To identify the cis-regulatory sequences, I initially compared various human ChIP-seq datasets for EOMES and components of the Nodal signalling pathway namely; the downstream effectors SMAD2/3, co-smad SMAD4, and co-factor FOXH1 in hESCs and Activin induced definitive endoderm differentiation (hDE) (Brown et al., 2011; Kim et al., 2011; Teo et al., 2011).

SMAD2/3, SMAD4, FOXH1 and EOMES bind to distal sites at the *Eomes* locus. Three distinct genomic regions are co-bound by these factors were identified:

enhancer 1a, 1b, and 3 (EH1a, EH1b and EH3 respectively) **(Figure 3.1A)**.

In undifferentiated hESCs SMAD2/3 binds EH1a, a region located 9 kb upstream of the transcriptional start site (TSS). Subsequently, upon Activin induced differentiation to hDE, EH1a is bound by SMAD2/3, EOMES and FOXH1. However, SMAD4 does not bind EH1a in either cell type.

EH1b is located 6 kb upstream of the TSS and is bound by SMAD2/3, SMAD4 and FOXH1 in hESCs. Upon differentiation to hDE, EH1b is co-bound by all four transcription factors.

EH3 is located 8kb downstream of the TSS. This region is bound by SMAD2/3 and SMAD4 in hESCs. After differentiation to hDE, EH3 becomes occupied by all four transcription factors.

Therefore, mediators of the Nodal signalling pathway associate with three distinct enhancer regions, EH1a, EH1b and EH3, at the *EOMES* locus during human definitive endoderm differentiation.

Interestingly these enhancers identified represent the same EOMES-SMAD2/3 co-bound peaks as described in Teo et al. (2011), moreover these peaks are associated with highly conserved regions amongst mammals, implying that EH1a, EH1b and EH3 are probably evolutionarily conserved cis-regulatory elements **(Figure 3.1B)**.

The corresponding putative enhancer regions in the mouse genome were identified by sequence conservation. At the time of writing, there are no relevant ChIP-seq data sets available in mouse definitive endoderm to evaluate whether these enhancers might be

specific to this tissue. However, data from other cell types, e.g. mESCs. I pluripotent cells *Eomes* has poised enhancers and a bivalent promoter (containing both active H3K4me3 and repressive H3K27me3 histone marks) to facilitate its rapid transcriptional activation in response to differentiation signals (Bernstein et al., 2006). This suggests that *Eomes*, like many other developmentally important genes, is primed for expression in mESCs.

To evaluate whether EH1a, EH1b and EH3 are poised enhancers in mESCs I analysed ENCODE datasets **(Figure 3.1C)**. Poised and active enhancers are characterised by DNaseI hypersensitive sites (DNaseHS) indicative of nucleosome depleted regions and the histone modification H3K4me1. Furthermore, the presence of H3K27ac, an active chromatin mark, p300 a histone acetyl transferase, and polII are diagnostic for active enhancers. Whereas, poised enhancers are associated with a repressive chromatin mark H3K27me3.

EH1a, EH1b and EH3 lack p300, polII and H3K27ac peaks, strongly suggesting these are not active in pluripotent cells. However, EH1b and EH3 are associated with H3K4me1, H3K27me3 and DNaseHS sites, indicating that these are poised enhancers in undifferentiated mESCs. EH1a contains a H3K4me1 peak and a DNaseHS site, but lacks any H3K27 modifications, because it contains neither active or repressive marks it seems to exhibit an intermediate status between poised and active (Zentner et al., 2011). Collectively, these observations suggest that the conserved cis-regulatory elements identified in humans correspond to poised enhancer regions in mESCs.

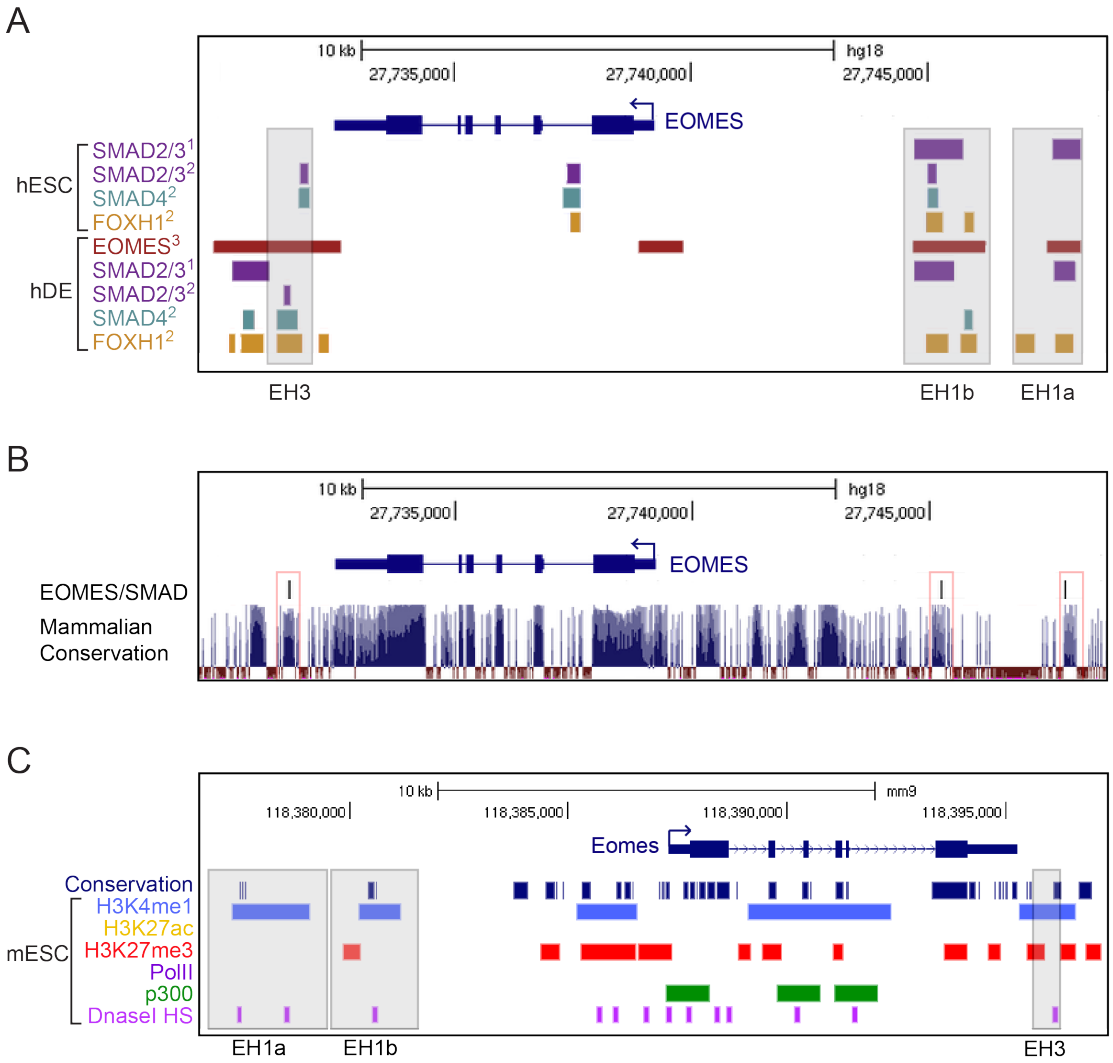
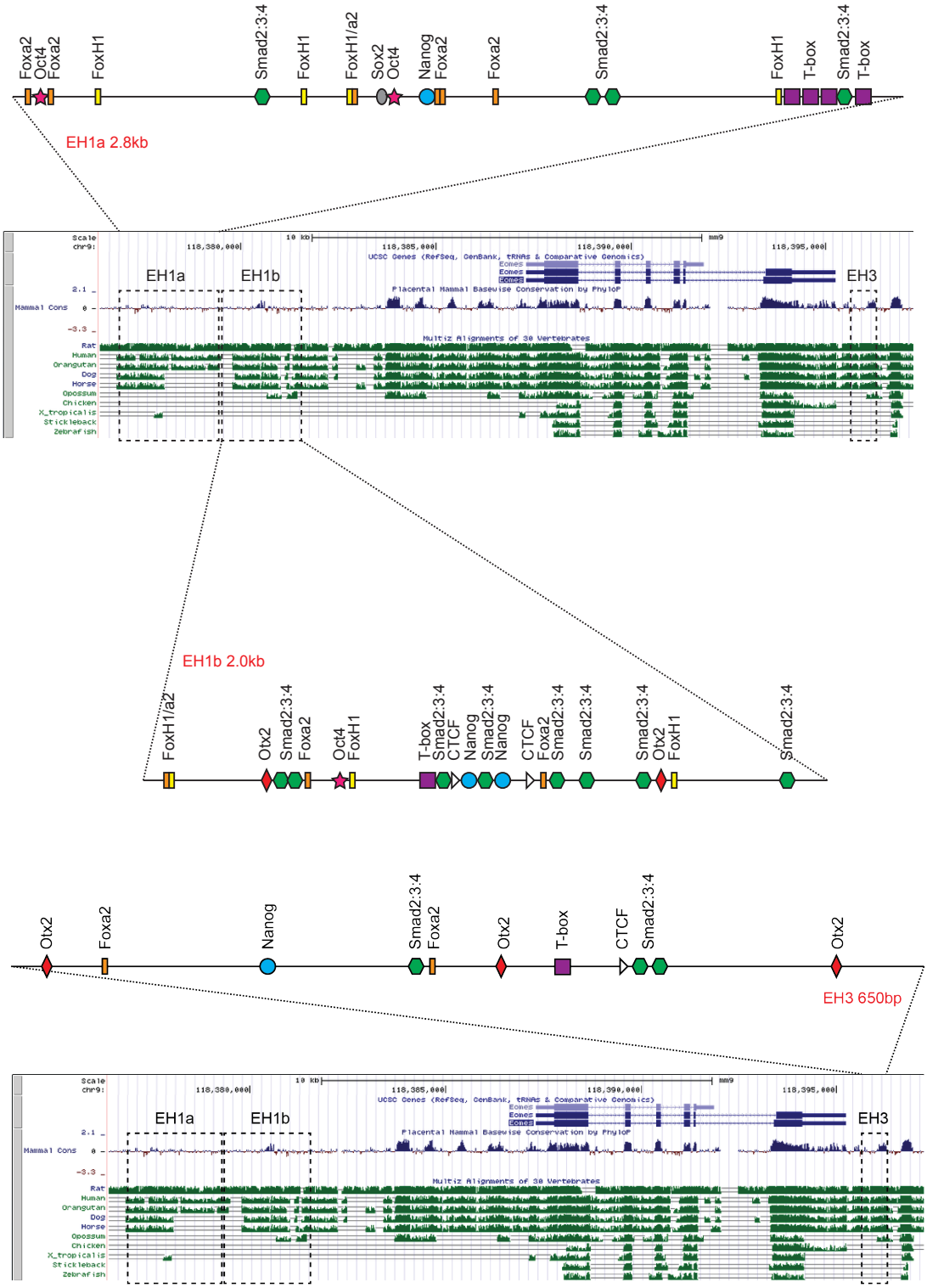


Figure 3.1: Identification of putative *Eomes* enhancers.
Cis-regulatory elements regulated by Nodal signalling in human definitive endoderm (hDE) cultures are highly conserved amongst vertebrates and are poised enhancers in mESCs. (A) Potential cis-regulatory elements are bound by EOMES and mediators of the Nodal signalling pathway in hESCs and hDE. (1=Brown et al., 2011, 2=Kim et al., 2011, 3=Teo et al., 2011). ChIP-seq data showing regions bound by SMAD2/3 (purple), SMAD4 (green), FOXH1 (orange) and EOMES (red) are represented by coloured bars and were aligned to the *EOMES* locus on the UCSC Genome browser Human Mar. 2006 (NCBI36/hg18) Assembly (http://genome.ucsc.edu/). Grey boxes indicate putative enhancers EH1a, EH1b, and EH3. (B) The centre of co-bound SMAD2/3 and EOMES peaks in hDE (Teo et al., 2011) were aligned to the *Eomes* locus as described above. Placental mammal base-wise conservation by PhyloP is shown below. Red boxes indicate regions of high conservation surrounding the binding sites. (C) ChIP-seq data in mouse ESCs from ENCODE project (http://genome.ucsc.edu/ENCODE/) demonstrate that EH1b and EH3 are poised enhancers whereas EH1a is an intermediate enhancer in this cell type. ChIP-seq peaks for H3K4me1 (blue), H3K27ac (yellow), H3K27me3 (red) histone modifications, PolII (purple) and p300 (green) binding, and DNaseI hypersensitive sites (pink) are shown for the *Eomes* locus on the UCSC Genome browser Mouse July 2007 (NCBI37/mm9) Assembly. Grey boxes indicate the enhancer regions.

To investigate the possible mechanism of activation of this locus in mouse, transcription factor binding motifs were identified in the EH1a, EH1b and EH3 regions **(Figure 3.2)**. Consensus sequence alignments were performed for transcription factors known to be important in the regulation of *Eomes*, either in their role as pluripotency factors to repress, and perhaps then initiate expression (e.g. Oct4, Nanog and Sox2), TGF-β signalling (e.g. Smad2/3/4 and FoxH1), endoderm differentiation (e.g. Foxa2, Otx2, and T-box factors Eomes and Brachyury) or chromatin looping (e.g. CTCF). DNA sequences were analysed for consensus sequences using JASPAR database (http://jaspar.genereg.net/) and STAMP tool (http://www.benoslab.pitt.edu/stamp/) with a cut-off score of 80 or an E-value $<8x10^{-5}$, respectively. Conservation between different vertebrates is shown on the UCSC genomic browser (http://genome.ucsc.edu/). Each mouse enhancer contains T-box and Smad2/3 consensus binding sites, indicating that this mode of regulation may be shared with the human enhancers.



Endoderm/Mesendoderm

- Foxa2: Forkhead: TGTTTACWYWNN
- Eomes/Brachyury: T-box: TTCACACCTAG
- Otx2: Homeobox: GGGATTAAT

Pluripotency

- Oct4: POU domain: CWTTGTNATGCAAAT
- Sox2: SRY-box: TNTGCATNACAAWGG
- Nanog: Homeobox: GGMAATGGNCC

TGF-β signalling

- Smad2:3:4: CAGAC
- FoxH1: Forkhead: TGTGGATT

Chromatin looping

- CTCF: CCTC-binding factor (zinc-finger): SYGCCMCCTRSTGG

Figure 3.2: Predicted transcription factor binding sites located within mouse *Eomes* enhancers.

Schematic diagram showing the transcription factor binding motifs identified within the mouse *Eomes* enhancers. Transcription factors involved in *Eomes* activation and those involved in pluripotency, endoderm formation, TGFβ signalling, and DNA looping were included in the analysis. DNA sequences were analysed for consensus sequences using JASPAR (http://jaspar.genereg.net/) and STAMP (http://www.benoslab.pitt.edu/stamp/) tools with a cut-off score of 80 or an E-value $<8x10^{-5}$, respectively. Conservation between different vertebrates is shown on the UCSC genomic browser (http://genome.ucsc.edu/)

3.3 Transgenic analysis of *Eomes* regulatory regions

Eomes is expressed dynamically during gastrulation. *Eomes* is expressed at E6.25 (Pre-streak: PrS) in the visceral endoderm and extra-embryonic ectoderm (ExE) and at reduced levels in the proximal-posterior epiblast, at E6.5 (Early and mid-streak: ES / MS) in the primitive streak, visceral endoderm and ExE, and at E7.5 (Late-streak: LS) in the anterior visceral endoderm (AVE), anterior primitive streak (APS), definitive endoderm and chorion (Ciruna and Rossant, 1999).

To test *Eomes* enhancer activity during gastrulation, I generated LacZ reporter constructs containing *Eomes* enhancer sequences cloned into a reporter expression plasmid, upstream of the minimal heat shock promoter 68 (HSP68), the LacZ reporter, and SV40 poly A termination sequence **(Figure 3.3A)**. Plasmids were microinjected into embryos at the 2-cell stage to generate transgenic mice[1]. Embryos were harvested at E6.5-E7.5 for transient transgenic analysis by X-gal staining. Alternatively, adult transgenics were bred to generate stable transgenic lines. Progeny were genotyped to detect transmission of the transgene. Positive transgenic F1 males were crossed to CD1 females and embryos harvested at E6.5 to E7.5 to assess reporter activity of the transgenic line by X-gal staining **(Figure 3.3B)**.

3.3.1 The EH1 regulatory region has enhancer activity in the primitive streak

Enhancer activity of the EH1 region was assessed in transient transgenic embryos. The EH1 genomic fragment 4,775 bp in length is located between 10.7 kb to 5.9 kb upstream of the mouse *Eomes* TSS, which encompasses both EH1a and EH1b enhancer regions **(Figure 3.3A)**. The EH1LacZ reporter was active in the proximal-

[1] This was done in collaboration with Ben Davies and his team at the Wellcome Trust Centre for Human Genetics transgenics core.

posterior epiblast at the PrS stage, and in the primitive streak at MS stages **(Figure 3.3C)**. There was no detectable LacZ staining in the ExE or visceral endoderm at these stages. These results demonstrate that EH1LacZ expression recapitulates endogenous *Eomes* expression, specifically within the primitive streak. Therefore, combined activities of EH1a and EH1b enhancer regions are sufficient to drive *Eomes* expression during gastrulation. Moreover, tissue specific EH1a and EH1b enhancer activity is confined to the primitive streak.

3.3.2 The EH3 regulatory region has enhancer activity in the primitive streak and anterior visceral endoderm

EH3 enhancer activity was similarly assessed during in transgenic embryos. The EH3 enhancer region is 696 bp in length and located 8.1 kb to 8.8 kb downstream of the mouse *Eomes* TSS. In transient transgenic embryos, EH3lacZ reporter activity was observed in the proximal-posterior epiblast at PrS stage, the site of the future streak **(Figure 3.3C)**.

Stable transgenic EH3LacZ mouse lines were also generated. Of 5 lines positive by PCR genotyping for carrying the stably inherited reporter fragment, LacZ expression was only detectable in 1 of the lines. As for transient transgenic embryos, LacZ reporter activity was detected in the proximal-posterior epiblast at PrS stage, and surprisingly also in the AVE. During early gastrulation, LacZ was expressed in the primitive streak, nascent mesendoderm and AVE. X-gal staining in the streak becomes restricted to the APS before being lost at LS stage. However, the LacZ reporter was not expressed in the ExE or chorion. In summary, EH3lacZ expression partially recapitulates sites of endogenous *Eomes* expression and is specific to the primitive streak and AVE. Hence, enhancer activity of the EH3 region is sufficient to drive transcription in appropriate sites of *Eomes* during gastrulation. Furthermore,

EH3 enhancer activity is tissue-specific, and is restricted to the primitive streak and anterior visceral endoderm. This finding represents the first description of an *Eomes* anterior visceral endoderm enhancer.

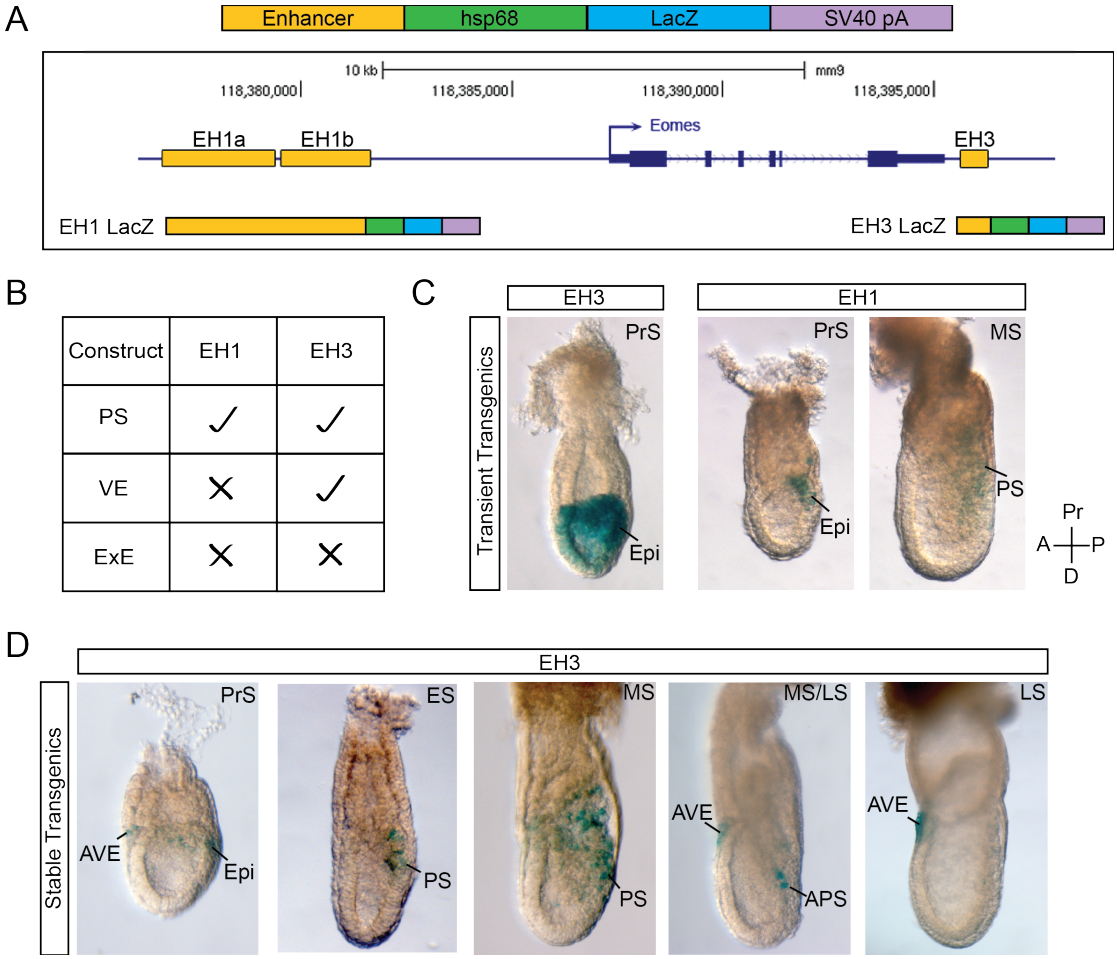


Figure 3.3: *Eomes* enhancer driven reporters recapitulate primitive streak and anterior visceral endoderm expression.
(A) Schematic representation of *Eomes* enhancer transgenic constructs. Enhancer regions EH1 (combination of EH1a and EH1b) and EH3 were used to generate transgenic constructs and are shown aligned to the *Eomes* locus on the UCSC genome browser (http://genome.ucsc.edu/) Mouse July 2007 (NCBI37/mm9) Assembly. (B) Table showing the expression domains of each transgenic at E6.5 and E7.5. (C) Representative images of LacZ expression for EH1 and EH3 LacZ constructs in transient transgenic embryos. Expression of both EH1 and EH3 LacZ recapitulate endogenous *Eomes* expression within the primitive streak. (D) Representative images of LacZ expression domains of EH3 stable transgenic line. EH3 recapitulates endogenous *Eomes* expression within the primitive streak and anterior visceral endoderm. PrS = pre-streak, ES=early streak, MS=mid-streak, LS=late streak. PS=primitive streak, APS=anterior primitive streak, VE=visceral endoderm, ExE=extra-embryonic endoderm, Epi=epiblast

Construct	Transient transgenics embryonic expression	Stable transgenic lines embryonic expression
EH1-LacZ	2 / 11	N/A
EH3-LacZ	4 / 37	1 / 5

Table 3.1: Number of expressing transgenics.
EH1-LacZ and EH3-LacZ embryonic expression in transient and stable transgenics. Numbers represent embryos at E6.0 – E8.0 with observed LacZ expression by X-gal staining compared to those positively genotyped for the LacZ transgene.

