ETS2 regulation of tissue size in early mouse embryogenesis

Di Hu

Jesus College



Thesis submitted for the degree of

Doctor of Philosophy

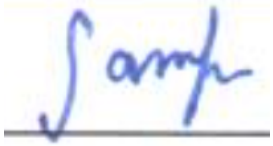
University of Oxford

Department of Physiology, Anatomy and Genetics

Trinity Term 2019

Declaration

I declare that the design, performance, analysis, and communication of the work presented in this thesis are my own unless otherwise stated.



Di Hu

Acknowledgements

First and foremost, I would like to thank Professor Shankar Srinivas for giving me the opportunity to learn about early mouse embryogenesis. The system is humbling. Anything it decides to tell me is fascinating and beautiful, and therefore motivates me to develop and refine skills to try to understand it better without bias.

I would like to thank Shankar and every member of the Srinivas laboratory who has helped me address questions and think critically about experimental design and interpretation. I would like to thank our collaborators Dr. Antonio Scialdone, Dr. Maria Solovey, and Professor Tristan Rodriguez; my transfer and confirmation of status examiners Professor Clive Wilson and Dr. Nicola Smart; my past mentors Professor Norman Rosenblum and Professor Sevan Hopyan; the WIMM FACS facility; and PSB. I would like to thank notable figures in the field who have come by my poster at conferences and meetings and provided helpful feedback: Professor Jaime Rivera Perez, Professor Takashi Hiiragi, Professor James Briscoe, Professor James Wells, Professor Kathy Niakan, and Professor Vasso Episkopou.

During my DPhil, I became a DJ and naturally formed incredible connections with those who are as passionate and excited about music as I am (about music and science). A lot of love and appreciation for Oli Ong, Christian Ash, Sam Smith, Natalie Carter, Angus Whitfield, Dejarne Bogle, Isaac Cole, Alec Jetha, Harry Stoodley, Sasha Stone, Christopher Sawyer, Emily Williams, Orlagh Dooley, Matthew Cutler, Tristan Hallis, Serena Pasion, Candice Hutchings, Peggy Gou, Abraham Delacey, Jack Rayner, Lola Lefebvre, Kiran Rogers, Riley Serola, Yusuf Zerdazi, Craig Fyffe, Theodore Ryan, Zak Watson, Lucas Contreras-Brown, Tim Hinson, Danny Voicu and many many more phenomenal individuals. I would like to thank my parents and family for keeping me grounded and inspiring me to always be kind and do good no matter what I choose to do. I apologise to all of those I have not thanked.

‘Show don’t tell’ – Jonas Lindstroem

Abstract

ETS2 regulation of tissue size in early mouse embryogenesis

Di Hu
Jesus College, University of Oxford
Thesis submitted for the degree of Doctor of Philosophy
Trinity Term 2019

ETS2 is a transcription factor with both oncogenic and tumour suppressive roles, yet the mechanisms underlying these opposing effects remain unclear. I have analysed an *Ets2*-loss-of-function mouse mutant that displays a diminished extra-embryonic ectoderm (ExE) on one hand and a thickened anterior visceral endoderm (AVE) on the other. The opposing effects of ETS2 on tissue size, relevant to its tumour suppressive and oncogenic properties, appear to manifest in this mutant. Here I showed that the ExE defect is due to increased peri-implantation apoptosis, verified by phenotypic rescue via apoptosis inhibition. I demonstrated that the AVE transforms from a simple to stratified epithelium due to increased proliferation and perpendicular cell divisions, reminiscent of neoplastic transformation. Inhibition of MEK, which acts upstream of ETS2, phenocopied the AVE defect in wild type (WT) embryos. To identify ETS2 transcriptional targets that cause these AVE defects, I conducted single-cell RNA-sequencing (scRNA-seq), which showed that all cell types in the *Ets2*-mutant retain normal character; and revealed upregulation of several WNT/β-CATENIN pathway agonists in the mutant AVE. Pharmacological activation of the WNT/β-CATENIN pathway in WT recapitulated the mutant AVE phenotype, supporting scRNA-seq results. 2C-HR-CRIPSR mediated misexpression of candidate WNT agonists in the VE of WT embryos was attempted to verify whether the novel ETS2-WNT axis of AVE tissue size regulation is relevant *in vivo*. Overall, I have demonstrated the opposing effects of ETS2 on tissue size in early mouse embryogenesis: ETS2 protects the peri-implantation ExE from apoptosis, while dampens WNT/β-CATENIN signalling in the AVE to prevent overproliferation and maintain simple epithelial integrity.

Abbreviations

A-P – anterior-posterior

AVE – anterior visceral endoderm

ChIP – chromatin immunoprecipitation

Co-IP – co-immunoprecipitation

DEGs – differentially expressed genes

DPT – diffusion pseudotime

DVE – distal visceral endoderm

EBs – embryoid bodies

ECM – extracellular matrix

EPC – ectoplacental cone

ESCs – embryonic stem cells

EmVE/Epi-VE – embryonic visceral endoderm

ExE – extra-embryonic ectoderm

ExVE/ExE-VE – extra-embryonic visceral endoderm

FACS – fluorescence activated cell sorting

FC – fold change

FDR – false discovery rate (or p-adj value)

GO – gene ontology

GRN – gene regulatory network

HA – homology arm

HCR – hybridisation chain reaction

ICM – inner cell mass

ISH – *in situ* hybridisation

LCM – laser capture microdissection

mTE – mural trophectoderm

pTE – polar trophectoderm

PVE – posterior VE

ParE – parietal endoderm

PrE – primitive endoderm

scRNA-seq – single-cell RNA-sequencing

TE – trophectoderm

TGCs – trophoblast giant cells

TSCs – trophoblast stem cells

UTR – untranslated region

VE – visceral endoderm

XEN – extra-embryonic endoderm

Table of contents

Declaration 2

Acknowledgements 3

Abstract 4

Abbreviations 5

Table of contents 7

List of figures 13

List of tables 21

1. Introduction 23

1.1 Mouse embryonic development from zygote to post-implantation 23

1.1.1 Formation of pre-implantation cell lineages 23

1.1.2 Peri-implantation changes 26

1.1.3 Post-implantation development and signalling interactions 27

1.2 ETS2 and ETS family of transcription factors 36

1.2.1 ETS domain family of transcription factors 36

1.2.2 MAPK signalling as upstream activator of ETS2 40

1.3 Role of ETS2 in mammalian development 43

1.3.1 Generation of the *Ets2*^{db1} mouse mutant 43

1.3.2 ETS2 is a core transcription factor for the mouse trophoblast lineage 44

1.3.3 Role of Ets2 in trophoblast stem cell self-renewal 47

1.3.4 Role of Ets2 in trophoblast stem cell induction from murine fibroblasts 48

1.3.5 ETS2 involvement in transdifferentiation to cardiac progenitor fate 49

1.4 ETS2 in cancer 50

1.4.1 Oncogenic roles of ETS2 50

1.4.2 Tumour suppressive functions of ETS2 52

1.5 Rationale for thesis project 54

1.6 Statement of aims 56

1.7 Hypothesis 56

2. Materials and methods 57

2.1 Animal husbandry 57

2.2 Mouse lines 57

2.3 Genotyping 59

2.3.1 DNA preparation by crude lysis 59

2.3.2 Genotyping PCR conditions 60

2.3.3 Quantitative PCR primers and conditions 61

2.4 Embryo dissections 63

2.5 Immunofluorescence staining of embryos and ESCs 63

2.5.1 Antibodies 63

2.5.2 Immunofluorescence sample preparation 64

2.6 Confocal microscopy 65

2.7 Embryo culture 66

2.8 Single-cell dissociation for scRNA-seq 66

2.9 *In silico* quantitative analysis and statistics 67

2.10 *Ets2*-mutant and control ESC derivation 70

2.11 Gene ontology and bioinformatics of scRNA-seq data 70

2.12 Cloning for 2C-HR-CRISPR 71

2.12.1 cDNA library preparation 71

2.12.2 Primers for amplifying top mutant AVE differentially expressed candidates 71

2.12.3 *Ttr* backbone plasmid linearisation 73

2.12.4 Ligation and Gibson Assembly to clone gene of interest into *Ttr* backbone 74

2.12.5 Biotinylated PCR 75

3. ETS2 protects the peri-implantation ExE from apoptosis 78

3.1 Introduction 78

3.1.1 The apoptotic pathway 78

3.1.2 Cell death at peri-implantation stages 83

3.1.3 ETS2 is a core regulator of ExE fate 84

3.2 Results 85

3.2.1 ETS2 is expressed in the ExE and its progenitors 85

3.2.2 Mutants maintained on a mixed genetic background displayed a milder ExE phenotype compared to pure 129S6/SvEvTac by morphometric analysis 87

3.2.3 Post-implantation mutant ExE is significantly reduced in size, yet retains signature marker gene expression 92

3.2.4 Reduced mutant ExE is sufficient for normal epiblast marker gene expression 94

3.2.5 Mutant epiblast fails to induce BRACHYURY-positive mesoderm at late post-implantation 95

3.2.6 Mutant ExE proliferation is normal from early to late post-implantation 98

3.2.7 Normal mutant ExE apoptosis at post-implantation 99

3.2.8 Peri-implantation mutant ExE is comparable to control 100

3.2.9 Increased apoptosis in the developing ExE at peri-implantation 102

3.2.10 Mutant ExE phenotypic rescue by apoptosis inhibition from peri- to post-implantation 104

3.3 Discussion 106

4. ETS2 regulates AVE proliferation, orientation of cell division, and cell density 109

4.1 Introduction 109

4.1.1 Maintenance of epithelial cell division rates and orientation 109

4.1.2 Characterisation of proliferation in the VE and AVE at post-implantation stages 110

4.1.3 ETS2 regulation of epithelial proliferation in intestinal tumorigenesis 111

4.2 Results 111

4.2.1 ETS2 protein is expressed in the VE and AVE at post-implantation stages 111

4.2.2 AVE migration is complete in mutants 114

4.2.3 Apical-lateral polarity is intact in mutants 115

4.2.4 LAMININ- and COLLAGEN IV-positive basement membrane is thickened beneath mutant AVE cells 136

4.2.5 Mutant AVE epithelial stratification and increased cell number by E6.5 142

4.2.6 Increased mutant AVE proliferation at E5.5 144

4.2.7 Increased proportion of perpendicular cell division in the mutant at E6.5 146

4.2.8 Increased mutant AVE cell crowding by E6.5 149

4.2.9 Mutant AVE phenotype is consistent regardless of genetic background 153

4.2.10 Mutant AVE phenocopy upon treatment with two independent inhibitors of known ETS2 upstream activator MEK 154

4.2.11 MEK-inhibitor-mediated mutant AVE phenocopy is still present upon ExE removal prior to culture 164

4.2.12 Additional evidence to support the independence of mutant ExE and AVE phenotypes 165

4.3 Discussion 166

5. ScRNA-seq to elucidate downstream effectors of ETS2 transcriptional activity 173

5.1 Introduction 173

5.1.1 Rationale for use of scRNA-seq to investigate the *Ets2*-mutant 174

5.1.2 ScRNA-seq biology and technology 174

5.1.3 Bioinformatics for analysis of misregulated genes and pathways 179

5.1.4 Two-cell homologous recombination (2C-HR)-CRISPR to validate scRNA-seq findings *in vivo* 182

5.2 Results 184

5.2.1 ScRNA-seq of pure 129 background E5.5 control versus mutant – mostly epiblast dataset 184

5.2.2 E5.5 mutant epiblast transcriptome is broadly unaltered 190

5.2.3 Mutant ExE at E5.5 retains signature marker expression by scRNA-seq 195

5.2.4 E5.5 mutant ExE expresses normal proliferation and apoptotic pathway regulators by scRNA-seq 197

5.2.5 ScRNA-seq of HHEX-GFP-positive AVE in control versus mutant 200

5.2.6 GO analysis of biological processes and signalling pathways misregulated in the mutant AVE 206

5.2.7 Top misregulated genes in the mutant AVE 211

5.2.8 Validation of scRNA-seq WNT pathway enrichment by antibody stain and pharmacological intervention 221

5.2.9 Validation of top misregulated genes from mutant AVE scRNA-seq by *in vivo* genetics using 2C-HR-CRISPR 231

5.3 Discussion 239

6. Technology development for investigating mechanistic underpinnings of early mouse embryogenesis 247

6.1 Introduction 247

6.1.1 ScRNA-seq of the VE lineage 247

6.1.2 ESC derivation and technology 249

6.2 Results 252

6.2.1 Enriching for VE cells by vital fluorescent labelling for scRNA-seq 252

6.2.2 Analysis of WT VE scRNA-seq data 255

6.2.3 *Ets2*-mutant ESC derivation 271

6.2.4 Using Pan-Caspase inhibition and scRNA-seq to investigate mechanisms of epiblast cell competition 273

6.3 Discussion 277

7. Closing remarks 281

8. References 285

Appendix A 297

List of figures

Chapter 1. Introduction

Figure 1. Pre-implantation mouse development

Figure 2. Pre- to post-implantation changes resulting from peri-implantation development

Figure 3. Early post-implantation cell types and morphologies

Figure 4. Interactions between embryonic and extra-embryonic tissues in the pre-gastrulation stage mouse embryo

Figure 5. Egg cylinder stage mouse embryo at different stages of DVE/AVE migration

Figure 6. Conserved domains in Ets2 and interacting partners with the ETS domain

Figure 7. ETS domain transcription factors are generally regulated by phosphorylation downstream of MAPK (ERK, JNK, p38) signalling cascades

Figure 8. MAPK signalling pathway

Figure 9. Expression pattern of *Ets2* by ISH from E3.5 to E7.75

Figure 10. Characterisation of the ExE defect in *Ets2*^{db1} mutant

Figure 11. ETS2 is required for ExE and A-P axis development

Figure 12. Opposing roles of ETS2 in cancer

Chapter 3. ETS2 protects the peri-implantation ExE from apoptosis

Figure 13. Extrinsic and intrinsic pathways of apoptosis

Figure 14. Peri-implantation stages of mouse development

Figure 15. ETS2 expression from pre- to post-implantation mouse embryogenesis

Figure 16. E5.5 morphometric comparison of control versus mutant in pure (Etss) versus mixed (Etex) genetic backgrounds

Figure 17. E6.5 morphometric comparison of control versus mutant in pure (Etss) versus mixed (Etex) genetic backgrounds

Figure 18. Reduced mutant ExE retains signature marker gene expression

Figure 19. Normal mutant epiblast marker gene expression by immunofluorescence

Figure 20. Absence of T induction in mutants at E6.5

Figure 21. Absence of T induction in mutants at E7.0

Figure 22. Proliferation is unperturbed in the mutant ExE at peri- and post-implantation

Figure 23. Apoptosis is normal in the mutant ExE at post-implantation stages

Figure 24. Mutant ExE is normal at peri-implantation

Figure 25. Increased apoptosis in the mutant ExE at peri-implantation stage E4.75

Figure 26. Increased apoptosis in the mutant ExE at peri-implantation stage E5.0

Figure 27. Mutant ExE phenotypic rescue by apoptosis inhibition from peri- to post-implantation

Chapter 4. ETS2 regulates AVE proliferation, orientation of cell division, and cell density

Figure 28. ETS2 expression at post-implantation stages (E5.5 to E6.5)

Figure 29. Thickened mutant AVE at E6.5

Figure 30. AVE migration is completed in mutants at late E5.5 and E6.5

Figure 31. PAR6B apical polarity marker expression at E5.5

Figure 32. PAR6B apical polarity marker expression at E6.5

Figure 33. PAR3 apical junctional marker expression at E5.5

Figure 34. PAR3 apical junctional marker expression at E6.5

Figure 35. ZO-1 tight junctional marker expression at E5.5

Figure 36. ZO-1 tight junctional marker expression at E6.5

Figure 37. SCB lateral junctional marker expression at E5.5

Figure 38. SCB lateral junctional marker expression at E6.5

Figure 39. ECAD lateral junctional marker expression at E5.5

Figure 40. ECAD lateral junctional marker expression at E6.5

Figure 41. AMOT expression at distal tip AVE stage E5.5

Figure 42. AMOT expression at migrating AVE stage E5.75

Figure 43. AMOT expression at migrated AVE stage E6.5

Figure 44. GATA6 VE marker expression at E5.5

Figure 45. GATA6 PrE/VE marker expression at E4.75 and E5.0

Figure 46. F-ACTIN and normalised intensity quantification at E6.5 to analyse thickened mutant AVE

Figure 47. E6.5 immunofluorescent detection of active actomyosin contractility using pMLC2

Figure 48. E6.5 AVE zoom of pMLC2 stain in control versus mutant

Figure 49. LAMININ basement membrane marker expression at E5.5

Figure 50. LAMININ basement membrane marker expression at E6.5

Figure 51. LAMININ expression in E7.5 WT showing upregulation in the anterior

Figure 52. COLLAGEN IV basement membrane marker expression at E6.5

Figure 53. Thickened mutant AVE is characterised by epithelial stratification

Figure 54. E6.5 mutant AVE cells have decreased volume and surface area

Figure 55. Increased mutant AVE cell number by E6.5

Figure 56. No significant difference in proliferation in the E5.0 mutant AVE

Figure 57. Increased proliferation in the mutant AVE at E5.5

Figure 58. Increased proportion of perpendicular cell divisions in mutant AVE cells at E6.5

Figure 59. Regionalised mutant AVE angle of cell division at E6.5

Figure 60. Regionalised AVE cell density quantification in control versus mutant at E5.5 and E6.5

Figure 61. YAP is excluded from the nucleus of AVE cells in both control and mutant at post-implantation

Figure 62. Thickened mutant AVE phenotype is consistent in both pure and mixed genetic backgrounds

Figure 63. Inhibition of known upstream signalling cascade of ETS2 in WT embryos

Figure 64. E5.5 to E6.5 PD03 (1 μM) MEK inhibitor culture of WT embryos resulted in mutant AVE phenotypic recapitulation

Figure 65. E5.5 to E6.5 U0126 (10 μM) MEK inhibitor culture of WT embryos resulted in mutant AVE phenotypic recapitulation

Figure 66. AVE zoom for U0126 (10 μM) MEK inhibitor treated highlighting the severity of the AVE thickening defect recapitulation

Figure 67. MEK inhibitor-mediated mutant AVE defect recapitulation in WT embryos can be titrated

Figure 68. Quantification of epiblast width morphometrics normalised to length in U0126 MEK inhibitor treated conditions compared to DMSO control

Figure 69. GATA6 expression at E6.5 in mutant versus control

Figure 70. MEK inhibitor-mediated recapitulation of mutant AVE PAR6B and GATA6 expression

Figure 71. MEK inhibitor-mediated partial recapitulation of mutant AVE LAMININ thickening phenotype

Figure 72. WT embryos with ExE removed then cultured in MEK inhibitor from E5.5 to E6.5 still develop AVE thickening

Figure 73. Mutant ExE and AVE defects may be independent, as rescue of ExE defect by FMK-mediated apoptosis inhibition still results in mutant AVE thickening by E6.5

Chapter 5. ScRNA-seq to elucidate downstream effectors of ETS2 transcriptional activity

Figure 74. Identification and characterisation of cell types in a heterogeneous population

Figure 75. Mouse developmental time window and components for 2C-HR-CRISPR microinjection

Figure 76. Workflow of single cell collection from phenotyped E5.5 control versus mutant embryos

Figure 77. Detection of *Ets2* in both control and mutants by Smart-seq2 due to retention of the long *Ets2* 3' UTR in the mutant allele

Figure 78. E5.5 pure 129 background scRNA-seq – heatmap clustering based on tissue-specific marker genes and t-SNE

Figure 79. Diffusion map to show the embryo stages and cell types recovered in relation to reference dataset from Mohammed et al., 2017

Figure 80. Cluster stability analysis additionally demonstrated absence of differences between control versus mutant epiblast at E5.5

Figure 81. E5.5 control versus mutant epiblast signature marker gene expression

Figure 82. Identification of mesoderm (*T*)-correlated genes that are also downregulated in the E5.5 mutant epiblast

Figure 83. Signature marker gene expression of E5.5 control versus mutant scRNA-seq data, showing no significant difference in mutants for all markers

Figure 84. GO_0008285 'negative regulation of proliferation' is unchanged between control versus mutant ExE scRNA-seq data

Figure 85. GO_0043065 'positive regulation of apoptosis' is unchanged between control versus mutant ExE scRNA-seq data

Figure 86. Workflow of scRNA-seq collection to purify for mutant and control AVE cells at E5.5 and E6.25 using HHEX-GFP fluorescent reporter

Figure 87. E5.5 and E6.25 HHEX-GFP-positive (mixed background) scRNA-seq – heatmap clustering based on tissue-specific marker genes and t-SNE

Figure 88. Cluster stability analysis additionally demonstrated absence of differences between control versus mutant AVE at E5.5 and E6.25

Figure 89. E5.5 and E6.25 control versus mutant AVE signature marker gene expression

Figure 90. Number of DEGs in the mutant AVE at E5.5 and E6.25 with FC > 2 and p-adj < 0.1

Figure 91. PANTHER and Reactome GO analyses of E6.25 mutant AVE recurrently show significant enrichment of the WNT pathway

Figure 92. Top enriched GO terms in E6.25 mutant AVE DEG list

Figure 93. E6.25 mutant AVE upregulated DEGs that are WNT downstream components

Figure 94. E6.25 mutant AVE upregulated DEGs that are WNT upstream components

Figure 95. E6.25 mutant AVE upregulated DEGs that are oncogenes

Figure 96. Validation of MOESIN upregulation in the mutant AVE at E6.0

Figure 97. *Msn* promoter analysis to search for conserved ETS binding sites

Figure 98. Mutant AVE phenotypic recapitulation upon culturing WT embryos in WNT agonist CHIR (3 μM) from E5.5 to E6.5, stained for ETS2

Figure 99. Mutant AVE phenotypic recapitulation with CHIR can be titrated whereby a higher dosage of CHIR (5 μM) led to a visibly more severe AVE thickening phenotype

Figure 100. Comparison of ETS2 expression in the AVE of CHIR (3 μM) versus MEKi (1 μM) treated WT embryos from E5.5 to E6.5

Figure 101. Mutant AVE phenotypic recapitulation upon culturing WT embryos in WNT agonist CHIR (3 μM) from E5.5 to E6.5, stained for total β-CATENIN and GATA6

Figure 102. Mutant AVE phenotypic recapitulation upon culturing WT embryos in WNT agonist CHIR (3 μM) from E5.5 to E6.5, stained for LAMININ

Figure 103. 2C-HR-CRISPR donor PCR template construct design with *Moesin* (5.2 kb total length)

Figure 104. E9.0 embryos utilised to extract RNA to make cDNA library

Figure 105. Genotyping for positive integration of 2C-HR-CRISPR targeting construct into the *Ttr* endogenous locus

Figure 106. Potential mutant AVE phenotypic recapitulation in WT embryos upon 2C-HR-CRISPR-mediated knock-in of *Msn* driven by VE-specific promoter TTR, stained for AMOT

Figure 107. E6.5 embryo genotyping for positive integration of *Xiap-Ttr* from 2C-HR-CRISPR microinjections

Figure 108. 2C-HR-CRISPR-mediated knock-in of *Xiap* into the *Ttr* endogenous locus

Chapter 6. Technology development for investigating mechanistic underpinnings of early mouse embryogenesis

Figure 109. PKH26 red fluorescent vital membrane labelling (0.1% concentration) of E5.5 embryos, imaged right away and following overnight culture

Figure 110. Claret far red fluorescent membrane labelling (0.1% concentration) of E5.5, E5.75 and E6.5 embryos

Figure 111. Unbiased clustering based on highly variable genes (HVGs) with t-SNE representation

Figure 112. Plotting normalised expression levels of *Hhex*, *Lefty1*, and *Cer1* AVE markers against each other demonstrate that the markers do not fully overlap, supporting heterogeneity in the AVE

Figure 113. Top 50 DEGs upregulated in the AVE compared to surrounding VE

Figure 114. JUN kinase (JNK) is known to be activated downstream of TGFβ signalling

Figure 115. Trajectory formation of AVE and VE cells to spatially position them along the anterior-posterior embryo axis at E6.25

Figure 116. Marker genes that overlap with trajectory DEGs are mostly in accordance with previous literature apart from *Mixl1*

Figure 117. Expected marker genes that do not overlap with trajectory DEGs

Figure 118. Utilising the validated trajectory plot to identify the top 10 trajectory DEGs

Figure 119. Top 20 markers for the AVE and VE based on trajectory pseudospatial analysis

Figure 120. TCF7L1 is slightly upregulated in the AVE compared to the surrounding VE at 5.5

Figure 121. TCF7L1 is upregulated in the AVE compared to the surrounding VE by E6.5

Figure 122. Top 50 trajectory DEGs highlighting receptors (blue) and ligands (orange)

Figure 123. Unbiased t-SNE of E5.5 and E5.75 scRNA-seq data

Figure 124. Antibody stain validation of mutant and control ESCs to confirm presence of pluripotency markers OCT4 and SOX2

Figure 125. Click-TUNEL assay to show absence of apoptotic cell death upon treatment with FMK Pan-Caspase inhibitor

Figure 126. Comparison of epiblast cell number in DMSO control versus FMK-treated

Figure 127. t-SNE analysis to show presence of a novel cluster upon apoptosis inhibition, suggesting recovery of cells that would otherwise have died by cell competition

List of tables

Chapter 2. Materials and methods

Table 1. Summary of transgenic mouse lines used in this thesis

Table 2. Embryo genotyping frequencies from E4.75 to E6.5

Table 3. Lysis buffer

Table 4. Standard PCR reaction

Table 5. Standard qPCR reaction

Table 6. Standard fast PCR mix

Table 7. Standard fast PCR reaction

Table 8. Technical summary of scRNA-seq sample preparation strategies and FACS collections for different subpopulations in the post-implantation stage mouse embryo

Table 9. *Ttr* backbone linearisation reaction

Table 10. Gibson Assembly of *Ttr* linearised backbone and gene of interest

Table 11. Standard Phusion High-Fidelity DNA Polymerase PCR reaction

Table 12. Biotinylated PCR reaction conditions

Chapter 5. ScRNA-seq to elucidate downstream effectors of ETS2 transcriptional activity

Table 13 and 14. Significantly up- and downregulated DEGs in the E5.5 mutant epiblast

Table 15. HHEX-GFP-positive AVE single cell recovery from FACS, and after Smart-seq2 QC analysis

Table 16. Significantly upregulated DEGs in the E5.5 mutant AVE

Table 17. Significantly downregulated DEGs in the E5.5 mutant AVE

Table 18. Top 62 significantly upregulated DEGs in the E6.25 mutant AVE

Table 19. Top 63-124 significantly upregulated DEGs in the E6.25 mutant AVE

Table 20. Top significantly upregulated DEGs in the E6.25 mutant AVE excluded from Table 18, 19

Chapter 6. Technology development for investigating mechanistic underpinnings of early mouse embryogenesis

Table 21. GO analysis of upregulated DEGs in the AVE using GOrilla

Table 22. Number of far red-positive live single VE cells collected at E5.5 and E5.75 from pure C57BL/6 background for scRNA-seq

Table 23. Number of mutant and heterozygous (het) control ESCs derived from *Ets2*-mutants in both the mixed (HHEX-GFP-positive) and pure 129 genetic backgrounds

1. Introduction

1.1 Mouse embryonic development from zygote to post-implantation

1.1.1 Formation of pre-implantation cell lineages

Cell lineage establishment and patterning in the blastocyst

The cellular and molecular basis for blastocyst lineage formation will be discussed for the three cell lineages at pre-implantation: inner cell mass (ICM), trophectoderm (TE), and primitive endoderm (PrE). Mammalian development begins with the fertilisation of the zygote followed by successive rounds of cell division. At approximately the 16-32 cell stage, the first cell fate decision occurs that results in the formation of the ICM and TE. These cells are spatially organised such that TE cells are on the outer layer and ICM cells are enclosed on the inside. The embryo then undergoes a morphogenetic change forming an asymmetric blastocyst with a fluid-filled cavity. Subsequently, a further lineage segregation occurs where a subset of the ICM differentiates into PrE, which lines the blastocoel cavity (Figure 1). The pluripotent epiblast is the predominant population that gives rise to the foetus, while the TE and PrE contribute to extra-embryonic tissues that support embryo development within the uterus, and also signal to the epiblast to coordinate embryo patterning (Rossant and Tam, 2009; Saiz and Plusa, 2013). *In vitro* counterparts may be derived from the blastocyst and self-renew in defined culture. These encompass: embryonic stem cells (ESCs) from the epiblast (Evans and Kaufman, 1981; Martin, 1981); trophoblast stem cells (TSCs) that preserve properties of TE (Tanaka et al., 1998); and extra-embryonic endoderm (XEN) stem cells that have resemblance to the PrE (Kunath et al., 2005). Core transcription factors specify lineage identities and conversions *in vivo*, as well as enable differentiation *in vitro*, e.g. ESC conversion to TSCs or XEN-like cells (Rossant and Tam, 2009).