3.4 Targeted deletion of *Eomes* enhancer regions

EH1 (a combination of EH1a and EH1b) and EH3 are sufficient to drive endogenous patterns of *Eomes* expression in the embryo. To test whether they are also necessary for correct expression of *Eomes* and proper embryonic development, I have generated mice harbouring deletions of the enhancers.

3.4.1 Generation of a ΔEH3 targeted deletion

A targeting vector to delete the 656bp EH3 region 8.2kb downstream of the TSS of *Eomes* by homologous recombination was transfected into mESCs **(Figure 3.4A)**. Successfully targeted clones were identified by Southern blot and PCR. Excised clones were then injected into blastocysts to produce chimeras. Germline chimeric mice were bred to generate mice harbouring the EH3 deletion ($Eomes^{\Delta EH3}$) **(Figure 3.4B,C)**.[2]

[2] The EH3 targeting vector was built by Pia Rantakari. Targeting in mESCs was carried out by Ita Costello. Screening of successfully targeted mESCs clones by Southern blot and PCR was done by Li Li. Elizabeth Robertson generated the germline chimeras by performing blastocyst injections.

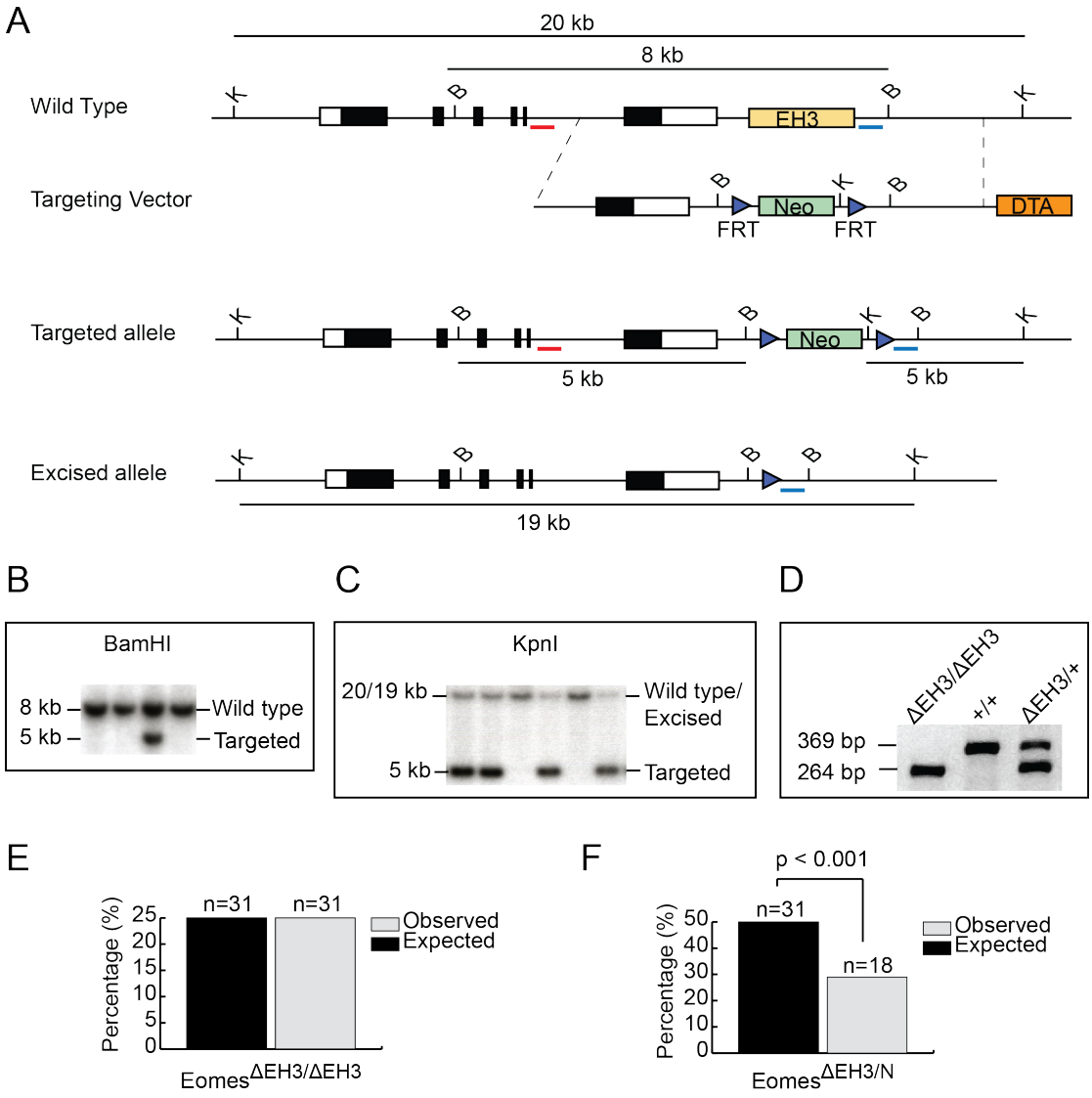


Figure 3.4: Generation of *Eomes*$^{\Delta EH3}$ mice, a mild hympomorph.
(A) Targeting strategy to delete the 656bp EH3 region by homologous recombination in mESCs. The EH3 targeting vector contains a FLP-recombinase recognition (FRT) site flanked Neomycin resistance positive selection cassette, 5' and 3' arms homologous to the *Eomes* locus, and a diphtheria toxin A (DTA) negative selection cassette. Southern blot probes, restriction digests and expected fragment sizes are indicated for the different alleles. B= BamHI K= KpnI. (B) Southern blot of targeted mESC clones. (C) Southern blot to identify excision of Neo cassette in targeted mESC clones. (D) PCR genotyping ΔEH3 allele in animals derived from *Eomes*$^{\Delta EH3/+}$ intercrosses. (E) No difference between observed and expected *Eomes*$^{\Delta EH3/\ \Delta EH3}$ pups born from *Eomes*$^{\Delta EH3/+}$ intercrosses. *Eomes*$^{\Delta EH3/\ \Delta EH3}$ mice are viable (Chi-square test, n.s., n=124). (F) Significant difference between observed and expected *Eomes*$^{\Delta EH3/N}$ pups derived from *Eomes*$^{\Delta EH3/\ \Delta EH3}$ x *Eomes*$^{N/+}$ mating (Chi-square test, p<0.001, n=62) indicating that *Eomes* $^{\Delta EH3}$ is a hypomorph.

Eomes$^{\Delta EH3/\Delta EH3}$ homozygotes generated from *Eomes*$^{\Delta EH3/+}$ intercross are viable and represented at 25%, the expected Mendelian ratio (Chi square value= *1.2, 0.95>P>0.9, n=124)*. In addition, *Eomes*$^{\Delta EH3/\Delta EH3}$ homozygotes are fertile and produce healthy adult offspring.

Heterozygous *Eomes* null (e.g. *Eomes*$^{GFP/+}$ or *Eomes*$^{\Delta exon2\text{-}5/+}$) mice are also viable, and display no phenotypic abnormalities (Arnold et al., 2008; Arnold et al., 2009). To test whether ΔEH3 could potentially be a hypomorphic allele, *Eomes*$^{\Delta EH3/\Delta EH3}$ mice were crossed to heterozygous *Eomes*^{GFP} mouse line (hereafter referred to as *Eomes* null; *Eomes*$^{N/+}$) (Arnold et al., 2009). At weaning, the *Eomes*$^{\Delta EH3/N}$ mice were significantly under-represented (29%) compared to their *Eomes*$^{\Delta EH3/+}$ littermates (71%) (Chi Square value=10.9, P<0.001, n=62). Therefore, a subset of *Eomes*$^{\Delta EH3/N}$ mice do not survive to weaning.

Genotype	Wild type	*Eomes*$^{\Delta EH3/+}$	*Eomes*$^{\Delta EH3/\Delta EH3}$	Total
Observed	36 (29%)	57 (46%)	31 (25%)	124 (100%)
Expected	31 (25%)	62 (50%)	31 (25%)	124 (100%)

Table 3.2: Genotypes of *Eomes* $^{\Delta EH3/+}$ intercross offspring.
Eomes $^{\Delta EH3/+}$ intercross offspring were genotyped by PCR. The table shows the observed and expected numbers and corresponding percentages of offspring at weaning age recovered from the intercross.

Genotype	*Eomes*$^{\Delta EH3/+}$	*Eomes*$^{\Delta EH3/N}$	Total
Observed	44 (71%)	18 (29%)	62 (100%)
Expected	31 (50%)	31 (50%)	62 (100%)

Table 3.3: Genotypes of *Eomes* $^{\Delta EH3/\Delta EH3}$ x *Eomes* $^{N/+}$ offspring
Eomes $^{\Delta EH3/\Delta EH3}$ x *Eomes*$^{N/+}$ offspring were genotyped by PCR. The table shows the observed and expected numbers and corresponding percentages of offspring at weaning age recovered from the mating.

3.4.2 Gastrulation abnormalities in Eomes$^{\Delta EH3/N}$ embryos

To further investigate the *Eomes*$^{\Delta EH3/N}$ lethal phenotype, I dissected embryos between E6.5 and E9.5 and performed whole-mount in situ hybridisation. *Eomes*$^{\Delta EH3/N}$ embryos display three distinct phenotypes, with approximately a third of embryos in each class. Approximately a third of *Eomes*$^{\Delta EH3/N}$ embryos are morphologically normal at E9.5 and can only be distinguished from control littermates by a small reduction in size. Severely affected mutants arrest at gastrulation due to a failure in anterior-poster (A-P) axis specification. A second class of embryos had problems in specification of anterior primitive streak (APS) derivatives such as definitive endoderm, which causes anterior truncations and ventral closure defects.

3.4.2.1 Class I embryos exhibit anterior visceral endoderm defects

At E6.5 severely affected *Eomes*$^{\Delta EH3/N}$ embryos, class I, express the AVE marker *Hex* at the distal tip, indicating that the AVE is specified but fails to migrate properly to the anterior side of the embryo **(Figure 3.5A,B)**. Subsequently, these embryos do not correctly orient the A-P axis and fail to form the primitive streak. During gastrulation these mutant embryos can clearly be distinguished from littermate controls due to their abnormal morphology; lack of primitive streak, an accumulation of disorganised mesenchymal cells in the epiblast cavity and constriction at the embryonic and extra-embryonic boundary. At E7.5 markers of mesoderm (*Brachyury*) and definitive endoderm (*Cer1* and *Foxa2*) are restricted proximally or are absent in these mutants **(Figure 3.5C-H)**.

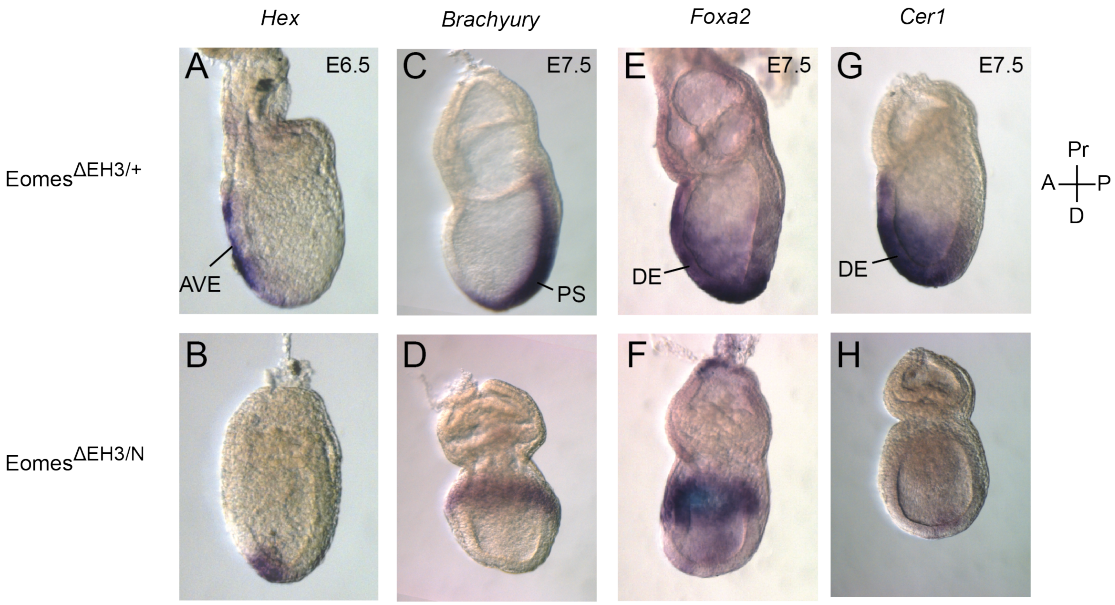


Figure 3.5: Class I *Eomes*$^{\Delta EH3/N}$ embryos display A-P axis patterning defects.

Whole-mount in-situ hybridisation of E6.5 – E7.5 embryos demonstrate defects in migration or specification of anterior visceral endoderm (AVE) resulting in failure to correctly specify the A-P axis. (A-B) The AVE marker *Hex* remains at distal tip of the embryo, rather than migrating anteriorly. (C-D) The expression of mesoderm marker *Brachyury* is restricted proximally rather than extending posteriorly. (E-H) *Foxa2* and *Cer1* expressed in the definitive endoderm (DE) are mis-localised proximally, or completely absent.

3.4.2.2 Class II embryos display anterior primitive streak defects

A second set of *Eomes*$^{\Delta EH3/N}$ embryos, class II, initiate gastrulation but exhibit defects with the APS, which gives rise to the definitive endoderm, node and notochord. At E7.5, Brachyury is expressed in the primitive streak and notochord. However, in the mutant embryos the Brachyury expression domain is restricted to the posterior region of the streak and fails to extend anteriorly **(Figure 3.6A,B)**. In addition, mutants specify Foxa2 positive definitive endoderm, but it does not migrate anteriorly from the streak and intercalate onto the surface of the embryo **(Figure 3.6C,D)**. Thus, the overlying AFP positive visceral endoderm is not displaced to the extra-embryonic region of the conceptus in these *Eomes*$^{\Delta EH3/N}$ embryos **(Figure 3.6E,F)**.

The APS is essential for providing trophic signals and to pattern underlying anterior structures. Thus, at E9.5 this class of *Eomes*$^{\Delta EH3/N}$ embryos have gross morphological abnormalities; failure in ventral closure, neural tube defects, reduction in head size, enlarged pericardial sac, lack of yolk sac vasculature and fused or malformed somites **(Figure 3.6H)**. Expression of *Otx2* in the anterior neurectoderm indicates that anterior structures are specified, but due to the lack trophic signals from the midline there is a great reduction in the size of the anterior domain.

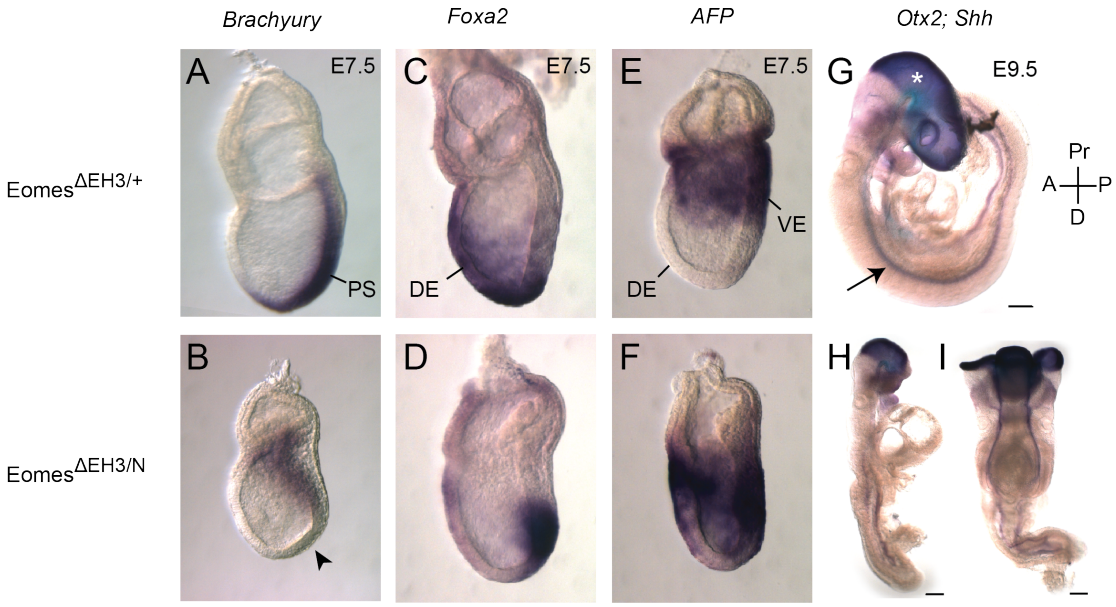


Figure 3.6: Class II *Eomes*$^{\Delta EH3/N}$ embryos have anterior primitive streak defects.

Whole-mount in-situ hybridisation at E7.5 demonstrates that a subset of mutants have anterior primitive streak (APS) defects resulting in gross morphological abnormalities at E9.5. (a-b) At E7.5, the primitive streak, marked by *Brachyury*, does not extend fully in the mutants, as indicated by the arrow head. (c-d) The *Foxa2*+ definitive endoderm (DE) does not properly migrate to the anterior side of the embryo. (e-f) Expression of visceral endoderm (VE) marker, *AFP*, overlying the embryo fails to be dispersed by the definitive endoderm in mutant embryos. (h) Expression of *Shh* (arrow) in the neural tube and *Otx2* (*) in the anterior neurectoderm in a wild type embryo at E9.5 (h-i) Examples of two mutant embryos at E9.5 (ventral and dorsal views). They exhibit a range of morphological abnormalities including reduced anterior domain, enlarged pericardial sac, malformed somites and ventral closure defects.

It has long been known *Eomes*$^{N/+}$ embryos are morphologically normal and express half the levels of *Eomes*. Deletion of the enhancer in *Eomes* $^{\Delta EH3/N}$ embryos must reduce the transcription of *Eomes* within the AVE and primitive streak to sub-threshold levels. To investigate if deletion of EH3 compromised the expression domains of *Eomes*, I performed whole-mount in situ hybridisation on wild type and *Eomes* $^{\Delta EH3/\Delta EH3}$ embryos at E6.5. *Eomes* mRNA expression in the mutants is present in the same domains within the primitive streak and ExE as in the control embryos **(Figure 3.7)**.

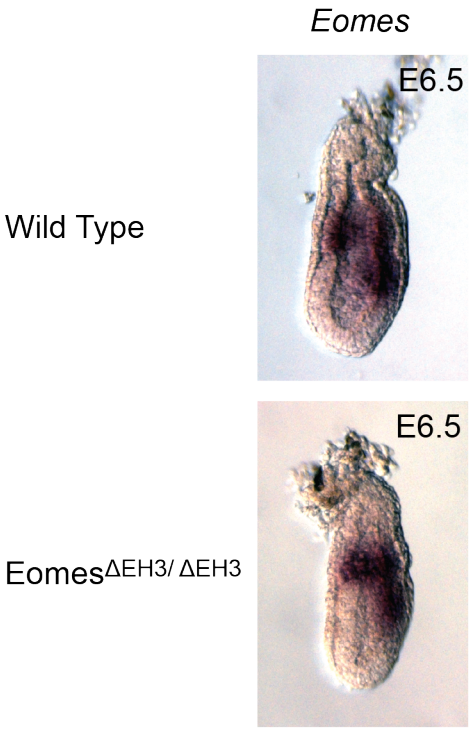


Figure 3.7: Loss of EH3 fails to perturb domains of *Eomes* expression in *Eomes* $^{\Delta EH3/\ \Delta EH3}$ embryos.
Whole-mount in situ hybridisation of *Eomes* transcripts at E6.5 shows that *Eomes* expression domains are unaltered in homozygous EH3 deletion mutants compared to wild type.

3.4.3 Deletion of EH3 in the context of Eomes^{GFP} reporter allele greatly diminishes expression

Deletion of the enhancer in *Eomes* $^{\Delta EH3/N}$ embryos must reduce the transcription of *Eomes* within the AVE and primitive streak to sub-threshold levels. To better quantify changes in gene expression caused by deletion of the enhancer EH3, I utilised an existing *Eomes*^{GFP} allele as a reporter of *Eomes* expression (Arnold et al., 2009). The enhancer EH3 was deleted from the same chromosome as *Eomes*^{GFP} by homologous recombination in heterozygous *Eomes*^{GFP} mouse embryonic stem cells. Successfully targeted mESCs were screened by Southern blot to determine which allele (GFP or wild type) contained the enhancer deletion. *Eomes*$^{GFP\ \Delta EH3/+}$ mESCs were used to generate mice.