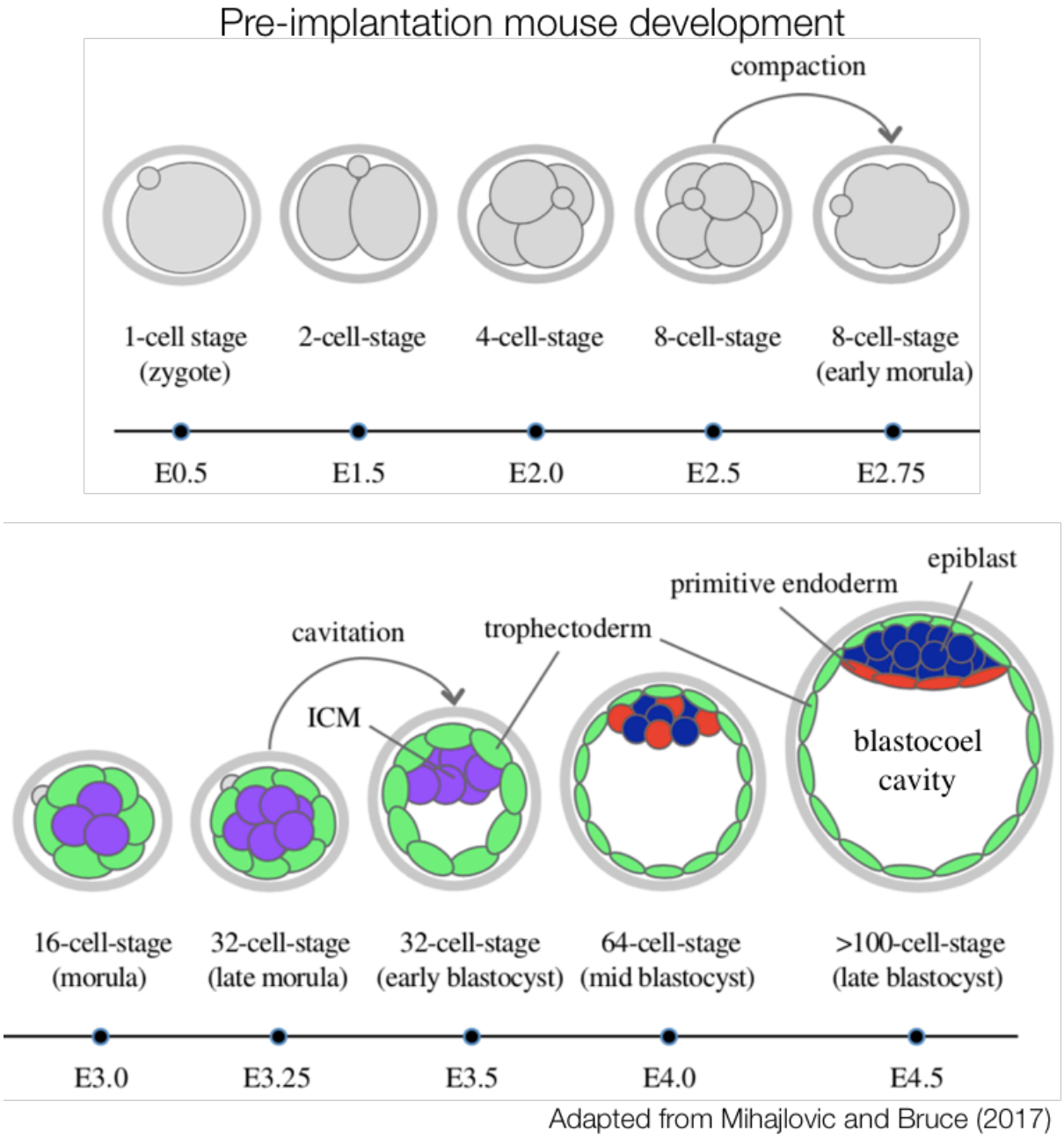


Figure 1. Pre-implantation mouse development
Temporal sequence of pre-implantation mouse embryogenesis with staging, number of cells, and cell lineages generated. TE – green; ICM – purple; PrE – red; pre-implantation epiblast – blue (adapted from Mihajlovic and Bruce, 2017).

Trophectoderm lineage establishment

The core transcription factor for TE establishment is CDX2. CDX2 expression *in vivo* commences at approximately the 8-cell stage and subsequently becomes elevated in outside cells of the TE (Dietrich and Hiiragi, 2007). Presence of CDX2 in ESCs is sufficient to convert them into trophoblast-like cells (Niwa et al., 2005). TEAD4 acts upstream of CDX2 in TE fate

specification (Nishioka et al., 2008). Although TEAD4 expression is not restricted to the TE, its transcriptional activity is dependent on presence of YAP transcriptional cofactor in the nucleus, which is regulated by the Hippo signalling pathway. Outside cell establishment is dependent on post-translation regulation of TEAD4/YAP transcription factor complex (Reddy and Irvine, 2008), which links apical-basolateral cell polarity and adhesion to positional sensing and cell fate specification (Biggins et al., 2015). This ultimately leads to transcription of *Cdx2* and TE formation. While TE-specific transcription factors CDX2 and EOMES are restricted to outside cells of the morula, ICM-specific pluripotency factors – SOX2 (Avilion et al., 2003), OCT4 (Nichols et al., 1998), and NANOG (Chambers et al. 2003) – are present in all cells of the embryo up to the morula stage and only display restricted expression after the blastocyst has been formed. Restricted expression of these pluripotency factors requires CDX2 (Niwa et al., 2005). As such, loss of CDX2 function *in vivo* results in retention of OCT4 and NANOG expression in the TE (Strumpf et al., 2005; Ralston and Rossant, 2008). Thus, lineage acquisition up to the blastocyst stage commences with TE-specific transcription factor activation, which in turn act to repress ICM fate in the outside cells that eventually lead to restriction of pluripotency factors in the ICM (Niwa et al., 2005). Subsequently, OCT4/SOX2/NANOG in the ICM repress genes of TE fate and vice versa to retain lineage restrictions (Boyer et al., 2005), and autoregulation of *Oct4* (Chew et al., 2005) and *Cdx2* (Xu et al., 1999) additionally maintain lineage identities.

Primitive endoderm lineage establishment

In the blastocyst, the ICM segregates into epiblast and PrE. Contrary to the first cell fate decision where cell position drives cell fate, this process seems to involve cell fate determination before cell positioning (Chazaud et al., 2006). ICM cells have been observed to express transcription factors of the PrE (*Gata4/6*), or epiblast genes (e.g. *Nanog*) in a ‘salt and

pepper' pattern in the ICM (Chazaud et al., 2006) before segregated positioning of the PrE and epiblast by E4.5 (Rossant et al., 2003) (Figure 1). Lineage tracing, chimera analysis, as well as gene expression profiling have demonstrated that cells in the ICM at E3.5 are distinct in their fate, restricting to one lineage or the other (Chazaud et al., 2006), and that there are distinct epiblast and PrE-specific expression profiles (Kurimoto et al., 2006). Live imaging suggests that lineage segregation of ICM into PrE and epiblast occurs at the 64-cell stage (Plusa et al., 2008) where NANOG (Mitsui et al., 2003) and SALL4 (Zhang et al., 2006) retain pluripotency in the epiblast. Forced GATA6 expression in ESCs is sufficient to induce differentiation towards PrE fate *in vitro* (Fujikura et al., 2002) The expression of PrE marker GATA6 requires GRB2, which is an upstream activator of MAPK signalling (Arman et al., 1998). Cell sorting that results in movement and segregation of the PrE and epiblast involves differences in cell adhesion and movement. Cells committed to PrE fate that have not reached their correct location in lining the blastocoel cavity undergo apoptosis (Plusa et al., 2008).

1.1.2 Peri-implantation changes

TE and PrE are epithelial cells with apical-basolateral polarity and a underlying basement membrane composed of LAMININ, COLLAGEN, and FIBRONECTIN (Brown, 2011). In the late blastocyst, the TE divides into polar (pTE) and mural (mTE) where the former is adjacent to the epiblast and the latter lines the blastocoel cavity. The process of implantation is mediated by the TE, which proliferates to give rise to progenitors of the placenta – extra-embryonic ectoderm (ExE) and ectoplacental cone (EPC) – as well as trophoblast giant cells (TGCs) (Stephenson et al., 2012). At E4.5, the PrE segregates into: 1. Parietal endoderm (ParE), which migrates from a position adjacent to the epiblast to one that is in contact with uterine tissue; and 2. Visceral endoderm (VE), which proliferates to envelop the ExE and epiblast (Figure 2), later contributing to the definitive endoderm (Kwon et al., 2008) and

endoderm of the visceral yolk sac (Rivera-Perez and Hadjantonakis, 2015). The post-implantation epiblast gives rise to both somatic cells and germ cells of the embryo proper. More details on the mechanisms of peri-implantation development are presented in 3.1.2, as they pertain directly to the findings in that Chapter.

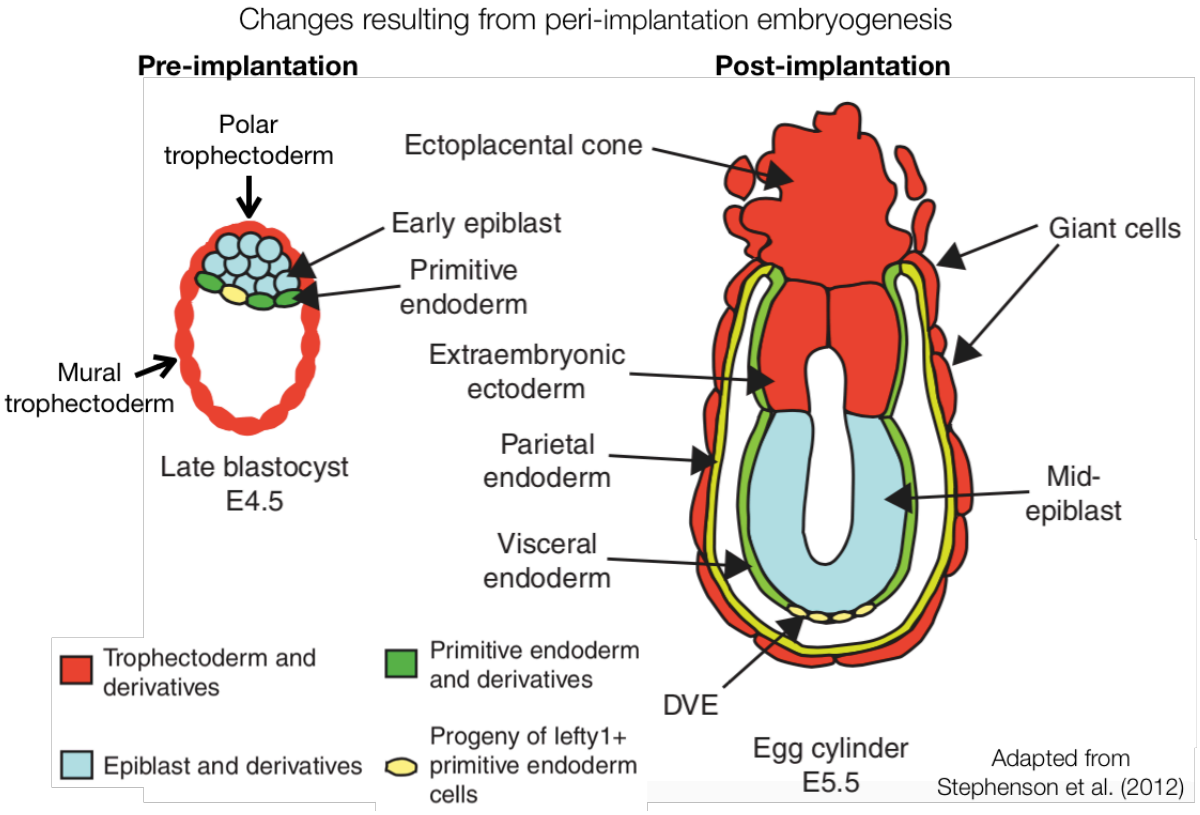


Figure 2. Pre- to post-implantation changes resulting from peri-implantation development
Late blastocyst stage E4.5 showing presence of segregated pre-implantation epiblast (blue) and PrE (green); a subset of the PrE expresses anterior visceral endoderm (AVE) marker *Lefty1*. Following implantation, the TE gives rise to ExE, EPC, and trophoblast giant cells (TGCs), while the PrE differentiates into ParE, and VE, a subset of which is the distal visceral endoderm (DVE) (E5.5) and AVE (E6.5) (adapted from Stephenson et al., 2012).

1.1.3 Post-implantation development and signalling interactions

Following blastocyst implantation, the embryo undergoes a period of rapid growth and expansion of the pluripotent epiblast into the cavity surrounded by the ParE (Snow, 1977). The epiblast takes on a cup-shape characterised by a columnar epithelium with a central pro-amniotic cavity. The epithelium subsequently becomes pseudostratified concomitant with epiblast growth. Following implantation and epiblast cavitation, the embryo takes on a

cylindrical shape from E5.0 to E7.5 (Pereira et al., 2011), which are referred to as 'egg cylinder stages'. The epiblast and ExE are both polarised epithelial tissues with a central lumen that are distinct, but in contact (Stephenson et al., 2012). The VE shares a basement membrane with the post-implantation epiblast and ExE that originates from the basement membrane between the pre-implantation PrE and epiblast (Rivera-Perez and Hadjantonakis, 2015). The monolayered VE epithelium enclosing the epiblast is known as the emVE or Epi-VE, and that adjacent to the ExE is known as the exVE or ExE-VE (Figure 3). The VE serves as the main site for nutrient, gas, and waste exchange prior to the formation of the maternal-embryonic circulation, where its ability for transcytosis and polarised absorption confers this function (Bielinska et al., 1999).

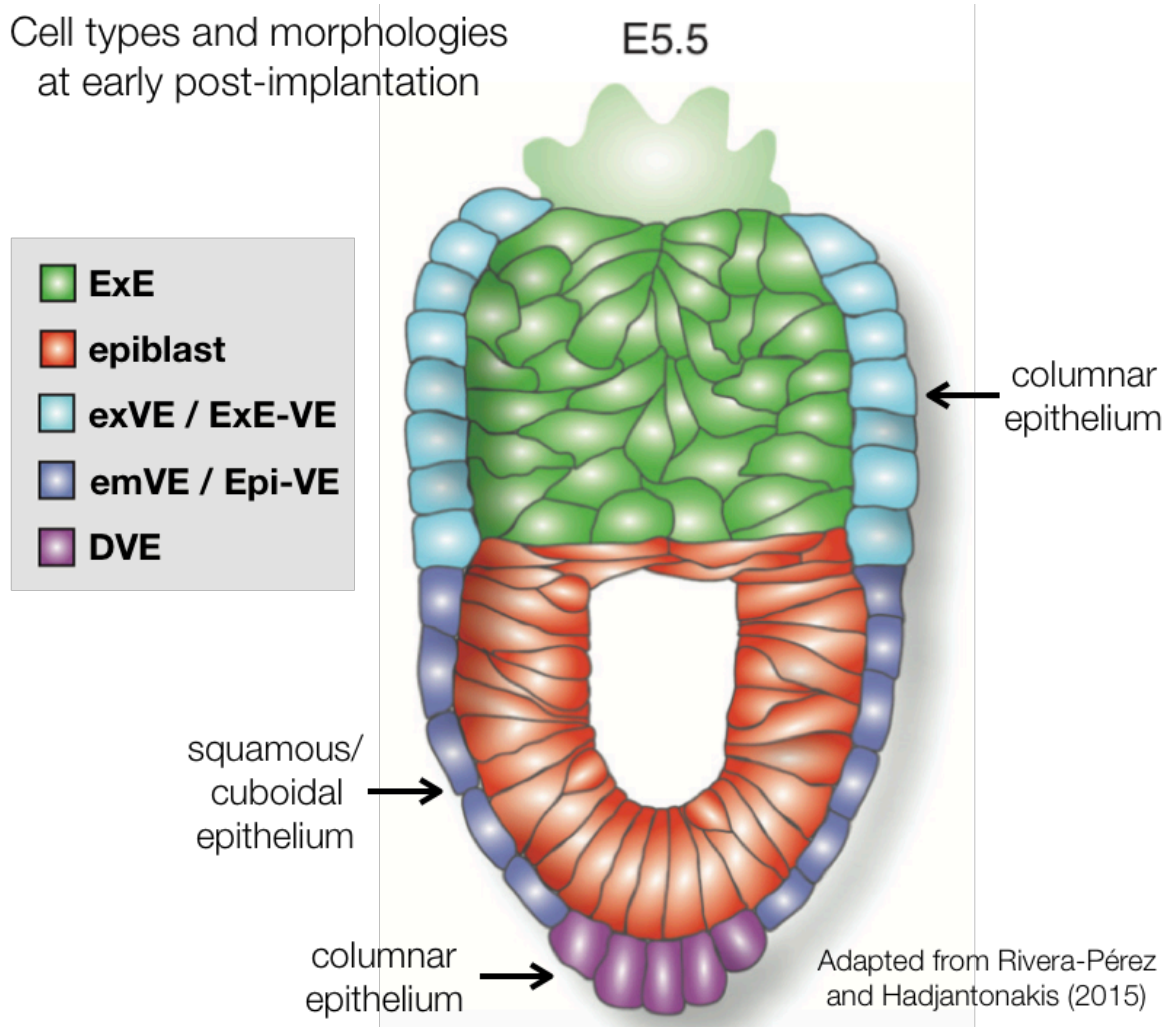


Figure 3. Early post-implantation cell types and morphologies
Diagram of E5.5 early post-implantation embryo with ExE, epiblast, and different VE subpopulations (adapted from Rivera-Perez and Hadjantonakis, 2015).

Extra-embryonic ectoderm signalling interactions

The ExE appears to be molecularly defined by at least two populations. Some marker genes are expressed throughout the ExE, including *Ets2* (Georgiades and Rossant, 2006), *Pem* (Chazaud et al., 2006), *Elf5* (Donnison et al., 2005), and *Wnt7b* (Kemp et al., 2007); while others are concentrated to the distal third of the ExE, including *Brachyury* (*T*) (Perea-Gomez et al., 2004), *Bmp4* (Winnier et al., 1995), and ESRRB (Srinivas laboratory unpublished data). The ExE is a source of molecular signals, secreting factors that influence the proximal epiblast to a greater extent than distal, thereby establishing morphogen gradients within the epiblast (Arnold and Robertson, 2009; Figure 4). The ExE plays a role in modulating Nodal signalling levels in the rest of the embryo by secreting NODAL processing enzymes SPC1 (FURIN) and SPC4 (PACE4) (Beck et al., 2002), which act to cleave the pre-protein form of NODAL to make a functional protein *in vivo*. SPC1/4 are essential, as their absence limits NODAL activation in the epiblast and result in gastrulation arrest. Active NODAL in the epiblast in turn participates in a feedback loop to promote *Bmp4* expression. BMP4 then signals to the epiblast in a proximal to distal gradient, and activates essential primitive streak marker *Wnt3* (Ben-Haim et al., 2006; Figure 4).

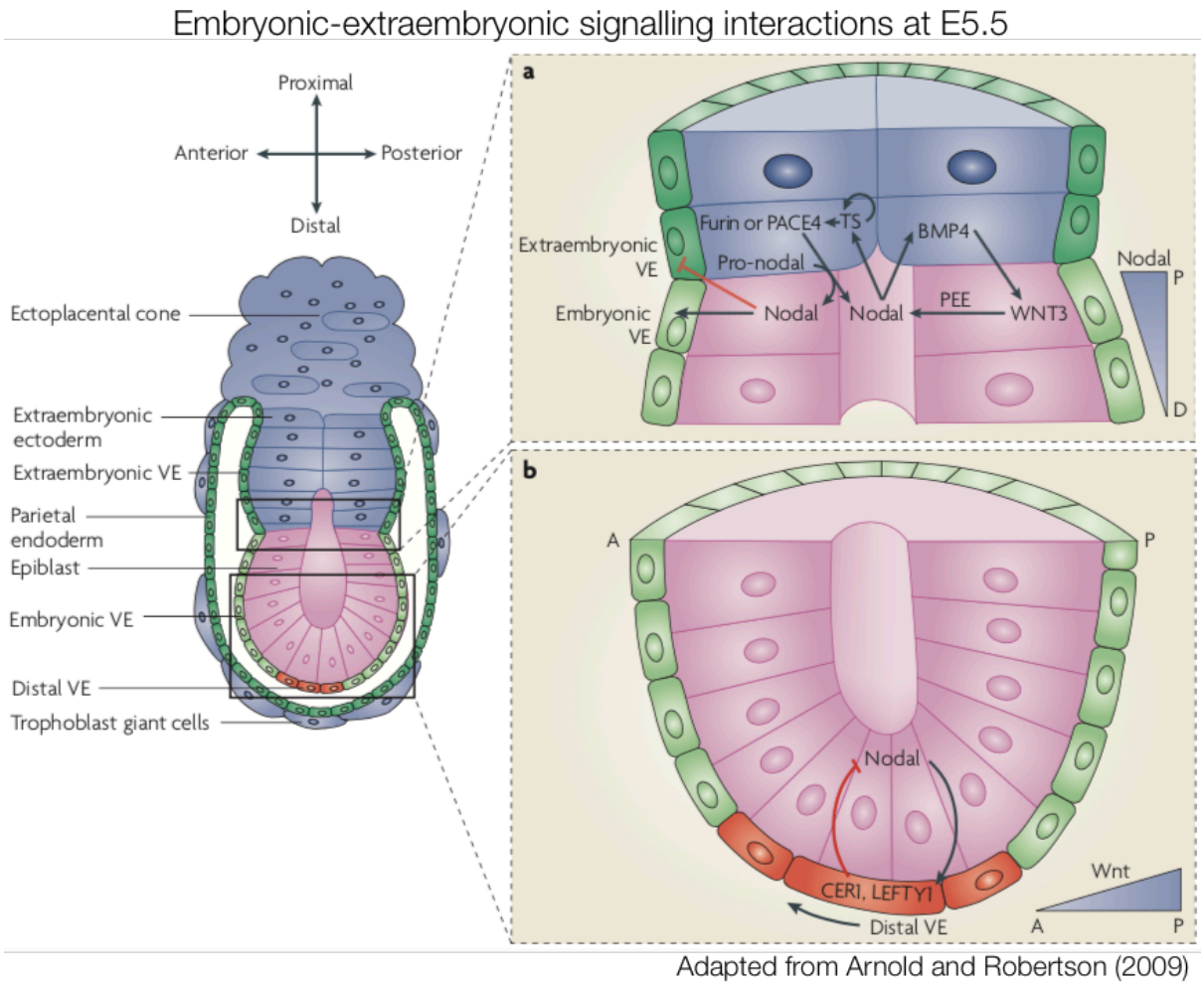


Figure 4. Interactions between embryonic and extra-embryonic tissues in the pre-gastrulation stage mouse embryo

At E5.5, a NODAL signalling gradient is present along the proximal distal axis of the epiblast.

(A) NODAL signals are high proximally, promoted by elevated levels of Nodal processing enzymes FURIN and PACE that emanate from the ExE. Active NODAL in the proximal epiblast upregulates BMP4 in the ExE, which signals back to the epiblast to result in *Wnt3* expression. Nodal proximal epiblast enhancer (PEE) has been shown to be a direct target of WNT/β-CATENIN signalling. Patterning within the VE is also thought to be dependent on signalling gradients established from the ExE and epiblast.

(B) A few cells at the distal tip (red) of the early post-implantation stage mouse embryo are distinct from surrounding VE and express extracellular inhibitors of WNT and NODAL signalling, such as *Cer1* and *Lefty1*, which downregulate NODAL signalling in the adjacent epiblast. This further enhances the proximal-distal NODAL gradient in the epiblast at E5.5 (adapted from Arnold and Robertson, 2009).

The visceral endoderm has morphological and molecular heterogeneities

The VE has a number of subcompartments defined by morphology (Figure 3) and molecular marker expression (Perea-Gomez et al., 2007; Yamamoto et al., 2009). A distinct VE subpopulation becomes evident at the distal tip of the egg cylinder at E5.5 known as the distal visceral endoderm (DVE), or anterior visceral endoderm (AVE). These cells are characterised by columnar morphology, i.e. elongated along the apical-basal epithelial axis,

appearing taller than their surrounding Epi-VE neighbours (Rivera-Perez et al., 2003; Srinivas et al., 2004). The AVE uniquely expresses several marker genes: transcription factor *Hhex*, and antagonists of the NODAL and WNT signalling pathways, such as *Lefty1, Cerl1, Dkk1, Sfrp1,* and *Sfrp5* (Pfister et al., 2007). To aid in the investigation of DVE/AVE dynamics and the cell types they give rise to, a number of fluorescent reporter lines have been made to enable *ex vivo* live imaging: *Hhex-GFP* (Rodriguez et al., 2001), *Cerl-GFP* (Mesnard et al., 2004), and *Lefty1-mVenus* (Takaoka et al., 2011). Live imaging of AVE cell movements revealed its migration up one side of the embryo (Figure 5; Srinivas et al., 2004), whereby the DVE contributes at least in part to the AVE (Takaoka et al., 2011). The AVE acts as a signalling centre to establish the anterior-posterior (A-P) body axis, localising *Wnt3* expression to the posterior VE (PVE) (Rivera-Perez and Magnuson, 2005). The PVE also distinctly expresses *Mixl1* (Robb et al., 2000), commencing at E5.5 and persisting to gastrulation. This population, however, remains stationary during gastrulation (Kwon et al., 2008). An additional population of the VE is the columnar epithelium that overlies the ExE with distinct morphology from the more squamous cells that surround the epiblast, with the exception of the AVE. These differences in epithelial morphology could be attributed to the physical environment brought about by both biochemical and mechanical cues (Hiramatsu et al., 2013). Genes that distinctly mark the ExE-VE from other VE populations at E5.5 to E6.5 include *Pem/Rhox5* (Chazaud et al., 2006), *Gata4* (Pfister et al., 2007), *Sox7* and *Sox17* (Kanai-Azuma et la., 2002). Another VE sub-region is that surrounding the epiblast that does not fall into the aforementioned categories. These cells are localised at the distal-posterior region of the VE at E5.5 and initially do not express AVE markers. They are subsequently found at the distal region of the VE around E5.75 when the AVE has migrated up one side of the embryo; then later undergo anterior shift, becoming a part of the AVE and expressing AVE markers (Rivera-Perez et al., 2003). Analysis and validation of single-cell RNA-sequencing (scRNA-seq) data can help elucidate new

markers for these different regions of the VE, as well as define different regions of the ExE and epiblast based on analysis of receptor-ligand interactions (Vento-Tormo et al., 2018; Cheng et al., 2019).