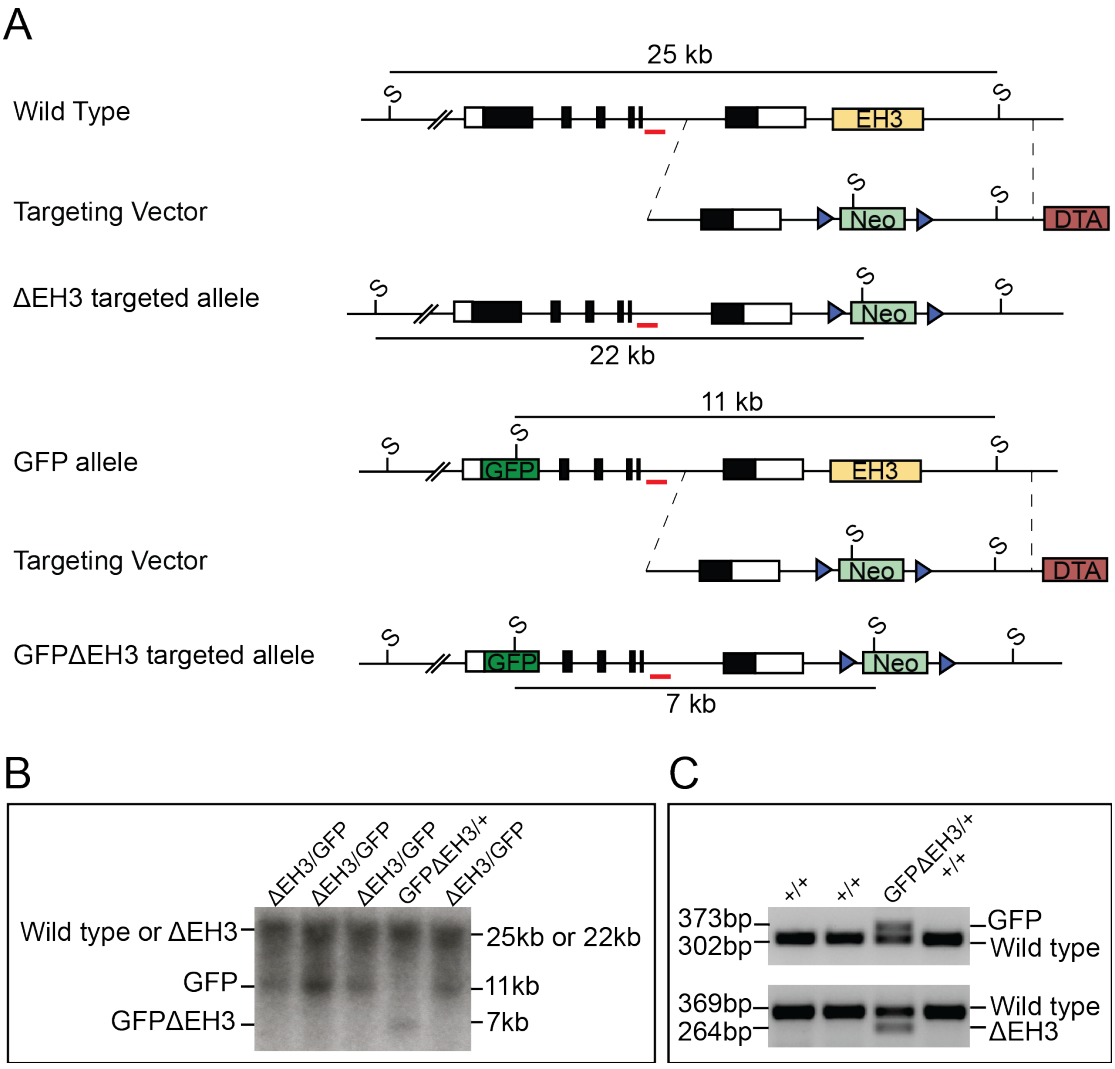


Figure 3.8: Deletion of EH3 in the context of an *Eomes*^{GFP} reporter allele. (A) Heterozygous *Eomes*$^{GFP/+}$ (Arnold et al., 2009) mESCs were re-targeted using the same construct and primary screening strategy to delete EH3 as shown in Figure 3.4. Southern blot strategy to distinguish between GFP and wild type ΔEH3 targeted alleles, and expected fragment sizes are indicated. S = SpeI. (B) Southern blot showing two different genotypes of successfully targeted clones; *Eomes*$^{GFPΔEH3/+}$ and *Eomes*$^{GFP/ΔEH3}$.

To quantify the effect of the EH3 deletion on the expression of the reporter, I analysed embryos harbouring the *Eomes*$^{GFP\Delta EH3/+}$ allele by immunofluorescence staining. Embryos were stained with an anti-GFP antibody, Phalloidin dye to stain F-actin cytoskeleton, and DAPI to label nuclei. Confocal microscopy images of E6.5 embryos reveals that GFP is expressed in the same domains as endogenous Eomes, and the *Eomes*^{GFP} allele; the ExE, primitive streak and nascent mesoderm, and the visceral endoderm **(Figure 3.9A,B)**. In agreement with the whole-mount in situ analysis, these results show that deletion of the enhancer fails to perturb domains of *Eomes* expression at E6.5. Although, the data suggests that expression of the GFP reporter may be weaker in all *Eomes* expressing tissues.

As the levels of *Eomes* are very dynamic within the primitive streak, and change dramatically within the space of a few hours, I instead used a cell differentiation to better quantify the effect of the EH3 deletion on the expression of the reporter. Flow analysis was performed to measure the levels of GFP in differentiated ESCs. Wild type (GFP negative), *Eomes*$^{GFP/+}$ and *Eomes*$^{GFP\Delta EH3/+}$ ESCs were grown without LIF as embryoid bodies for 4 days. Embryoid bodies were then dissociated and subjected to flow cytometry analysis. The intensity of GFP expression is greatly reduced in the *Eomes*$^{GFP\Delta EH3/+}$ cells, but is not diminished to negative-control levels **(Figure 3.9C)**. The percentage of positive GFP cells in wild type cells was 0%, *Eomes*$^{GFP/+}$ cells were 96% GFP positive. The percentage of *Eomes*$^{GFP\Delta EH3/+}$ cells positive for GFP was reduced to 39% **(Figure 3.9D)**. Therefore, deletion of the enhancer, EH3, from the *Eomes*^{GFP} reporter significantly reduces the levels of GFP, but does not completely abolish expression.

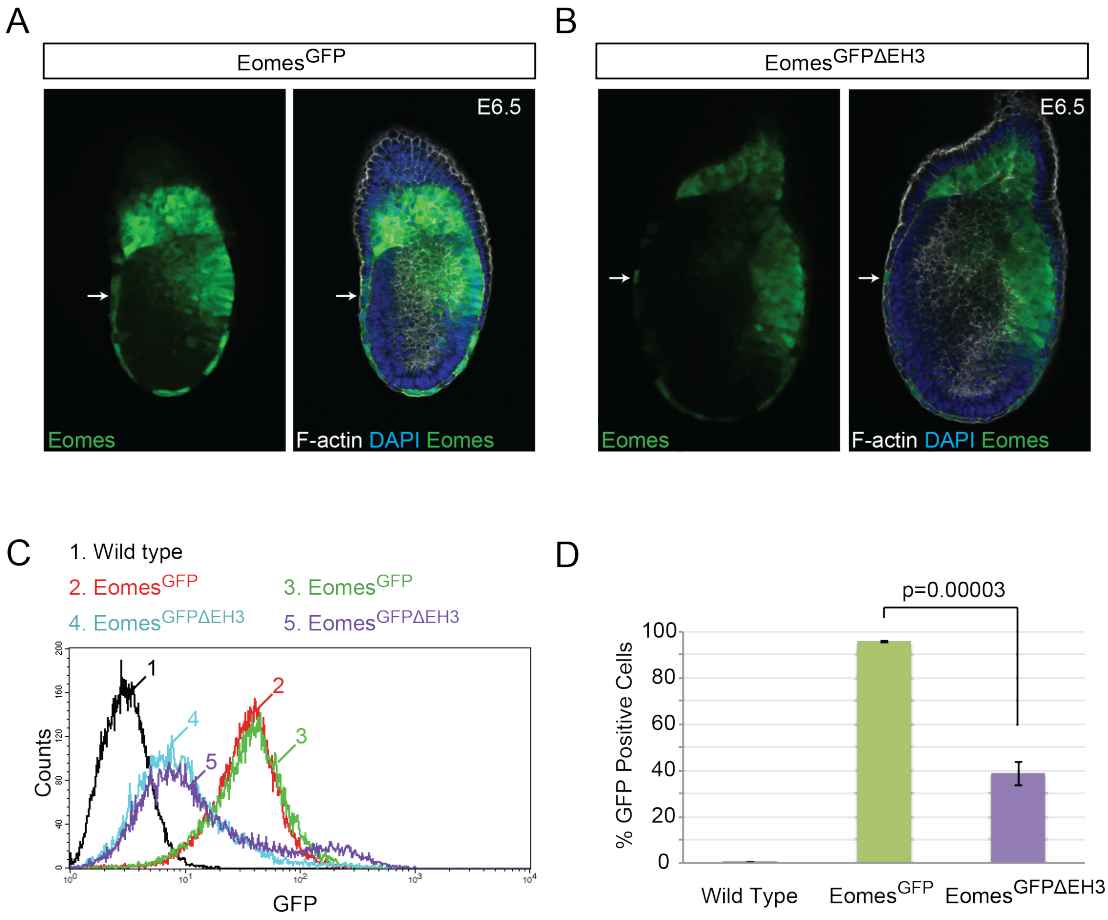


Figure 3.9: Loss of the EH3 enhancer profoundly reduced the level of GFP reporter gene expression.

(A,B) Confocal microscopy images of *Eomes*^{GFP} and *Eomes*^{GFPΔEH3} reporter expression at E6.5. Embryos were staining using anti-GFP antibody, DAPI (DNA) and Phalloidin (F-actin) stains. Arrows indicate expression in the AVE. Domains are expression are conserved but appear to be expressed at lower levels at E6.5 (C) Histogram of flow cytometry data comparing GFP intensity of *Eomes*$^{GFP/+}$ and *Eomes*$^{GFPΔEH3/+}$ ESCs on day 4 of EB differentiation. GFP expression is reduced upon removal of the enhancer. Representative data from two independent experiments with one wild type, two *Eomes*^{GFP} and two *Eomes*^{GFPΔEH3} clones. (D) Quantification of percentage of GFP positive cells from two independent flow cytometry experiments. *Eomes*^{GFPΔEH3} results in a significant reduction in the percentage of GFP cells (Student t-test, p=0.00003).

3.4.3.1 The EH1b enhancer is dispensable for proper Eomes expression in vivo

A targeting vector to delete the 2,019bp EH1b region 7.9kb – 5.9kb upstream of the *Eomes* TSS by homologous recombination was transfected into mESCs **(Figure 3.10A)**. Successfully targeted clones were identified by Southern blot, and were used to generate EH1b deleted (*Eomes*$^{\Delta EH1b}$) mice **(Figure 3.10B,C)**.

Eomes $^{\Delta EH1b/+}$ mice were intercrossed and the genotypes of their offspring are summarised in **Table 3.3**. Mice homozygous for the deletion, *Eomes* $^{\Delta EH1b/\ \Delta EH1b}$, were viable and observed in Mendelian ratios (22%, n=23/106, Chi-square test, p=0.4). *Eomes* $^{\Delta EH1b/\ \Delta EH1b}$ mice are indistinguishable from wild type littermates and can be maintained as a homozygous strain.

To investigate whether this enhancer deletion leads to a hypomorphic phenotype similar to the deletion of EH3, *Eomes* $^{\Delta EH1b/\ \Delta EH1b}$ were mated with *Eomes*$^{N/+}$ mice. The genotypes of offspring generated from this cross are shown in **Table 3.4**. *Eomes*$^{\Delta EH1b/N}$ are present in Mendelian ratios (48% n=33/69, Chi square test, p=0.7) and are indistinguishable from control littermates. Therefore, the enhancer EH1b is dispensable for proper *Eomes* expression *in vivo*.

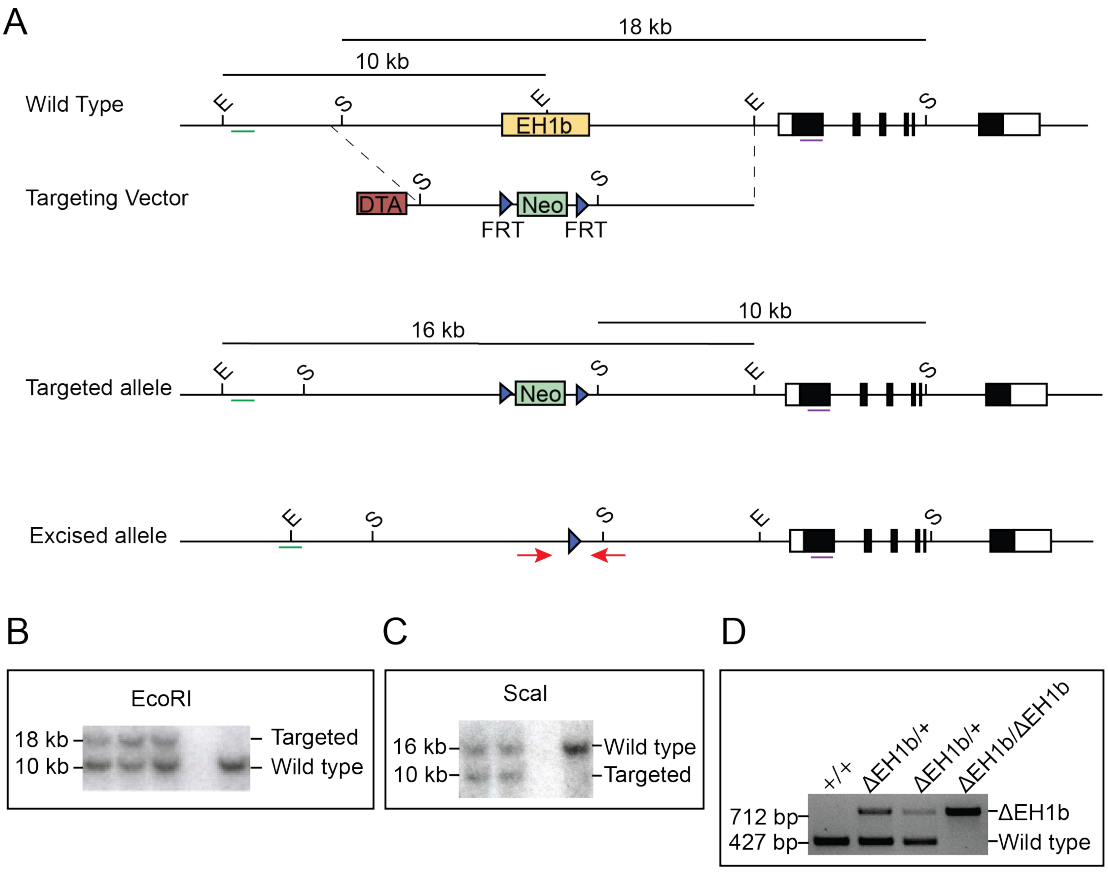


Figure 3.10: Generation of *Eomes*$^{\Delta EH1b}$ mice.
(A) Targeting strategy to delete the 2019bp EH1b region by homologous recombination in mESCs. Targeting vector contains FLP-recombinase recognition (FRT) site flanked Neomycin resistance positive selection cassette, 5' and 3' arms homologous to the *Eomes* locus, and a diphtheria toxin A (DTA) negative selection cassette. Southern blot restriction digest indicated with probes and expected fragment sizes. (B,C) Southern blot of successfully targeted mESC clones. D) PCR genotyping of *Eomes*$^{\Delta EH1b}$ mice.

Genotype	Wild type	*Eomes*$^{\Delta EH1b/+}$	*Eomes*$^{\Delta EH1b/\Delta EH1b}$	Total
Observed	41 (39%)	42 (40%)	23 (22%)	106 (100%)
Expected	26.5 (25%)	53 (50%)	26.5 (25%)	106 (100%)

Table 3.4: Genotypes of *Eomes* $^{\Delta EH1b/+}$ intercross offspring.
Eomes $^{\Delta EH1b/+}$ intercross offpring were subjected to PCR genotyping. The table shows the observed and expected numbers and corresponding percentages of genotypes of offspring at weaning age recovered from the cross.

Genotype	*Eomes*$^{\Delta EH1b/+}$	*Eomes*$^{\Delta EH1b/N}$	Total
Observed	36 (52%)	33 (48%)	69 (100%)
Expected	34.5 (50%)	34.5 (50%)	69 (100%)

Table 3.5: Genotypes of *Eomes* $^{\Delta EH1b/\ \Delta EH1b}$ x *Eomes* $^{N/+}$ offspring
Eomes $^{\Delta EH1b/\ \Delta EH1b}$ x *Eomes*$^{N/+}$ offspring were subjected to PCR genotyping. The table shows the observed and expected numbers and corresponding percentages of genotypes of offspring at weaning age recovered from the cross.

3.5 Discussion

This present work provides insight into the regulation of the *Eomes* gene during early embryonic development. Enhancers located -10kb, -7kb upstream, and +8kb relative to the *Eomes* TSS are highly conserved amongst vertebrates, and are poised in mouse ESCs for subsequent activation during differentiation.

3.5.1 The primitive streak enhancer EH1b is dispensable for proper regulation of Eomes

Here, I have demonstrated that the combined activity of EH1a and EH1b (EH1) is sufficient to drive transgenic reporter expression within the primitive streak. Generation of stable transgenic lines for EH1LacZ would help to identify if there are additional sites of expression at different embryonic stages. However, gene-targeting experiments demonstrate that EH1b is dispensable for proper *Eomes* expression in the embryo.

In humans, the homologous EH1b region has long been cited as an archetypal poised developmental enhancer in ESCs (Buecker and Wysocka, 2012; Rada-Iglesias et al., 2011). In addition, much of the mechanistic insight into TGF-β signalling activation of *Eomes* during human mesendoderm differentiation has been shown to act via this enhancer (Beyer et al., 2013; Estaras et al., 2015; Funa et al., 2015; Kartikasari et al., 2013; Teo et al., 2011; Tsuneyoshi et al., 2012).

Recently, EH1b enhancer function has also been shown in mouse definitive endoderm. The authors of this study demonstrated that recruitment of the histone demethylase Jmjd3 was required for the removal of H3K27me3 repressive chromatin marks during definitive endoderm differentiation. This de-repression allows activation

of the enhancer, and subsequent looping with the *Eomes* promoter to initiate transcription (Kartikasari et al., 2013).

Analysis of differentiating ESCs using ChIP-seq and 3C provides great insight into mechanisms of transcriptional activation and the prediction of potential enhancers. However, as I have demonstrated, enhancer function still needs to be experimentally validated by reporter and deletion experiments in the embryo.

Perhaps, the surprising lack of phenotype is due to redundancy between the different enhancers. Compound deletions of the enhancers would reveal whether other enhancers were shadowing EH1b.

3.5.2 The anterior visceral endoderm and anterior primitive streak enhancer EH3 is required for correct embryonic patterning

By contrast, removal of the EH3 region leads to a profound reduction in *Eomes* expression. This enhancer is sufficient to drive transgenic reporter expression in both the primitive streak and the anterior visceral endoderm (AVE). The different classes of phenotypes of *Eomes* $^{\Delta EH3/N}$ mutants correspond to these domains of expression.

Previous work has shown that *Eomes* is required for the maintenance and migration of the AVE (Nowotschin et al., 2013). Here, I demonstrate that when the levels of *Eomes* are lowered in *Eomes*$^{\Delta EH3/N}$ embryos, defects in the AVE results in a failure to correctly specify the anterior-posterior axis. Furthermore, *Eomes* is required in the primitive streak for the expression of anterior primitive streak (APS) markers and for the formation of definitive endoderm (Arnold et al., 2008). In this present study, a second class of *Eomes*$^{\Delta EH3/-}$ embryos exhibit mis-specification of APS structures, such

as the definitive endoderm, node, and midline. The lack of APS progenitors in these embryos results in faults in anterior neurectoderm patterning and ventral closure defects.

The range of developmental abnormalities in *Eomes*$^{\Delta EH3/N}$ embryos is consistent with *Eomes* being a downstream target of Nodal (Brennan et al., 2001; Nowotschin et al., 2013), and also mediating Nodal signalling (Picozzi et al., 2009; Teo et al., 2011). Intriguingly, the effect of lowering *Eomes*, by deletion of the EH3 enhancer, phenocopies a decrease in Nodal signalling, loss of FoxH1 (a Smad2/3/4 co-factor), and compound *Eomes* and Nodal heterozygotes (Arnold et al., 2008; Hoodless et al., 2001; Norris et al., 2002; Yamamoto et al., 2001). In human definitive endoderm cultures the EH3 region is co-bound by FOXH1, SMAD2/3 and EOMES. The similarity of the phenotypes seems to suggest that FoxH1 and Smad2/3 may also converge on this enhancer in mouse. Future experiments aim to address whether FoxH1 is directly regulating *Eomes*, via the EH3 enhancer, or indirectly through auto regulation of *Nodal* expression.

EH3 enhancer function specifically in the AVE and APS suggests a similar mechanism of activation, mediated by high levels of Nodal signalling, which is distinct from *Eomes* regulation in the cardiac mesoderm and extra-embryonic ectoderm (ExE). The cascade of transcription factor expression in the AVE is very similar to the gene regulatory network that later governs definitive endoderm differentiation, with *Eomes* playing pivotal roles in both tissues. Perhaps the gene regulatory networks in endoderm tissues, including the activation of *Eomes* by Nodal, share an evolutionary origin. By contrast, initiation of *Eomes* expression in the ExE is independent of Nodal (Brennan et al., 2001). Separate mechanisms may act to control the expression of *Eomes* in this tissue and may comprise of distinct enhancers,

regulated by transcription factors specific to the trophoblast lineage, such as Ets2 and Elf5 (Donnison et al., 2005; Polydorou and Georgiades, 2013). Attempts to find trophectoderm enhancers of *Eomes* have so far failed to find robust drivers of reporter expression (Pernaute et al., 2010a).

Chapter 4: Generation of an *Eomes* driven improved Cre-recombinase mouse line

4.1 Introduction

Eomes expressing cells within the primitive streak have been shown to give rise to both cardiac mesoderm and definitive endoderm lineages in fate mapping studies (Costello et al., 2011). However, the *Eomes*^{Cre} mouse line used in this study gave mosaic recombination of the Rosa^{26R} reporter locus and differences between littermates in fate-mapping experiments. Whole-mount in-situ hybridisation revealed robust levels of Cre mRNA expression; therefore the stability or translation efficiency of the *Eomes*^{Cre} must be problematic. Unreliable recombination efficiency has been seen in other Cre mouse lines and can cause major difficulties in interpreting inconsistent phenotypes of conditional gene deletion experiments (Heffner et al., 2012).

I aim to design a more robust *Eomes* driven Cre allele to produce a finer and more precise fate map for *Eomes* positive cells. In addition, the new allele could be used as a tool to reliably delete conditional alleles in a tissue specific manner.

4.2 Targeting of *Eomes* improved Cre allele

The *Eomes*^{Cre} allele was generated by recombining a Cre-recombinase targeting vector into the first start codon of the endogenous locus (Costello et al., 2011). Exon1 is partially deleted and replaced with a bacterial Cre recombinase. The variable Cre recombinase activity may be caused by poor translational efficiency due to a lack of a polyA signal immediately after the Cre. To produce a more robust allele, I used an alternative strategy utilising a β-globin polyA tail to increase stability and translation of the mRNA, and an improved Cre (iCre) (Shimshek et al., 2002), which is a

mammalian codon-optimised version of the prokaryotic Cre recombinase **(Figure 4.1A)**.

The *Eomes*^{iCre} targeting construct was transfected into mESCs and successfully targeted clones were screened by Southern blot, and selected for blastocyst injection[3] **(Figure 4.1B,C)**. High percentage chimeras were obtained and bred to produce heterozygous mice that were identified by PCR genotyping **(Figure 4.1D)**.

To assess whether expression of *iCre* mRNA reflected endogenous *Eomes* mRNA expression whole-mount in situ hybridisation was performed on *Eomes*$^{iCre/+}$ embryos. At E6.5, *iCre* mRNA is observed in the primitive streak, which recapitulates endogenous *Eomes* expression **(Figure 4.2A-D)**. However, *iCre* mRNA is not detectable within the extra-embryonic ectoderm (ExE) or chorion **(Figure 4.2E-H)**.

[3] Chimeras were generated by Elizabeth Robertson.

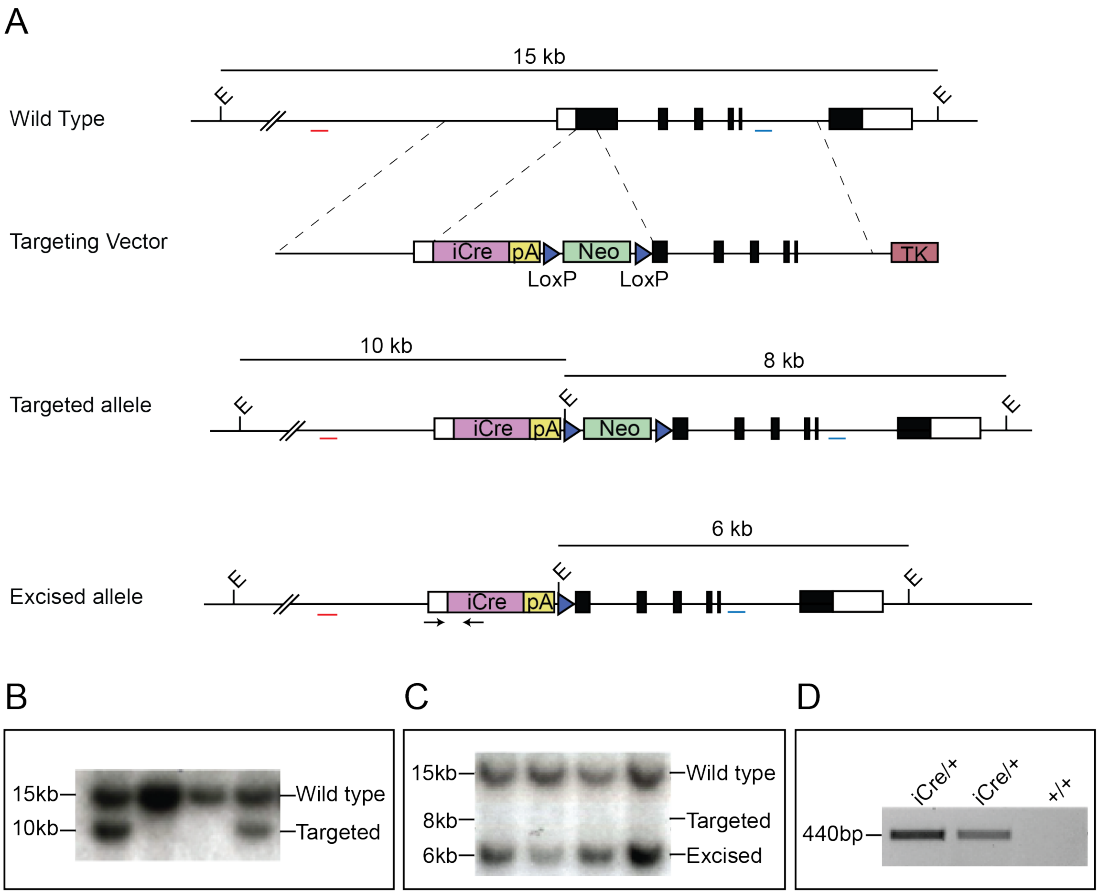


Figure 4.1: Generation of *Eomes*^{iCre} mice.
(A) The targeting vector introduces a mammalian codon-improved Cre (iCre), β-globin polyA cassette into the endogenous *Eomes* initiator methionine. LoxP flanked PGK-Neo positive drug selection cassette and TK negative selection cassette. (B) Southern blot showing wild type (15kb) and targeted (10kb) alleles. (C) Southern blot after Cre-mediated excision of the Neo drug selection cassette showing excised targeted (6kb) and wild type (15kb) alleles and a loss of a targeted allele (8kb). (D) PCR genotyping of *Eomes*^{iCre} mice; iCre band size is 440bp. Arrows indicate PCR primers. E = EcoRI.

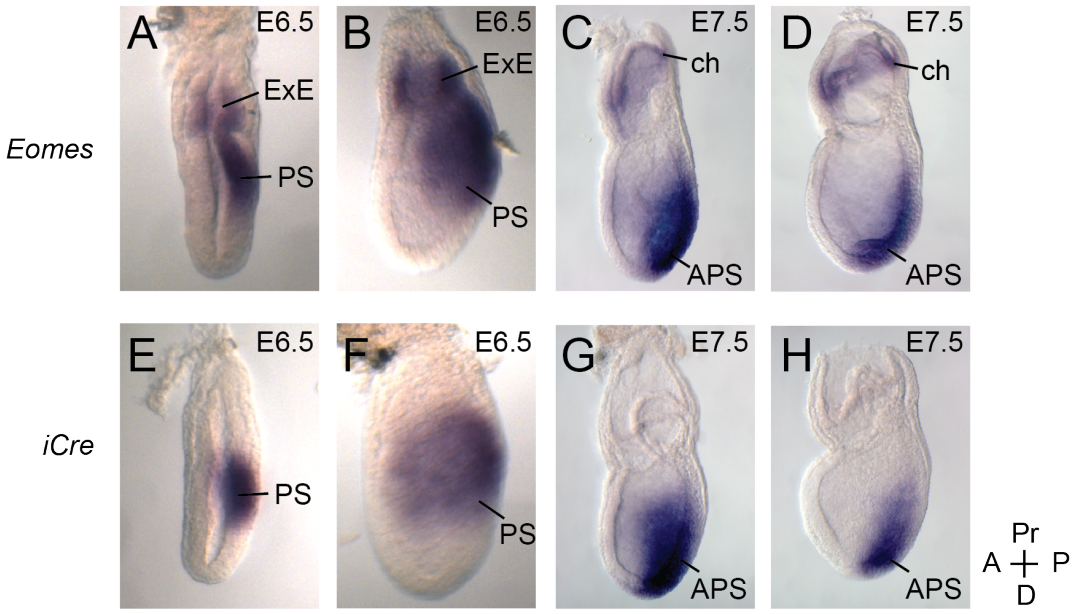


Figure 4.2: Assessing the fidelity of *iCre* expression from the *Eomes*^{iCre} reporter allele.

Whole-mount in situ hybridisation comparing *Eomes* and *iCre* mRNA expression in *Eomes*$^{iCre/+}$ embryos at E6.5. (A-D) *Eomes* is expressed in the extra-embryonic ectoderm (ExE) and the primitive streak (PS) at E6.5, and in the anterior primitive streak (APS) and chorion (ch) at E7.5 in wild type embryos. (E-H) By contrast, *Eomes* driven iCre expression is restricted to the primitive streak at E6.5 and the APS at E7.5.

4.3 Fate mapping using the *Eomes* improved Cre allele

To test if the new *Eomes* driven iCre allele generated more robust and reliable Cre-mediated deletion, I generated a fate-map of *Eomes* expressing cells. *Eomes*$^{iCre/+}$ mice were crossed to the Rosa26 LacZ reporter mouse line (Rosa^{26R}) (Soriano, 1999), and E7.5 and E8.5 embryos were stained with X-gal and then paraffin embedded, sectioned and counterstained with phenol red **(Figures 4.3A-E and 4.4A-E)**.

At E7.5, the ExE, primitive streak, visceral and definitive endoderm and nascent mesoderm all consist of LacZ positive cells **(Figure 4.4A-E)**. As expected, the ectoplacental cone (EPC) derived from *Eomes* positive trophectoderm also contains LacZ expressing cells **(Figure 4.3E)**. Marked cells are absent from the extra-embryonic endoderm and anterior epiblast cells. This is consistent with fate-map from the previous *Eomes*^{Cre} allele and known domains of *Eomes* expression. Therefore, at E7.5 the *Eomes*^{iCre} appears to faithfully map the progeny of *Eomes* positive cells.

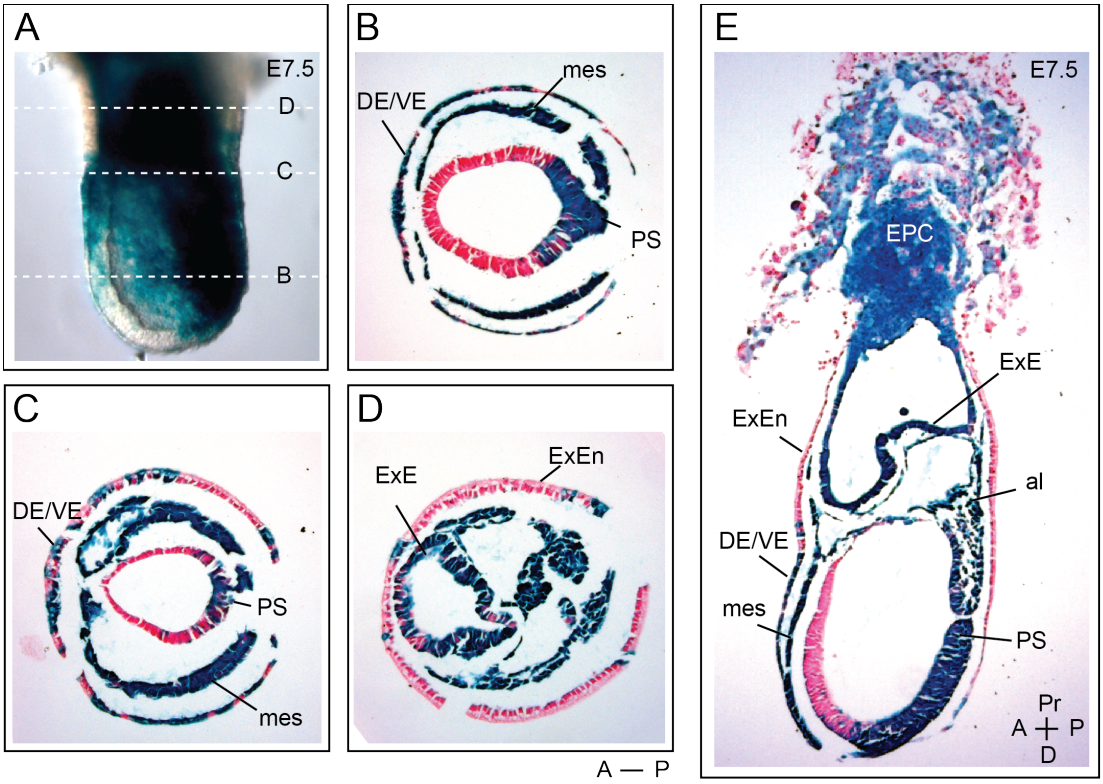


Figure 4.3: Fate mapping studies of *Eomes*^{iCre} at E7.5.
Fate map generated from *Eomes*^{iCre} x Rosa^{26R} cross. Progeny of cells that transiently expressed *Eomes*^{iCre} identified via X-gal staining. Sections were counter-stained with nuclear fast red. (A) Whole-mount stained E7.5 embryo. (B) Mid-sagittal section of E7.5 embryo. (C-F) Transverse sections of E7.5 embryo. Dotted lines in (A) indicate plane of section. Positive cells reside in the ectoplacental cone (EPC), chorion, mesoderm (mes), definitive and visceral endoderm (DVE/VE), allantois (al), primitive streak (PS). Marked cells are absent from the ectoderm and extra-embryonic endoderm (ExEn).

Surprisingly, by E8.5 all primitive streak derivatives are labelled; both extra-embryonic and embryonic mesoderm and the definitive endoderm. *Eomes*+ cells in the primitive streak contribute to extra-embryonic mesoderm of the yolk sac blood islands, allantoic mesoderm, and amniotic mesoderm. *Eomes*+ cells in the primitive streak contribute to the embryo proper and include the cardiac mesoderm, head mesenchyme, somitic mesoderm and the definitive endoderm derived gut tube **(Figure 4.4A-E)**. The neurectoderm of the neural folds and the endoderm layer of the yolk sac are devoid of marked cells. The fate map suggests that $Eomes^{iCre}$ expressing progenitor cells within the primitive streak give rise to both extra-embryonic and embryonic mesoderm, and definitive endoderm.

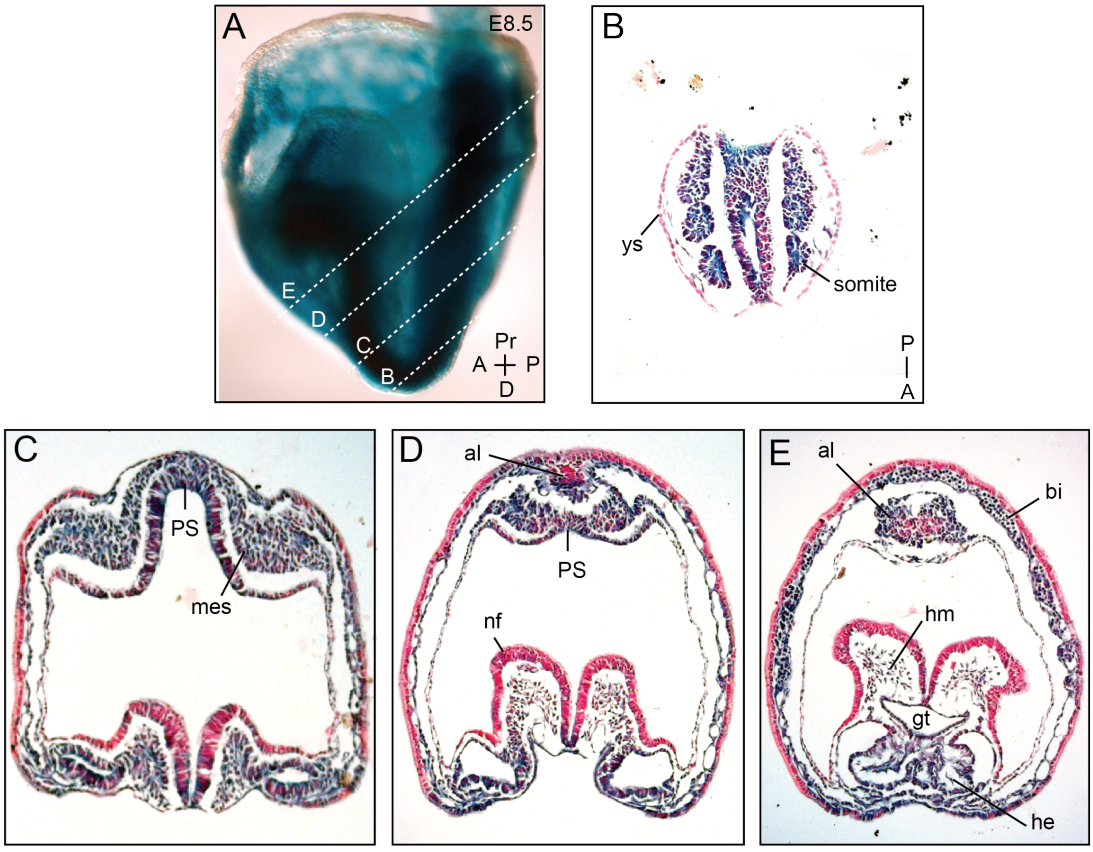


Figure 4.4: Fate mapping studies of *Eomes*^{iCre} at E8.5.
Fate map generated from *Eomes*^{iCre} x Rosa^{26R} cross. Progeny of cells that expressed *Eomes*^{iCre} stained blue with X-gal. Sections were counter-stained with nuclear fast red. (G) Whole-mount stained E8.5 embryo. Dotted lines indicate transverse plane of sections shown in (H-I) At E8.5 progeny of *Eomes*^{iCre} expressing cells are reside in definitive endoderm derived gut tube (gt) and all mesodermal subtypes; somites, posterior mesoderm (mes), heart (he), head mesenchyme (hm), allantois (al) and yolk sac blood islands (bi). Unlabelled cells reside in the endoderm layer of the yolk sac (ys) and the neural folds (nf).

4.4 Discussion

I have designed a new *Eomes*^{iCre} allele that is able to recombine Rosa^{26R} locus in cells expressing *Eomes* at E7.5; the ectoplacental cone (EPC), extra-embryonic ectoderm (ExE), visceral endoderm overlying the embryo and the primitive streak. Interestingly, at later stages the fate map labels all derivatives of the primitive streak; definitive endoderm and both extra-embryonic and embryonic mesoderm, suggesting that all of these cell types originated from *Eomes*^{iCre} expressing progenitors within the streak.

Previously, fate maps of *Eomes* cells generated using two independent *Eomes* Cre alleles demonstrate that *Eomes* progenitors from the primitive streak give rise to the head mesenchyme, cardiac mesoderm, vasculature, definitive endoderm and the mesodermal component of the yolk sac blood islands (Costello et al., 2011; Pimeisl et al., 2013). This is in agreement with the fate map I have generated with the *Eomes*^{iCre} allele. However, in contrast to *Eomes*^{iCre} fate map, in both of these studies paraxial and lateral plate mesoderm are negative for *Eomes*+ derived cells.

The new *Eomes*^{iCre} allele is a very efficient and robust at recombining the Rosa^{26R} reporter allele and does not show any mosaicism in a similar way to the previous *Eomes*^{Cre}. This would make it ideal for conditional deletion studies. However, due to the broad expression pattern of the iCre allele, at such early embryonic development it may not provide useful as a conditional deletion strategy for the majority of genes.

4.4.1 EomesiCre progeny reside in the mesoderm component of the yolk sac blood islands

Lineage tracing of *Eomes*+ cells to the mesodermal component of the yolk sac blood

islands has been demonstrated by previous Cre alleles. (Costello et al., 2011; Pimeisl et al., 2013). Yolk sac blood islands are derived from the earliest mesodermal cells types that exit the streak and migrate to the extra-embryonic visceral endoderm. Here, they form blood islands, comprising of mesodermal derived endothelial, haematopoietic and smooth muscle cells. *Eomes* is expressed in the primitive streak at the onset of gastrulation, at the time that this lineage is allocated. It would be interesting to investigate whether *Eomes* is essential in the specification of this cell type; this is something I hope to address in Chapter 5.

4.4.2 *Eomes^{iCre}* fate-map labels posterior mesoderm

Eomes has previously been shown to be required for cardiac mesoderm specification. However, the requirement of *Eomes* in more posterior mesoderm cell types has not been demonstrated. RNAi knock-down of *Eomes* in mESCs suggest that it is involved in the earliest steps of mesendoderm formation, and that it may also be necessary for lateral plate mesoderm formation, but dispensable for paraxial mesoderm (Izumi et al., 2007). Furthermore, *Eomes* null mESCs are able to colonise posterior mesoderm tissues and form mesoderm subtypes within teratomas (Arnold et al., 2008; Russ et al., 2000). Therefore, although *Eomes^{iCre}* progeny reside within posterior mesoderm, *Eomes* is unlikely to play a role in the differentiation of this tissue.

However, *Eomes* is absolutely required for morphogenetic movements of all mesoderm through the primitive streak (Russ et al., 2000). Graded levels of the closely related T-box family member, Brachyury, have been shown in influence the timing of exit from the primitive streak in chimera experiments. Cells with higher levels of Brachyury tend to exit the streak prematurely and are nearly absent from posterior mesoderm (Wilson and Beddington, 1997). Perhaps, Eomes also acts in a

dosage-dependent way to regulate the timing of ingression through the primitive streak, and thus allocate cell fates. Cells with lower levels of Eomes may reside in the streak longer, contributing to more posterior mesodermal cell types. If this were true, previous Cre alleles that were not as sensitive as *Eomes*^{iCre} would not detect contributions to these lineages.

Alternatively, cells within the streak may stochastically express a number of lineage specifiers, including *Eomes*, before they become allocated to a particular fate. Perhaps the high stability and translational efficiency of *Eomes*^{iCre} is such that even low expressing cells also recombine the Rosa^{26R} locus, leading to an ectopic fate map. Similarly, *Blimp1*^{iCre} mice (unpublished observations from our lab) also produce an ectopic fate map due to the high efficiency of iCre in recombining the Rosa^{26R} locus in stochastically low expressing *Blimp1* epiblast cells.

4.4.3 Eomes^{iCre} mRNA cannot be detected in the ExE

Whole-mount in situ hybridisation is unable to detect *iCre* mRNA expression within the ExE of *Eomes*^{iCre} E6.5 embryos, but *iCre* mRNA is detected within the normal domains in the primitive streak. It is possible that insertion of the improved Cre-recombinase and the β-globin polyA may disrupt transcriptional regulation of the locus specifically within the ExE (perhaps by disrupting an enhancer) whilst regulation within the primitive streak is unaffected.

Despite the absence of *iCre* mRNA at E6.5 the ExE is positively labelled at E7.5. In the blastocyst *Eomes* is expressed throughout the trophectoderm, which gives rise to both the EPC and ExE. Therefore, positively labelled cells in the ExE at E7.5 is likely due to earlier *Eomes* trophectoderm expression. Whether subsequently there are

sufficient low levels of iCre protein are translated in the ExE to recombine the Rosa^{26R} locus, is unknown.

Chapter 5: Investigating redundancy between *Eomes* and *Brachyury* during early mesoderm formation

5.1 Introduction

During gastrulation different mesodermal lineages are specified depending on the time and site of ingression through the streak (Lawson et al., 1991). At early-streak stages, the first mesoderm lineage to emerge is the extra-embryonic lineage. This includes cells fated to become erythrocytes that differentiate within the yolk-sac blood islands (Kinder et al., 1999). As gastrulation proceeds, this is followed by ingression of cardiac, cranial, lateral plate, paraxial and axial mesoderm lineages (Arnold and Robertson, 2009; Kinder et al., 2001).