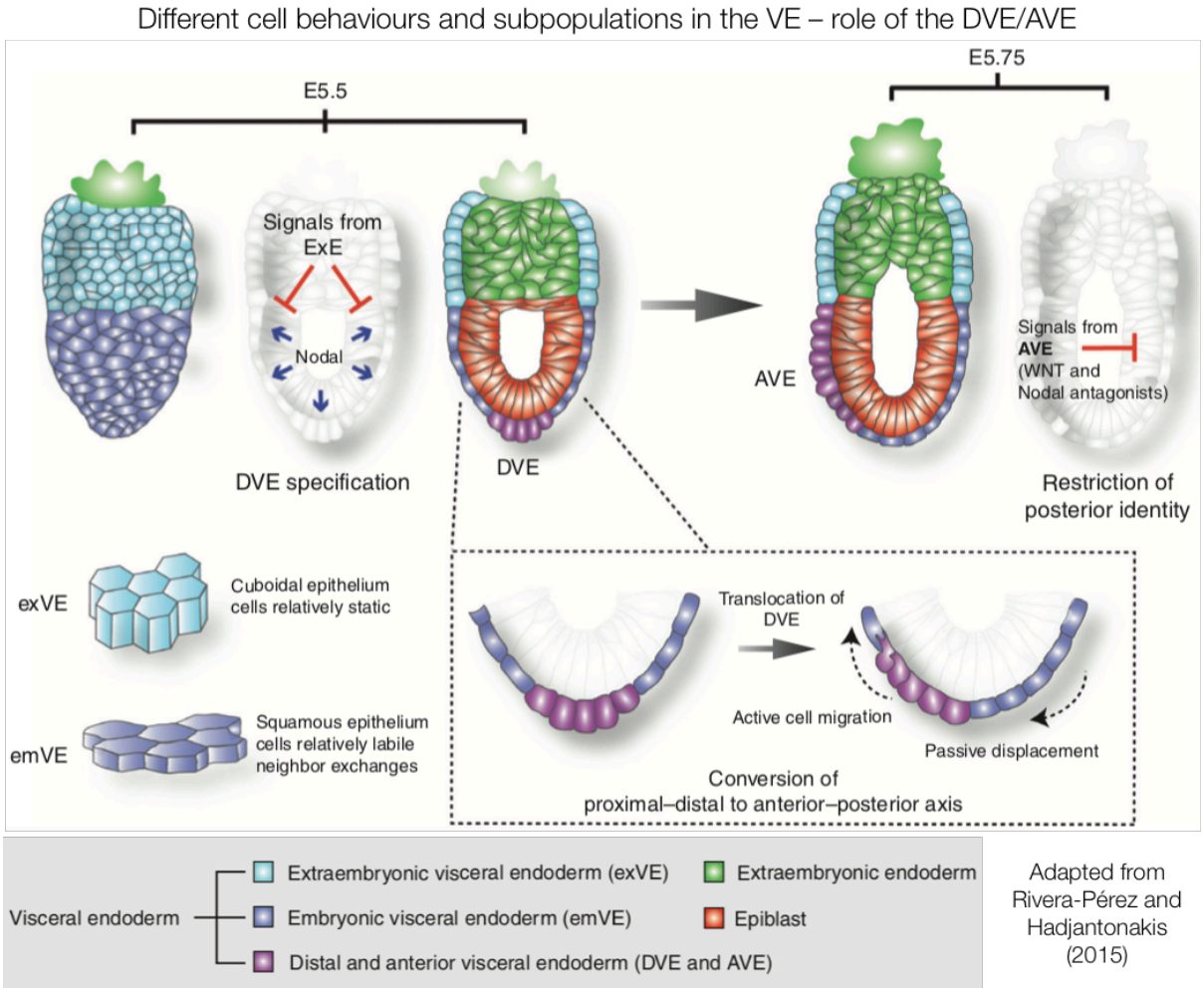


Figure 5. Egg cylinder stage mouse embryo at different stages of DVE/AVE migration
The VE is comprised of different cell populations that vary in morphology, localisation, and behaviour. DVE/AVE anterior shift (E5.5 to E5.75) is important to restrict WNT signalling to the posterior VE and prepare the embryo for gastrulation (adapted from Rivera-Perez and Hadjantonakis et al., 2015).

DVE/AVE induction

The predominant evidence for AVE induction at the distal tip of the embryo implicates the correct levels of NODAL signalling (Lu and Robertson, 2004). The ExE has been suggested to play a critical role in the restricted induction of the AVE at the distal tip through establishing signalling gradients that prevent DVE/AVE induction throughout the Epi-VE (Rodriguez et al., 2005; Richardson et al., 2006). WNT/β-CATENIN antagonist/degradative complex

component APC has also been shown to be important for localised expression of DVE markers, where loss-of-function mutants have defective specification of the DVE (Huelsken et al., 2000; Chazaud and Rossant, 2006).

AVE dynamics and function

The AVE becomes localised to the anterior of the embryo opposite the site of primitive streak induction (Figure 5). At E5.5 DVE/AVE cells initiate their movement from a latero-distal position, then migrate to the boundary between the epiblast and ExE (Rivera-Perez et al., 2003; Srinivas et al., 2004), thereby establishing the A-P axis of the mouse embryo. The AVE by E6.5 has reached the anterior-proximal aspect of the epiblast, adjacent to the cranial endodermal furrow (Rivera-Perez and Hadjantonakis, 2015).

The AVE is a heterogeneous population displaying dynamic behaviour. Clonal studies have demonstrated that at E5.75, a stage when the DVE has just reached the ExE-epiblast boundary and is known as the AVE, the population is mostly comprised of HHEX-GFP positive cells from the distal tip (Rivera-Perez et al., 2003; Srinivas et al., 2004). 12 hours later at E6.25, initially non-HHEX-GFP-expressing cells at E5.5 also migrate up the anterior aspect of the epiblast and join the 'first wave' of AVE cells (Takaoka et al., 2011). The E6.5 AVE therefore consists of descendants of the original DVE population, which have spread laterally along the ExE-epiblast boundary, as well as non-DVE cells that now express AVE markers covering the anterior of the epiblast. The AVE is initially positioned next to progenitors of the extra-embryonic mesoderm, which contribute to visceral yolk sac, chorion, and amnion structures (Rivera-Perez and Hajantonakis, 2015). Subsequently, the AVE lies adjacent to anterior embryonic structures, i.e. anterior neurectoderm at mid- to late-streak embryonic stages (Lawson and Pedersen, 1992).

AVE cell movement dynamics

AVE and VE cells undergo cellular rearrangements during AVE anterior shift whilst retaining epithelial integrity. Passive displacement of DVE cells is unlikely to be a driving mechanism for AVE formation, as DVE to AVE transition takes place rapidly in a 5-7 hour time frame (Rivera-Perez et al., 2003; Srinivas et al., 2004), while cell cycle doubling time in the VE during that period is 10 hours (Snow, 1977; Stuckey et al., 2011), which cannot explain the rate and extent of AVE translocation. Proliferation rates are consistent throughout the surrounding Epi-VE at E5.5 and E5.75, removing the possibility of a burst in posterior proliferation to account for the DVE anterior shift (Stuckey et al., 2011). Live imaging studies characterising AVE anterior shift have captured morphological dynamics in conjunction with genetic perturbations that affect the F-ACTIN cytoskeletal machinery, together supporting active migration. AVE cells have been described to extend basal filopodial projections in an angular bin that is oriented toward the direction of AVE migration, (Srinivas et al., 2004; Omelchenko et al., 2014), particularly evident in the first DVE cells that begin to move. The functional significance of these filopodial extensions is not precisely known, but have been postulated to be a result of AVE cells moving toward a chemoattractant in the prospective anterior (Srinivas et al., 2004), or away from a chemorepellant (Kimura-Yoshida et al., 2005). However, presence of filopodia, in addition to its involvement with cell movement, may also play a role in enabling cells to sense their environment, as well as contribute to cellular adhesion (Mattila and Lappalainen, 2008). Thus, presence of filopodial processes alone is insufficient to ascertain presence of active AVE cell migration.

AVE cells have additionally been described to undergo neighbour exchange with cells of the surrounding VE, where cell movements are restricted to the Epi-VE (Trichas et al., 2011). Time-lapse live imaging studies have demonstrated that AVE cells stereotypically stop their migration at the ExE-epiblast boundary (Srinivas et al., 2004), potentially attributed to

ExE-VE versus Epi-VE differences in protein expression and cell behaviour, resulting in only the Epi-VE being permissive to cell movement (Trichas et al. 2011). Furthermore, Trichas et al., 2011 provided evidence for the interplay between PCP and NODAL signalling in demarcating these regions of differing cell behavior in the Epi- versus ExE-VE, specifically contributing to cellular rearrangements in the Epi-VE. Assessment of the Epi-VE by mathematical modelling, along with experimentation in mutants with defective AVE migration, have demonstrated a role for multicellular rosette formation, defined by five or more cells sharing a central apex. These structures appear to stabilise the 'disequilibrium' arising from epithelial dynamics to enable coherent, collective AVE migration (Trichas et al., 2012). Future investigation in whether and potentially how the AVE undergoes active cell movement would benefit from genetic and imaging studies of the entire VE, as opposed to focusing on the AVE only.

Signalling from the AVE

NODAL inhibitors CER1 and LEFTY1, which mark the AVE, have been shown to be important in localised primitive streak induction, as their absence results in an ectopic primitive streak. This defect can be rescued by removing one copy of *Nodal* (Perea-Gomez et al., 2002), demonstrating that NODAL signalling dosage is important for A-P axis formation. The AVE has also been proposed to serve as an antagonist of the WNT pathway. Mutants of genes that agonise WNT signalling have been reported to display AVE distal tip arrest, including loss of function mutants of *Cripto* (Ding et al., 1998), *Otx2* (Zakin et al., 2000), and *Lpp3* (Escalante-Alcalde et al., 2003). These mutants show ectopic primitive streak induction, and the latter is characterised by presence of radial *Wnt3* expression. Evidence for the role of the AVE as a source of WNT antagonists have been presented by *in vitro* studies implicating secreted Wnt inhibitors of various Sfrps, which directly bind and sequester away Wnt; as well as Dkk1 that

act by blocking LRP receptors to enable AVE cells to have a Wnt antagonistic effect (Kawano and Kypta, 2003). Overall, genetic studies to date have predominantly supported a role for the AVE in serving as a signalling centre to inhibit NODAL and WNT, preventing activation of these pathways in the anterior and localising their activity to the posterior.

The ExE and AVE (encompassing the DVE) extra-embryonic lineages are the main cell types of focus in this thesis, as they are particularly sensitive to loss of transcription factor ETS2.

1.2 ETS2 and ETS family of transcription factors

1.2.1 ETS domain family of transcription factors

The ETS family of transcription factors presents a paradigm for achieving context-dependent transcriptional outputs as a result of their ability to interact with a wide range of transcriptional cofactors. The ETS family name originated from avian erythroblastosis virus E26 that contained the v-ets 'E twenty-six' oncogene (Nunn et al., 1983), and the family has been defined by presence of an ETS DNA-binding domain (Karim et al., 1990). The family is unique to metazoans, presumably due to its regulation by metazoan-specific signalling cascades (Degnan et al., 1993). Another domain that can classify a protein in the ETS family is sequence homology to the Pointed signalling domain, originally identified in Drosophila melanogaster as Pnt (Klambt, 1993). The Pnt domain is conserved amongst only a subset of the ETS family. In addition to its signalling functions, this domain can have transcriptional repressive roles (Fenrick et al., 1999). As a result of Pnt and DNA-binding domain homology, ETS family proteins have overlapping functions in a myriad of biological processes, such as embryonic development of fruitflies, worms, frogs, and mice, as well as in cancer (Sharrocks, 2001). There are 29 ETS family members in humans (Hollenhorst et al., 2011) that all share an

85-amino acid winged helix-turn-helix DNA binding domain. This domain recognises the conserved ETS DNA binding sequence GGA(A/T).

In cancer, ETS proteins may be present in its native form, but expressed at aberrant levels; occur as an abnormal fusion protein; or interact with proteins commonly mutated in cancer, such as p53 (Martinez, 2016). ETS family misregulation in cancer perturbs the physiological response to normally tightly regulated signalling pathways for growth, thereby resulting in oncogenesis (Dittmer et al., 1998). Members of the ETS family interact with a variety of different transcriptional activators and repressors (Figure 6; Sharrocks et al., 1997). A combination of context-dependent signalling activity, presence of cofactors, and epigenetic state of the cell type can dictate the specific transcriptional response to ETS proteins. ETS transcription factors are autoregulated, i.e. their DNA-binding capability is normally disabled until the correct signals are present. These include phosphorylation in the Pnt signalling domain and binding interaction with transcriptional cofactors (Sharrocks et al., 2001).

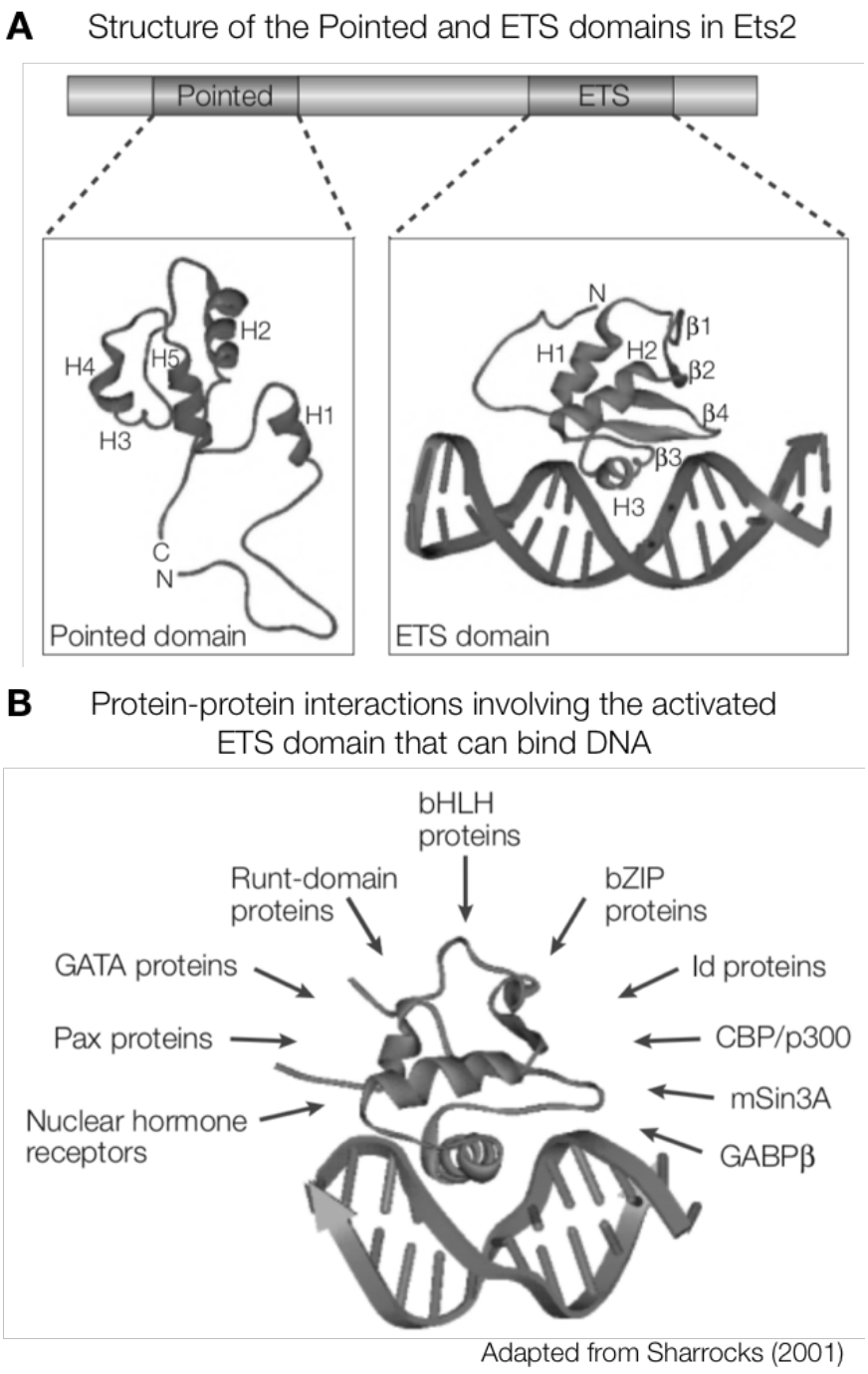


Figure 6. Conserved domains in Ets2 and interacting partners with the ETS domain
(A) Structure of the Pointed (Pnt) and ETS domain of Ets2, showing the relative positioning of the helices (H) and/or β-strands (β) in the protein domains.
(B) Potential protein-protein interactions with the active ETS domain that has bound to DNA (adapted from Sharrocks, 2001).

Signalling pathways that trigger ETS family transcriptional activation

MAPK signalling is the predominant cascade linked to ETS family regulation (Wasylyk et al., 1998). All members of the TCF ETS subfamily are direct targets of ERK-mediated

phosphorylation of MAPK signalling (Sharrock et al., 1997), as well as are responsive to p38 MAPK and JNK cascades (Whitmarsh et al., 1995; Price et al., 1996). Thus, the TCF subfamily integrates cellular responses to growth (MAPK) and stress (p38 and JNK) signals. Ets2 is a known direct target of MAPK whereby phosphorylation by the MAP kinase ERK potentiates Ets2 activation (McCarthy et al., 1997), and simultaneously results in nuclear export of Ets2 Repressor Factor (ERF) (Figure 7; Le Gallic et al., 1999).

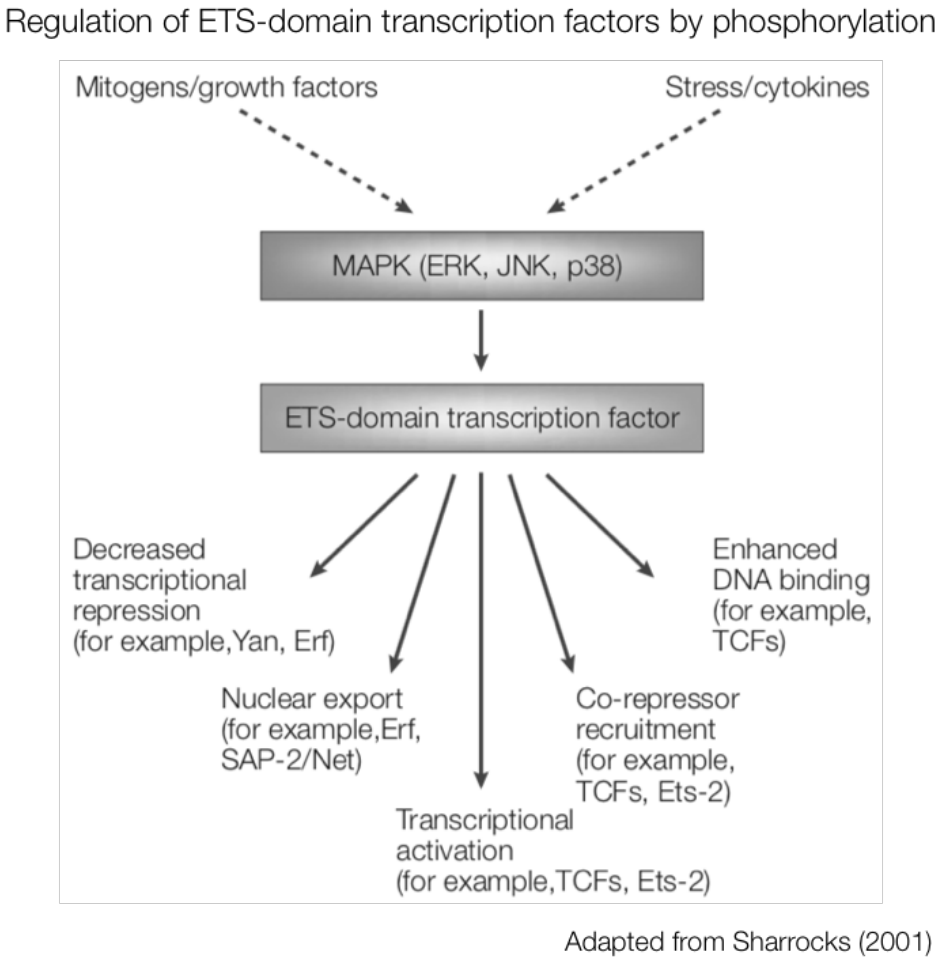


Figure 7. ETS domain transcription factors are generally regulated by phosphorylation downstream of MAPK (ERK, JNK, p38) signalling cascades
Extrinsic signals are relayed via MAPKs that lead to activation of ERK, JNK, or p38, which in turn phosphorylate ETS domain transcription factors (adapted from Sharrocks et al., 2001).

Biological functions of ETS family transcription factors

The ETS domain family of transcription factors have been implicated in a diversity of biological processes. They are known to regulate proliferation upon ERK activation (Treisman,

1994); prevent apoptosis in the case of Ets2 (Sevilla et al., 1999); and affect cellular senescence (Ohtani, 2001). For the latter, Ets2 has been shown to interact with the Id family of transcriptional repressors, which in turn regulates transcription of p16 that promotes cell cycle block and entry into senescence.

In development, ETS domain proteins are well known to regulate haematopoiesis (Scott et al., 1994), as well as vasculogenesis (including a role in the lymphatic system – Ayadi et al., 2001) and angiogenesis (Lelievre et al., 2001). ETS proteins also play a role in neuronal biology downstream of FGF signalling during brain development (Roehl and Nusslein-Volhard, 2001).

Overall, the ETS domain family is unified by presence of the ETS DNA-binding domain. The mechanism by which ETS family transcription factors can affect a diversity of biological outcomes would be interesting to uncover along with their signalling context and nuclear cofactors.

1.2.2 MAPK signalling as upstream activator of ETS2

The link between signal pathway activation and alteration in gene expression is a fundamental feature of biological regulation. Ras/MAPK signalling is prevalent in development and cancer (Tetsu and McCormick, 2017). The pathway involves activation of Ras GTPase by a cell surface receptor tyrosine kinase, which in turn activates Raf, MEK, and ERK MAPK (Figure 8; Zhang and Liu, 2002). Upon activation, MAPK may be imported to the nucleus to affect a variety of transcription factors by phosphorylation. In doing so, MAPK influences the stability, nuclear localisation, co-regulator binding, and DNA binding of a number of transcription factors (Whitmarsh and Davis, 2000), thus regulating gene expression.

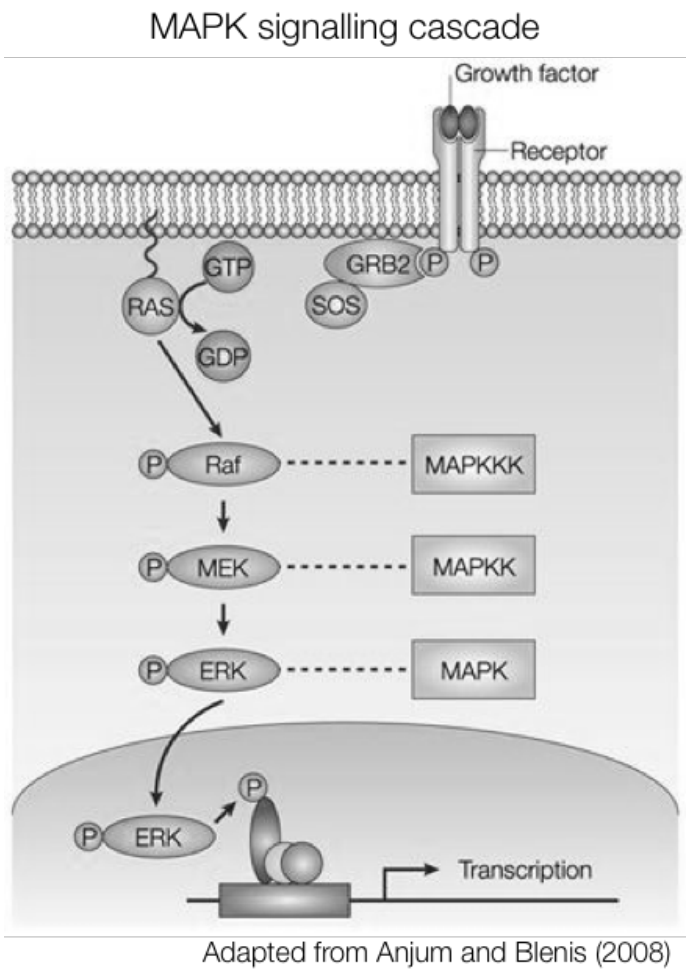


Figure 8. MAPK signalling pathway
The MAPK signalling cascade begins with growth factor binding to cell surface receptor tyrosine kinase, which results in receptor phosphorylation and initiation of sequential activation of Ras, Raf, MEK, and ERK. This in turn activates nuclear transcription factors (adapted from Anjum and Blenis, 2008).

A number of transcription factors of the ETS family are activated by phosphorylation from the Ras/MAPK cascade (Foulds et al., 2004). Upon ERK-mediated phosphorylation of Ets2 in Drosophila, Ets2 is able to persistently turn on the transcription of genes containing Ras-responsive elements (RRE) (Fowles et al., 1998). The mechanism by which Ras/MAPK pathway triggers the activation of the Ets2 Pnt signalling domain was demonstrated by Foulds et al., 2004. Threonine 72 (T72) is a highly conserved ERK MAPK phosphorylation site in the Pnt domain of Ets2. The Pnt domain is able to directly promote transactivation of Ets2 via recruitment of MAPK by containing a docking region for MAPK ERK2 (Seidel and Graves, 2002). Foulds et al., 2004 probed for proteins that would preferentially bind to phosphorylated

versus unphosphorylated Ets2 via affinity chromatography of nuclear extracts using HeLa cells with ERK2-phosphorylated versus mock-treated Ets2. Through this unbiased screen, they found the increased binding of CBP and p300 to phosphorylated Ets2. Mutagenesis experiments deleting different domains of Ets2 revealed the necessary requirement for both the phospho-acceptor site T72, as well as the entire Pnt domain in enabling Ets2 interaction with its transcriptional cofactors. Specifically, CBP/p300 can superactivate the transcriptional activity of Ets2 via direct interaction in the nucleus. Overall, Foulds et al., 2004 showed how Ras/MAPK signalling activates Ets2 via ERK2-mediated phosphorylation of T72, as well as subsequently promoting the binding of transcriptional coactivator CBP/p300. This mechanism links extracellular MAPK-activating cues to robust transcription of Ets2 target genes.

Methods to identify authentic Ets2 targets

The following methods may be employed to identify and validate authentic targets of Ets2 transcriptional activity to gain mechanistic insight into how it regulates cellular processes in normal and abnormal conditions. 1. RNA-based approaches (identification of differentially expressed genes via inspecting mRNA, with RNA-sequencing being a high throughput approach); 2. Mouse models (by evaluating gain of function and knockout phenotypes); 2. Chromatin immunoprecipitation (ChIP) – based on transcription factor occupancy of a promoter and enriching for transcription factor bound chromatin by immunoprecipitation using a specific antibody (Orlando, 2000); 3. DNA-based approaches for identification of sequence specific interaction; and 4. Whole genome PCR to identify direct target gene promoter and enhancer sequences for known DNA binding proteins (Watson et al., 2000). ChIP may be conducted using several antibodies simultaneously to elucidate synergistic interactions of Ets2 and tissue-specific cofactors. Co-localisation staining studies may then be employed to validate

presence of Ets2 co-expression with tentative transcriptional target. (Sementchenko and Watson, 2000)

Overall, Ets2 activates target gene transcription upon its MAPK-mediated phosphorylation at T72. Due to the context dependency of Ets2 transcriptional cofactors and activity, a combination of DNA, RNA, chromatin, and genetic-based approaches may be employed to elucidate authentic Ets2 targets that confer its transcriptional output.

1.3 Role of ETS2 in mammalian development

1.3.1 Generation of the *Ets2* ^{db1} mouse mutant

Yamamoto et al., 1998 generated a targeted deletion of the conserved DNA binding domain of *Ets2* to make the *Ets2* ^{db1}-mutant (also referred to as '*Ets2*-mutant'). Homozygous mutants for this deletion displayed reduced growth, particularly of the trophoblast, accompanied by decreased *Mmp9* expression concomitant with aberrant accumulation of extracellular matrix (ECM). Mutants displayed decreased proliferation of the EPC, and lethality by around E9.0.