The T-box transcription factors *Eomes* and *Brachyury* are essential for specification of mesendoderm and mesoderm during gastrulation, and have partially overlapping expression within the primitive streak (Beddington et al., 1992; Ciruna and Rossant, 1999; Herrmann, 1991). Conditional deletion of *Eomes* in the epiblast has revealed essential roles in specification of the cardiac mesoderm and definitive endoderm (Arnold et al., 2008; Costello et al., 2011). Analysis of *Brachyury* mutants revealed a critical role in node morphogenesis and posterior mesoderm formation and allantoic vasculogenesis (Conlon et al., 1995; Inman and Downs, 2006a). In addition, both transcription factors are involved in delamination of mesendoderm cells within the streak (Arnold et al., 2008; Russ et al., 2000; Wilson et al., 1995). Furthermore, they have both been shown to activate similar targets during mesoderm differentiation, including *Mesp1* (Costello et al., 2011; David et al., 2011; van den Ameele et al., 2012).

Brachyury has long been used alongside Flk1 as a marker of the haemangioblast; a progenitor with the potential to differentiate into yolk sac endothelial and haematopoietic cells (Huber et al., 2004). A recent study revealed functional haematopoietic targets of Brachyury, although they were largely associated with later waves of haematopoiesis or allantoic vasculogenesis (Lolas et al., 2014). Fate-mapping experiments reveal that the mesodermal component of the yolk-sac blood islands, the first site of primitive haematopoiesis in the embryo, is derived from *Eomes*+ cells (Costello et al., 2011). Whether either of these T-box factors is necessary for specification of haematopoietic mesoderm progenitors that migrate to the yolk sac has not been investigated.

Eomes and *Brachyury (T)* are T-box transcription factors that contain the T-box binding domain that binds the T-half site consensus sequence; TCACACCT. Specificity of targets appears to lie predominantly within the T-box domain although N-terminal and C-terminal regions of the protein are able to bind to other factors to form complexes that may further drive specificity (Conlon et al., 2001; Naiche et al., 2005; Showell et al., 2004).

The aim of this study is to find out to what extent Brachyury and Eomes proteins are co-expressed during mesoderm differentiation. Furthermore, I hope to discover whether they act redundantly to specify haematopoietic progenitors that arise first from the early streak, and differentiate within the yolk sac blood islands.

5.2 Comparison of *Eomes* and *Brachyury* T-box transcription factors

In order to identify similarities between Eomes and Brachyury, I performed a Clustal Omega alignment of the mouse Eomes protein (707 amino acid; UniProt ID: O54839) and mouse Brachyury protein (436 amino acid; UniProt ID: P20293).

The percentage identity was calculated for each region; N-terminal (24%), T-box DNA binding domain (51%) and C-terminal (15%) **(Figure 5.1A).** The amino acid sequence of the T-box domain is the most highly conserved region between the two factors. However, they differ in a crucial residue that makes contact with the DNA and confers binding specificity (Conlon et al., 2001) **(Figure 5.1B)**. The consensus binding sequences of both factors are highly similar, and comparable to the T-half site sequence; TCACACCT **(Figure 5.1C)**.

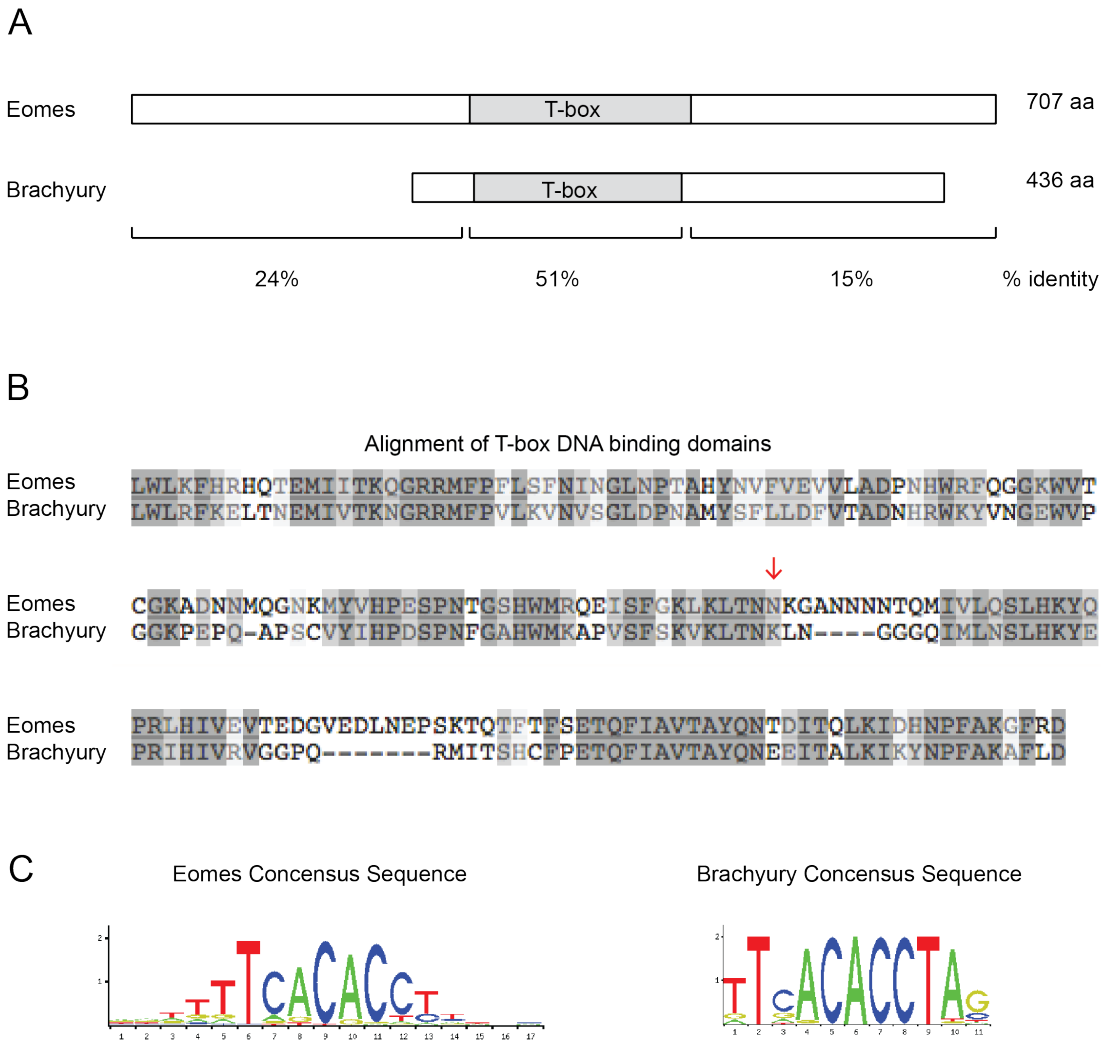


Figure 5.1: Sequence comparison of Eomes and Brachyury domain structure and protein coding information.
(A) Schematic representation of Eomes and Brachyury T-box transcription factors. Amino acid sequences of each protein (Mouse Brachyury UniProt ID:P20293, Mouse Eomes UniProt ID: O54839) were aligned using Clustal Omega and the similarity between the N-terminal, T-box domain, and C-terminal are indicated by percentage identity. (B) UniProt Alignment of DNA binding domain (T-box). Dark grey residues are fully conserved. Medium grey residues indicate residues with strongly similar properties (>0.5 in the Gonnet PAM 250 matrix). Light grey residues indicate residues with weakly similar properties (=<0.5 in the Gonnet PAM 250 matrix). Red arrow indicates lysine (K) and asparagine (N) residues that contribute to T-box specificity of Brachyury and Eomes, respectively. (C) Transcription factor consensus sequence for binding sites available on JASPAR database (http://jaspar.genereg.net/).

5.3 Dynamics of *Eomes* and *Brachyury* expression during early embryonic development and mesoderm specification

To act redundantly in mesoderm allocation *Eomes* and *Brachyury* would have to be expressed in the same spatiotemporal domains. To investigate the kinetics and possible overlap of expression between *Eomes* and *Brachyury* mRNA whole-mount in situ-hybridisation was performed on gastrulating embryos. At mid-streak (MS) stages both *Eomes* and *Brachyury* are expressed in the primitive streak **(Figure 5.2A,D)**. *Eomes* expression domain is much broader than *Brachyury*, persisting in the migrating mesoderm wings. At late-streak (LS) to late bud (LB) stages *Eomes* expression becomes progressively more restricted to the anterior primitive streak (APS) **(Figure 5.2 B,C)**. However, *Brachyury* continues to be robustly expressed throughout the primitive streak and notochord **(Figure 5.2E,F)**. Additionally, *Eomes* mRNA is detected in the extra-embryonic ectoderm (ExE) and later in the chorion (ch) during gastrulation **(Figure 5.2A-C)**.

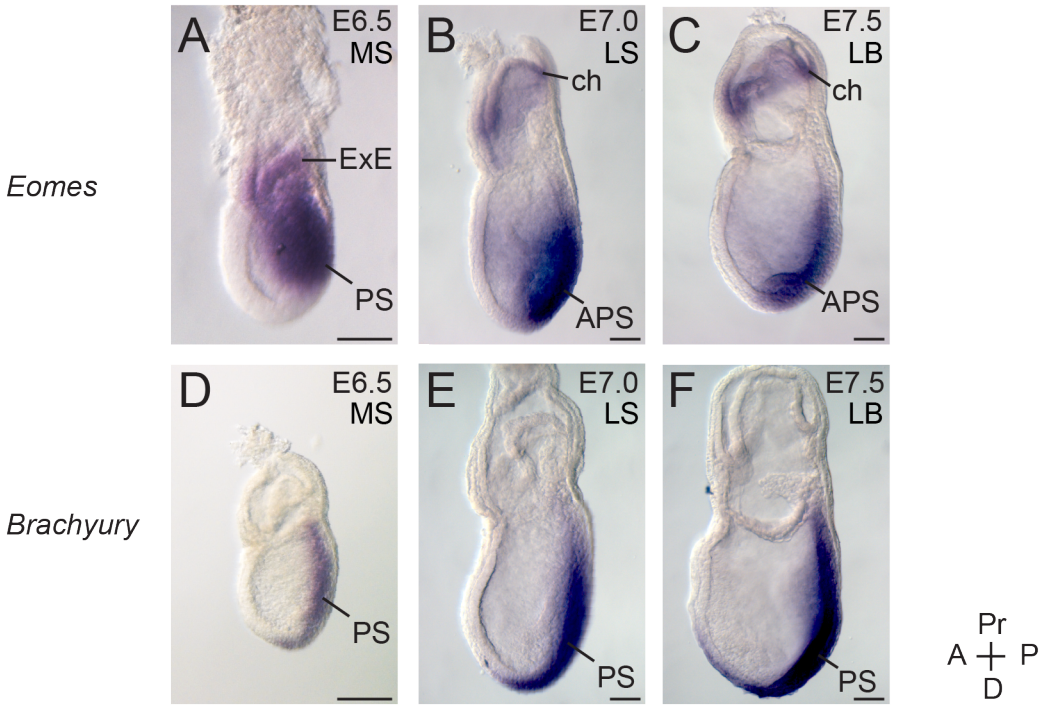


Figure 5.2: Partially overlapping mRNA expression domains of *Eomes* and *Brachyury* within the primitive streak

Whole-mount in situ hybridisation of *Eomes* and *Brachyury* mRNA expression during gastrulation. (A-C) Between E6.5 – E7.5 *Eomes* is expressed in the primitive streak (PS), and becomes restricted to the anterior primitive streak (APS). (D-F) Between E6.5 – E7.5 *Brachyury* is expressed in the full extent of the primitive streak. Scale bars,100µm.

To further investigate whether Eomes and Brachyury proteins are co-expressed, E6.5 embryos were subjected to immunofluorescence staining followed by confocal imaging. At early streak (ES) stage Eomes is robustly detected in the full extent of the streak and nascent mesoderm, whereas Brachyury is only expressed in a very small number of cells the proximal-posterior epiblast **(Figure 5.3A-C)**. Brachyury was undetectable at slightly early stages, and in contrast to mRNA expression, radial expression of Brachyury protein at the extra-embryonic-embryonic boundary was not observed (Thomas et al., 1998). These results suggest that Eomes is expressed prior to Brachyury and they are not co-expressed within the early streak.

By mid-streak (MS) stage the streak is regionalised into Eomes low / Brachyury high in the posterior primitive streak (PPS), and Eomes high / Brachyury high in the anterior primitive streak (APS). Eomes protein is at the highest levels in the APS and in nascent mesoderm. Whereas Brachyury is expressed throughout the primitive streak and is also highly expressed in nascent mesoderm emerging from the PPS. Therefore, Eomes and Brachyury are co-expressed in the anterior region of the mid-streak, and Brachyury expression predominates in the posterior region of the mid-streak.

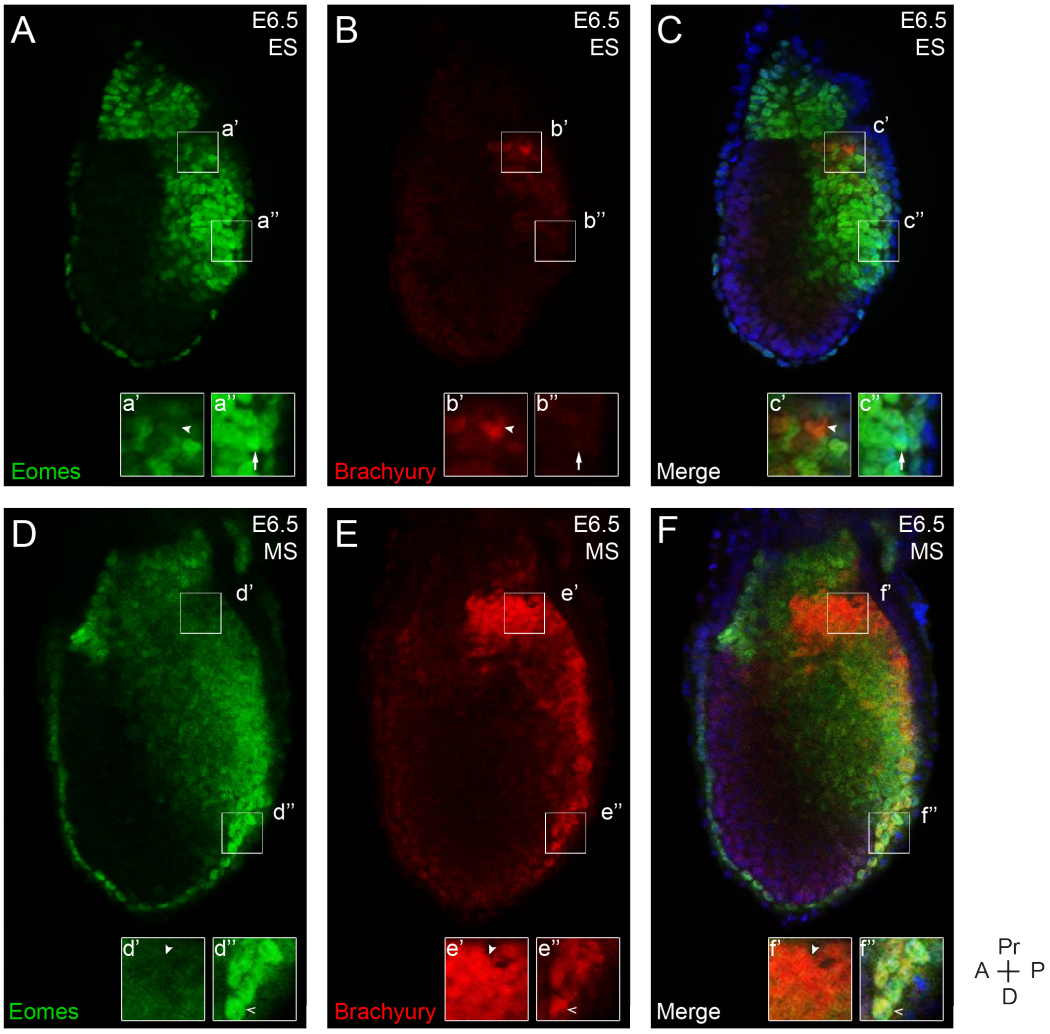


Figure 5.3: Unique and overlapping Eomes and Brachyury cell populations in the early embryo.

Confocal images of embryos stained with anti-Eomes and anti-Brachyury antibodies. (A-C) At early-streak (ES) stage Eomes is robustly expressed within the primitive streak and nascent mesoderm. A small cluster of cells express Brachyury in the most proximal-posterior epiblast. At ES stage, expression of Eomes and Brachyury does not overlap (a'-c',a''-c''). (D-F) By mid-streak (MS) stages the streak is regionalised into Eomes low / Brachyury high in the posterior primitive streak, and Eomes high / Brachyury high in the anterior primitive streak. Boxed regions are shown inset at 2x magnification. Arrows indicate an Eomes high / Brachyury low cell. Closed arrowheads indicate an Eomes low / Brachyury high cell. Open arrowheads indicate a cell with overlapping high Eomes and Brachyury expression.

In vitro differentiation of mESCs as embryoid bodies (EBs) can be used to model gastrulation and study specification of the earliest mesoderm populations to transit the streak; namely haematopoietic and cardiac mesoderm (Choi et al., 1998; Kouskoff et al., 2005). To investigate whether embryonic regionalised patterns of Eomes and Brachyury could be replicated *in vitro*, immunofluorescence staining and confocal microscopy were performed on day 4 EBs. In agreement with embryo data, EBs are regionalised into Eomes high or Brachyury high expressing cells, and a small number of Eomes high / Brachyury high co-expressing cells **(Figure 5.4A)**.

To further assess the kinetics of mesoderm specification within the EBs, mRNA expression was measured by qRT-PCR between d0 – d4 of EB differentiation **(Figure 5.4B)**. In mESCs (d0) mesoderm markers are not expressed. *Eomes* is the first to be strongly expressed at d2.5, and expression levels are maintained at a high level between d3.0 and d4.0. *Brachyury* is expressed slightly later, initially at low levels on d2.5 and reaching high levels of expression between d3.0 – d4.0. Similarly, *ER71* (alternatively known as Etv2, a haematopoietic mesoderm transcription factor) (Lee et al., 2008) is initially expressed at low levels on d2.5, moderately at d3.0, and reaches a high level between d3.5- d4.0. *Mesp1* (a haematopoietic and cardiac mesoderm transcription factor) (Chan et al., 2013; Saga et al., 2000) expression is induced last, with moderate expression at d3.0, peaking at d3.5 before decreasing to moderate levels at d4.0. Thus, EBs seem to recapitulate the kinetics of early mesoderm development in the embryo; *Eomes* is expressed prior to *Brachyury*, and haematopoietic mesoderm (*ER71*) emerges before cardiac mesoderm (*Mesp1*) (Kouskoff et al., 2005).

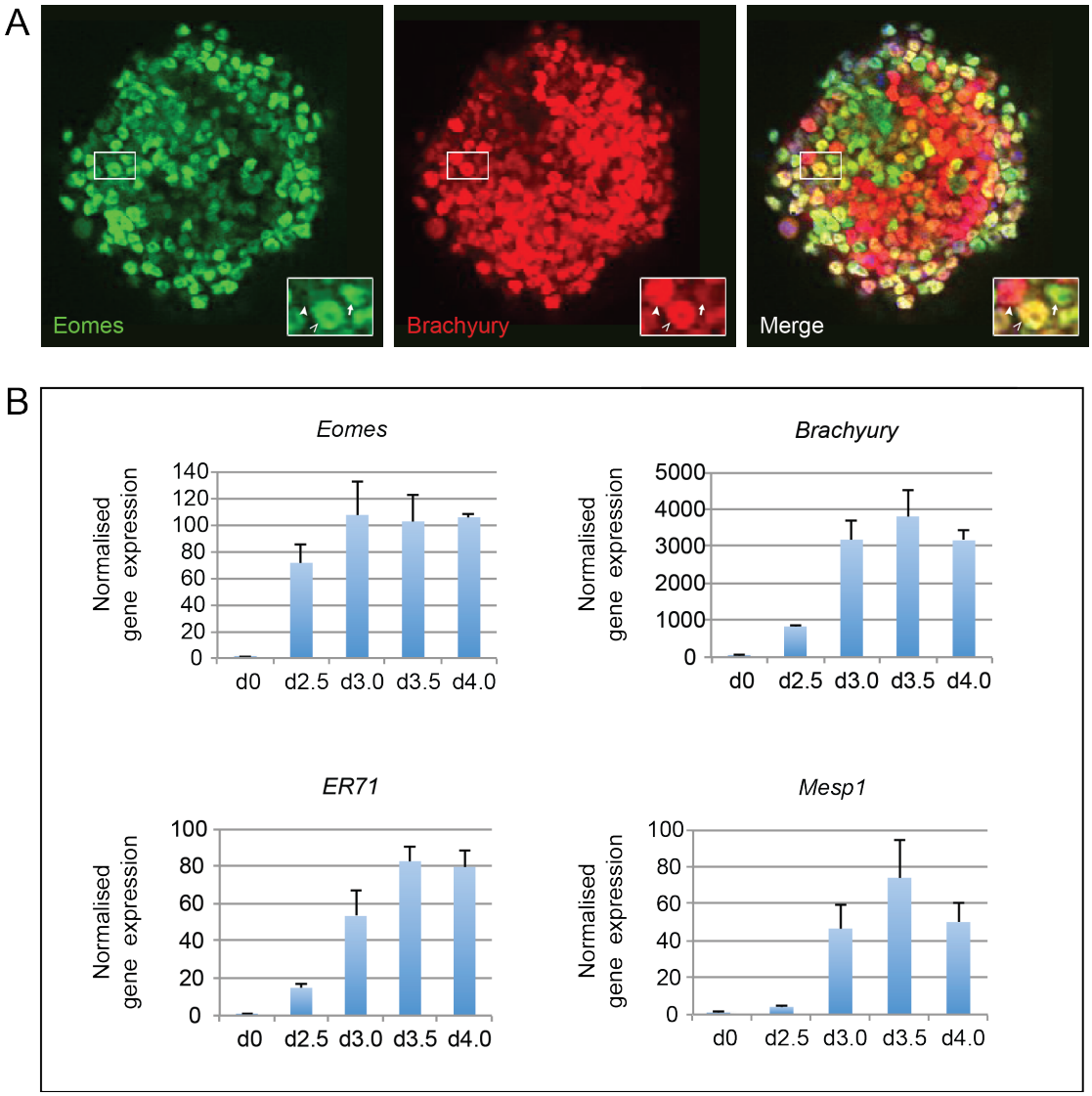


Figure 5.4: Kinetics of mesoderm markers during mESC differentiation.
(A) Confocal images of mESCs differentiation for 4 days as embryoid bodies (EBs) stained with anti-Eomes and anti-Brachyury antibodies. EBs are regionalised into Eomes high or Brachyury high expressing cells, and Eomes high / Brachyury high co-expressing cells. Boxed regions are shown inset at 2x magnification. Arrow indicates an Eomes high / Brachyury low cell. Closed arrowhead indicates an Eomes low / Brachyury high cell. Open arrowhead indicates a cell with overlapping high Eomes and Brachyury expression. (B) qRT-PCR analysis of the kinetics of mesoderm markers during wild type EB differentiation. *Eomes* (cardiogenic mesoderm and endoderm), *Brachyury* (posterior mesoderm), *ER71* (aka Etv2, haematopoietic mesoderm) and *Mesp1* (cardiac mesoderm). Mean ± SEM (n=3). Gene expression normalised to *Gapdh*.

5.4 *Eomes* is required for haematopoietic progenitor specification

Both *Eomes* and *Brachyury* have been implicated in haematopoietic mesoderm fate allocation. *Eomes*+ cells in the primitive streak give rise to haematopoietic and endothelial components of the yolk sac blood islands in fate mapping experiments (Costello et al., 2011). Brachyury+/Flk1+ cells isolated from E7.5 embryos represent a sub-population of mesoderm with high *in vitro* differentiation potential for haematopoietic and endothelial cell types (Huber et al., 2004). However, their roles in haematopoietic specification have not been elucidated.

To assess whether *Eomes* or *Brachyury* are necessary for the specification of haematopoietic mesoderm, I performed EB differentiation on wild type, *Eomes* null (Arnold et al., 2008) and *Brachyury* null (Rashbass et al., 1991) mESCs. Day 4 EBs were subjected to flow cytometry analysis for the surface markers Flk1 (Vegfr2/KDR) and Pdgfrα to discern between mesoderm subsets **(Figure 5.5A)**. The different populations are enriched for particular progenitors: Pdgfrα-/Flk1+ haematopoietic, Pdgfrα+/Flk1+ cardiac, and Pdgfrα+/Fkl1– paraxial mesoderm (Liu et al., 2012).

Eomes null EBs display a dramatic decrease in Flk1+ mesoderm populations (15% compared with 67% in wild type). Specifically, the decrease in Pdgfrα+/Flk1+ population (5% compared with 21% in wild type, p=0.0008) supports the role *Eomes* plays in cardiac specification (Costello et al., 2011). *Eomes* is also required for robust generation of Pdgfrα-/Flk1+ haematopoietic progenitors (10% compared with 25% in wild type, p=0.008). The loss of Flk1+ mesoderm in *Eomes* null EBs is associated with a concomitant increase in paraxial mesoderm progenitors (54% compared with 14% in wild type, p=0.00006), suggesting that *Eomes* may not only specify cardiac and haematopoietic mesoderm, but also inhibits excessive paraxial mesoderm differentiation. Surprisingly, *Brachyury* null cells showed no significant changes in

mesoderm populations when compared with wild type cells. In summary, these results suggest that *Eomes* is necessary for specification of haematopoietic mesoderm. However, *Brachyury* is dispensable for haematopoietic mesoderm specification.

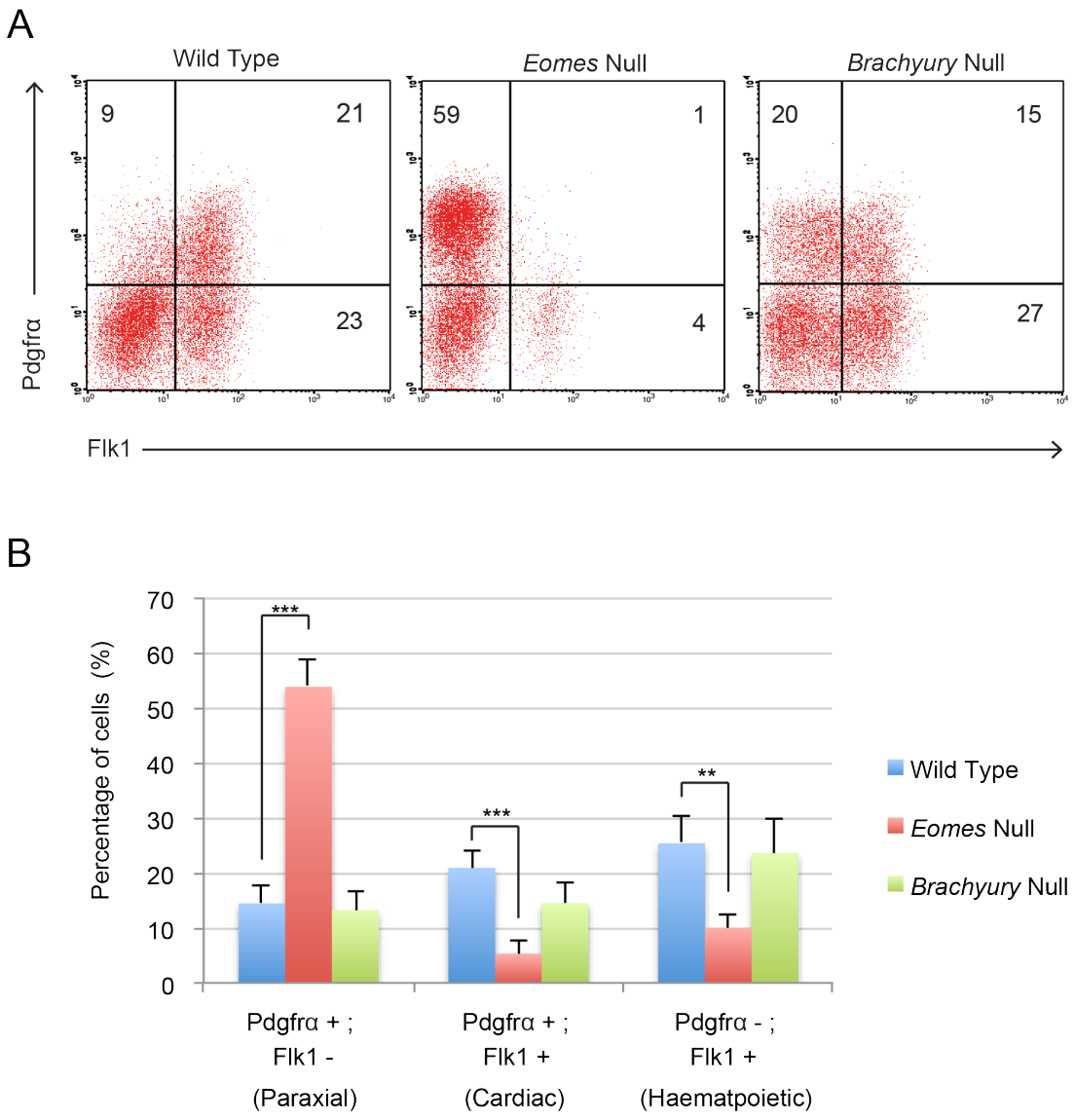


Figure 5.5: *Eomes* is required for specification of cardiac and haematopoietic progenitors *in vitro*

(A) FACS profile (B) and quantification of Pdgfrα and Flk1 mesoderm populations in d4 EBs. Average percentage of mesoderm populations shown for three independent experiments with a total of 2 wild type clones, 3 *Eomes* null clones, and 2 *Brachyury* null clones each. *Eomes* null cells have a significant reduction in presumptive cardiac and haematopoietic mesoderm, with a concomitant increase in presumptive paraxial mesoderm compared to wild type and *Brachyury* null cells. Student's t-test, n=21, ** $p<0.01$ *** $p<0.001$. Error bars represent SEM.

To confirm the haematopoietic specification defect in *Eomes* null cells I performed two functional assays to assess their differentiation potential (Kennedy et al., 1997). Flk1+ cells collected from d3.5 EBs were differentiated to form blast colonies (haematopoietic and endothelial potential). *Eomes* null cells formed significantly fewer blast colonies than wild type cells (13 colonies compared with 91 colonies for wild type, p=0.02) **(Figure 5.6A)**. In addition, d6 EBs were dissociated into single cells and differentiated to form haematopoietic colonies (primitive or definitive erythroid and granulocyte / macrophage cell types). *Eomes* null cells were severely compromised in generating any haematopoietic colonies **(Figure 5.7B)**. Together, this data confirms that *Eomes* null cells are functionally unable to differentiate to haematopoietic cell types *in vitro* due to the early defect in haematopoietic mesoderm specification.

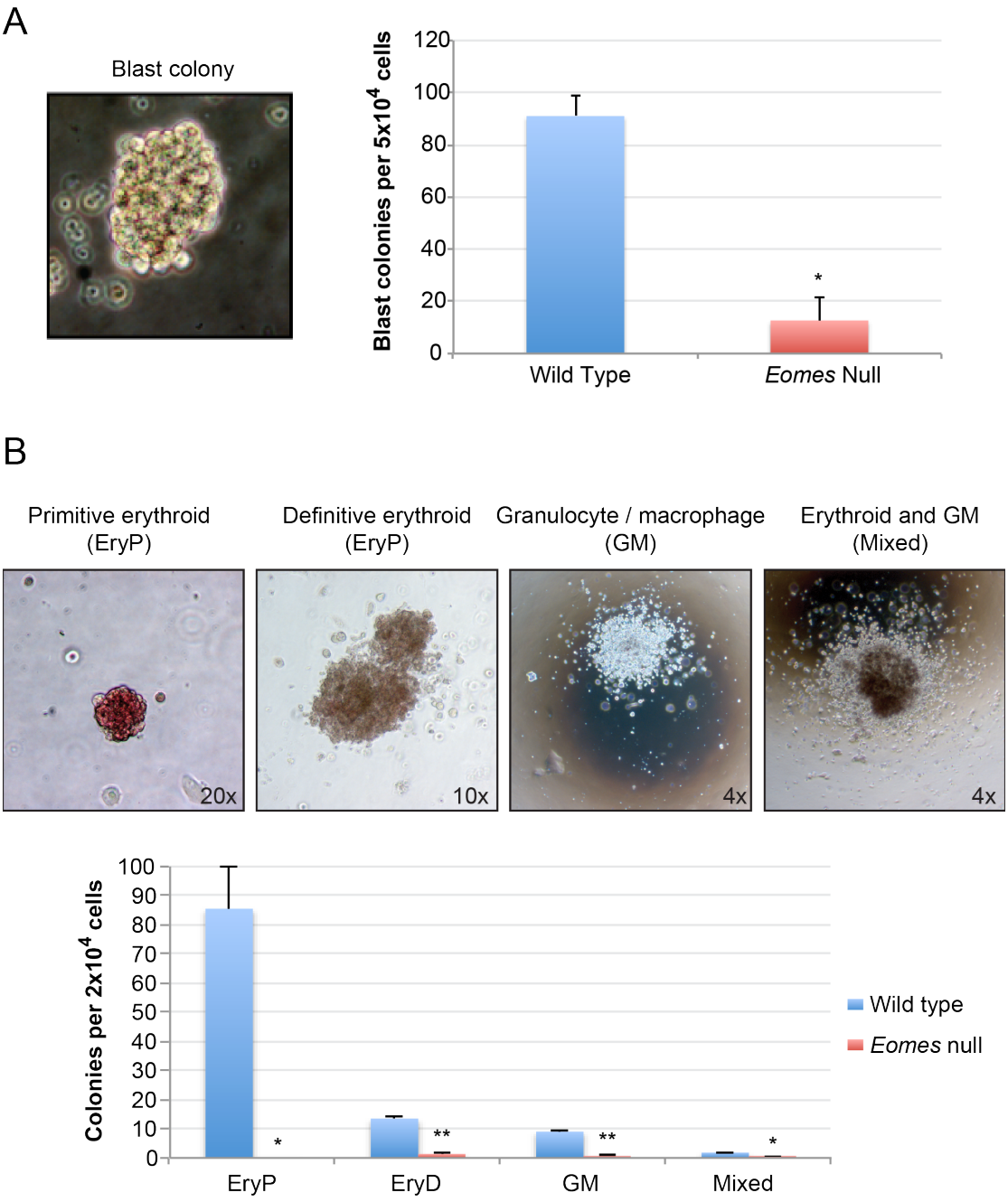


Figure 5.6: *Eomes* is essential for haematopoietic differentiation *in vitro*
(A) Flk1+ cells isolated from d3.5 wild type and *Eomes* null EBs were subjected to a blast colony-forming assay. Average number of blast colonies from triplicate experiments with 2 wild type and 2 *Eomes* null clones are shown. *Eomes* null cells produce significantly fewer blast colonies than wild type cells. (B) Cells from d6 wild type and *Eomes* null EBs were dissociated and subjected to a colony-forming assay. Average number of EryP, EryD, GM and Mixed colonies from triplicate experiments with 2 wild type and 2 *Eomes* null clones are shown. Student's t-test. P-values indicated as * $p<0.05$ ** $p<0.01$. Error bars represent SEM

Conditional ablation of *Eomes* in the epiblast causes a defect in EMT within the primitive streak. Although cells aggregate within the primitive streak, some mesoderm subtypes are still specified (Arnold et al., 2008). In order to investigate whether *Eomes* is necessary for haematopoietic specification, I analysed markers of haematopoietic mesoderm in these mutant embryos. Female mice homozygous for an *Eomes* conditional allele ($Eomes^{CA/CA}$) were crossed to male mice harbouring an *Eomes* null allele ($Eomes^{N/+}$) and epiblast specific Sox2 driven Cre (*Sox2-Cre)* with or without a *Flk1-LacZ* transgene. This cross generated $Eomes^{N/CA}$*;Sox2-Cre* embryos, ablating *Eomes* in the epiblast (referred to as $Eomes^{\Delta epi}$) with or without the *Flk1-LacZ* transgene (Arnold et al., 2008; Ema et al., 2006; Hayashi et al., 2002).

Flk1 (*Vegfr2 / KDR*) is essential for the differentiation of endothelial and haematopoietic cell types in the yolk sac blood islands and formation of the endocardium (Shalaby et al., 1995). In control embryos, Flk1-LacZ is expressed in the extra-embryonic mesoderm and lateral plate mesoderm at E7.5. **(Figure 5.7A-C)**. In $Eomes^{\Delta epi}$ mutants, Flk1-LacZ is lost at E7.5. However, in some mutants at later stages (approximately at late head fold: LHF) a few cells aggregated within the primitive streak express Flk1-LacZ. This is in agreement with studies in *Eomes* null EBs where Flk1+ mesoderm is dramatically reduced but not completely abolished. This Flk1-LacZ expression may represent Flk1+ skeletal muscle, or other extra-embryonic mesoderm cell types (e.g. allantoic, amniotic or chorionic) (Ema et al., 2006).

ER71 (Etsrp71, Etv2) is also essential for the formation of endothelial and haematopoietic cell types in the yolk sac blood islands (Lee et al., 2008). Whole-mount in situ hybridisation for *ER71* was performed on E7.5 embryos. In wild type embryos, *ER71* is expressed in the extra-embryonic yolk sac mesoderm and allantoic mesoderm. Interestingly, expression of *ER71* is completely absent in $Eomes^{\Delta epi}$

embryos. These data suggest that *Eomes* is required for the specification of haematopoietic and endothelial mesoderm fate in the primitive streak.

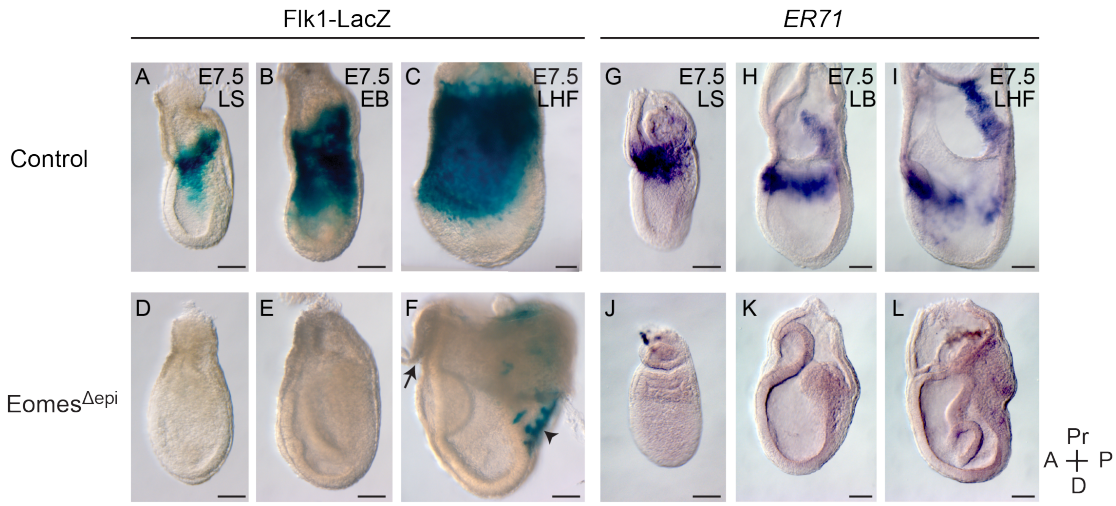


Figure 5.7: ***Eomes*** **is required for specification of haematopoietic mesoderm** ***in vivo***

Comparison of haematopoietic mesoderm marker expression in wild type and epiblast specific conditional deletion of *Eomes* (*Eomes*$^{\Delta epi}$) embryos. (A-F) X-gal stained whole-mount E7.5 embryos of wild type and mutant embryos that harbour the Flk1-LacZ transgene. (A-C) Flk1-LacZ is expressed in extra-embryonic mesoderm and lateral plate mesoderm in wild type embryos. (D-E) Flk1-LacZ expression is completely abolished at early stages in *Eomes*$^{\Delta epi}$ embryos. At later stages, there is no Flk1-LacZ expression in the yolk sac (arrow). However, Flk1-LacZ is expressed in a few mesoderm cells trapped within the primitive streak (arrow head). (G-L) Whole-mount in situ hybridisation of ER71 in E7.5 wild type and mutant embryos. (G-I) ER71 is expressed in extra-embryonic yolk sac mesoderm and allantoic mesoderm in wild type embryos. (J-L) Expression of ER71 is completely lost in *Eomes* $^{\Delta epi}$ embryos. Scale bars,100μm

5.5 *Brachyury* cDNA expressed from the endogenous *Eomes* locus

Despite being partially co-expressed the same cells as *Eomes* during gastrulation and EB differentiation, *Brachyury* is unable to rescue defects in haematopoietic and cardiac specification in *Eomes* mutants. One possibility is that *Brachyury* expression occurs slightly later than *Eomes*, and after these mesoderm fates would have been specified. Another, not mutually exclusive possibility is that *Brachyury* is unable to activate the same target genes to specify haematopoietic or cardiac mesoderm.

In order to investigate these possibilities further, I devised a substitution experiment, replacing *Eomes* with *Brachyury* protein coding information. *Brachyury* (*T*) cDNA was introduced into the start methionine of the endogenous *Eomes* locus by homologous recombination in mESCs **(Figure 5.8A)**. Successfully targeted clones were identified by Southern blot **(Figure 5.8B,C)** and were used to generate $Eomes^{T}$ mice that were identified by PCR genotyping **(Figure 5.8D)**.

The $Eomes^{T}$ allele removes the majority of the *Eomes* first exon, and therefore functions as an *Eomes* null allele (Arnold et al., 2009). By using this strategy, *T* cDNA expression is regulated by a complete set of endogenous *Eomes* enhancers and promoter, and therefore expressed in the same spatiotemporal domains as wild type *Eomes*. In order to confirm that the expression of $Eomes^{T}$ reflected the expression of endogenous *Eomes* I differentiated $Eomes^{T/+}$ mESCs for 6 days as EBs and analysed mRNA expression by qRT-PCR **(Figure 5.8E)**. $Eomes^{T}$ mRNA expression perfectly replicates the kinetics of wild type *Eomes* expression, both of which peak in expression at d3.0 of EB differentiation. In contrast, the peak of *Brachyury (T)* expression occurs later, at d4.0 of EB differentiation. However, the "T" primer pair cannot distinguish *T* cDNA expressed from the wild type or $Eomes^{T/+}$ locus, so it is

possible that the difference in kinetics is even more dramatic.

Genotyping of *Eomes*$^{T/+}$ intercross offspring did not identify any homozygous *Eomes*$^{T/T}$ mice **(Table 5.1)**. This implies that *Eomes*$^{T/T}$ mice are lethal during embryonic development.

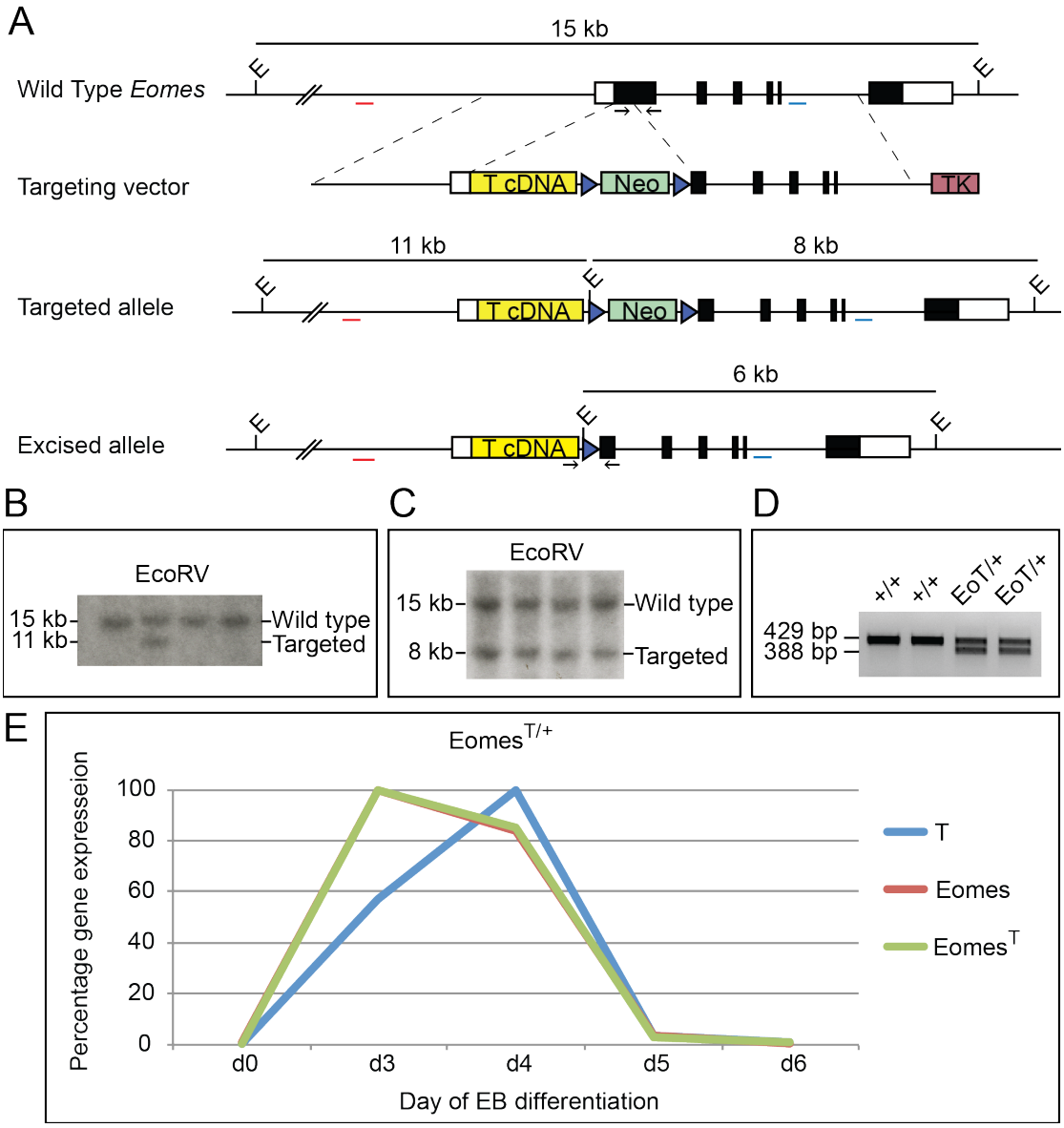


Figure 5.8: Strategy for replacing *Eomes* with *Brachyury* coding sequence

(A) Targeting strategy to replace *Eomes* exon 1 with Brachyury (T) cDNA and SV40 polyA cassette by homologous recombination to generate an *Eomes*^{T} allele. The targeting vector includes a positive selection Neomycin cassette flanked by LoxP sites, and a negative selection TK cassette. E = EcoRV. Arrows indicate PCR primers to identify successful Cre-mediated excision of Neomycin cassette and for PCR genotyping. (B, C) Southern blot to identify successfully targeted clones. (D) PCR genotyping of *Eomes*^{T} mice. (E) Expression analysis of *Eomes*$^{T/+}$ mESCs differentiated for 6 days as embryoid bodies (EBs). Expression of *Eomes*^{T} (green) follows the same kinetics as the wild type *Eomes* (red) expression, rather than *Brachyury* (*T,* blue).