Yamamoto et al., 1998 targeted murine *Ets2* via homologous recombination in ESCs using a vector designed to replace a part of the three exons of the *Ets2* DNA binding domain with pMC1NeoA selectable marker (Henkel et al., 1996). Mutant *Ets2* lacked DNA binding activity and the signals required for nuclear translocation (Boulukos et al., 1989). Yamamoto et al., 1998 validated their *Ets2* ^{db1} ESC targeting in a number of ways. They used RNA protection assay to verify lack of wild type (WT) RNA in ESCs with two targeted alleles. They, however, noted that these cells expressed comparable levels of slightly larger RNA that could hybridise to a probe, which recognised the 5' end of the *Ets2* mRNA. This indicated presence of a fusion transcript via transcription through the pMC1neo promoter through the large 3' untranslated region (UTR) of *Ets2* RNA (Watson et al., 1990), and was also shown to have neo

coding sequence. This fusion transcript was preferentially localised to the cytoplasm attributed to absence of DNA binding domain mediated nuclear translocation. Although the allele is not null, it is loss of function due to absence of the DNA binding domain. This was verified by fractionation of nuclear and cytoplasmic compartments in which mutant Ets2 was only found in the cytoplasm.

1.3.2 ETS2 is a core transcription factor for the mouse trophoblast lineage

The earliest reported role for ETS2 is in the post-implantation mouse ExE (Georgiades and Rossant, 2006). *Ets2* has been shown to be expressed, by *in situ* hybridisation (ISH), as early as E5.5 in the ExE. This expression pattern persists until around E7.75 when *Ets2* is additionally expressed in the primitive streak (Figure 9). Georgiades and Rossant, 2006 suggested a requirement for ETS2 in activating ExE signature marker gene expression, i.e. of *Bmp4, Cdx2, ErrB, Spc4* and *Sox2* (Figure 10). This was demonstrated using the *Ets2* ^{db1}-mutant (Figure 10 – 'type 1'). Though the ExE tissue is markedly reduced in this mutant, there is still a small piece of tissue present at the right location and respecting embryonic compartment boundaries. This was demonstrated by double ISH using epiblast marker *Oct4* and EPC marker *Mash2* where unstained tissue was found between these two tissues (Figure 10F).



Figure 9. Expression pattern of *Ets2* by ISH from E3.5 to E7.75
(A-E) *Ets2* was first reported to be expressed at E5.0 in the ExE, and remains restricted to the trophoblast up to E6.75. By E7.75, *Ets2* is also expressed in the primitive streak (asterisk) (adapted from Georgiades and Rossant, 2006).

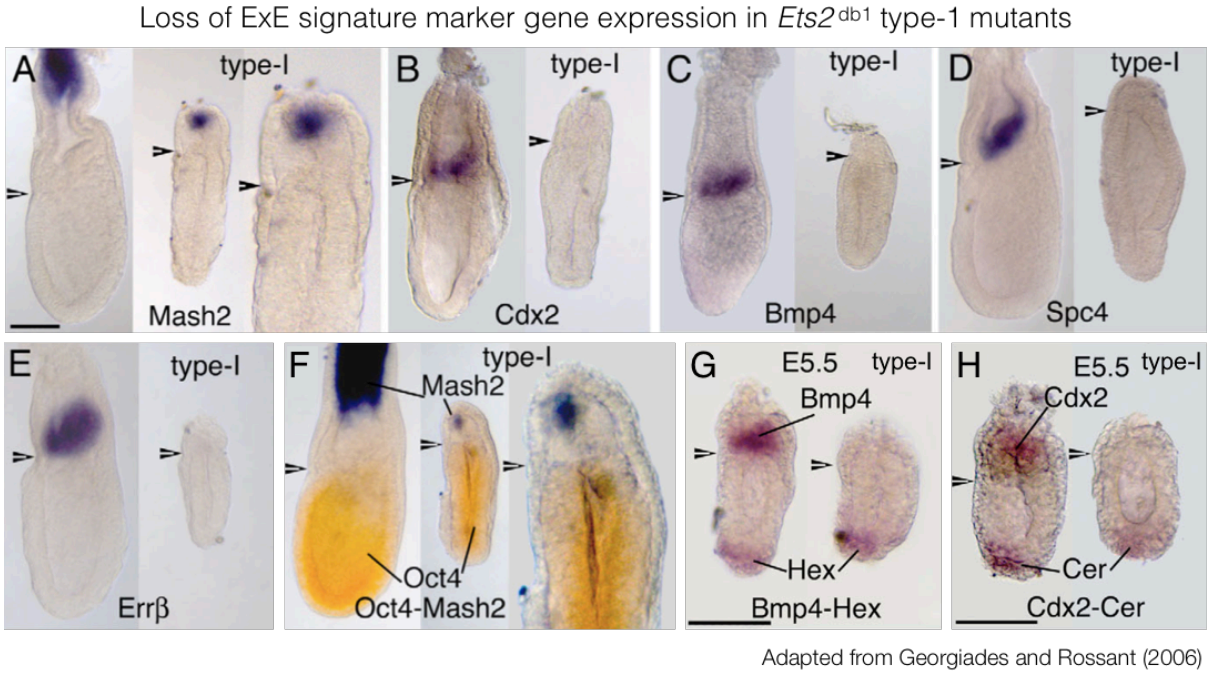


Figure 10. Characterisation of the ExE defect in *Ets2*^{db1} mutant
(A, F) E*ts2*^{db1} (type 1) mutant retains a small ExE visualised by presence of tissue between *Mash2*-positive EPC and *Oct4*-positive epiblast.
(B-E) Type 1 mutants do not express signature markers of the ExE, such as *Cdx2, Bmp4, Spc4,* and *Errβ* at E6.75
(G, H) Type 1 mutants at E5.5 express AVE markers *Hex* and *Cer*, but do not express ExE markers *Bmp4* and *Cdx2* (adapted from Georgiades and Rossant, 2006). Scale bars: 100 µm.

In addition to a defective ExE, Georgiades and Rossant, 2006 also found an abnormal AVE. AVE cells retained signature marker expression *Cer1* and *Hhex* by ISH, but displayed incomplete migration to the epiblast-ExE boundary, shifting only approximately one third of the way up the epiblast (Figure 11C). The AVE also appeared abnormally thickened upon

inspection by brightfield microscopy. This defect consequently led to absence of primitive streak induction. Therefore, the authors concluded defective A-P patterning due to a role for ETS2 in affecting the complete anterior shift of the AVE. As both the ExE and AVE defects could not be rescued upon providing *Ets2*-mutants with WT epiblast via tetraploid aggregations, ETS2 activity was thought to be restricted to extra-embryonic lineages. Since only the ExE expressed *Ets2* by ISH, Georgiades and Rossant, 2006 suggested that ETS2 functions in the ExE, possibly downstream of epiblast signals such as FGF4, to affect the patterning of the VE/AVE.

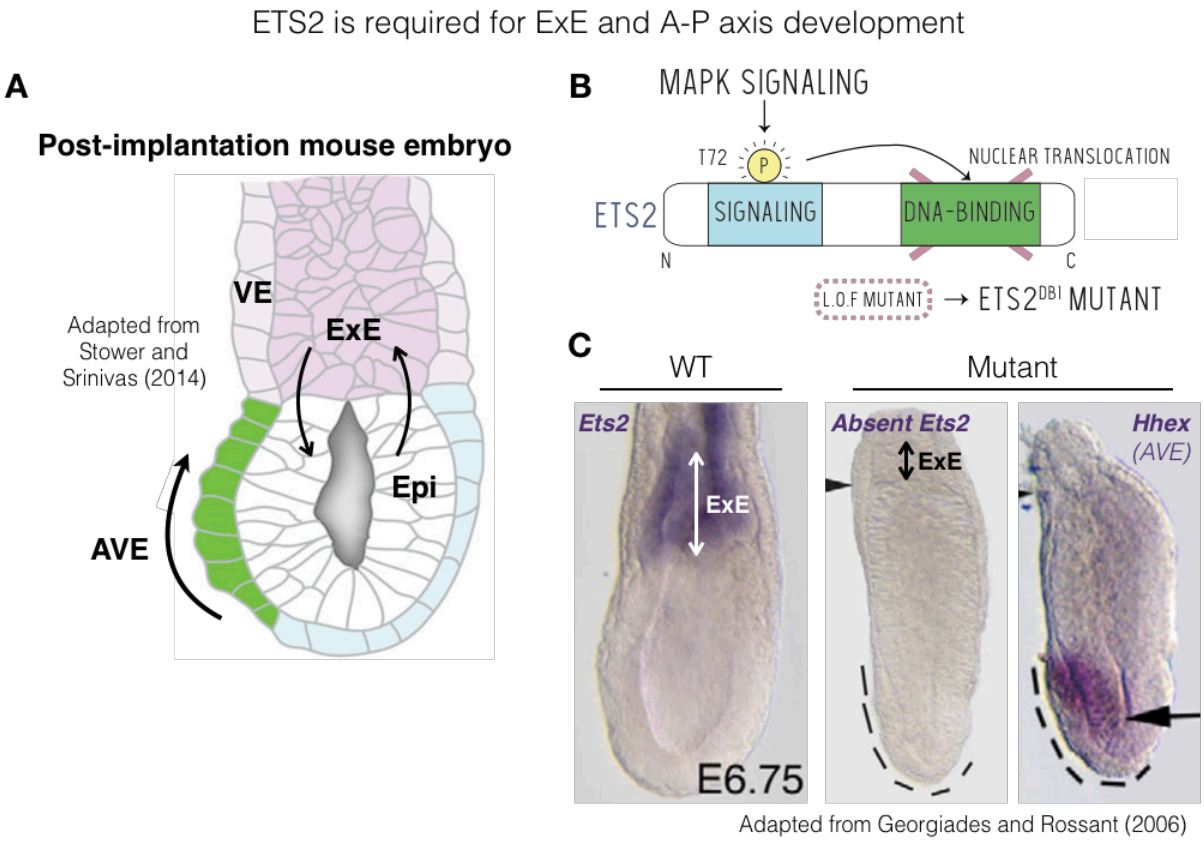


Figure 11. ETS2 is required for ExE and A-P axis development
(A) Schematic of post-implantation stage mouse embryo relevant to the characterised *Ets2*-mutant. Signalling interactions are known to occur between the epiblast and ExE, and the AVE undergoes directed migration to establish the A-P axis (adapted from Stower and Srinivas, 2014).
(B) ETS2 transcription factor schematic showing the N-terminal Pointed signalling domain that becomes phosphorylated at T72, which results in ETS2 activation and nuclear translocation. *Ets2* DNA-binding domain has been deleted in the *Ets2* ^{db1}-mutant.
(C) ISH of *Ets2* in WT versus *Ets2*-mutant showing presence of expression in the ExE of WT, but absence in the drastically reduced ExE of the mutant. The AVE in the mutant, marked by *Hhex*, appears as an abnormal thickening and does not complete its anterior shift at E6.75 when AVE migration would be complete in WT (adapted from Georgiades and Rossant, 2006).

Similar to the *Ets2*-mutant, loss of another Ets family member *Elf5* led to the complete absence of ExE tissue and retention of the EPC (Donnison et al., 2005). *Elf5* was found to be expressed in the ExE, and mutants lacked ExE signature marker expression of *Bmp4, Cdx2, Furin,* and *Spc4*. This manifested in the inability to derive TSCs from *Elf5*-mutants, similar to *Ets2*-mutants (Wen et al., 2007). Donnison et al., 2005, reported the presence of AVE induction in *Elf5*-mutants, but absence of gastrulation. Although the extent of AVE migration was not characterised, *Elf5*-mutants displayed a localised thickening in the VE, potentially suggesting a similar phenotype between *Ets2*- and *Elf5*-mutants. Functional redundancy of ETS2 and ELF5 was reported in Donnison et al., 2015, where the compound heterozygous (het) mouse model of $Ets2^{+/-}$;$Elf5^{+/-}$ displayed a reduced ExE phenotype, whereas het mice for either transcription factor were aphenotypic. This suggests dosage sensitivity of ETS2 and ELF5-mediated transcriptional outputs in the ExE.

1.3.3 Role of Ets2 in trophoblast stem cell self-renewal

Ets2 has been shown to be required for the self-renewal of TSCs *in vitro* (Wen et al., 2007). This was demonstrated using a conditional knockout TSC model of *Ets2* where its Cre-mediated inactivation resulted in defective TSC proliferation. Wen et al., 2007 found *Ets2*-deficient TSCs to turn on differentiation programmes similar to conditions of growth factor withdrawal. In particular, *Cdx2, Eomes, Pace4,* and *Errβ* were shown to be sensitive to loss of Ets2 in TSCs. Fgf4 signalling mediated by ERK2 is critical for TSC self-renewal. Fgf4-ERK2 could be upstream of Ets2, which in turn can activate a number of genes for TSC self-renewal. Fgf receptor-dependent ERK activation has been shown *in vivo* to be required in the ExE, which is where TSCs are present (Corson et al., 2003).

Wen et al., 2007 did not detect changes in apoptosis gene expression. However, this pathway is mostly regulated post-transcriptionally (Taylor et al., 2008). Thus, changes in

apoptosis may be more suitably assessed by observing changes in protein, or cellular characteristics that are hallmarks of apoptosis. ChIP was further conducted in TSCs in accordance with Hollenhorst et al., 2004 to verify the physical association of Ets2 and the highly conserved regulatory region of *Cdx2*. This was validated by the ability for Ets2 to activate minimal reporters driven by *Cdx2* regulatory region. However, removal of *Ets2* still enabled low levels of *Cdx2* mRNA expression in TSCs, as well as in the colon (Wen et al., 2007). This suggests that Ets2 is not the sole activator of *Cdx2* expression.

1.3.4 Role of Ets2 in trophoblast stem cell induction from murine fibroblasts

Reprogramming technology involves the use of a set of core signalling and transcription factors that can reset the epigenetic state and maintain gene expression profiles distinctive of a particular stem cell type. TSCs originate at the first cell fate decision and eventually contribute to the embryonic portion of the placenta. Embryonic and extra-embryonic lineages in mouse are strictly segregated by an epigenetic barrier that is still challenging to overcome when converting ESCs into TSCs (Latos and Hemberger, 2016). Kubaczka et al., 2015 deduced that reprogramming of postnatal mouse fibroblasts into TSCs can be achieved through the transient expression of four transcriptional factors: Tfap2c, Eomes, Gata3, and Ets2. These induced TSCs (iTSCs) were characterised to have overcome the epigenetic barrier that segregates extra-embryonic lineages and somatic cells. This was supported by similarity between TSCs (derived from the blastocyst) and iTSCs in morphological features, gene expression, as well as patterns of DNA methylation. In addition, iTSCs are able to sustain transgene independent self-renewal, as well as be able to give rise to extra-embryonic lineages and successfully chimerise in host placentas. As Ets2 is one of the four factors sufficient to reprogram fibroblasts into bona fide TSCs comparable to the blastocyst-derived counterpart, this further supports the importance of Ets2 in significantly contributing to the transcriptional and epigenetic programme required for

TSC fate and maintenance. Interestingly, Ets2 but not Elf5 was deduced as important for efficient iTSC conversion. Ets2 may also have a general role in mediating transdifferentiation, since its overexpression with MESP1 is sufficient for direct conversion of human skin fibroblast cells into cardiac progenitors (Islas et al., 2012).

1.3.5 ETS2 involvement in transdifferentiation to cardiac progenitor fate

Islas et al., 2012 demonstrated that use of core cardiac transcription factor Mesp1 alone was insufficient to reprogram human dermal fibroblasts into cardiac progenitors. In Ciona development, Ci-Mesp (Mesp1 orthologue) is expressed in pre-cardiac lineages, following which Ets2 expression is activated and serves to define pre-cardiac mesoderm by preparing the cells to become responsive to cardiac fate inducing signals (Davidson et al., 2006). This developmental observation was adopted to reprogram human skin fibroblasts with MESP1 and ETS2. Islas et al., 2012 found that neither MESP1 nor ETS2 alone were sufficient for conversion into cardiac progenitors, but that forced expression of both transcription factors could successfully reprogram human fibroblasts into cardiac progenitors. These reprogrammed cells expressed core cardiac transcription factors, as well as displayed sarcomeres and calcium transients. ETS2 alone is able to induce transcription of core cardiac transcription factors NKX2.5, GATA4, ISL1, MESP1, and MEF2; however, is unable to activate SMA/ACTA2, SM22α, and affect terminal cardiogenesis. Addition of MESP1 enables differentiation to acquire classic cardiomyocyte characteristics. Upon constitutive overexpression of ETS2 in human dermal fibroblasts, cells with endothelial characteristics could also be generated, characterised by expression of PECAM1 and KDR (Islas et al., 2012). In addition, ETS2 could also lead to enriched transcriptional programme of ACTIVIN/NODAL pathway members (*Tdgf1*, *Cited2*, *Lefty1*, *Acvr2b*), which is essential for cardiac progenitor fate. Human fibroblasts reprogrammed with MESP1 and ETS2 appear to resemble early to intermediate

stages of cardiomyocyte maturation and require further factors to enable maturation into terminally differentiated cardiomyocytes (Kattman et al., 2011).

Overall, ETS2 is a core transcription factor for trophoblast fate, as seen in the severely reduced ExE and the inability to derive TSCs from the *Ets2* ^{db1} homozygous mouse mutant. Ets2 is also required for self-renewal of TSCs and reprogramming to TSC fate for *in vitro* studies. ETS2 in conjunction with MESP1 is sufficient to reprogram human fibroblast cells into cardiac progenitors capable of forming cardiomyocytes. ETS2 has additionally been implicated in ECM remodelling via activating transcription of *Mmp3/9/13* described in trophoblast invasion to form the placenta (Yamamoto et al., 1998), as well as in bone mineralisation to achieve mature osteocytes from differentiating osteoblasts (Vary et al., 2000).

1.4 ETS2 in cancer

1.4.1 Oncogenic roles of ETS2

In human cancers, ETS2 has been found to be oncogenic predominantly in the stromal compartment, as demonstrated by its role in the mammary tumour microenvironment to drive oncogenesis (Tynan et al., 2005). Using a condition *Ets2*-flox allele, loss of *Ets2* in mammary tumour cells was found to have no effect, whereas transplantation of mammary tumours in a host with stroma restricted in ETS2 activity limited the early growth of the tumour.

Trimboli et al., 2009 elucidated a PI3K-ETS2 axis of signalling in mammary stromal fibroblasts where PTEN plays a critical tumour suppressive role for mammary epithelial cancer development. The mechanism by which tumour stromal cells in mammary epithelia fuel tumour formation was investigated by genetic inactivation of *Pten* in stromal fibroblasts of mouse mammary glands (Trimboli et al., 2009). This tissue-specific loss of *Pten* led to marked increase in all phases of tumorigenesis, including initiation, progression, and malignant

mammary epithelial cancer spread. Classic hallmarks of cancer were present, such as ECM remodelling, angiogenesis, and infiltration of immune cells via elevated chemokine and cytokine production. Mechanistically, Trimboli et al., 2009 found that loss of *Pten* resulted in elevated expression, activation (via upregulation of T72 phosphorylation), and nuclear recruitment of ETS2 to target promoters to drive cancer formation. Inactivation of *Ets2* in *Pten*-deleted stroma was sufficient to suppress tumour growth and progression via restoration to normal mammary stroma that lacked oncogenic features. Blinded histopathological analysis of *Pten*-stroma-deleted mice and human breast cancer samples were indistinguishable. Pten is known to regulate multiple signalling pathways, e.g. Ras, Akt, JNK (Georgescu, 2010), which could cooperate to affect ETS2 transcriptional activity. The ability for loss of *Ets2* in the mammary stroma to rescue the tumorigenic effects of *Pten* absence, demonstrates the oncogenic role of ETS2 in the mammary stroma. This observation is in line with the finding by Tynan et al., 2005 for the non-cell autonomous effect of ETS2 in mouse mammary tumours.

Wallace et al., 2013 further demonstrated the additional role of Ets2 in tumour stromal fibroblasts in mediating angiogenesis to promote mammary epithelial cancer progression. PyMT and ErbB2 breast cancer mouse models with loss of *Ets2* in stromal fibroblasts led to reduced tumour size (Figure 12A). This was characterised by decreased blood vasculature associated with changes in gene expression profile of breast tumour endothelial cells. ETS2 was sufficient to elicit *de novo* blood vessel formation *in vivo* without tumour cells. Thus, ETS2 in the tumour stroma is sufficient to induce angiogenesis and affect changes in endothelial cells to support tumour growth and progression.

In the context of pancreatic metastatic transformation, ETS2 was found to promote immune cell recruitment via regulating chemokine production (Pitaressi et al., 2016). This results in pancreatic acinar-to-ductal metaplasia, which initiates pancreatic tumour formation. ETS2 was found to be upregulated in smooth muscle actin-positive fibroblasts in pancreatic

stromal cells of human pancreatic ductal adenocarcinoma tissue. Removal of *Ets2* from stromal fibroblasts led to decreased proliferation and increased apoptosis of pancreatic epithelial cells, as well as resulted in markedly reduced immune cell infiltration in the stroma. ChIP confirmed ETS2 direct activation of immune cell recruiting chemokines *Ccl3/4* and *Cxcl4/5/10*. Pitaressi et al., 2016 demonstrated the activation of ETS2 in both mouse and human fibroblasts during acinar-to-ductal metaplasia of the pancreatic epithelium. Loss of *Ets2* in stromal fibroblasts decreased hallmarks of this oncogenic transformation, and decreased cytokine and chemokine production, resulting in decreased immune cell infiltration. These specific targets of ETS2 have been implicated in elevated tumour proliferation in a number of cancers (Kawamura et al., 2012). The precise mechanism and downstream consequences by which ETS2 activation promotes stroma-induced oncogenesis of epithelial tumours can help inform and rationalise targeted therapeutic strategies.

1.4.2 Tumour suppressive functions of ETS2

Munera et al., 2011 demonstrated the tumour suppressive role of ETS2 in intestinal epithelia. They found that condition removal of *Ets2* in intestinal epithelial cells resulted in increased number of *Ets2*-deficient colonic crypts, which is the intestinal stem cell compartment, in both young and adult mice. These *Ets2*-loss-of-function crypts were characterised by increased proliferation (Figure 12B), and also featured increased crypt fission resulting in large regions of *Ets2*-deficient crypt compartments. These crypts were further found to have a thickened mucosal layer and increased goblet cell number, with more densely packed epithelium. Mice with loss of intestinal epithelial *Ets2* were found to have significantly increased tumour formation upon treatment with tumour inducing agents. Munera et al., 2011 demonstrated an intestinal epithelial cell autonomous role for ETS2 in repressing tumour

formation, potentially contributing to the tumour protective properties of increased *Ets2* dosage in individuals with an extra copy of chromosome 21 (Sussan et al., 2008).

The WNT signaling cascade is known to be involved in colon tumor formation. *Ets2* is located on human chromosome 21 and has been demonstrated to be a target of WNT signaling in intestinal stem cells and colorectal cancer cells (van de Wetering et al., 2002). *Ets2* is expressed in the intestinal stem cell crypt compartment of both humans and mice (Segditsas et al., 2008), and could be a direct WNT target supported by presence of TCF binding sites in the promoter of *Ets2*. ETS2 may also be regulated by ASCL2, which is an intestinal stem cell marker and a known direct target of WNT signaling (van der Flier et al., 2009). Overall, ETS2 appears to moderate the levels of proliferation in response to WNT pathway activation in intestinal stem cells (Garcia et al., 2009).

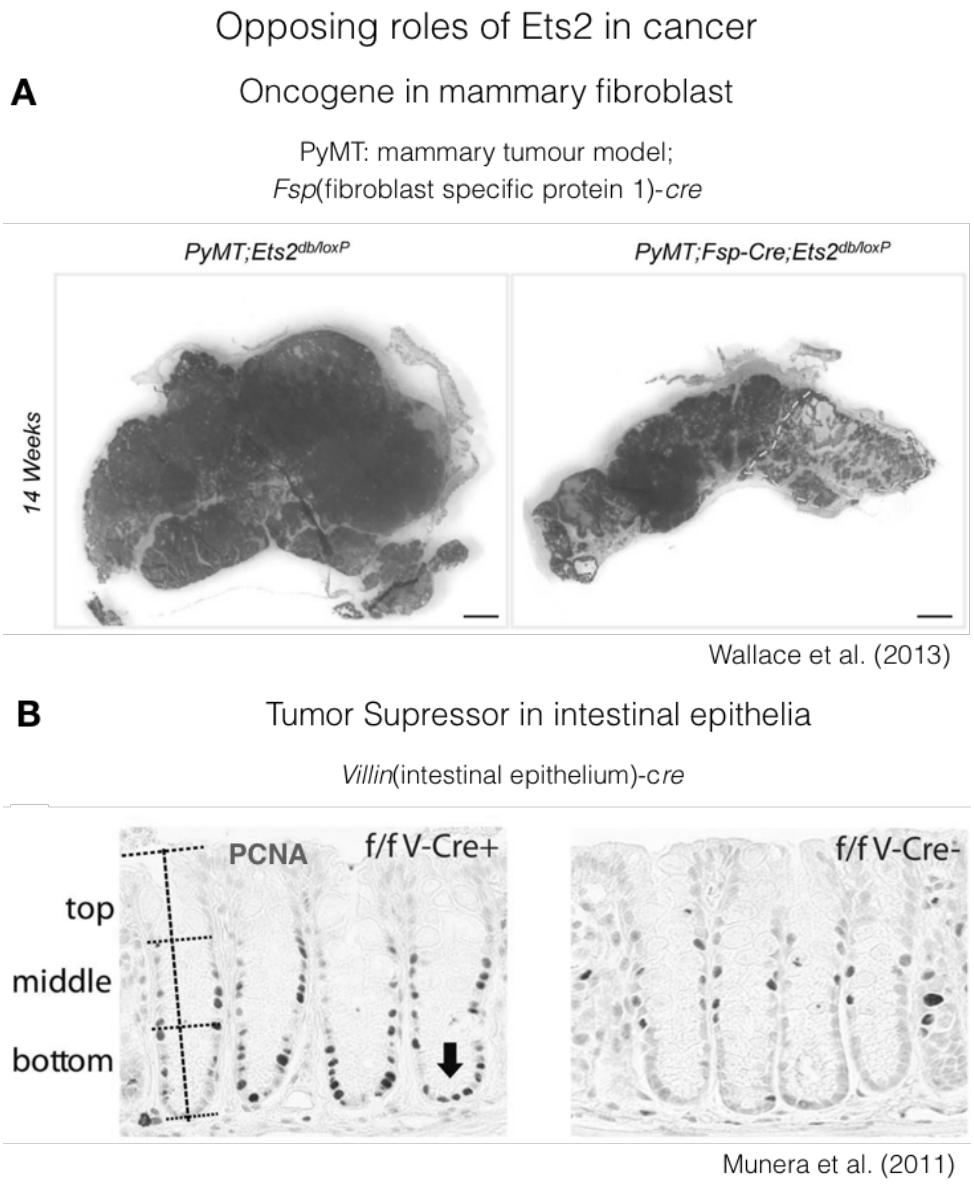


Figure 12. Opposing roles of ETS2 in cancer
(A) Loss of *Ets2* in fibroblasts of mammary tumour model PyMT leads to reduced tumour size (right; scale bar: 2 mm) (adapted from Wallace et al., 2013).
(B) *Ets2* loss in intestinal epithelia results in increased proliferation (visualised by PCNA), particularly at the lower third of the intestinal crypt, which is where stem cells reside (adapted from Munera et al., 2011).

1.5 Rationale for thesis project

ETS2 acts as a tumour suppressor in a cell-autonomous fashion in intestinal epithelia (Munera et al., 2011), yet the transcriptional effectors of ETS2 in epithelial cells have not been characterised in any context. Absence of ETS2 in the murine gut leads to increased proliferation and more rapid cell cycle rates in the crypt-based stem cell compartment, as well as crypt fission that indicates a competitive advantage. ETS2 may be regulating WNT/β-

CATENIN pathway induced proliferation in intestinal stem cells; however, the ETS2-WNT axis of signalling is unclear. In early mouse embryogenesis, the AVE epithelium was shown to be thickened in the *Ets2* ^{db1} loss-of-function mutant (Georgiades and Rossant, 2006), but proliferation was not characterised, and neither was the transcriptional basis for this observation. The ExE in the *Ets2* ^{db1}-mutant was drastically reduced, lacking marker gene expression. The opposing tissue size phenotypes in the *Ets2* ^{db1}-mutant makes it a suitable model for studying how transcription factor ETS2 may suppress or promote tissue growth, related to its tumour suppressive and oncogenic effects in cancer. However, the origins and cell biological bases for the ExE and AVE phenotypes in the *Ets2* ^{db1}-mutant have not been explored.