Genotype	Wild type	*Eomes*$^{T/+}$	*Eomes*$^{T/T}$	Total
Observed	4 (22%)	14 (78%)	0 (0%)	18 (100%)
Expected	4.5 (25%)	9 (50%)	4.5 (25%)	18 (100%)

Table 5.1: Genotypes of *Eomes*$^{T/+}$ intercross offspring.
Eomes$^{T/+}$ intercross offspring were subjected to PCR genotyping. The table shows the observed and expected numbers and corresponding percentages of genotyping of offspring at weaning age recovered from

5.6 Discussion

5.6.1 Eomes is expressed in the early streak and is essential for haematopoietic mesoderm formation

Here, I have shown that Eomes protein is expressed prior to Brachyury during gastrulation, and at early streak stages their expression domains do not overlap. Brachyury is expressed in a very small number of cells in the posterior-proximal embryo, while Eomes is robustly expressed within the streak and nascent mesoderm.

Haematopoietic mesoderm is the first cell type to ingress through the primitive streak at early streak stages and migrates to the yolk sac (Kinder et al., 1999). Previously *Eomes*+ cells progeny have been shown to reside in the haematopoietic and endothelial component of the yolk sac (Costello et al., 2011). In this study, analysis of *Eomes* mutant embryos and null mESCs demonstrates that *Eomes* plays an essential for the specification and differentiation of haematopoietic mesoderm. However, the molecular mechanism remains to be resolved.

The VEGF receptor, *Flk1* is essential for the formation of haematopoietic and endothelium in the yolk sac (Shalaby et al., 1995). Interestingly, Eomes has been shown to bind to *Flk1* in differentiated mEpiSCs and hESCs (Teo et al., 2011). Intriguingly, Brachyury also occupies this site during hESCs differentiation (Faial et al., 2015). I have demonstrated that *Flk1* expression is greatly reduced in *Eomes* mutant embryos and differentiated null mESCs. In agreement with this, *Eomes* shRNA knockdown experiments have show a similar decrease of Flk1+ mesoderm (Izumi et al., 2007). Feasibly, Eomes may specify haematopoietic fate by direct activation of *Flk1*.

Furthermore, Eomes has been shown to directly activate the bHLH transcription factor, *Mesp1* during cardiac mesoderm specification (Costello et al., 2011; van den Ameele et al., 2012). Recently, *Mesp1* has also been implicated in specification of haematopoietic mesoderm via the direct activation *ER71* (Chan et al., 2013; Shi et al., 2015). *ER71* is essential for haematopoietic and endothelial formation in the yolk sac (Lee et al., 2008), and it's expression is completely abolished in *Eomes* mutants. Perhaps Eomes specifies haematopoietic fate by indirect activation of *ER71* via Mesp1.

How can a role for *Eomes* in specification of both haematopoietic and cardiac mesoderm be reconciled? Elegant experiments have shown that *Mesp1* expression specifies haematopoietic or cardiac mesoderm in a time-dependent fashion. Early induction of *Mesp1* expression in differentiating mESCs biases differentiation to haematopoietic fates, whereas later induction biases differentiation to cardiac fates (Chan et al., 2013). It would be important to investigate whether such mechanisms in fate determination also exist for *Eomes*, and if this is dependent or independent of *Mesp1*.

5.6.2 Eomes and Brachyury are co-expressed in the anterior region of the mid-streak

At mid-streak stages Brachyury expression predominates in the posterior primitive streak, whereas Eomes and Brachyury protein expression overlap in the anterior primitive streak (APS). APS derivatives include definitive endoderm, node, notochord and axial mesoderm. In Activin induced hESCs differentiation, BRACHYURY and EOMES have both been shown to be sufficient in regulating similar endoderm genes (e.g. CER1, FOXA2, SOX17) (Faial et al., 2015; Teo et al., 2011). *EOMES* is

essential for endoderm differentiation, however *BRACHYURY* is dispensable.

Do Eomes and Brachyury interact with each other to specify different fates within the APS? They may act co-operatively, perhaps as placeholders, or within the same complex to activate downstream genes. Alternatively, they could act in opposition with one another either by competitively binding similar genomic sites and co-factors, or inhibiting the others expression. More sophisticated *in vitro* differentiation protocols coupled with ChIP-seq will hopefully be able to answer these critical questions.

5.6.3 Eomes - Brachyury substitution experiment

Eomes null mESCs are unable to differentiate into haematopoietic and cardiac mesoderm. This specification defect is also seen *in vivo*, when *Eomes* is conditionally ablated in the epiblast. Eomes and Brachyury are able to activate to similar genes in *Xenopus* development and BMP induced hESC mesoderm differentiation (Faial et al., 2015; Gentsch et al., 2013). In *Xenopus*, Eomes and Xbra act co-operatively to pattern mesoderm (Gentsch et al., 2013). However, expression analysis I have performed in embryos and EBs shows that Eomes precedes Brachyury protein expression during the earliest stages of mouse mesoderm formation. Therefore, Brachyury function cannot compensate for the loss of *Eomes* if it is not expressed at the right stage.

I generated an $Eomes^{T}$ allele that replaces the first exon of *Eomes* with *Brachyury (T)* protein coding information. $Eomes^{T}$ mRNA expression reflects the same kinetics as *Eomes* during mESC differentiation. This allele will be a useful tool to ask if *Brachyury* can rescue early mesoderm specification defects seen in *Eomes* mutants. However, homozygous $Eomes^{T}$ mice are not viable, suggesting it is embryonic lethal.

Further experiments are required to fully characterise the expression domains of Eomes^{T} protein in the embryo and to find out when the lethality of *Eomes*$^{T/T}$ mice occurs. It is possible that Brachyury cannot rescue Eomes function in the trophectoderm, therefore the embryos will fail to implant. To circumvent this problem Eomes^{T} mice could be crossed with a condition *Eomes* allele (Eomes^{CA}), to generate Eomes$^{T/CA}$ mice. The epiblast specific Sox2.Cre allele would then be used to excise the conditional allele (generating Eomes$^{T/\Delta epi}$), leaving Eomes function in the trophectoderm intact. Eomes $^{T/\Delta epi}$ mice could then be analysed to assess whether Brachyury can rescue Eomes function in the embryo proper.

Chapter 6: General Discussion

6.1 Aims, key results and future work

6.1.1 Identification and targeted deletion of Eomes enhancers

Eomes is expressed in the anterior visceral endoderm (AVE), primitive streak and extra-embryonic ectoderm (ExE) during gastrulation (Ciruna and Rossant, 1999). I aimed to characterise cis-regulatory regions that control *Eomes* expression during this important developmental time frame. I identified regions by bioinformatic analysis of published data and confirmed enhancer activity using transgenic reporters and targeted deletions generated with homologous recombination.

I was able to demonstrate that the EH1 region (5kb region, -11kb to -6kb from the TSS) that compromises two enhancers identified by bioinformatics (EH1a and EH1b) recapitulates *Eomes* expression within the primitive streak. Surprisingly, the previously reported enhancer within this region, EH1b (2kb region from -8kb to -6kb from the TSS) (Kartikasari et al., 2013) is dispensable for proper expression in the embryo. This may be due to compensation from the neighbouring EH1a enhancer. Future work will confirm if the entire EH1 region is necessary for primitive streak expression by deletion of both EH1a and EH1b enhancers.

Intriguingly, I discovered an enhancer, EH3 (650bp region +8kb from the TSS) that controls expression in both the AVE and primitive streak but not the ExE. EH3 is necessary for robust expression of *Eomes*, and its removal leads to a range of hypomorphic phenotypes corresponding to known expression domains of *Eomes* expression. Problems in AVE specification result in defects with A-P axis formation and positioning of the primitive streak (Class I mutants). Later problems in anterior

primitive streak (APS) specification result in a failure of ventral closure and defects in anterior neurectoderm patterning (Class II mutants).

The similarity of the EH3 phenotype to that seen when components of the *Nodal* signalling pathway are reduced implies that Nodal signalling controls this enhancer. I aim test if *Nodal* directly regulates this enhancer by performing Smad2 ChIP at this site in Activin differentiated mESCs.

Future experiments will hopefully validate these enhancers *in vitro*. Firstly, I aim to demonstrate that enhancer activity can drive luciferase reporter expression in response to Nodal signalling. Secondly, I hope to perform 3C (chromatin-conformation capture) to reveal higher-level chromatin structure of the *Eomes* locus such as enhancer-promoter interactions.

6.1.1.1 Generation of an Eomes driven improved-Cre recombinase mouse line

The $Eomes^{Cre}$ allele gave mosaic recombination in *Eomes* expressing cells (Costello et al., 2011). I aimed to generate a new *Eomes* driven Cre allele to produce a robust fate map, and to use for efficient lineage-specific excision of conditional alleles. I created this by introducing a mammalian optimised improved Cre (iCre) and β-globin polyA tail cassette into the endogenous *Eomes* locus.

Cells that expressed low levels of *Eomes*, expressed sufficient amounts of the highly stable iCre mRNA to mediate Cre-excision. This resulted in the generation of a very sensitive fate map of *Eomes*+ progeny. Unexpectedly, the fate map revealed *Eomes*+ contributions to all mesoderm and endoderm lineages including the yolk sac

endothelium and haematopoietic cells. However, it remains unclear if low levels of *Eomes* expression are stochastic or actually important for cell fate allocation within the primitive streak. This is a point I will discuss later in this chapter.

6.1.2 Investigating redundancy between Eomes and Brachyury during early mesoderm formation

Eomes and Brachyury are T-box transcription factors that are expressed within the primitive streak during gastrulation, and are both implicated in allocation of mesoderm cell fates. I aimed to examine the kinetics of expression, requirements and redundancy of *Eomes* and *Brachyury* during early mesoderm differentiation. I accomplished this by immunofluorescence staining and analysis using *Eomes* mutant embryos and *Eomes* and *Brachyury* null mESCs. Furthermore, to address the issue of redundancy I generated an allele that expressed *Brachyury* cDNA from the endogenous *Eomes* locus.

Analysis of gastrulating embryos confirmed that Eomes is expressed prior to Brachyury during early mesoderm formation. I determined that *Eomes* is necessary for haematopoietic mesoderm (Pdgfrα-/Flk1+) and cardiac mesoderm (Pdgfrα+/Flk1+) fates, but not paraxial mesoderm (Pdgfrα+/Flk1-) formation in mESCs differentiation. In contrast, *Brachyury* is dispensable for haematopoietic and cardiac mesoderm formation. *Eomes* null cells were unable to differentiate into haemangioblast and haematopoietic cell types in culture. Additionally, upon conditional deletion of *Eomes* in the epiblast there was a dramatic loss of endothelial and haematopoietic markers. Thus, collectively this data validates my novel finding of the essential role *Eomes* plays in haematopoietic specification.

Endogenous *Brachyury* expression in *Eomes* null cells and embryos cannot rescue haematopoietic or cardiac specification. Hence, despite DNA binding similarities of these T-box factors, they do not act redundantly to specify these early mesoderm fates. This may be partly due to the later expression of Brachyury protein in the primitive streak, which occurs after the ingression of haematopoietic and cardiac fated cells.

When *Brachyury (T)* is expressed from the endogenous *Eomes* locus ($Eomes^{T}$) it cannot rescue the embryonic lethal *Eomes* null phenotype. However, this may be due to an inability to rescue *Eomes'* early function in the trophectoderm. The most immediate experiments will analyse when $Eomes^{T}$ mutants die, at implantation (i.e. trophectoderm abnormalities) or later (i.e. gastrulation defects).

Future efforts will also aim to generate $Eomes^{T/CA}$; Sox2-Cre mediated conditional deletion in the epiblast to reveal whether earlier *Brachyury* expression can rescue *Eomes* function in the primitive streak, bypassing the role of *Eomes* in the trophectoderm and visceral endoderm lineages. Equally, this question of redundancy in mesoderm allocation could be addressed *in vivo* by $Eomes^{T/-}$ mESC differentiation.

The molecular mechanism of how *Eomes* controls specification of haematopoietic mesoderm remains to be investigated. It would be crucial to identify the direct targets of Eomes by ChIP during mESC differentiation to haematopoietic mesoderm. Moreover, dissecting in what contexts (i.e. extrinsic signalling cues) *Eomes* can specify either haematopoietic or cardiac fates will further our understanding of how lineage decisions during early embryo development are made.

6.2 Outstanding questions

T-box factors often exert their effects in a graded manner (Naiche et al., 2005). In *Xenopus*, Eomes has been shown to exhibit dosage dependent effects on mesoderm specification (Ryan et al., 1996). This has not tested been directly during mouse mesoderm differentiation. However, observations throughout my thesis indicate that there are critical threshold levels of *Eomes* that govern cell fate allocation during early mouse embryonic development.

In *C. elegans*, it has been demonstrated that partially penetrant phenotypes are caused by stochastic fluctuations in gene expression. Variation in gene expression disturbs cell specification because cell fate decisions are thresholded within a critical developmental time window (Raj et al., 2010). Subtly lowering the levels of *Eomes* by removal of the AVE and primitive streak enhancer, EH3, demonstrates that there are critical thresholds of *Eomes* required for cell specification. $Eomes^{\Delta EH3/-}$ embryos fall into three categories of varying severity; AVE defects, APS defects or undistinguishable from wild type. Presumably, without the EH3 enhancer to ensure properly controlled expression, *Eomes* levels fluctuate stochastically, and those cells that do not reach the required threshold have defects in cell fate specification (e.g. AVE or APS). Interestingly, cardiac mesoderm is correctly specified in all $Eomes^{\Delta EH3/-}$ mutants recovered at E9.5, implying that this tissue has a lower threshold requirement of *Eomes* within the primitive streak. Consistent with this, Eomes activation of *Mesp1/2* in the cardiac lineage is transient due to a *Mesp1/2* negative autoregulatory feedback loop (Bondue et al., 2008; Costello et al., 2011; Kitajima et al., 2000; Saga et al., 1999), suggesting that fluctuating levels rather than sustained expression of *Eomes* may be sufficient to specify cardiac mesoderm.

The highest levels of Nodal signalling are required to specify the APS (Vincent et al.,

2003). *Nodal* and *Eomes* converge together in auto-regulatory loop to reinforce high expression in the APS (Arnold et al., 2008). It is tempting to speculate that EH3, discovered based on EOMES/SMAD2/3 occupancy in human Activin induced definitive endoderm, is crucial for maintaining the very high levels of *Eomes* in the APS required for proper endoderm formation.

In the absence of Nodal signalling *Eomes* has been shown to specify cardiac mesoderm, and not endoderm (Costello et al., 2011; van den Ameele et al., 2012). However, here I have demonstrated that *Eomes* is also essential for specification haematopoietic mesoderm. Haematopoietic and cardiac mesoderm are the first cells to emerge from the primitive streak. *Eomes* is also required for EMT in the streak (Arnold et al., 2008). I propose that *Eomes* may specify these fates in a time-sensitive manner, similar to its direct target *Mesp1* (Chan et al., 2013), perhaps by controlling the time of epiblast cell ingression.

Aims to identify the molecular roles of *Eomes* during mesoderm differentiation have utilised overexpression systems, focusing on specification of cardiac mesoderm. David et al. (2011) overexpressed *Eomes* in mESCs and failed to observe induction of *Mesp1* mRNA, during differentiation. However, enforced expression of *Eomes* in mESCs initiates a conversion to trophectoderm fate, so this strategy is unlikely to give rise to any mesoderm populations.

Two studies were able to demonstrate that transient induction of *Eomes* expression between d2 – d4 of EB differentiation causes an increase in mesoderm specification. Despite differing culture conditions, they found approximately a 3-fold increase of Pdgfrα+ or Flk1+ mesoderm populations upon *Eomes* induction (van den Ameelee 12% → 36%, Lindsley 6% → 18%) (Lindsley et al., 2008; van den Ameele et al.,

2012).

The key question of whether Eomes elicits dose-dependent responses to pattern mesoderm can be addressed by utilising similar inducible expression systems. Induction of *Eomes* expression at different time points, intensities and duration with subsequent analysis of cell specific markers would reveal whether *Eomes* specifies either haematopoietic or mesoderm fates in a time or dosage dependent manner. Or alternatively, if *Eomes* acts to specify an early immature mesoderm progenitor, which is later specified into either haematopoietic or cardiac mesoderm.

Furthermore, inducible systems could be used for the reciprocal experiment. Conditional deletion of *Eomes* could be performed by administration of tamoxifen in cells that harbour an *Eomes* conditional allele $Eomes^{CA/CA}$ and $Roasa^{26Cre\text{-}ER}$ (tamoxifen inducible Cre) alleles. This strategy could be used to analyse the loss of *Eomes* function at different time points during mesoderm differentiation. Thus, further define the time-critical window of *Eomes* expression for cell fate decisions.

Ultimately, it would be useful to evaluate whether levels of *Eomes* expression regulate the timing of exit from the streak. In chimeras, *Eomes* null cells tend to accumulate in the posterior primitive streak. *Eomes* null fail to contribute to the definitive endoderm and cardiac mesoderm. However, *Eomes* null cells are able to contribute to some mesoderm derivatives, mainly posterior mesoderm (Arnold et al., 2008). This implies that *Eomes* acts cell autonomously to control the exit and specification of both early-primitive streak mesoderm derivatives and late APS derivatives but not late PPS derivatives. This is in agreement with expression patterns. It would be interesting to see if mESCs that contain additional copies of *Eomes* on a BAC transgene, and therefore express abnormally high levels, would exit the streak

prematurely in chimeras, as has been shown for *Brachyury* (Wilson and Beddington, 1997).

6.3 Implications

Identification of these new *Eomes* enhancers during gastrulation improves our understanding of how this critical gene is controlled, and may lead to the discovery of novel upstream regulators.

In addition, the discovery that *Eomes* is involved in the specification of haematopoietic mesoderm holds great potential. Eomes+ cells isolation could provide more efficient derivation of haematopoietic cells in culture to study blood-related disorders and develop cell therapies.

References

Ahituv, N., Zhu, Y., Visel, A., Holt, A., Afzal, V., Pennacchio, L.A., and Rubin, E.M. (2007). Deletion of ultraconserved elements yields viable mice. PLoS biology *5*, e234.

Arnold, S.J., Hofmann, U.K., Bikoff, E.K., and Robertson, E.J. (2008). Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. Development *135*, 501-511.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol *10*, 91-103.

Arnold, S.J., Sugnaseelan, J., Groszer, M., Srinivas, S., and Robertson, E.J. (2009). Generation and analysis of a mouse line harboring GFP in the Eomes/Tbr2 locus. Genesis *47*, 775-781.

Baron, M.H. (2013). Concise Review: early embryonic erythropoiesis: not so primitive after all. Stem Cells *31*, 849-856.

Beddington, R.S., Rashbass, P., and Wilson, V. (1992). Brachyury--a gene affecting mouse gastrulation and early organogenesis. Dev Suppl, 157-165.

Beddington, R.S., and Robertson, E.J. (1998). Anterior patterning in mouse. Trends in genetics : TIG *14*, 277-284.

Beddington, R.S., and Robertson, E.J. (1999). Axis development and early asymmetry in mammals. Cell *96*, 195-209.

Behringer, R., Gertsenstein, M., Nagy, K., and Nagy, A. (2013). Manipulating the mouse embryo : a laboratory manual, fourth edition, Fourth edition. edn (Cold Spring Harbor Laboratory Press).

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E.J., and Constam, D.B. (2006). The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. Dev Cell *11*, 313-323.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C.*, et al.* (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144-155.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Beyer, T.A., Weiss, A., Khomchuk, Y., Huang, K., Ogunjimi, A.A., Varelas, X., and Wrana, J.L. (2013). Switch enhancers interpret TGF-beta and Hippo signaling to control cell fate in human embryonic stem cells. Cell reports *5*, 1611-1624.

Bollag, R.J., Siegfried, Z., Cebra-Thomas, J.A., Garvey, N., Davison, E.M., and Silver, L.M. (1994). An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. Nat Genet *7*, 383-389.

Bondue, A., Lapouge, G., Paulissen, C., Semeraro, C., Iacovino, M., Kyba, M., and Blanpain, C. (2008). Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. Cell Stem Cell *3*, 69-84.

Brennan, J., Lu, C.C., Norris, D.P., Rodriguez, T.A., Beddington, R.S., and Robertson, E.J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. Nature *411*, 965-969.

Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C.H., Lim, B., Vardy, L., Dunn, N.R., Trotter, M., Pedersen, R.*, et al.* (2011). Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm

progenitors. Stem Cells *29*, 1176-1185.
Buecker, C., and Wysocka, J. (2012). Enhancers as information integration hubs in development: lessons from genomics. Trends in genetics : TIG *28*, 276-284.
Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327-339.
Chan, S.S., Shi, X., Toyama, A., Arpke, R.W., Dandapat, A., Iacovino, M., Kang, J., Le, G., Hagen, H.R., Garry, D.J.*, et al.* (2013). Mesp1 patterns mesoderm into cardiac, hematopoietic, or skeletal myogenic progenitors in a context-dependent manner. Cell Stem Cell *12*, 587-601.
Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. Development *125*, 725-732.
Ciruna, B.G., and Rossant, J. (1999). Expression of the T-box gene Eomesodermin during early mouse development. Mech Dev *81*, 199-203.
Conlon, F.L., Fairclough, L., Price, B.M., Casey, E.S., and Smith, J.C. (2001). Determinants of T box protein specificity. Development *128*, 3749-3758.
Conlon, F.L., Wright, C.V., and Robertson, E.J. (1995). Effects of the TWis mutation on notochord formation and mesodermal patterning. Mech Dev *49*, 201-209.
Costello, I., Pimeisl, I.M., Dräger, S., Bikoff, E.K., Robertson, E.J., and Arnold, S.J. (2011). The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. Nat Cell Biol *13*, 1084-1091.
Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.
Dahle, O., Kumar, A., and Kuehn, M.R. (2010). Nodal signaling recruits the histone demethylase Jmjd3 to counteract polycomb-mediated repression at target genes. Sci Signal *3*, ra48.
David, R., Jarsch, V.B., Schwarz, F., Nathan, P., Gegg, M., Lickert, H., and Franz, W.M. (2011). Induction of MesP1 by Brachyury(T) generates the common multipotent cardiovascular stem cell. Cardiovascular research *92*, 115-122.
de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev *26*, 11-24.
DiTacchio, L., Bowles, J., Shin, S., Lim, D.S., Koopman, P., and Janknecht, R. (2012). Transcription factors ER71/ETV2 and SOX9 participate in a positive feedback loop in fetal and adult mouse testis. J Biol Chem *287*, 23657-23666.
Donnison, M., Beaton, A., Davey, H.W., Broadhurst, R., L'Huillier, P., and Pfeffer, P.L. (2005). Loss of the extraembryonic ectoderm in Elf5 mutants leads to defects in embryonic patterning. Development *132*, 2299-2308.
Ema, M., Takahashi, S., and Rossant, J. (2006). Deletion of the selection cassette, but not cis-acting elements, in targeted Flk1-lacZ allele reveals Flk1 expression in multipotent mesodermal progenitors. Blood *107*, 111-117.
Estaras, C., Benner, C., and Jones, K.A. (2015). SMADs and YAP Compete to Control Elongation of beta-Catenin:LEF-1-Recruited RNAPII during hESC Differentiation. Molecular cell *58*, 780-793.
Faial, T., Bernardo, A.S., Mendjan, S., Diamanti, E., Ortmann, D., Gentsch, G.E., Mascetti, V.L., Trotter, M.W., Smith, J.C., and Pedersen, R.A. (2015). Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. Development *142*, 2121-2135.
Feng, X.H., Zhang, Y., Wu, R.Y., and Derynck, R. (1998). The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for smad3

in TGF-beta-induced transcriptional activation. Genes Dev *12*, 2153-2163.
Fong, Y.W., Cattoglio, C., Yamaguchi, T., and Tjian, R. (2012). Transcriptional regulation by coactivators in embryonic stem cells. Trends in cell biology *22*, 292-298.
Fukuda, M., Takahashi, S., Haramoto, Y., Onuma, Y., Kim, Y.J., Yeo, C.Y., Ishiura, S., and Asashima, M. (2010). Zygotic VegT is required for Xenopus paraxial mesoderm formation and is regulated by Nodal signaling and Eomesodermin. The International journal of developmental biology *54*, 81-92.
Funa, N.S., Schachter, K.A., Lerdrup, M., Ekberg, J., Hess, K., Dietrich, N., Honore, C., Hansen, K., and Semb, H. (2015). beta-Catenin Regulates Primitive Streak Induction through Collaborative Interactions with SMAD2/SMAD3 and OCT4. Cell Stem Cell *16*, 639-652.
Gaarenstroom, T., and Hill, C.S. (2014). TGF-beta signaling to chromatin: How Smads regulate transcription during self-renewal and differentiation. Seminars in cell & developmental biology.
Garg, V., Kathiriya, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., Hirayama-Yamada, K., Joo, K.*, et al.* (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature *424*, 443-447.
Gentsch, G.E., Owens, N.D., Martin, S.R., Piccinelli, P., Faial, T., Trotter, M.W., Gilchrist, M.J., and Smith, J.C. (2013). In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. Cell reports *4*, 1185-1196.
Ginis, I., Luo, Y., Miura, T., Thies, S., Brandenberger, R., Gerecht-Nir, S., Amit, M., Hoke, A., Carpenter, M.K., Itskovitz-Eldor, J.*, et al.* (2004). Differences between human and mouse embryonic stem cells. Dev Biol *269*, 360-380.
Gordon, S.M., Chaix, J., Rupp, L.J., Wu, J., Madera, S., Sun, J.C., Lindsten, T., and Reiner, S.L. (2012). The transcription factors T-bet and Eomes control key checkpoints of natural killer cell maturation. Immunity *36*, 55-67.
Guzman-Ayala, M., Ben-Haim, N., Beck, S., and Constam, D.B. (2004). Nodal protein processing and fibroblast growth factor 4 synergize to maintain a trophoblast stem cell microenvironment. Proc Natl Acad Sci U S A *101*, 15656-15660.
Hayashi, S., Lewis, P., Pevny, L., and McMahon, A.P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. Mech Dev *119 Suppl 1*, S97-S101.
Heffner, C.S., Herbert Pratt, C., Babiuk, R.P., Sharma, Y., Rockwood, S.F., Donahue, L.R., Eppig, J.T., and Murray, S.A. (2012). Supporting conditional mouse mutagenesis with a comprehensive cre characterization resource. Nature communications *3*, 1218.
Herrmann, B.G. (1991). Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. Development *113*, 913-917.
Hodge, R.D., Kowalczyk, T.D., Wolf, S.A., Encinas, J.M., Rippey, C., Enikolopov, G., Kempermann, G., and Hevner, R.F. (2008). Intermediate progenitors in adult hippocampal neurogenesis: Tbr2 expression and coordinate regulation of neuronal output. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 3707-3717.
Hoodless, P.A., Pye, M., Chazaud, C., Labbe, E., Attisano, L., Rossant, J., and Wrana, J.L. (2001). FoxH1 (Fast) functions to specify the anterior primitive streak in the mouse. Genes Dev *15*, 1257-1271.
Huber, T.L., Kouskoff, V., Fehling, H.J., Palis, J., and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature *432*, 625-630.

Inman, K.E., and Downs, K.M. (2006a). Brachyury is required for elongation and vasculogenesis in the murine allantois. Development *133*, 2947-2959.
Inman, K.E., and Downs, K.M. (2006b). Localization of Brachyury (T) in embryonic and extraembryonic tissues during mouse gastrulation. Gene Expr Patterns *6*, 783-793.
Izumi, N., Era, T., Akimaru, H., Yasunaga, M., and Nishikawa, S. (2007). Dissecting the molecular hierarchy for mesendoderm differentiation through a combination of embryonic stem cell culture and RNA interference. Stem Cells *25*, 1664-1674.
Jagannathan-Bogdan, M., and Zon, L.I. (2013). Hematopoiesis. Development *140*, 2463-2467.
Kartikasari, A.E., Zhou, J.X., Kanji, M.S., Chan, D.N., Sinha, A., Grapin-Botton, A., Magnuson, M.A., Lowry, W.E., and Bhushan, A. (2013). The histone demethylase Jmjd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation. The EMBO journal *32*, 1393-1408.
Keller, G.M., Webb, S., and Kennedy, M. (2002). Hematopoietic Development of ES Cells in Culture. Methods in molecular medicine *63*, 209-230.
Kennedy, M., Firpo, M., Choi, K., Wall, C., Robertson, S., Kabrun, N., and Keller, G. (1997). A common precursor for primitive erythropoiesis and definitive haematopoiesis. Nature *386*, 488-493.
Kim, S.W., Yoon, S.J., Chuong, E., Oyolu, C., Wills, A.E., Gupta, R., and Baker, J. (2011). Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. Dev Biol *357*, 492-504.
Kinder, S.J., Loebel, D.A., and Tam, P.P. (2001). Allocation and early differentiation of cardiovascular progenitors in the mouse embryo. Trends Cardiovasc Med *11*, 177-184.
Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. Development *126*, 4691-4701.
King, T., Beddington, R.S., and Brown, N.A. (1998). The role of the brachyury gene in heart development and left-right specification in the mouse. Mech Dev *79*, 29-37.
Kitajima, S., Takagi, A., Inoue, T., and Saga, Y. (2000). MesP1 and MesP2 are essential for the development of cardiac mesoderm. Development *127*, 3215-3226.
Kouskoff, V., Lacaud, G., Schwantz, S., Fehling, H.J., and Keller, G. (2005). Sequential development of hematopoietic and cardiac mesoderm during embryonic stem cell differentiation. Proc Natl Acad Sci U S A *102*, 13170-13175.
Kulaeva, O.I., Nizovtseva, E.V., Polikanov, Y.S., Ulianov, S.V., and Studitsky, V.M. (2012). Distant activation of transcription: mechanisms of enhancer action. Molecular and cellular biology *32*, 4892-4897.
Kwon, G.S., Viotti, M., and Hadjantonakis, A.K. (2008). The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. Dev Cell *15*, 509-520.
Labbe, E., Silvestri, C., Hoodless, P.A., Wrana, J.L., and Attisano, L. (1998). Smad2 and Smad3 positively and negatively regulate TGF beta-dependent transcription through the forkhead DNA-binding protein FAST2. Molecular cell *2*, 109-120.
Lagha, M., Bothma, J.P., and Levine, M. (2012). Mechanisms of transcriptional precision in animal development. Trends in genetics : TIG *28*, 409-416.
Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891-911.
Lee, D., Park, C., Lee, H., Lugus, J.J., Kim, S.H., Arentson, E., Chung, Y.S., Gomez, G.,

Kyba, M., Lin, S., *et al.* (2008). ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. Cell Stem Cell *2*, 497-507.
Lindsley, R.C., Gill, J.G., Murphy, T.L., Langer, E.M., Cai, M., Mashayekhi, M., Wang, W., Niwa, N., Nerbonne, J.M., Kyba, M., *et al.* (2008). Mesp1 coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. Cell Stem Cell *3*, 55-68.
Liu, F., Kang, I., Park, C., Chang, L.W., Wang, W., Lee, D., Lim, D.S., Vittet, D., Nerbonne, J.M., and Choi, K. (2012). ER71 specifies Flk-1+ hemangiogenic mesoderm by inhibiting cardiac mesoderm and Wnt signaling. Blood *119*, 3295-3305.
Loh, K.M., and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. Cell Stem Cell *8*, 363-369.
Lolas, M., Valenzuela, P.D., Tjian, R., and Liu, Z. (2014). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. Proc Natl Acad Sci U S A *111*, 4478-4483.
Mao, C.A., Kiyama, T., Pan, P., Furuta, Y., Hadjantonakis, A.K., and Klein, W.H. (2008). Eomesodermin, a target gene of Pou4f2, is required for retinal ganglion cell and optic nerve development in the mouse. Development *135*, 271-280.
Matsumoto, N., Kubo, A., Liu, H., Akita, K., Laub, F., Ramirez, F., Keller, G., and Friedman, S.L. (2006). Developmental regulation of yolk sac hematopoiesis by Kruppel-like factor 6. Blood *107*, 1357-1365.
Mesnard, D., Donnison, M., Fuerer, C., Pfeffer, P.L., and Constam, D.B. (2011). The microenvironment patterns the pluripotent mouse epiblast through paracrine Furin and Pace4 proteolytic activities. Genes Dev *25*, 1871-1880.
Miller, S.A., Huang, A.C., Miazgowicz, M.M., Brassil, M.M., and Weinmann, A.S. (2008). Coordinated but physically separable interaction with H3K27-demethylase and H3K4-methyltransferase activities are required for T-box protein-mediated activation of developmental gene expression. Genes Dev *22*, 2980-2993.
Minguillon, C., and Logan, M. (2003). The comparative genomics of T-box genes. Brief Funct Genomic Proteomic *2*, 224-233.
Naiche, L.A., Harrelson, Z., Kelly, R.G., and Papaioannou, V.E. (2005). T-box genes in vertebrate development. Annu Rev Genet *39*, 219-239.
Neijts, R., Simmini, S., Giuliani, F., van Rooijen, C., and Deschamps, J. (2014). Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. Dev Dyn *243*, 88-98.
Norris, D.P., Brennan, J., Bikoff, E.K., and Robertson, E.J. (2002). The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. Development *129*, 3455-3468.
Norris, D.P., and Robertson, E.J. (1999). Asymmetric and node-specific nodal expression patterns are controlled by two distinct cis-acting regulatory elements. Genes Dev *13*, 1575-1588.
Nowotschin, S., Costello, I., Piliszek, A., Kwon, G.S., Mao, C.A., Klein, W.H., Robertson, E.J., and Hadjantonakis, A.K. (2013). The T-box transcription factor Eomesodermin is essential for AVE induction in the mouse embryo. Genes Dev *27*, 997-1002.
Ong, C.T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat Rev Genet *12*, 283-293.
Padron-Barthe, L., Temino, S., Villa Del Campo, C., Carramolino, L., Isern, J., and Torres, M. (2014). Clonal analysis identifies hemogenic endothelium and not hemangioblasts as the source of the blood-endothelial common lineage in the mouse embryo. Blood.
Papaioannou, V.E. (2014). The T-box gene family: emerging roles in

development, stem cells and cancer. Development *141*, 3819-3833.
Pearce, E.L., Mullen, A.C., Martins, G.A., Krawczyk, C.M., Hutchins, A.S., Zediak, V.P., Banica, M., DiCioccio, C.B., Gross, D.A., Mao, C.A.*, et al.* (2003). Control of effector CD8+ T cell function by the transcription factor Eomesodermin. Science *302*, 1041-1043.
Perea-Gomez, A., Vella, F.D., Shawlot, W., Oulad-Abdelghani, M., Chazaud, C., Meno, C., Pfister, V., Chen, L., Robertson, E., Hamada, H.*, et al.* (2002). Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. Dev Cell *3*, 745-756.
Pereira, L.A., Wong, M.S., Lim, S.M., Sides, A., Stanley, E.G., and Elefanty, A.G. (2011). Brachyury and related Tbx proteins interact with the Mixl1 homeodomain protein and negatively regulate Mixl1 transcriptional activity. PLoS One *6*, e28394.
Pernaute, B., Canon, S., Crespo, M., Fernandez-Tresguerres, B., Rayon, T., and Manzanares, M. (2010a). Comparison of extraembryonic expression of Eomes and Cdx2 in pregastrulation chick and mouse embryo unveils regulatory changes along evolution. Dev Dyn *239*, 620-629.
Pernaute, B., Cañon, S., Crespo, M., Fernandez-Tresguerres, B., Rayon, T., and Manzanares, M. (2010b). Comparison of extraembryonic expression of Eomes and Cdx2 in pregastrulation chick and mouse embryo unveils regulatory changes along evolution. Dev Dyn *239*, 620-629.
Picozzi, P., Wang, F., Cronk, K., and Ryan, K. (2009). Eomesodermin requires transforming growth factor-beta/activin signaling and binds Smad2 to activate mesodermal genes. J Biol Chem *284*, 2397-2408.
Pimeisl, I.M., Tanriver, Y., Daza, R.A., Vauti, F., Hevner, R.F., Arnold, H.H., and Arnold, S.J. (2013). Generation and characterization of a tamoxifen-inducible Eomes(CreER) mouse line. Genesis *51*, 725-733.
Polydorou, C., and Georgiades, P. (2013). Ets2-dependent trophoblast signalling is required for gastrulation progression after primitive streak initiation. Nature communications *4*, 1658.
Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.
Raj, A., Rifkin, S.A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. Nature *463*, 913-918.
Rashbass, P., Cooke, L.A., Herrmann, B.G., and Beddington, R.S. (1991). A cell autonomous function of Brachyury in T/T embryonic stem cell chimaeras. Nature *353*, 348-351.
Rivera-Perez, J.A., and Magnuson, T. (2005). Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. Dev Biol *288*, 363-371.
Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C.*, et al.* (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. Nature *404*, 95-99.
Ryan, K., Garrett, N., Mitchell, A., and Gurdon, J.B. (1996). Eomesodermin, a key early gene in Xenopus mesoderm differentiation. Cell *87*, 989-1000.
Saga, Y., Kitajima, S., and Miyagawa-Tomita, S. (2000). Mesp1 expression is the earliest sign of cardiovascular development. Trends Cardiovasc Med *10*, 345-352.
Saga, Y., Miyagawa-Tomita, S., Takagi, A., Kitajima, S., Miyazaki, J., and Inoue, T. (1999). MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. Development *126*, 3437-3447.
Sasaki, H., and Hogan, B.L. (1996). Enhancer analysis of the mouse HNF-3 beta gene: regulatory elements for node/notochord and floor plate are independent

and consist of multiple sub-elements. Genes to cells : devoted to molecular & cellular mechanisms *1*, 59-72.
Schmierer, B., and Hill, C.S. (2007). TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. Nat Rev Mol Cell Biol *8*, 970-982.
Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L., and Schuh, A.C. (1995). Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature *376*, 62-66.
Shen, M.M. (2007). Nodal signaling: developmental roles and regulation. Development *134*, 1023-1034.
Shi, X., Zirbes, K.M., Rasmussen, T.L., Ferdous, A., Garry, M.G., Koyano-Nakagawa, N., and Garry, D.J. (2015). The transcription factor Mesp1 interacts with cAMP-responsive element binding protein 1 (Creb1) and coactivates Ets variant 2 (Etv2) gene expression. J Biol Chem *290*, 9614-9625.
Shimshek, D.R., Kim, J., Hubner, M.R., Spergel, D.J., Buchholz, F., Casanova, E., Stewart, A.F., Seeburg, P.H., and Sprengel, R. (2002). Codon-improved Cre recombinase (iCre) expression in the mouse. Genesis *32*, 19-26.
Showell, C., Binder, O., and Conlon, F.L. (2004). T-box genes in early embryogenesis. Dev Dyn *229*, 201-218.
Sinha, S., Abraham, S., Gronostajski, R.M., and Campbell, C.E. (2000). Differential DNA binding and transcription modulation by three T-box proteins, T, TBX1 and TBX2. Gene *258*, 15-29.
Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet *21*, 70-71.
Srinivas, S., Rodriguez, T., Clements, M., Smith, J.C., and Beddington, R.S. (2004). Active cell migration drives the unilateral movements of the anterior visceral endoderm. Development *131*, 1157-1164.
Sun, X., Meyers, E.N., Lewandoski, M., and Martin, G.R. (1999). Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev *13*, 1834-1846.
Takaoka, K., and Hamada, H. (2012). Cell fate decisions and axis determination in the early mouse embryo. Development *139*, 3-14.
Takaoka, K., Yamamoto, M., and Hamada, H. (2011). Origin and role of distal visceral endoderm, a group of cells that determines anterior-posterior polarity of the mouse embryo. Nat Cell Biol *13*, 743-752.
Tam, P.P., and Loebel, D.A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. Nat Rev Genet *8*, 368-381.
Taube, J.H., Allton, K., Duncan, S.A., Shen, L., and Barton, M.C. (2010). Foxa1 functions as a pioneer transcription factor at transposable elements to activate Afp during differentiation of embryonic stem cells. J Biol Chem *285*, 16135-16144.
Teo, A.K., Arnold, S.J., Trotter, M.W., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., and Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. Genes Dev *25*, 238-250.
Thomas, P.Q., Brown, A., and Beddington, R.S. (1998). Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. Development *125*, 85-94.
Tremblay, K.D., Hoodless, P.A., Bikoff, E.K., and Robertson, E.J. (2000). Formation of the definitive endoderm in mouse is a Smad2-dependent process. Development *127*, 3079-3090.
Tsuneyoshi, N., Tan, E.K., Sadasivam, A., Poobalan, Y., Sumi, T., Nakatsuji, N., Suemori, H., and Dunn, N.R. (2012). The SMAD2/3 corepressor SNON maintains pluripotency through selective repression of mesendodermal genes in human ES cells. Genes Dev *26*, 2471-2476.

van den Ameele, J., Tiberi, L., Bondue, A., Paulissen, C., Herpoel, A., Iacovino, M., Kyba, M., Blanpain, C., and Vanderhaeghen, P. (2012). Eomesodermin induces Mesp1 expression and cardiac differentiation from embryonic stem cells in the absence of Activin. EMBO reports *13*, 355-362.
Vincent, S.D., Dunn, N.R., Hayashi, S., Norris, D.P., and Robertson, E.J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. Genes Dev *17*, 1646-1662.
Wardle, F.C., and Papaioannou, V.E. (2008). Teasing out T-box targets in early mesoderm. Curr Opin Genet Dev *18*, 418-425.
Wilson, V., and Beddington, R. (1997). Expression of T protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. Dev Biol *192*, 45-58.
Wilson, V., Manson, L., Skarnes, W.C., and Beddington, R.S. (1995). The T gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. Development *121*, 877-886.
Wolf, X.A., Klein, T., Garcia, R., Hyttel, P., and Serup, P. (2012). Identification of a conserved cis-acting region driving expression of mouse Eomesodermin to the primitive streak, node, and definitive endoderm. Gene Expr Patterns *12*, 85-93.
Wolpert, L., Tickle, C., Jessell, T.M., Lawrence, P.A., Meyerowitz, E.M., and Smith, J. (2011). Principles of development, Fourth edition. edn (Oxford: Oxford University Press).
Wu, M.Y., and Hill, C.S. (2009). Tgf-beta superfamily signaling in embryonic development and homeostasis. Dev Cell *16*, 329-343.
Yamaguchi, T.P., Takada, S., Yoshikawa, Y., Wu, N., and McMahon, A.P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes Dev *13*, 3185-3190.
Yamamoto, M., Meno, C., Sakai, Y., Shiratori, H., Mochida, K., Ikawa, Y., Saijoh, Y., and Hamada, H. (2001). The transcription factor FoxH1 (FAST) mediates Nodal signaling during anterior-posterior patterning and node formation in the mouse. Genes Dev *15*, 1242-1256.
Yasuhiko, Y., Haraguchi, S., Kitajima, S., Takahashi, Y., Kanno, J., and Saga, Y. (2006). Tbx6-mediated Notch signaling controls somite-specific Mesp2 expression. Proc Natl Acad Sci U S A *103*, 3651-3656.
Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881-894.
Yuan, P., Han, J., Guo, G., Orlov, Y.L., Huss, M., Loh, Y.H., Yaw, L.P., Robson, P., Lim, B., and Ng, H.H. (2009). Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev *23*, 2507-2520.
Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res *21*, 1273-1283.