In summary, ETS2 has been shown to have opposing effects on tissue size and therefore progression of cancer. Yet, the mechanisms for these different functions of ETS2 remain unclear. The *Ets2* ^{db1}-mutant appears to manifest the opposing effects of ETS2 on tissue size: reduced ExE and thickened AVE. However, these phenotypes were characterised at a stage when the defects had been evident. Therefore, the origins for their abnormal development remain unknown. The cell biological changes in these tissues that characterise and result in these defects were unexplored. Finally, the downstream transcriptional targets of ETS2 in mediating these defects have not been elucidated. These questions require high resolution technologies, and high throughput sequencing to uncover transcriptional changes from low cell numbers in a tissue-specific fashion from early embryonic tissues.

1.6 Statement of aims

Aim 1: Characterise the embryological and cellular bases for the *Ets2* ^{db1}-mutant phenotype.

Aim 2: Characterise the transcriptional defects that result in the mutant phenotypes using scRNA-seq.

Aim 3: Find the mechanistic bases for the mutant phenotype via testing hypotheses arising from scRNA-seq data using pharmacological intervention and *in vivo* genetics.

1.7 Hypothesis

I hypothesise that the ExE and AVE phenotypes seen in the *Ets2* ^{db1}-mutant initiate prior to, or during early post-implantation development, as the anomalies are already evident at post-implantation. I hypothesise that the ExE and AVE defects arise due to different mechanisms and functions of ETS2. These could be a result of tissue-specific transcriptional cofactors (Plotnik and Hollenhorst, 2017). Characterising the starting point and progression of the ExE and AVE phenotypes may inform pharmacological strategies. These can result in phenotypic rescue if the perturbed pathway may be corrected before the defect arises. Conserved ETS2 upstream signals are likely to be applicable in early mouse embryogenesis whereby MAPK signalling leads to ETS2 activation. Thus, the mutant phenotypes may be recapitulated in WT upon inhibiting these upstream signals. I hypothesise that scRNA-seq capturing the initiation and progression of a particular phenotype will reveal differentially expressed genes and pathways downstream of ETS2 transcriptional activity, which may be validated *in vivo*.

2. Materials and methods

2.1 Animal husbandry

All mice were maintained on a 7am to 7pm light-dark cycle, housed and treated in accordance with the Animals (Scientific Procedures) Act 1986 of the Home Office. Noon on the day of finding a vaginal plug was designated E0.5. All mice were maintained under PPL 30/3420 (Srinivas: Morphogenesis during early mammalian development), and regulated procedures were conducted under the PIL IC9BDE3E3 (Ms Di Hu). Pregnant females were sacrificed by cervical dislocation and verified by cardiac puncture – Schedule 1 of Animals (Scientific Procedures) Act 1986. No distinction was made between female and male embryos in the analysis.

2.2 Mouse lines

Mouse nomenclature used was in accordance with MGI guidelines (http://www.informatics.jax.org/mgihome/nomen/gene.shtml).

Summary of transgenic mouse lines used in this thesis

Mouse line	Reference	Genetic background from breeding and maintenance
Ets2 ^{db1} (also known as *Ets2* ^{tm1Rgo} on Jax)	Yamamoto et al., 1998	Maintained as **pure** 129S6/SvEvTac background (original strain: 50% 129S2/SvPas – a substrain of the pure 129 inbred background; 50% FVB/N). 8 generations of pure 129 maintenance as at January 2019 (99.8% 129).
'Ets2-Hhex-GFP'	*Hhex-GFP:* Rodriguez et al., 2001	*Ets2* ^{db1} line crossed with *Hhex*-GFP reporter (Rodriguez et al., 2001), which has been maintained on a **mixed** background (CBA/J, C57BL/6) with some of the original 129S6/SvEvTac. 4 generations of crossing into mixed background as at January 2019 (6.25% 129).

Table 1. Summary of transgenic mouse lines used in this thesis

Ets2 ^{db1}-mutant (also referred to as '*Ets2*-mutant') was obtained from The Jackson Laboratory deposited by Dr. Robert G. Oshima. This targeted allele (also referred to as *Ets2* ^{tm1Rgo} on Jax) is loss-of-function where a portion of the three exons that encode the *Ets2* DNA binding domain

has been replaced by pMC1NeoA (Yamamoto et al., 1998). Heterozygous (het) *Ets2* $^{db1/+}$ mice were maintained by crossing to pure 129S6/SvEvTac strain (for 8 generations as at January 2019 – 99.8% 129 background). Homozygous (hom) *Ets2* ^{db1}-mutants, lethal by E9.0 (Yamamoto et al., 1998), were derived from het-het matings. *Ets2*-mutants from pre- to post-implantation were present in Mendelian frequencies (Table 2).

Stage	WT	Het	Mutant
E4.75	17 (25.4%)	35 (52.2%)	15 (22.4%)
E5.0	38 (24.8%)	74 (48.4%)	41 (26.8%)
E5.5	69 (23.9%)	147 (50.9%)	73 (25.2%)
E6.5	71 (26.7%)	130 (48.9%)	65 (24.4%)

Table 2. Embryo genotyping frequencies from E4.75 to E6.5
Approximately at Mendelian frequencies

An *Ets2-Hhex-GFP* line (also referred to as '*Ets2-Hhex*') was derived to fluorescently label the AVE in *Ets2*-mutants. The *Hhex-GFP* line was by generated by random insertion where cis-acting regulatory elements drive *Hhex-GFP* expression (Rodriguez et al., 2001), and was shown to label the AVE in a mosaic fashion (Srinivas et al., 2004). *Hhex-GFP* mice were of mixed background with CBA/J and C57BL/6. The *Ets-Hhex* line was generated by crossing *Ets2*-het male with *Hhex-GFP*-hom female. *Ets2*-het/*Hhex-GFP*-het female offspring were subsequently crossed with *Hhex-GFP*-hom males. Brother-sister crosses of *Ets2*-het/*Hhex-GFP*-hom offspring (with GFP levels determined by qPCR) and *Ets2*-WT/*Hhex-GFP*-hom were then conducted to maintain the line (for 4 generations as at January 2019 – 6.25% 129 background).

2.3 Genotyping

2.3.1 DNA preparation by crude lysis

DNA was extracted from ear biopsies conducted in accordance with standard procedures. Crude lysis was sufficient to achieve unequivocal results for all genotyping conducted in this thesis. This involved submerging each ear biopsy in 150 µl lysis buffer (Table 3) and 0.5 µl (0.33%) Proteinase K at 55°C anywhere between 3 hours to overnight with a brief vortex at maximum speed in between. Proteinase K was then heat inactivated at 95°C for 15 minutes. Samples were subsequently spun down at 10,000 rpm, 4ºC. For embryo genotyping, the only difference was use of 15 µl of lysis buffer and 0.05 µl (0.33%) Proteinase K, and 55°C incubation time was shortened to 1.5-2 hours.

Lysis buffer

Component	Stock concentration	Volume	Final concentration
50mM Tris HCl (pH 8-8.5)	1M	500 µl	0.05M
1mM EDTA	0.5M	20 µl	0.001M
0.5% TWEEN 20		50 µl	
H_2O		9.43ml	
Total		10 ml	

↓

Syringe filter, aliquot, and store at 4°C

Table 3. Lysis buffer

2.3.2 Genotyping PCR conditions

Ets2 PCR primers and conditions

Primers for *Ets2* ^{db1}-mutant (db1) and WT (WT) alleles (5'-3' sequence; Georgiades and Rossant, 2006):

Ets2_db1_WT forward: CGTCCCTACTGGATGACAGCGG

Ets2_WT reverse: TGCTTTGGTCAAATAGGAGCCACTG

Ets2_db1 reverse: AATGACAAGACGCTGGGCGG

Expected bands: 220 bp WT, 200 bp mutant – resolved using 2% agarose gel. Ets2 PCR conditions for standard PCR reaction (Table 4): 94°C for 3 minutes; 35 cycles of 94°C for 30 seconds, 64°C for 30 seconds, 74°C for 5 seconds; followed by 72°C for 5 minutes.

Standard PCR reaction

Component	Volume with primer 3	Volume without primer 3
H_2O	17.625 µl	18.125 µl
10X Buffer	2.5 µl (Sigma)	2.5 µl (Sigma)
Primer 1	0.5 µl (10 µM)	0.5 µl (10 µM)
Primer 2	0.5 µl (10 µM)	0.5 µl (10 µM)
Primer 3*	0.5 µl (10 µM)	
$MgCl_2$	0.5 µl (25 mM)	0.5 µl (25 mM)
dNTP	0.625 µl (10 mM)	0.625 µl (10 mM)
Taq DNA polymerase	0.25 µl (Sigma D1806)	0.25 µl (Sigma D1806)
DNA	2 µl	2 µl
Total	25 µl	25 µl

*Primer 3 was used when genotyping for WT and mutant alleles involve a shared forward or reverse primer (applicable to *Ets2* ^{db1} genotyping) and work with the same PCR conditions

Table 4. Standard PCR reaction
Final concentrations were in accordance with standard protocol (Sigma D1806)

2.3.3 Quantitative PCR primers and conditions

QPCR was employed to identify *Hhex-GFP* homozygosity. This genotyping was conducted using DNA extracted from ear notches prepared with standard crude lysis. Control DNA samples used included: WT, blank (no DNA control), and known *Hhex-GFP*-het and -hom samples. QPCR enabled amplification of a small (<100 bp) region of *Hhex-GFP* (Rodrigues et al., 2001), whereby amplification of homozygous samples would be twice that of heterozygous, normalized to GAPDH control. Standard qPCR conditions with SYBR green master mix (Applied Biosystems 4309155; Table 5) and delta delta CT method for quantification with experimental triplicates (Livak and Schmittgen, 2001) were used throughout.

Hhex qPCR genotyping conditions: Forward primer – AGTCCGCCCTGAGCAAAGA, reverse primer – TCCAGCAGGACCATGTGATC; product size: 55 bp. GAPDH served as housekeeping gene control; forward primer – AGAGACGGCCGCATCTTC, reverse primer – CCCAATACGGCCAAATCCGT; product size: 89 bp.

qPCR reaction with SYBR™ Green PCR Master Mix (Applied Biosystems 4309155)

Component	Volume
H_2O	6.5 µl
SYBR Green Mix	12.5 µl
Forward primer	0.5 µl (10 µM)
Reverse primer	0.5 µl (10 µM)
DNA	2 µl
Total	22 µl

(Per single reaction)

Table 5. Standard qPCR reaction
Final concentrations were in accordance with standard protocol (Applied Biosystems 4309155)

2.3.4 Fast PCR genotyping of Reichert's membrane (E5.5, E6.5)

2X GeneAmp Fast PCR Master Mix (Life Technologies, Catalog #: 4359187) and associated standard protocol was used.

Steps:

1. Set heat block to 95°C.
2. Dissect out Reichert's membrane from E5.5/E6.5 embryo (away from other embryos to avoid contamination). Try to keep the membrane as 1 piece.
3. Pick up each membrane right away after dissection and place in 5 µl of lysis buffer (making sure to transfer as little media into the buffer as possible).
5. Briefly vortex membrane in lysis buffer, then place at 95°C for 10 minutes. Remove from heat and vortex at maximum speed. Place the tubes back in 95°C for another 10 minutes.
6. Prepare fast genotyping PCR mix (Table 6):

Component	20 µl Reaction	Final Concentration
ddH_2O	to 20 µl	
2x GeneAmp Fast PCR Master Mix	10 µl	1X
10 µM forward primer	0.4 µl	0.2 µM
10 µM reverse primer	0.4 µl	0.2 µM
Template DNA	5 µl	<200 ng/rxn

Table 6. Standard fast PCR mix

7. Run fast PCR reaction (Table 7):

Step	Temperature (°C)	Time
1. Initial denaturation	95	10 seconds
2. Denature	94	1 second
3. Anneal/Extend	62-72 (depending on primer T_m)	~25 seconds/kb (e.g. 62°C for 13 seconds for *Ets2* genotyping)
35 PCR cycles for steps 2. & 3.		
Final Extension	72	10 seconds
Hold	4	indefinitely

Table 7. Standard fast PCR reaction

8. Add loading dye, load the entire reaction in the gel (2% gel for *Ets2* genotyping).

2.4 Embryo dissections

Peri- and post-implantation embryo dissections were conducted in M2 media (Sigma M7167) using forceps to recover decidua from the uterus and embryo from decidua. Hypodermic needles (30.5G) were used to remove Reichert's membrane. Standard post-implantation dissection protocol was used (Srinivas, 2010). Pre-implantation embryos were flushed from the uterus using M2 media.

2.5 Immunofluorescence staining of embryos and ESCs

2.5.1 Antibodies

Primary antibodies

Mouse anti-Oct-3/4 (Santa Cruz sc-5279), mouse anti-alpha-Tubulin (gift from Professor Elizabeth Robertson), mouse anti-Ets2 (Santa Cruz sc-365666) (recognises ETS2 DNA-

binding domain, which is deleted in mutants), mouse anti-active-β-Catenin (Millipore 05-665), rabbit anti-Laminin (Sigma L9393), rabbit anti-Collagen IV (AbD Serotec 2150-1470), mouse anti-Yap (Santa Cruz sc-10199), mouse anti-Errβ (Perseus Proteomics PP-H6705-00), rabbit anti-Cdx2 (CST 12306), rabbit anti-phospho-Histone H3 (Ser10) (Millipore 06-570), rabbit anti-Cleaved Caspase-3 (Asp175) (CST 9661), rabbit anti-phospho-Erk1/2 (Thr202/Tyr204) (CST #9101), rabbit anti-ZO-1 (Zymed 61-7300), rabbit anti-Par3 (Millipore 07-330), rabbit anti-phospho-Myosin Light Chain 2 (Ser19) (CST 3671), rabbit anti-β-Catenin (CST 9562), rabbit anti-Par6b (Santa Cruz sc-67393), rabbit anti-Scribble (Santa Cruz sc-28737), rabbit anti-Tcf7l1 (Proteintech 14519-1-AP), rabbit anti-Moesin (Proteintech 16495-1-AP), rabbit anti-Sox2 (Millipore AB5603), goat anti-Gata6 (R&D AF1700), goat anti-Brachyury (Santa Cruz sc-17745), goat anti-Amot (Santa Cruz sc-82491), rat anti-E-cadherin (Sigma U3254).

Secondary antibodies

Alexa Fluor 555 donkey anti-rabbit IgG (Invitrogen A-31572), Alexa Fluor 555 donkey anti-mouse IgG (Invitrogen A-31570), Alexa Fluor 488 donkey anti-rabbit IgG (Abcam ab150073), Alexa Fluor 488 donkey anti-goat IgG (Invitrogen A-11055), Alexa Fluor 633 donkey anti-goat IgG (Invitrogen A-21082), Phalloidin-Atto 647N (Sigma 65906), Click-iT Plus TUNEL Assay, Alexa Fluor 594 (Thermo Fisher Scientific).

2.5.2 Immunofluorescence sample preparation

Embryos were fixed with 4% PFA in 1x PBS for 10 minutes (pre-implantation), 15 minutes (peri-implantation), 20 minutes (E5.5/ESC colonies), or 30 minutes (E6.5) at room temperature with gentle rocking. They were then washed 3x 5 minutes using 0.1% Triton X-100 in 1x PBS (PBS-Tr); permeabilised in 0.25% Triton X-100 in 1x PBS for the same length of time as fixation; and washed in 0.1% Tween 20 in 1x PBS (PBS-Tw) for 3x 5 minutes. Embryos were

placed in 50 µl block overnight at 4°C. Block was made beforehand using 0.1% PBS-Tw, 3% BSA (Sigma A2153), and 2.5% donkey serum (Sigma D9663); filter-sterilised block was stored at -20°C for up to 6 months. On day 2 of the protocol, primary antibody was added in 1:100 dilution for all antibodies directly to the block followed by gentle mixing, then left to incubate overnight at 4°C. The following day, embryos were washed 4x 10 minutes in PBS-Tw, then placed in 50 µl of PBS-Tw with secondary antibodies at 1:100 dilution overnight at 4°C. Embryos were lastly washed for 4x 10 minutes in PBS-Tw, then mounted in Vectashield mounting media with DAPI (Vector Labs H-1200). For antibodies raised in mouse, the Mouse on Mouse (M.O.M.) Basic Kit (Vector Laboratories BMK-2202) was used to rid of non-specific maternal IgG fluorescent signal; where co-stained with non-mouse antibodies, the protocol was identical as that of mouse antibodies apart from use of 50% regular block and 50% M.O.M. block instead of 100% M.O.M. block.

2.6 Confocal microscopy

Samples processed for immunofluorescence were imaged on Zeiss LSM 880 using 40X/1.1NA water objective lens. Airyscan was conducted using 63X/1.4NA oil objective lens. Excitation wavelengths include: 405 nm for DAPI, 488 nm for EGFP and Alexa Fluor 488, 543 nm for Alexa Fluor 555 and 594, and 633 nm for Phalloidin-Atto 647N. Entire embryos were imaged at z-intervals of 0.8 µm for pre-implantation to E5.5 embryos, and 1.0 µm for E6.5 embryos. Z-slices and 3D opacity renderings of z-stacks were visualised using Volocity Software (Improvision, UK).

2.7 Embryo culture

Embryo cultures in inhibitor

Embryo culture media consisted of 50% Knockout Serum Replacement (Invitrogen), 50% CMRL (Invitrogen), and L-glutamine supplement, equilibrated at 37°C and 5% CO_2 for 2 hours prior to use. For Caspase inhibitor experiments, control and mutant embryos were dissected at E5.0, and cultured at 37°C and 5% CO_2 for 12 hours to E5.5 in 10 μM Pan-Caspase inhibitor Z-VAD-FMK (R&D Systems) (Bowling et al., 2018), or the equivalent DMSO percentage (0.125%). For MEK inhibitors – PD0325901 (PD03; Stemgent 04-006), U0126 (InvivoGen tlrl-u0126) – and Wnt agonist/GSK3 inhibitor CHIR99021(CHIR; Stemgent 04-0004) treatments, embryos were dissected at E5.5 and cultured to E6.5 in the same base medium, where DMSO control percentages ranged from 0.01-0.05%.

2.8 Single-cell dissociation for scRNA-seq

Individual E5.5, or E6.5 embryos were placed in 100 μl of TrypLE (Invitrogen 12563011) dissociation reagent at 37°C for different times depending on stage or cell type of interest: 3.5 minutes – E5.5 VE, 4.5 minutes – E6.5 VE, 6 minutes – E5.5 entire embryo. The embryo was then mouth-pipetted up and down once or twice for gentle mechanical dissociation using a glass capillary 10-15% larger than the size of the embryo at room temperature in TrypLE. Glass capillaries were pulled under flame, then carved to the appropriate size and smoothness using a diamond cutter. The dissociated embryo in the entire volume of TrypLE was then transferred to a microfuge tube, and neutralised with equal volume heat inactivated FBS, followed by centrifugation at 1000x RCF at 4°C for 3 minutes. For entire E6.5 embryos, 200 μl of StemPro Accutase (Invitrogen A1110501) was used at 37ºC for 12 minutes, tapping every 2 minutes. Then the same FBS neutralisation and 4ºC centrifugation steps were used. After removal of supernatant, cells were resuspended to 100 μl of 1x DPBS or HBSS with 1%

FBS. Samples were placed on ice, then transported by bike for ~15 minutes to the FACS facility at the Weatherall Institute of Molecular Medicine (WIMM). 10 minutes prior to FACS, 2X Live/Dead stain (Abcam ab115347), or DAPI was added. All FACS experiments were conducted on the SH800 cell sorter by facility operators Paul Sopp, or Kevin Clark, apart from E5.5 and E5.75 WT VE FACS conducted by myself.

Technical summary of scRNA-seq sample preparation to purify for specific cell types of the post-implantation mouse embryo

Stage/ Tissue	Sample preparation and dissociation conditions	FACS setup used
E5.5 epiblast	6 minutes in TrypLE at 37°C, gentle mouth pipetting (2-3 times), gentle tapping	Live cells: FITC (green) Dead cells: PE (red)
E5.5 AVE	3.5 minutes in TrypLE at 37°C, gentle mouth pipetting (2-3 times), gentle tapping	Live cells: DAPI-negative AVE cells: HHEXGFP-positive (FITC)
E6.25 AVE	4.5 minutes in TrypLE at 37°C, gentle mouth pipetting (2-3 times), gentle tapping	Live cells: DAPI-negative AVE cells: HHEXGFP-positive (FITC)
E5.5/E5.75 VE	Labelling for 5 minutes at room temperature with CellVue Claret far red dye (0.1% concentration); 3.5 minutes in TrypLE at 37°C, gentle mouth pipetting (2-3 times), gentle tapping	Live cells: DAPI-negative VE cells: far red-positive
E6.25 VE	Labelling for 5 minutes at room temperature with CellVue Claret far red dye (0.1% concentration); 4.5 minutes in TrypLE at 37°C, gentle mouth pipetting (2-3 times), gentle tapping	Live cells: DAPI-negative VE cells: far red-positive
E5.75 to E6.5 cultured	Removed EPC; disaggregated embryonic portion via treatment with Accutase for 12 minutes at 37°C, tapping every 2 minutes	Live cells: DAPI-negative

Table 8. Technical summary of scRNA-seq sample preparation strategies and FACS collections for different subpopulations in the post-implantation stage mouse embryo

2.9 *In silico* quantitative analysis and statistics

Morphometrics

Morphometrics of ExE and Epi length were measured on Volocity (in µm) using the longest length, often a central section, in the 3D stack. Cells were counted with the entire 3D stack using Volocity in a semi-automated fashion, where errors in automated detection were

manually corrected. Absolute ExE cell number was quantified using CDX2 signature marker, along with DAPI to visualise nuclei and F-ACTIN cell outline to clearly discern individual cells.

ExE proliferative and apoptotic indices

ExE proliferative index was quantified by taking the number of phospho-Histone H3 (PHH3)-positive ExE cells, or ExE cells with alpha-Tubulin-positive metaphase spindles and mitotic nuclei (marked by DAPI), divided by the total number of ExE cells. ExE apoptotic index was quantified by taking the number of Cleaved Caspase-3 (CC3)-positive, or Click-TUNEL-positive ExE cells divided by the total number of ExE cells.

AVE proliferative index and cell density

Similarly, AVE cell number was quantified using AMOT as ubiquitous marker (Shimono and Behringer, 2003) along with DAPI and F-ACTIN to ascertain cell boundaries. 3D opacity renderings of 5 or more views of the embryo was taken, and counting was conducted on Photoshop with counter tool. The cells at the boundary of each angle was highlighted and matched with the precise cells on Volocity followed by embryo rotation to the next angular view to ensure cells were not double counted, or missed. Quantification of mutant 3D surface view led to a significantly increased result in mutant AVE cell number. However, both control and mutant AVE cells were subsequently quantified by scrolling through the entire z-stack (0.8 μm intervals) of the AVE to account for the stratification of the mutant AVE. AVE proliferative index was quantified by taking the number of PHH3-positive AVE cells, or AVE cells with alpha-Tubulin-positive metaphase spindles and mitotic nuclei (marked by DAPI), divided by the total number of AVE cells.

AVE cell density was measured by quantifying the number of AMOT-positive AVE nuclei from 3 representative central sections spaced 10 μm apart, divided by the distance those nuclei span. The distance measurement was conducted using the freehand tool () on Fiji, tracing the path of the shared basement membrane between the epiblast and AVE cells counted. This is the standard method used to quantify cell crowding, demonstrated to validate increased cell density enabled by epithelial pseudostratification – a developmental strategy to accommodate more cells over a small surface area (Lee and Norden, 2013).

Statistics

In all cases (unless stated otherwise), hierarchical ANOVA was employed as the statistical test, as more than one measurement was taken per embryo, i.e. presence of multiple factors nested within the main factor – criteria for this statistical test. A Tukey test was added when more than 2 conditions were compared, e.g. control versus mutant at E5.5 and E6.5.

Orientation of cell division quantification

AVE angle of cell division was quantified using alpha-Tublin to visualise mitotic spindles; DAPI to mark mitotic nuclei; and AMOT to distinguish AVE cells. For each cell division considered, the embryo was resliced in Volocity to ensure that the two poles of the mitotic spindle were in the same plane. This removed the need to conduct the measurement in 3D (Luxenberg et al., 2010). Once the embryo was appropriately resliced for the cell division of interest, the image was opened in Fiji, and the cell division angle was measured by taking the angle created from a line drawn tangent to the basement membrane and another drawn across the poles of the centrosomes visualised using alpha-Tublin. Data was presented in roseplot format using the 'rose' functionality within the 'circular' package of R-studio. Statistical comparisons were conducted in a biologically-relevant manner, segregating the angular bins

into 0-30° = planar, 30-60° = oblique, 60-90° = perpendicular (Luxenberg et al., 2010). In additional to hierarchical ANOVA and Tukey statistical tests, the Fisher's exact test was also employed to appropriately account for angular bins compared. This test is pertinent to categorical data that involves classification in two different ways, i.e. the precise angle and associated bin in this case.

Imaris 3D outlining to confirm presence of mutant AVE epithelial stratification was conducted following standard procedure.

2.10 *Ets2*-mutant and control ESC derivation

Protocol for efficient ESC derivation in medium with inhibitors was kindly provided by Professor Jennifer Nichols (updated version from Nichols and Ying, 2006).

2.11 Gene ontology and bioinformatics of scRNA-seq data

Gene ontology (GO) was conducted using the following publicly available softwares:

1. Panther GO (http://geneontology.org/)
2. Reactome (https://reactome.org/PathwayBrowser/#TOOL=AT)
3. Princeton GO::TermFinder (https://go.princeton.edu/cgi-bin/GOTermFinder)
4. g:Profiler (https://biit.cs.ut.ee/gprofiler/gost)
5. GOrilla (http://cbl-gorilla.cs.technion.ac.il/).

Promoter analysis was conducted using Mulan Dcode (https://mulan.dcode.org/).

2.12 Cloning for 2C-HR-CRISPR

2.12.1 cDNA library preparation

RNA was extracted from E9.0 embryos, which were verified to express top differentially expressed (DE) candidates in the mutant AVE (reference: E9.0 scRNA-seq data from Pijuan-Sala et al., 2019). Embryos were dissected in M2. 3 embryos were pooled per Eppendorf tube. Media was removed to the extent possible, then samples were snap frozen in liquid nitrogen and placed in -80ºC overnight. The following day, RNA extraction was conducted using RNeasy Mini Kit (50) (Qiagen 74104) adhering to standard protocol for harvesting RNA from animal tissues. Vortexing at maximum speed was sufficient to completely dissolve embryos in buffer RLT provided. Followed RNA extraction the iScript select cDNA synthesis kit (Biorad 170-8896) was used, precisely following the oligo(dT) protocol, using 1 µg of starting RNA. This kit was able to successfully generate full length cDNA libraries with the RNase H+ MMLV reverse transcriptase, verified by the ability to amplify full length products for mutant AVE top DE candidates tested (1488 nt – *Xiap*, 1731 nt – *Msn*).

2.12.2 Primers for amplifying top mutant AVE differentially expressed candidates

Primers were designed to amplify full length cDNA for genes of interest, as well as that plus homology arms (HA) to the *Ttr* backbone (underlined). The former was used to verify primer efficacy, and the latter served to enable subsequent cloning into the *Ttr* backbone. *Hprt* housekeeping gene was used as positive control to check for presence of cDNA in case the full-length product could not be amplified from the cDNA library generated. Primers were designed to amplify both short and long *Hprt* product, providing further flexibility to evaluate the cDNA library.

Hprt long – 574 nt

Hprt long forward: CCGACCCGCAGTCCCAGCGTCGTG

Hprt long/short shared reverse: GGGCATATCCAACAACAAACTTGTC

Hprt short – 223 nt

Hprt short fw: GGAGATGATCTCTCAACTTTAACTGG

Same PCR conditions as for gene of interest.

Moesin – 1731 nt product (including stop codon)

Msn forward: CCGAAGACGATCAGTGTGCG

Msn reverse: CTACATGGACTCAAACTCATC

Msn-TTR HA forward:

GTCGAGGAGAATCCTGGCCCACCGAAGACGATCAGTGTGCG

Msn-TTR HA reverse:

GAGGAAGAGTCGAAGGGAAGCCTACATGGACTCAAACTCATC

Msn PCR conditions: 94°C for 3 minutes; 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 2 minutes; followed by 72°C for 5 minutes.

Xiap – 1488 nt product

Xiap forward: ACTTTTAACAGTTTTGAAGGAACTAG

Xiap reverse: CTAAGACATAAAAATTTTTTGCTTGAACG

Xiap-TTR HA forward:

GTCGAGGAGAATCCTGGCCCAACTTTTAACAGTTTTGAAGGAACTAG

Xiap-TTR HA reverse:

GAGGAAGAGTCGAAGGGAAGCCTAAGACATAAAAATTTTTTGCTTGAACG

Xiap PCR conditions: 94°C for 3 minutes; 35 cycles of 94°C for 30 seconds, 64°C for 30 seconds, 72°C for 1.5 minutes; followed by 72°C for 5 minutes.

PCR without *Ttr* HA was diagnostic and therefore regular Taq polymerase (Sigma D1806) was sufficient. For cDNA with *Ttr* HA, the Phusion High-Fidelity DNA Polymerase (NEB M0530) was used to increase chances of 100% accurate product, verified by sequencing, which could then be appropriately used for subsequent steps. Standard conditions were employed using less than 250 ng template cDNA according to protocol.

2.12.3 *Ttr* backbone plasmid linearisation

The *Ttr2k* (also referred to as '*Ttr'*) backbone plasmid provided by Dr. Satish A. Jayaram (adapted from Kwon and Hadjantonakis, 2009) was linearised via the following steps so that the gene of interest with *Ttr* HA could be cloned in.

Component	**µl**
DNA (1µg total)	4
10x CutSmart buffer (NEB B7204S)	2
SmaI restriction enzyme	1
ddH_2O	13
total	20

Table 9. *Ttr* backbone linearisation reaction
Final concentrations were in accordance with standard protocol (NEB B7204S).

1. Incubate above at 25ºC for 2 hours.
2. Add 1 µl CIP (NEB M0290) and incubate for another 1 hour.
3. Run product on 1% gel, expected size: 3.5-4kb (use 1kb ladder).
4. Gel purify PCR product using QIAquick Gel Extraction Kit (Qiagen 28115).

2.12.4 Ligation and Gibson Assembly to clone gene of interest into *Ttr* backbone

1. Ligation reaction using Gibson Assembly Master Mix (NEB E2611) – following standard protocol.

Component	Volume (µl)
Insert (e.g. 12ng/µl)	6
Vector (e.g. 14.8ng/µl)	4
2X enzyme mix	10
Total	20

Table 10. Gibson Assembly of *Ttr* linearised backbone and gene of interest
Final concentrations were in accordance with standard protocol (NEB E2611).

Bacterial transformation

2. Incubate reaction at 50ºC for 20 minutes.
3. Take out DH5α competent bacteria (ThermoFisher 18265017) from -80ºC at the 20 minute mark.
4. Incubate reaction for another 10 minutes, for a total of 30 minutes at 50ºC.
5. Chill reaction on ice for 2 minutes.
6. Add 3 µl of ligation reaction to bacteria and tap Eppendorf tube.
7. Let bacteria and reaction rest for 15 minutes on ice.
8. Tap Eppendorf tube.
9. Let bacteria and reaction sit for another 15 minutes on ice.
10. Heat shock bacteria at 42ºC for 1 minute.
11. Add 1 ml of cold plain LB.
12. Incubate for 37°C on shaker for 1 hour.
13. Spin down Eppendorf at 8000x RPM for 5 min at room temperature.
14. Remove supernatant leaving 100µl.

15. Resuspend pellet in the 100 µl of remaining media.

16. Plate on LB-amp agar using a sterile, bent pipet.

Day 2: colony expansion and miniprep

1. Pick 2-4 colonies with sterile pipet tip; 1 colony per tip.
2. Drop each tip in individual 5 ml round bottom Falcon vials with 3ml LB amp.
3. Place in 37°C shaker for 12 hours.
4. Spin down, pellet, perform miniprep using standard procedure from QIAprep Spin Miniprep Kit (50) (Qiagen 27104).
5. Sequence miniprep product using primers that would amplify the inserted gene of interest and both sit in the *Ttr* backbone:

Forward primer that sits at junction between Td-Tomato and 2A-peptide, which directly precedes the gene of interest:

2APJ_forward: CGAGCTGTACAAGGGCAGTGG

Reverse primer that sits in the 3' HA of TTR

TTR_3HA_reverse: CCATGTCTGGATCGCTCACAG

2.12.5 Biotinylated PCR

Perform biotinylated PCR in accordance with Gu et al., 2018 using Phusion High-Fidelity DNA Polymerase (NEB M0530).

1. Biotinylated PCR reaction:

Biotinylated forward primer for TTR backbone*: GAGGCTGCGCTTTGTCATGG

Biotinylated reverse primer for TTR backbone*: TGGATTCTCCAGCACCCTGG

*Designed by Dr. Satish A. Jayaram

Component	µl
Phusion DNA polymerase	0.25
5x buffer	5
Forward primer	1.25
Reverse primer	1.25
10mM dNTP	0.5
Template DNA	1
ddH_2O	15.75
Total	25

Table 11. Standard Phusion High-Fidelity DNA Polymerase PCR reaction
Final concentrations were in accordance with standard protocol (NEB M0530).

2. Biotinylated PCR conditions:

Step	Temperature (ºC)	Time/repeats
Heated lid	110	
Denaturation	95	3 min
Start cycle		34x
Denaturation	95	30s
Annealing	66.5	30s
Elongation	72	6 min
End cycle		
Final elongation	72	10 min
Hold	4	∞

Table 12. Biotinylated PCR reaction conditions

3. Run PCR product on 1% gel using TAE buffer and TAE loading dye.

4. Perform gel extraction with QIAquick Gel Extraction Kit (Qiagen 28115), attending to sterility (though can be conducted at the bench).
5. Elute gel extraction product using RNA injection elution buffer provided by Dr. Satish A. Jayaram, prepared according to Gu et al., 2018.
6. 20 µl of 20 ng/µl final product was handed to Dr. Satish A. Jayarm who added equal volume guide RNA.

The following components were prepared by Dr. Satish A. Jayaram:

7. Cas9-mSA plasmid was obtained from Addgene #103882, and Cas9-mSA RNA was prepared using mMESSAGE mMACHINE kit (Life Technologies 1340M), aliquoted and stored at -80ºC.
8. Sequence of guideRNA for TTR targeting: ACACAAATACCAGTCCAGCG
9. GuideRNA was prepared using EnGen sgRNA Synthesis Kit (E3322), aliquoted and stored at -80ºC.
10. The final combination of: 1. 20 µl (20 ng/µl) biotinylated PCR product, 2. Equal volume guideRNA, and 3. Cas9-mSA was passed to Dr. Jonathan Godwin for microinjection at the mouse embryonic 2-cell stage.

3. ETS2 protects the peri-implantation ExE from apoptosis

3.1 Introduction

3.1.1 The apoptotic pathway

Background in the apoptotic pathway is relevant for the understanding and interpretation of ExE tissue reduction in the *Ets2*-mutant.

Caspases are the predominant regulators of apoptosis

Apoptotic cell death is a means of removing unwanted cells in a controlled manner. The process results in cellular fragmentation followed by phagocytic clearance. The signalling effectors of the apoptotic process is mediated by sequential activation of cysteine proteases of the Caspase family that can act on a wide range of protein substrates in the cell for regulated degradation. This feature is important to minimise damage to surrounding cells and prevent unwanted immune response. As a result, the process is contained and regulated, allowing surrounding cells to carry forth their normal functions (Ichim and Tait, 2016).

Sequential activation of Caspases results in irreversible commitment towards cell death. Caspase protease activity is necessary for the phenotypic and molecular features of the pathway (Taylor et al., 2008). Caspases are grouped into: 1. Initiators that first become activated (including Caspase-2, -8, -9, -10) downstream of stress stimuli; and 2. Executioners (including Caspase-3, -6, -7) that act on cellular substrates for degradation and are the ultimate effectors of apoptosis. A biochemical hallmark of apoptosis is the inversion of phosphatidylserine from normal inward orientation in the lipid bilayer to outwards, which may then be recognised by phagocytes for cell clearing. Other proteins exposed to the cell surface to mediate clearing are Annexin I and Calreticulin (Elmore, 2007). Morphological hallmarks of end stage apoptosis include cell size reduction, cell condensation, and nuclear fragmentation (Kerr et al., 1972).

Extrinsic and intrinsic apoptotic pathways

Caspase activation occurs via two pathways: 'extrinsic death receptor' and 'intrinsic mitochondrial'. The extrinsic pathway for apoptosis is activated via death receptor interaction with its ligand. These receptors are members of the TNF family, with common cognate ligand-receptor pairs being FasL-FasR and TNFα-TNFR1, which share a common intracellular death domain. Upon ligand-receptor interaction, the death domain recruits adaptor protein FADD, which in turn binds and cleaves Procaspase-8. Caspase-8 subsequently activates execution Caspases, such as Caspase-3 (Li et al., 1998). The intrinsic pathway can respond to multiple stress signals, such as growth factor reduction, endoplasmic reticulum (ER) stress, and DNA damage (Czabotar et al., 2014). These stimuli can all lead to mitochondrial outer membrane permeabilisation (MOMP) (Tait et al., 2010). MOMP may also be triggered when Caspase-8 cleaves BID to make the activated form: truncated BID (tBID) (Li et al., 1998), which is the level at which extrinsic and intrinsic apoptotic pathways overlap (Luo et al., 1998). MOMP results in the release of Cytochrome c that leads to Caspase activation via cytoplasmic Apoptosome (Tait et al., 2010). MOMP can be inhibited via anti-apoptotic BCL-2 protein interaction with pro-apoptotic BH3-only proteins (such as PUMA), and activated pro-apoptotic MOMP effectors BAX, BAK, and BOK (Llambi et al., 2011) (Figure 13). Competitive inhibition of these interactions forms the basis of BH3-mimetic therapeutics that favour apoptosis (Oltersdorf et al., 2005).

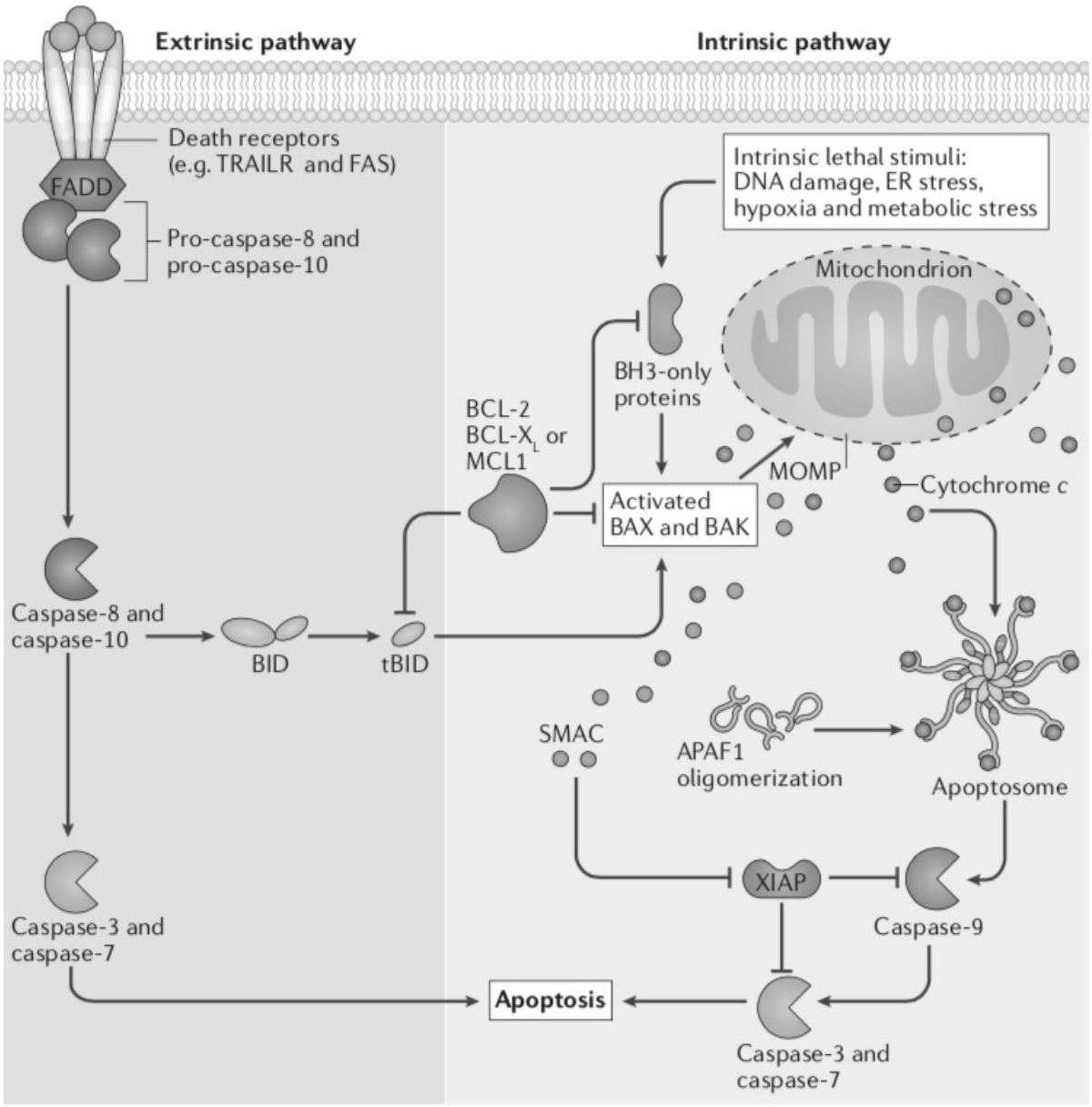


Figure 13. Extrinsic and intrinsic pathways of apoptosis
The extrinsic pathway of apoptosis (left) involves death receptor engagement with an extrinsic ligand/stimulus that results in activation of initiator Caspases (8/10) through adaptor-mediated interaction with the FADD domain that is made accessible by receptor dimerisation. Initiator Caspase activation subsequently triggers the cleavage of effector Caspases-3 and -7, which acts on multiple cellular substrates to affect cell-wide apoptosis. In the intrinsic mitochondrial pathway (right), intracellular stimuli, e.g. cellular stress and nutrient deprivation, leads to MOMP. This is mediated by activation of BH3-only proteins, which activate BAX and BAK to trigger MOMP. BCL-2 family proteins antagonise this process, thus serving as anti-apoptotic factors. MOMP enables the release of SMAC and Cytochrome c into the cytoplasm. Cytochrome c subsequently interacts with APAF1 to assemble the Apoptosome that activates Caspase-9. Caspase-9 proceeds to activate effector Caspases-3/7, which ultimately execute apoptosis (Wolf et al., 1999). SMAC release from MOMP acts to inhibit XIAP, an anti-apoptotic factor. Extrinsic and intrinsic apoptotic pathways converge at the level of tBID, which is cleavage-activated by Caspase-8 (adapted from Ichim and Tait, 2016).

Assays for detecting apoptosis

As there are many steps to the apoptotic cascade, detection of apoptosis can be made at various levels of the pathway, with some methods being more specific than others. Due to overlapping morphological features between apoptosis and necrosis, multiple assays should be employed to verify presence of apoptosis specifically. In particular, cells in culture have a high likelihood of undergoing necrosis (Elmore, 2007). Ideally if the experimental design permits, one assay can by employed to detect early initiation stages, whereas another can be used to visualise later execution stages of apoptosis. Pros and cons of different assays mean that some assays may be appropriate in one context, but not another (Watanabe et al., 2002). The kinetics of the system is also important to consider, as Caspases can be expressed over a narrow time window. Cells undergoing apoptosis in any context can fragment and disappear quickly, on the order of 2-3 hours. This further increases the chance of false negatives if the assay were to be conducted too early or late (Elmore, 2007).

Apoptosis detection can be made via: 1. Assessment of phenotypic nuclear fragmentation (e.g. visualised by DAPI), 2. Detection of fragmented DNA (TUNEL assay), and 3. Detection of activated Caspases (Cleaved Caspase-3 stain for activated executioner Caspase). Using a combination of these assays could be helpful for evaluating apoptosis within a developing tissue.

Cleaved Caspase-3

Upon cleavage activation, Aspartic acid 175 residue in Caspase-3 becomes accessible. This residue is specifically recognised by the Cleaved Caspase-3 antibody. Activated Caspase-3 is the predominant mediator of apoptotic DNA fragmentation (Wolf et al., 1999) into pieces as small as 200 bp (Williams et al., 1974).

Click chemistry-based TUNEL assay

DNA fragmentation is a characteristic hallmark of apoptosis. Terminal deoxynucleotidyl Transferase (TdT) dUTP Nick-End Labelling (TUNEL) is an established assay for detecting DNA fragments. The Click-iT Plus TUNEL Assay (Life Technologies C10618) harnesses click chemistry for improved sensitivity of DNA fragment labelling detection. The approach makes use of a modified dUTP called EdUTP that has the addition of a small alkyne moiety. This EdUTP can be readily incorporated at 3'-OH ends of fragmented DNA via the TdT enzyme. Signal detection is mediated by a click chemistry-mediated copper-catalysed covalent reaction between Alexa Fluor picolyl azide dye and the alkyne. The small alkyne moiety within the modified EdUTP enables more efficient incorporation by the TdT enzyme, thereby making this assay more sensitive than those that use other modified nucleotides. Another major advantage of this assay is the relatively mild conditions required for the reaction to take place, suitable for whole mount samples. The approach preserves cell morphology, permits other co-stains (e.g. Phalloidin), and allows retention of signal from fluorescent reporters like GFP for end stage analysis.

3.1.2 Cell death at peri-implantation stages

Peri-implantation stages of development span late pre-implantation E4.75 to early post-implantation E5.25 (Figure 14).

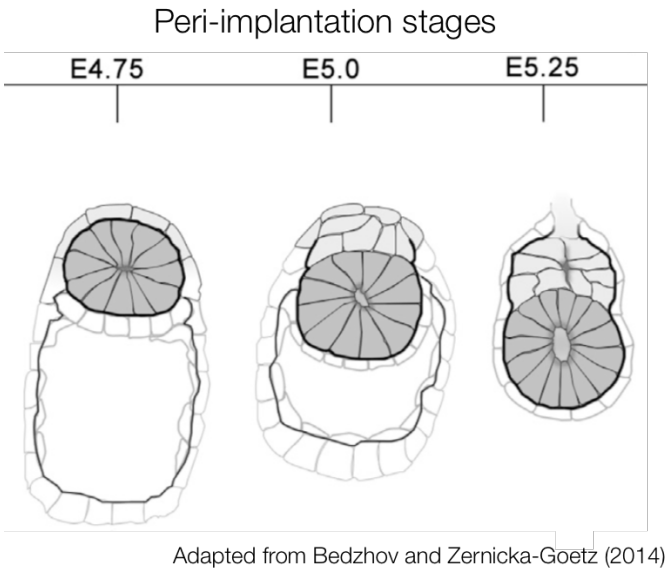


Figure 14. Peri-implantation stages of mouse development

Cell death during peri-implantation stages have been minimally characterised due to difficulty of embryo access at these stages, and the rapid succession of events that demand high numbers for stage-specific analysis. Bedzhov and Zernicka-Goetz, 2014 showed that peri-implantation cavitation of the epiblast occurs by a mechanism that is not dependent on apoptosis, but rather is mediated by an ECM-triggered polarisation signal to form the epiblast rosette. Apoptosis of the implanting ExE has not been characterised in the literature; it remains unclear the extent to which there is any apoptosis in the ExE.

In development, apoptosis is thought to be the default mechanism (Raff et al., 1993) that is constantly being repressed by survival factors. These can come from the ECM (Frisch and Francis, 1994), adjacent cells (Hermiston and Gordon, 1995), as well as occur in the form of growth signals. In the pre-implantation blastocyst, TE cells rarely undergo apoptosis (Brison and Schultz, 1998). The ICM on the other hand can eliminate cells that fail to translocate to the right position by apoptosis during the process of PrE and epiblast segregation (Plusa et al.,

2008). Upon implantation, interaction between the mTE and uterus results in apoptosis at the interface. This allows the embryo to invade the maternal uterine tissue. The process is mediated by activation of the extrinsic apoptotic pathway involving TNF1 activation that leads to Caspase-3 cleavage and irreversible apoptosis (Joswig et al., 2003; Parr et al., 1987). Concomitant with this, the maternal stroma undergoes rapid proliferation to encase the embryo and form the decidua (Nancy et al., 2012).

At post-implantation, the IGF pathway promotes embryo survival and proliferation via PI3K/AKT signalling that activates pro-survival mTOR (Yang and Guan, 2007). Depletion of oncogene MDM2 causes embryo lethality by massive cell elimination following implantation. MDM2 functions by sequestering P53 to prevent P53-mediated apoptosis in the absence of apoptotic signals. Loss of MDM2 triggers P53-mediated apoptosis without stimuli, resulting in embryonic lethality by E5.5. P53-null embryos are largely aphenotypic until E13.5 (Sah et al., 1995). Therefore, in early embryogenesis, P53-mediated apoptosis needs to be suppressed to prevent early embryo death, while its absence is inconsequential.

In this chapter, I examine the underlying cause for the observed ExE reduction in the *Ets2*-mutant.

3.1.3 ETS2 is a core regulator of ExE fate

The contents in this section were reviewed in detail in 1.3.2 to 1.3.4.

The mouse ExE gives rise to the definitive placenta, essential for supporting the embryo throughout pregnancy (Rossant and Cross, 2001). Proper ExE formation is crucial for establishing the feto-maternal interface. The ExE originates from the TE lineage of the pre-implantation blastocyst. At E4.75, the TE mediates implantation into the uterus and proliferates to give rise to the ExE concomitant with egg cylinder formation (Stephenson et al., 2012). ETS2 is a core transcription factor for patterning the ExE (Georgiades and Rossant 2006,

Donnison et al., 2015). *Ets2* is expressed in the ExE along with another ETS-family member *Elf5*. The two transcription factors were shown, by compound *in vivo* genetics, to cooperate synergistically to pattern the developing ExE (Donnison et al., 2015). *In vitro*, Ets2 is required for self-renewal of TSCs, a model of the extra-embryonic lineage (Wen et al., 2007; Odiatis and Georgiades, 2010). Furthermore, Ets2 but not Elf5, is one of the four factors sufficient to reprogram fibroblasts into iTSCs (Kubaczka et al., 2015). However, the extent to which these *in vitro* observations reflect the requirement for ETS2 *in vivo* during ExE development remains unclear. A targeted *Ets2*-loss-of-function mouse mutant with DNA-binding domain deleted (*Ets2* ^{db1}-mutant) was previously found to be lethal by E9.0 due to placentation failure (Yamamato et al., 1998). At post-implantation, this mutant was shown to have a drastically reduced ExE that lacks core signature marker expression by ISH, and impairment of A-P embryonic axis development (Georgiades and Rossant 2006). However, the cause of ExE reduction was not explored. I therefore aimed to: 1) Describe the reduced mutant ExE in higher resolution by confocal immunofluorescence; 2) Quantitatively characterise mutant ExE tissue growth and death rates; and 3) Pinpoint the origin of the diminished mutant ExE phenotype that is apparent at post-implantation.

3.2 Results

3.2.1 ETS2 is expressed in the ExE and its progenitors

To investigate when and where ETS2 may be present and playing a role, I conducted immunofluorescent characterisation at pre- and post-implantation stages. I used an antibody against the DNA-binding domain of ETS2 such that mutants lacking this domain could serve as a negative control. ETS2 expression in the morula was weak and diffuse, lacking evident nuclear signal (Figure 15A). Evaluation of *Ets2*-null morulae would help clarify whether the observed signal is real. ETS2 protein was detectable with nuclear localisation in the pre-

implantation early blastocyst (Figure 15B). Progressing through pre-implantation development, ETS2 becomes increasingly upregulated and distinctly nuclear in the TE lineage (co-stained with CDX2) as compared to other lineages. This was apparent by the late blastocyst stage (Figure 15C). At post-implantation, ETS2 was highly expressed in the ExE (Figure 15D). ETS2 signal was not detectable in *Ets2*-null negative controls at post-implantation when CDX2-positive ExE is markedly reduced (Figure 15E).

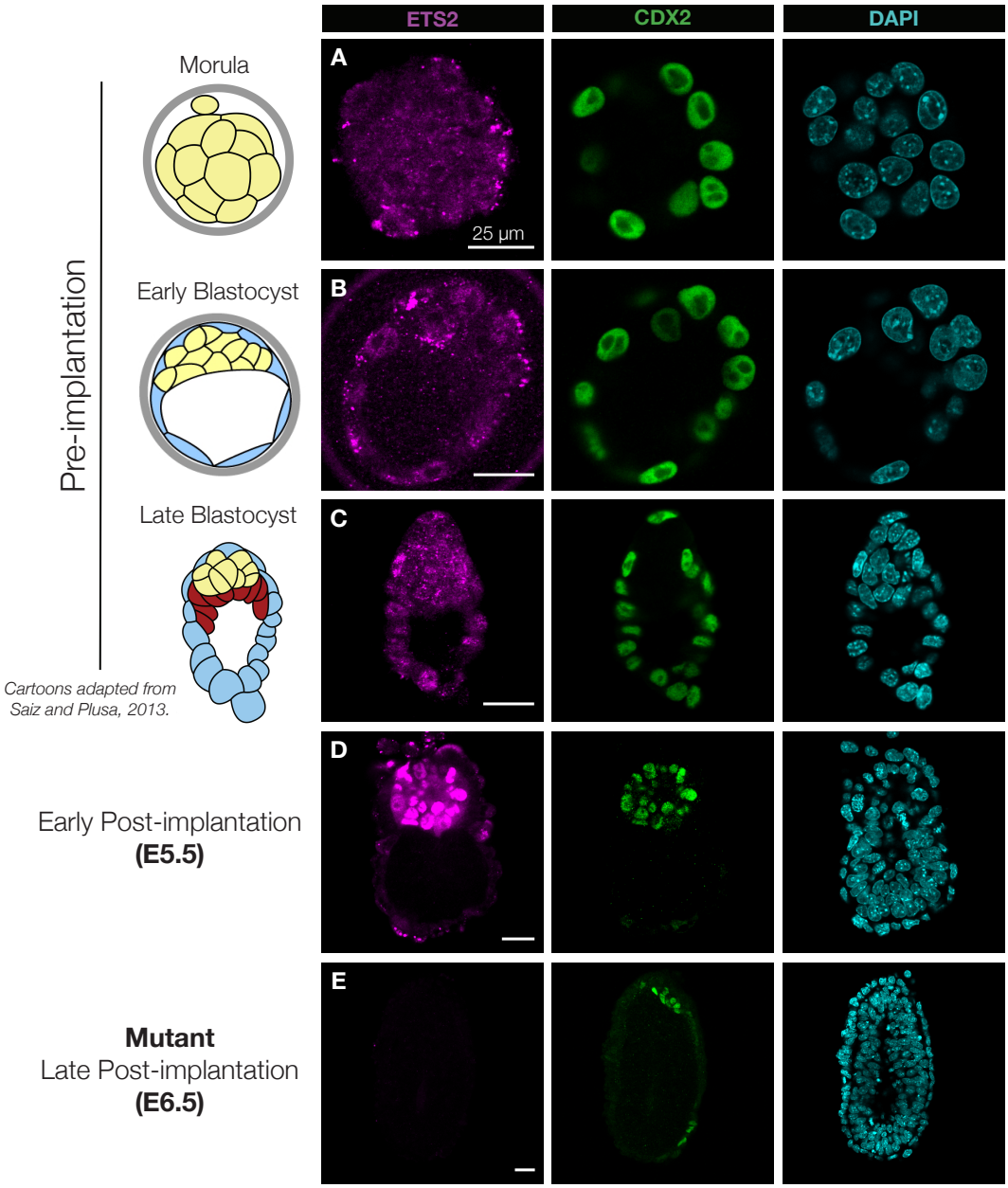


Figure 15. ETS2 expression from pre- to post-implantation mouse embryogenesis
Representative mid-sagittal sections of whole mount immunofluorescent stains of ETS2 (magenta); pre-implantation TE lineage and post-implantation ExE marker CDX2 (green); DAPI to mark nuclei (cyan).
(A) Morula (~E3.0). (B) Early blastocyst (~E3.5). (C) Late blastocyst (E4.5). (D) Early post-implantation (E5.5). (E) Late post-implantation (E6.5). Scale bars: 25 µm.

3.2.2 Mutants maintained on a mixed genetic background displayed a milder ExE phenotype compared to pure 129S6/SvEvTac by morphometric analysis

Georgiades and Rossant, 2006 reported two classes of mutant phenotypes in their *Ets2*-strain maintained on the outbred ICR genetic background. Type 1 mutants had a more severe ExE phenotype: i.e. markedly reduced ExE, lacking marker gene expression upon inspection of 5 independent ExE markers by ISH – *Cdx2*, *Bmp4*, *Spc4*, *Esrrβ*, *Sox2* – and was found to occur at 11.3% frequency at E6.75. A small amount of 'ExE tissue' was still found to be present between the epiblast (visualised by *Oct4*) and EPC (visualised by *Mash2*), which did not express ExE-specific marker genes. Type 2 mutants, occurring at 8.2% frequency, had a morphologically normal ExE by visual inspection and retained ExE signature gene expression (same markers tested as above), although morphometric analysis was not conducted.

Here, I maintained the *Ets2* mutation on both pure and mixed genetic backgrounds. For the pure breeding strategy, *Ets2* $^{db1/+}$ male or female (imported as 50% FVB/N and 50% 129S6/SvEvTac – referred to as 129) was crossed with pure 129/*Ets2* $^{+/+}$ female or male, respectively. These were maintained by backcrossing into the pure 129 background for 8 generations (as at January 2019). A mixed background *Ets2*-mutant line was also made since the *Hhex-GFP* reporter (Rodriguez et al., 2001), maintained as a mixture of CBA/J and C57BL/6J, was desirable for fluorescently labelling the AVE to aid analysis of the mutant AVE phenotype. This strain was generated by the following strategy: 1. *Ets2* $^{db1/+}$ male was crossed with *Hhex-GFP*-homozygous female; 2. *Ets2* $^{db1/+}$/*Hhex-GFP*-heterozygous female offspring (with relative copy number of *Hhex-GFP* genomic transgene determined by qPCR analysis) were subsequently crossed with *Ets2* $^{+/+}$/*Hhex-GFP*-homozygous males; 3. Brother-sister crosses of *Ets2* $^{db1/+}$/*Hhex-GFP*-homozygous and *Ets2* $^{+/+}$/*Hhex-GFP*-homozygous offspring were then conducted to maintain the line (for 4 generations as at January 2019). The 129 percentage of the mixed background mice used for ExE characterisation was either 12.5% or

6.25% with the remaining genetic background composition being a mix of CBA/J and C57BL/6 from the *Hhex-GFP*-strain.

Because the type 1 mutant reduced ExE phenotype reported by Georgiades and Rossant, 2006 was evident by E6.75, the defect origins would have occurred prior to that stage. Therefore, I conducted my initial characterisation at both E5.5 and E6.5 to assess mutant ExE defect progression across post-implantation.

Quantitative comparison of the ExE in mixed versus pure genetic background

In my characterisation, I did not observe type 1 and type 2 mutants based on the ExE phenotype. ExE mutant phenotypic severity was consistent within a genetic background, and mutants occurred at expected 25% Mendelian frequency. However, the ExE phenotype appeared to vary in severity depending on genetic background. I therefore quantitatively characterised these differences to assess: 1) Whether they were statistically significant amongst strains and between mutant versus control within strains; 2) Which strain to use for detailed characterisation and functional studies; and 3) What the implications of ExE phenotypic differences were for the growth and patterning of the rest of the embryo in the *Ets2*-mutant.

To investigate whether different genetic backgrounds vary in severity of the ExE phenotype, I compared tissue morphometrics across backgrounds and post-implantation stages. To do so, I generated mutant embryos through *Ets2*$^{db1/+}$ x *Ets2*$^{db1/+}$ crosses, and used wild type (WT) and heterozygous littermates as controls. I stained for F-ACTIN to visualise cell and tissue morphology, and conducted whole mount confocal microscopy from which I analysed mid-sagittal sections for morphometric measurements. Specifically, I took the longest distance for each morphometric measure such that a consistent reading could be made across embryos (method detailed in 2.9). I quantified the following metrics at both E5.5 and E6.5: ExE length, epiblast length, and epiblast A-P width. The latter two metrics were used to obtain a normalised

measure for each embryo, to correct for variability in embryo size. 'Etss' refers to pure while 'Etex' refers to mixed genetic background. Hierarchical ANOVA and Tukey statistical tests were used to account for multiple measurements taken per embryo and comparison across multiple groups (i.e. genotypes and genetic backgrounds), respectively. The differences within backgrounds, or differences that are of note to highlight across backgrounds, are in bold (Figure 16).

At E5.5, ExE length was equally reduced in pure and mixed genetic backgrounds. Epiblast length was unchanged in WT and mutant. ExE length normalised to that of the epiblast was significantly reduced in mutants of both genetic backgrounds. For epiblast width, the mixed background had a decreasing trend in the mutant, but was not significant, whereas pure background mutant epiblast width was comparable to control. ExE length normalised to epiblast width was also significantly reduced in both backgrounds, but to a lesser degree in the mixed background (Figure 16).

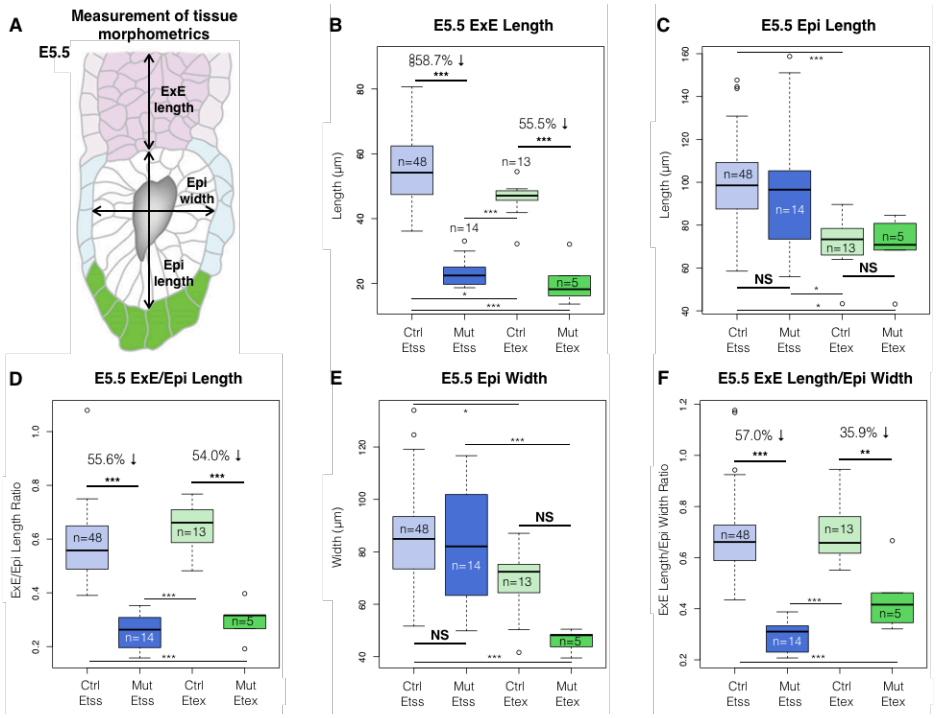


Figure 16. E5.5 morphometric comparison of control versus mutant in pure (Etss) versus mixed (Etex) genetic backgrounds

(A) Schematic of how tissue morphometric measurements were conducted (adapted from Stower and Srinivas, 2014). Morphometrics of (B) ExE length, (C) Epi length, (D) ExE length normalised to Epi length, (E) Epi width, (F) ExE length normalised to Epi width. (Statistical tests: hierarchical ANOVA and Tukey. NS = not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

At E6.5, ExE length was markedly reduced in the pure background, but was decreased to a lesser extent in the mixed background. Notably, the mixed background mutant ExE was significantly greater in length than the pure, indicating a milder, but still significant, reduced ExE phenotype in the mixed background. E6.5 epiblast length was normal in mutants of both genetic backgrounds. Again, mutant ExE length normalised to epiblast length in the mixed background was significantly greater than the pure. E6.5 A-P epiblast width was significantly reduced for both genetic backgrounds, suggesting that the defect in epiblast tissue morphometrics was consistent regardless of background (Figure 17).

Interestingly, analysis of E6.5 ExE length normalised to epiblast width per embryo revealed an insignificant difference between mixed background mutant versus control, while there was a significant decrease in pure background mutant versus control. ExE length normalised to epiblast width in the mixed background was significantly greater than that of the pure (Figure 17).

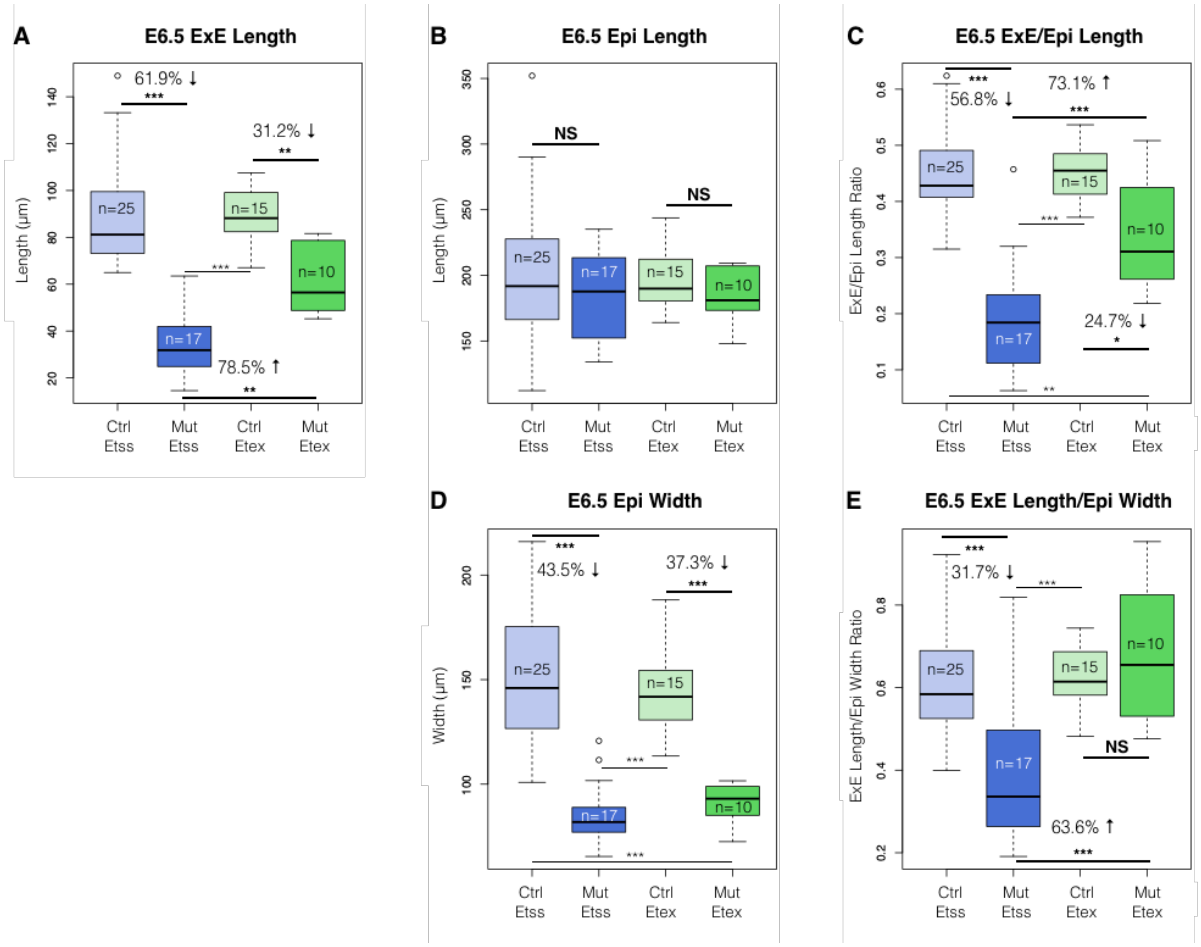


Figure 17. E6.5 morphometric comparison of control versus mutant in pure (Etss) versus mixed (Etex) genetic backgrounds

(A) ExE length. (B) Epi length. (C) ExE length normalised to Epi length. (D) Epi width. (E) ExE length normalised to Epi width. (Statistical tests: Hierarchical ANOVA and Tukey. NS = not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

Overall, detailed comparison of tissue morphometrics revealed a milder ExE phenotype in the mixed background as compared to the pure, while the decrease in A-P epiblast width was consistent in both genetic backgrounds. The latter may suggest background-independent defects in the A-P axis, which is elaborated on in Chapter 4. I decided to conduct all subsequent ExE characterisation using the pure genetic background to investigate the cause of the severe ExE reduction and how it affects embryo patterning.

3.2.3 Post-implantation mutant ExE is significantly reduced in size, yet retains signature marker gene expression

Similar to what was reported in Georgiades and Rossant, 2006 for type 1 mutants, I found the mutant ExE to be visibly smaller compared to control. Despite being reduced in size, the mutant ExE retained its signature marker expression CDX2, EOMES, and ESRRB by immunofluorescence (Figure 18). In contrast, these markers, as well as *Bmp4*, *Spc4*, and *Sox2*, were not detectable in type 1 mutants by ISH (Georgiades and Rossant, 2006).

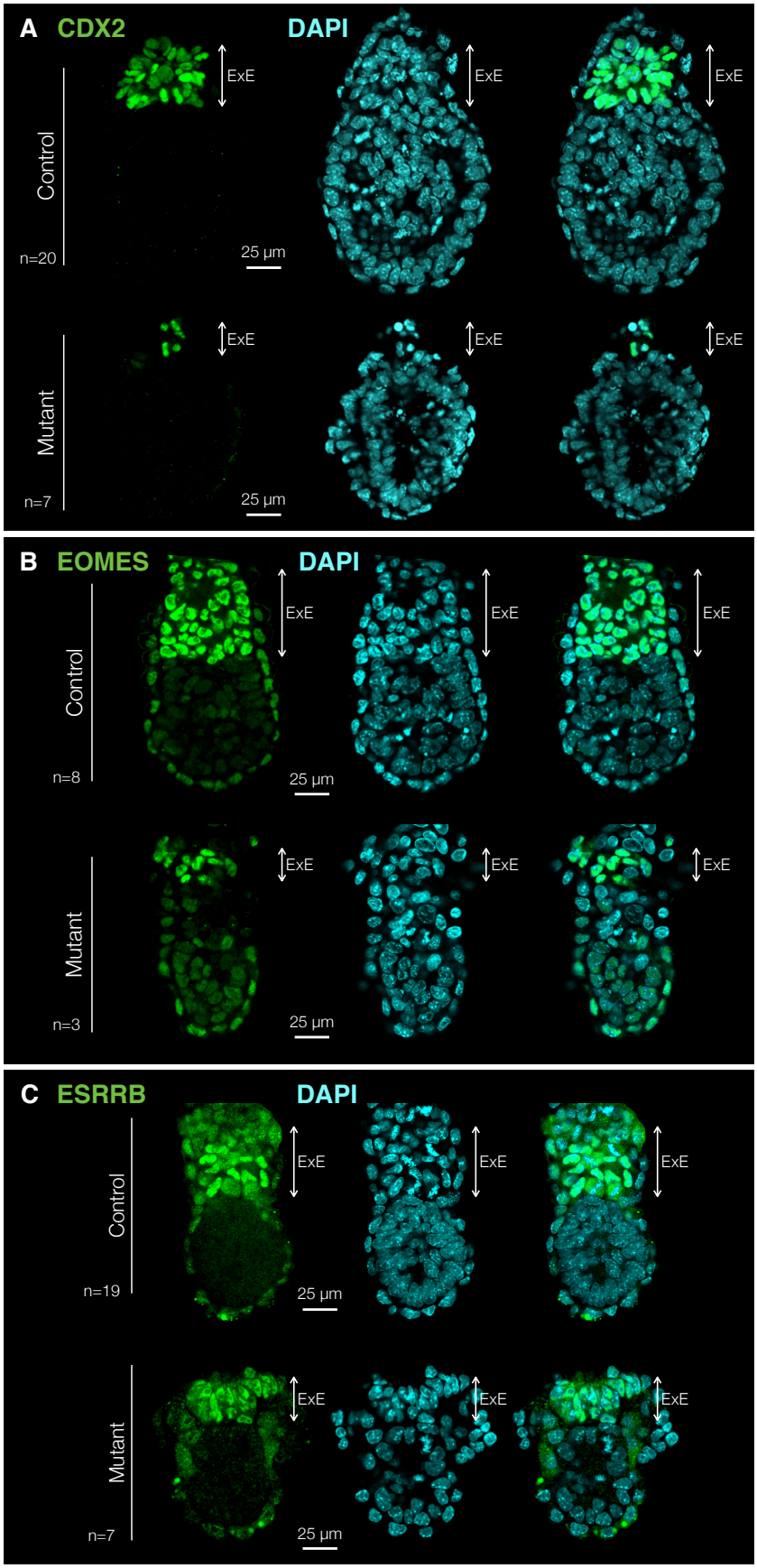


Figure 18. Reduced mutant ExE retains signature marker gene expression
Immunofluorescent characterisation of ExE signature markers at early post-implantation (E5.5). (A-C) CDX2, EOMES, ESRRB (green); DAPI (cyan). Scale bars: 25 µm.

3.2.4 Reduced mutant ExE is sufficient for normal epiblast marker gene expression

Given the known roles of the ExE in patterning the epiblast and establishing proximal-distal morphogen gradients (Ang and Constam, 2004), which in turn pattern the A-P axis (Arnold et al., 2008), I examined whether mutants had defective epiblast by analysis of its signature marker OCT4. In section 3.2.2, I identified that the reduced ExE had no effect on epiblast length morphometrics, but may contribute to differences in late post-implantation epiblast A-P width. Here, I characterised mutant epiblast identity by staining for signature marker OCT4. Mutant OCT4 in the epiblast was normal in expression and localisation at post-implantation stages, and respected compartment boundaries with the ExE (visualised by CDX2 co-stain) (Figure 19).

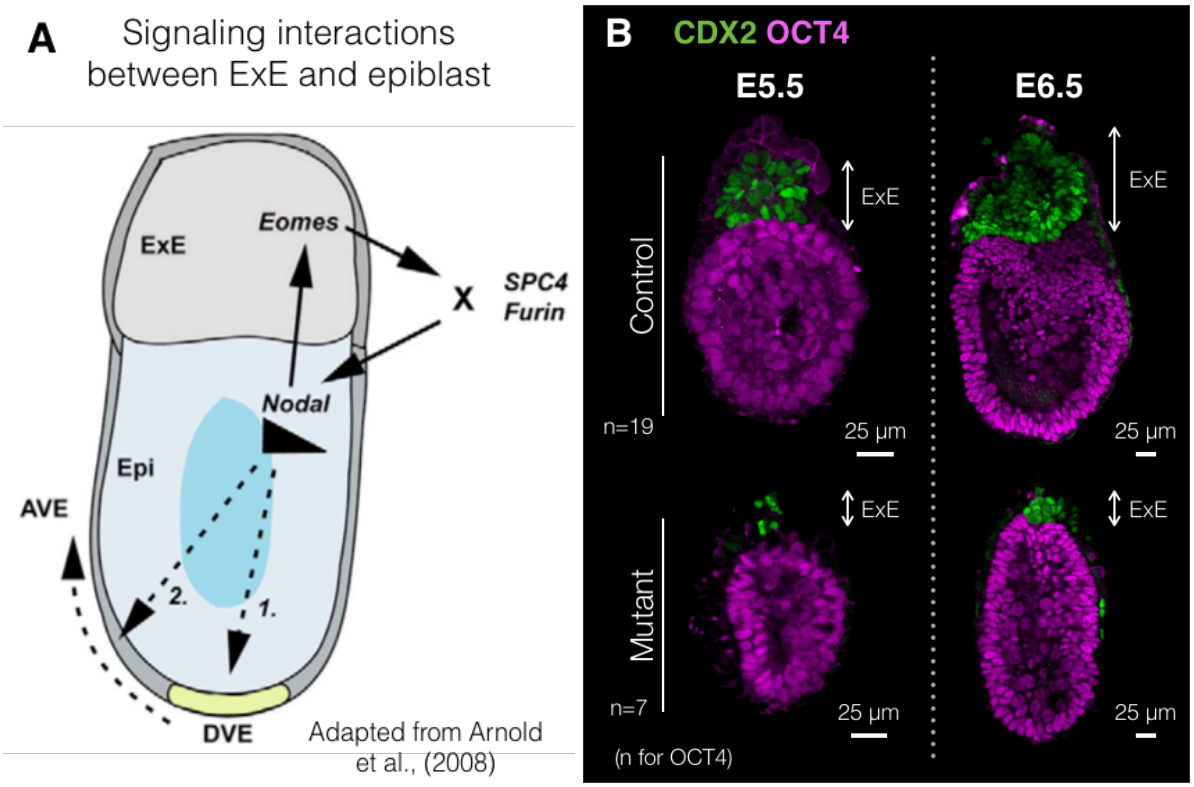


Figure 19. Normal mutant epiblast marker gene expression by immunofluorescence
(A) Known signalling interactions between the ExE and epiblast, which affect AVE migration (adapted from Arnold et al., 2008).
(B) Representative section from whole mount immunofluorescence of OCT4 (epiblast – magenta) and CDX2 (ExE – green) in control versus mutant at E5.5 and E6.5. Scale bars: 25 µm.

3.2.5 Mutant epiblast fails to induce BRACHYURY-positive mesoderm at late post-implantation

To gauge whether the mutant epiblast affects normal development at late post-implantation, I checked whether mutants undergo gastrulation by staining for mesoderm marker BRACHYURY (T) at E6.5 (Figure 20) and E7.0 (Figure 21). I found absence of T-positive mesoderm in mutants while localised T expression was present in littermate controls. Posterior VE CDX2 expression was present in control as expected (Rayon et al. 2016), and CDX2 expression in the VE was localised to one side of the embryo in mutants (although it is difficult to ascertain whether this is indeed posterior VE expression in the absence of an anterior or posterior marker). Whether the inability to induce T-positive mesoderm in mutants is due to defective AVE migration is investigated in Chapter 4.

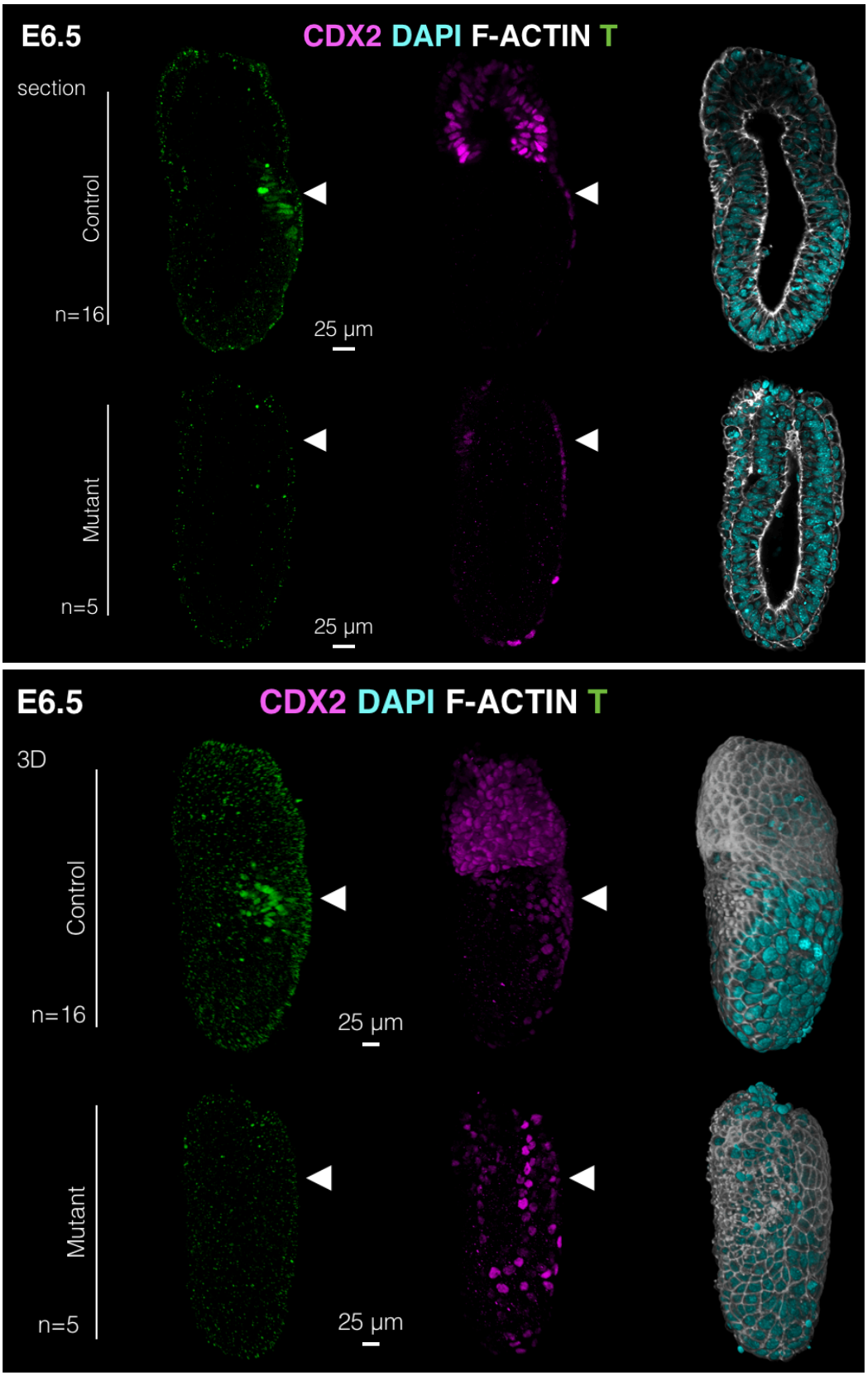


Figure 20. Absence of T induction in mutants at E6.5
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescent images at E6.5 of CDX2 marking ExE (magenta) with presence of relatively lower levels of posterior VE nuclear signal in control and possibly mutant. Absence of T-positive-mesoderm (green) in mutants while control litter mates displayed T induction. White arrowhead – (presumptive) posterior; F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

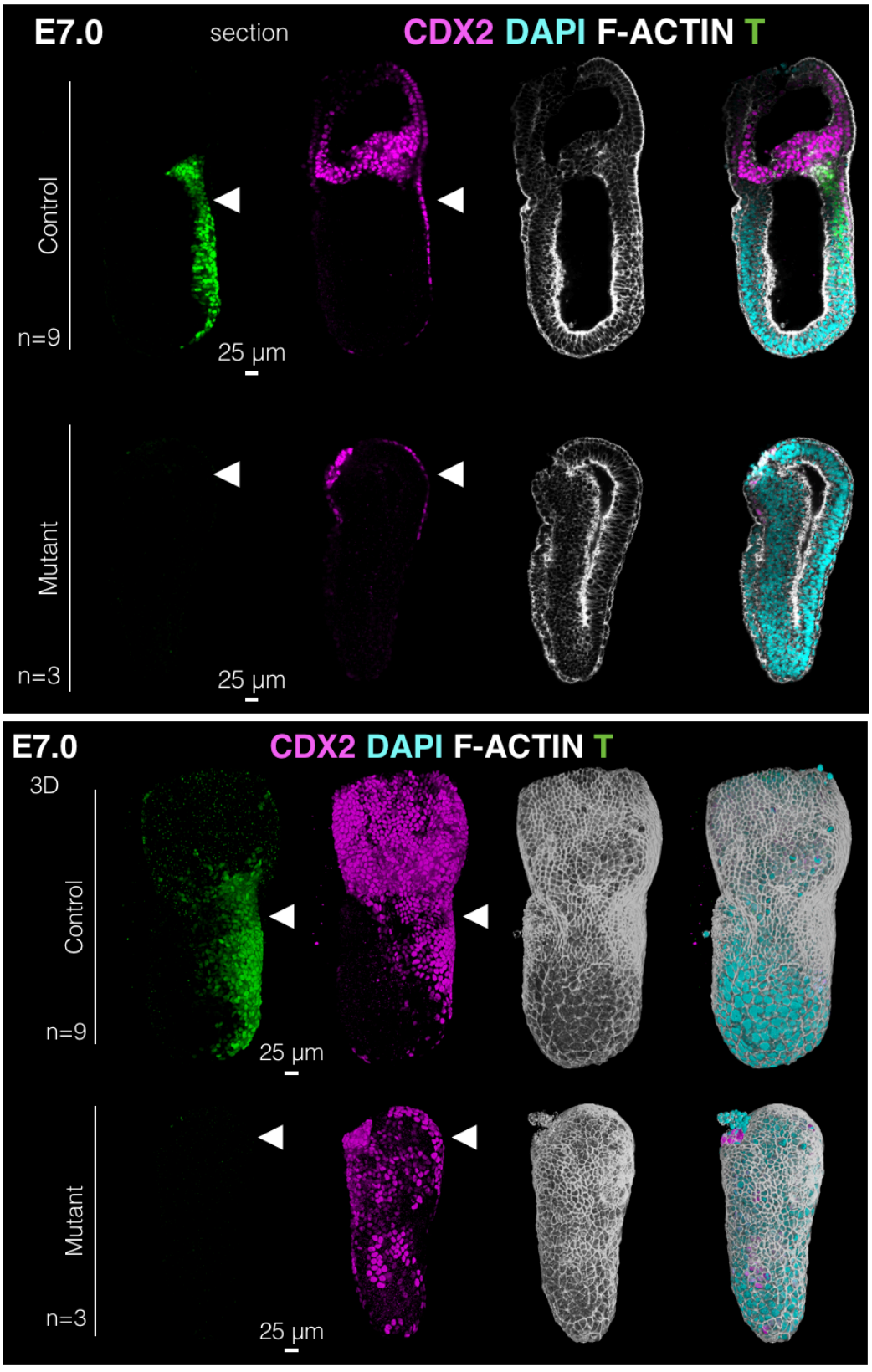


Figure 21. Absence of T induction in mutants at E7.0
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescent images at E7.0 of CDX2 marking ExE (magenta) with presence of relatively lower levels of posterior VE nuclear signal in control and possibly mutant. Absence of T-positive-mesoderm (green) in mutants while control litter mates displayed T induction. White arrowhead – (presumptive) posterior; F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

3.2.6 Mutant ExE proliferation is normal from early to late post-implantation

To investigate whether the decrease in ExE length is due to aberrant proliferation, I used phospho-Histone H3 (PHH3) to mark cells undergoing active mitosis (Kim et al., 2017) (Figure 22A). I visualised proliferative cells in mutant versus littermate control embryos by conducting whole mount confocal immunofluorescence at E5.0, E5.5, and E6.5, where I captured the entire 3D volume of each embryo and stained for: PHH3, DAPI to visualise mitotic nuclei, and F-ACTIN to distinguish cell outlines (stain not shown). I quantified proliferative index by taking the number of PHH3-positive ExE cells divided by total number of ExE cells identified using DAPI and F-ACTIN to demarcate individual cells and cell/tissue morphology, respectively. I detected no difference in proliferation between control versus mutant at peri- and post-implantation stages (Figure 22B). As expected in normal development, peri-implantation (E5.0) ExE proliferation – a period of rapid embryo growth as the ExE becomes established from pTE progenitors (Saiz and Plusa, 2013) – was significantly higher than post-implantation (E5.5) in both control and mutant.

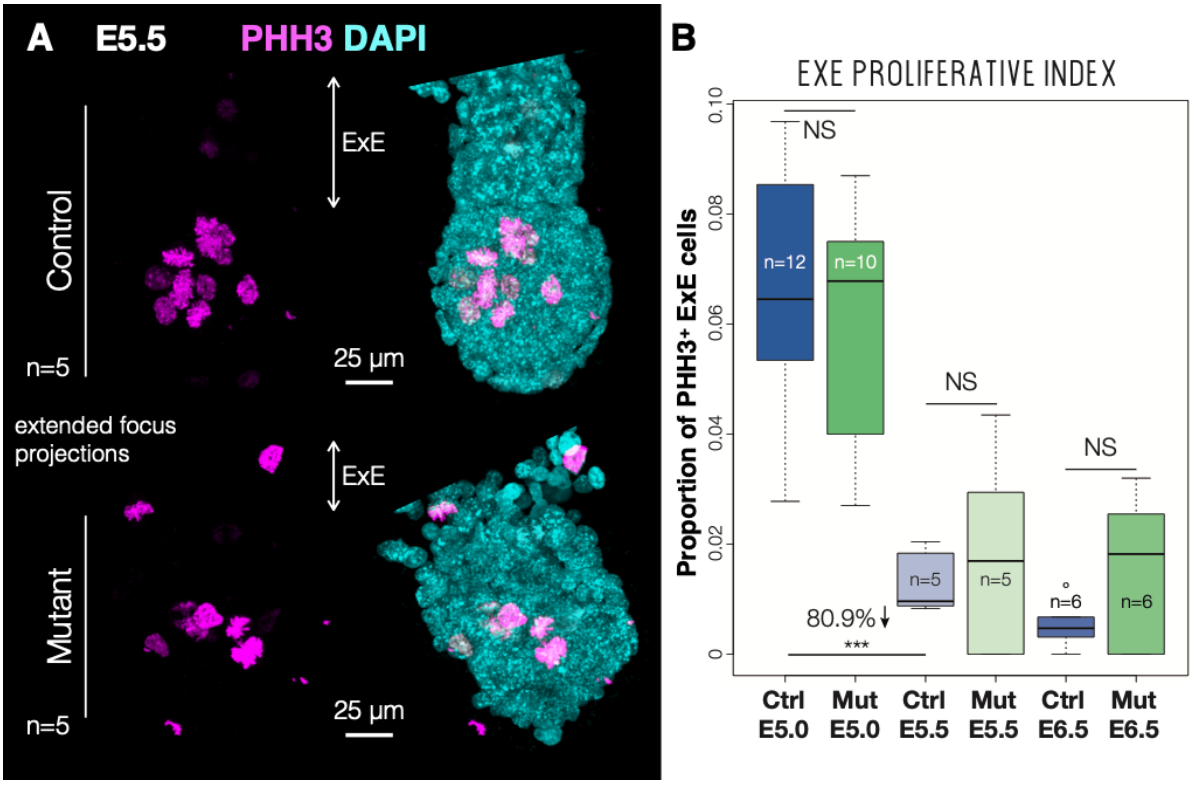


Figure 22. Proliferation is unperturbed in the mutant ExE at peri- and post-implantation
(A) Extended focus projections of representative E5.5 control versus mutant embryos showing no visible differences in proliferation (PHH3 – magenta; DAPI – cyan). The entire whole mount image stack was considered for the analysis.
(B) Quantification of proliferative index at E5.0 (peri-implantation), E5.5 (early post-implantation), and E6.5 (late post-implantation). (Statistical tests: hierarchical ANOVA and Tukey; NS = not significant; *** $p < 0.001$). Scale bars: 25 µm.

3.2.7 Normal mutant ExE apoptosis at post-implantation

As absence of differential proliferation could not explain the diminished ExE phenotype, I next looked at apoptosis. To test whether the post-implantation mutant ExE has aberrant death rates, I conducted whole mount immunofluorescence in control versus mutant embryos staining for CC3, DAPI, and F-ACTIN. I quantified ExE apoptotic index by taking the number of CC3-positive ExE cells divided by total number of ExE cells at E5.5 and E6.5. I detected no significant difference in post-implantation ExE apoptosis in the mutant ExE (Figure 23). This and previous section 3.2.6 suggested that the reduced post-implantation mutant ExE has normal cell growth and death rates.

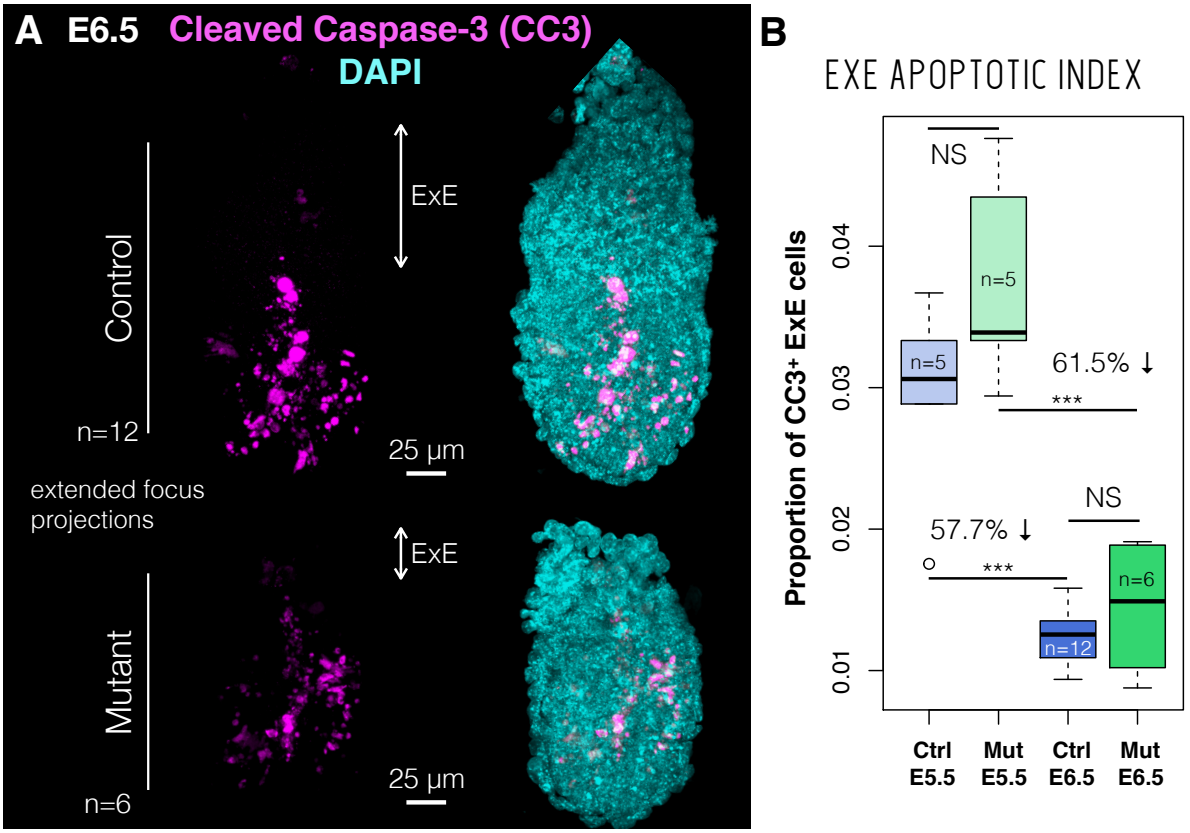


Figure 23. Apoptosis is normal in the mutant ExE at post-implantation stages
(A) Extended focus projections of representative post-implantation (E6.5) control versus mutant embryos where there was no difference in ExE apoptosis (CC3 – magenta; DAPI – cyan).
(B) Quantification of apoptotic index at E5.5 (early post-implantation), and E6.5 (late post-implantation). (Statistical tests: hierarchical ANOVA and Tukey. NS = not significant; *** $p < 0.001$). Scale bars: 25 µm.

3.2.8 Peri-implantation mutant ExE is comparable to control

Since the mutant ExE was already reduced at post-implantation, I next analysed the earlier peri-implantation stage E5.0 to trace the origin of the defect. Quantification of ExE cell number revealed unchanged metrics between mutant versus control at E5.0, whereas ExE tissue was markedly reduced in the E5.5 mutant compared to E5.0 mutant (Figure 24B). Furthermore, ExE morphometric analysis demonstrated normal ExE length in mutants at E5.0, but markedly reduced mutant ExE at E5.5 (Figure 24C). This was also the case for ExE length normalised to that of the epiblast (Figure 24E), where epiblast length was normal at peri- and post-implantation stages (Figure 24D). Together, these data suggested that mutant ExE tissue disappearance occurred between E5.0 and E5.5, i.e. the peri- to post-implantation time window.

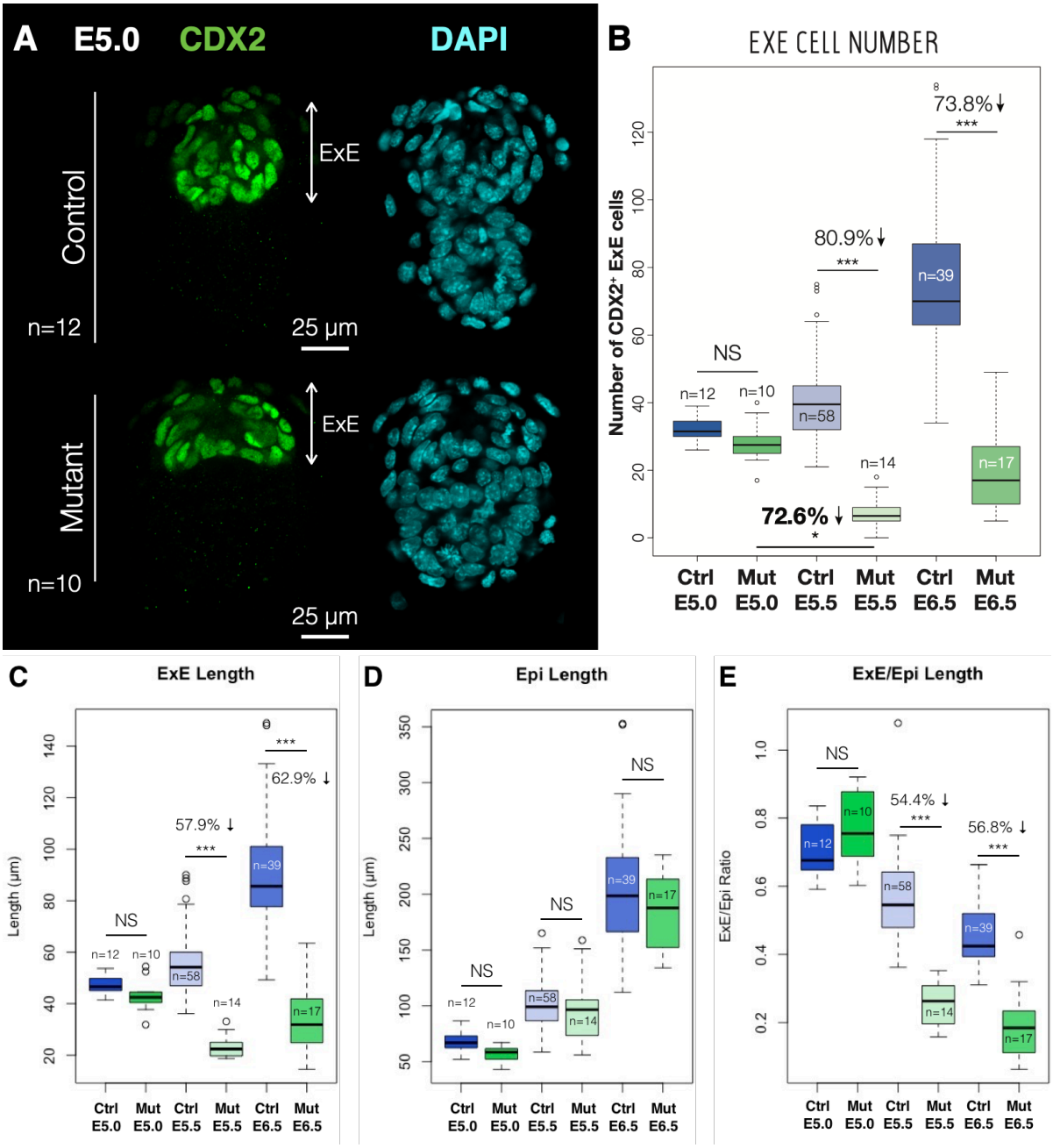


Figure 24. Mutant ExE is normal at peri-implantation

(A) Representative mid-sagittal section from whole mount immunofluorescence of CDX2 (ExE – green) and DAPI (nuclei – cyan) in control versus mutant at E5.0.

(B) Quantification of CDX2-positive ExE cell number from peri- to post-implantation highlighting a significant reduction in mutant ExE cell number from E5.0 to E5.5.

(C-E) Morphometric measurements of ExE length, Epi (epiblast) length, and ExE/Epi length at peri- (E5.0) and post-implantation (E5.5, E6.5).

(Statistical tests: hierarchical ANOVA and Tukey. NS = not significant; * $p < 0.05$; *** $p < 0.001$). Scale bars: 25 µm.

3.2.9 Increased apoptosis in the developing ExE at peri-implantation

As mutant ExE tissue diminishes from peri- to post-implantation and there were no differences in the proliferative index of the peri- and post-implantation mutant ExE, I next assessed whether there was elevated apoptosis. To do so, I quantified the apoptotic index at early and late peri-implantation using two assays to detect different stages of the apoptotic cellular process. At E4.75 early peri-implantation, I used CC3, an indicator of activated executioner Caspase. This event would precede the cleavage of DNA, which executioner Caspases act on (reviewed in section 3.1.1). I quantified apoptotic index in the pTE (ExE progenitors) by taking the number of CC3-positive pTE cells divided by total number of pTE cells and found a significant increase in CC3 apoptotic index of the mutant pTE (Figure 25). At later peri-implantation stage E5.0, I used the Click-TUNEL assay as a sensitive approach to detect and visualise DNA fragmentation downstream of Caspase-3 cleavage activation. I quantified apoptotic index in the E5.0 ExE by taking the number of Click-TUNEL-positive ExE cells divided the total number of ExE cells and found a significant increase in the mutant (Figure 26). Collectively, these data show that there is elevated apoptosis in the peri-implantation mutant ExE.

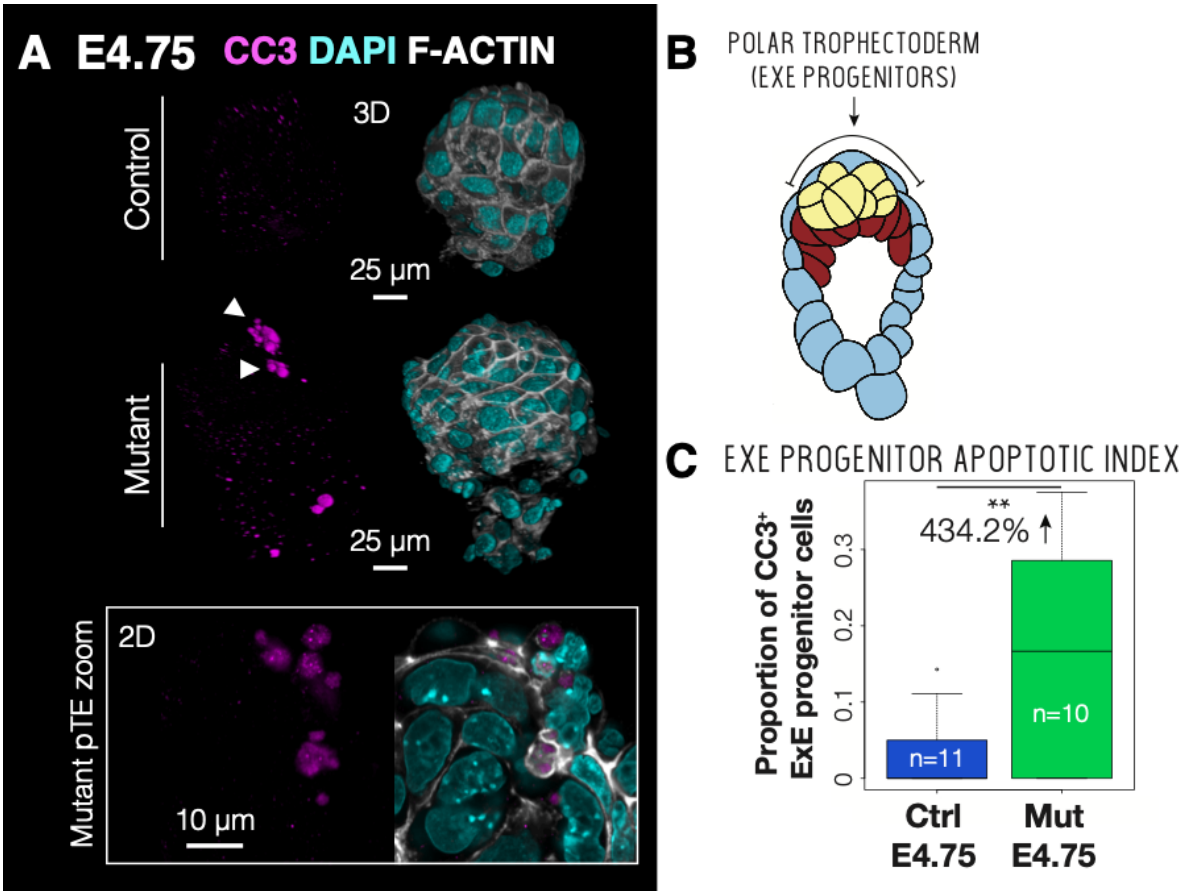


Figure 25. Increased apoptosis in the mutant ExE at peri-implantation stage E4.75

(A) 3D opacity rendering of mutant E4.75 early peri-implantation control versus mutant stained for CC3 (magenta), F-ACTIN (gray), DAPI (cyan). Zoom-in of a section of mutant pTE cells displaying hallmarks of apoptosis, i.e. fragmented nuclei overlaid with CC3 signal. Scale bars: 25 µm, or 10 µm as indicated.

(B) Schematic of E4.75 embryo showing location of the pTE (ExE progenitors; adapted from Saiz and Plusa, 2013).

(C) Quantification of CC3-apoptotic index in the pTE of control versus mutant showing significantly increased apoptosis in mutant. (Statistical test: hierarchical ANOVA; *** $p < 0.001$).

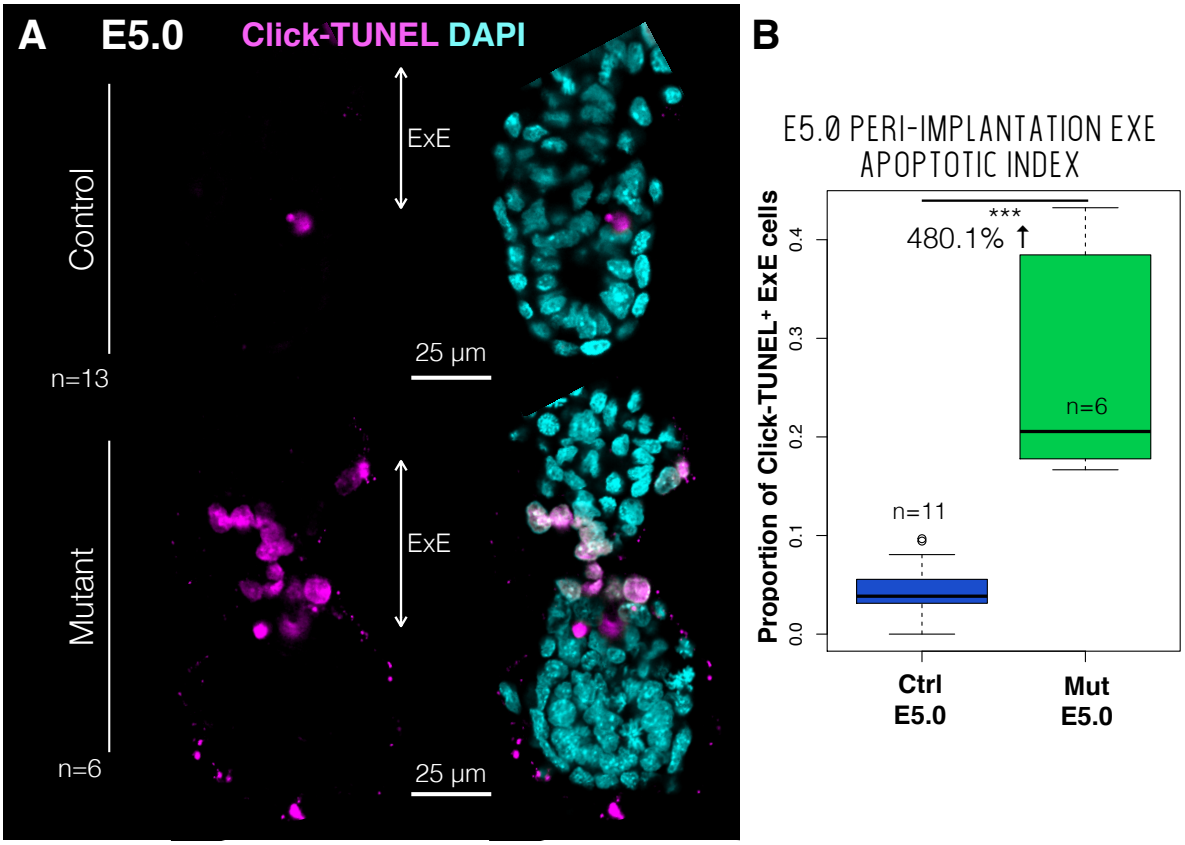


Figure 26. Increased apoptosis in the mutant ExE at peri-implantation stage E5.0
(A) Representative mid-sagittal section from whole mount immunofluorescence of E5.0 control versus mutant stained with the Click-TUNEL assay to detect fragmented nuclei (magenta); F-ACTIN (gray); and DAPI (cyan). Scale bars: 25 µm.
(B) Quantification of Click-TUNEL-apoptotic index in peri-implantation ExE in control versus mutant showing significantly elevated apoptosis in mutant. (Statistical test: hierarchical ANOVA; *** $p < 0.001$).

3.2.10 Mutant ExE phenotypic rescue by apoptosis inhibition from peri- to post-implantation

To validate whether apoptosis is indeed the causal mechanism for ExE tissue disappearance in the *Ets2*-mutant, I next tested whether I could rescue the mutant ExE phenotype by treating mutants with an apoptosis inhibitor from peri- to post-implantation. I used the Pan-Caspase Inhibitor Z-VAD-FMK (R&D systems; also referred to as 'FMK'), a synthetic inhibitor that works to block the Caspase family of cysteine proteases, which act in both the initiation and execution phases of apoptosis (Taylor et al., 2008). I used the inhibitor at the lower end of the recommended concentration (as peri-implantation embryos are relatively small): 10 µM versus 0.05% DMSO control from E5.0 to E5.5. I subsequently

quantified ExE cell number using CDX2 to mark ExE cells. Quantifications were conducted blinded to genotype and condition. Upon FMK treatment, I found a significant recovery of ExE cells to near control levels (Figure 27). I tested a titration curve with this inhibitor where 5 μM gave no effect and 25 μM gave an effect, but with excessive cell death in the DMSO-only condition. DMSO control mutants were similar to non-cultured mutants with very few ExE cells present. Together, these data suggest that increased apoptosis is the mechanism for *Ets2*-mutant ExE tissue disappearance. However, close inspection of the data suggest that the rescue is not complete, as numbers of CDX2-positive cells have not been restored to match control counterparts. A large proportion of the more proximal ExE in mutants treated with FMK Pan-Caspase inhibitor is CDX2-low while control proximal ExE is CDX2-high (Figure 27A – arrowhead).

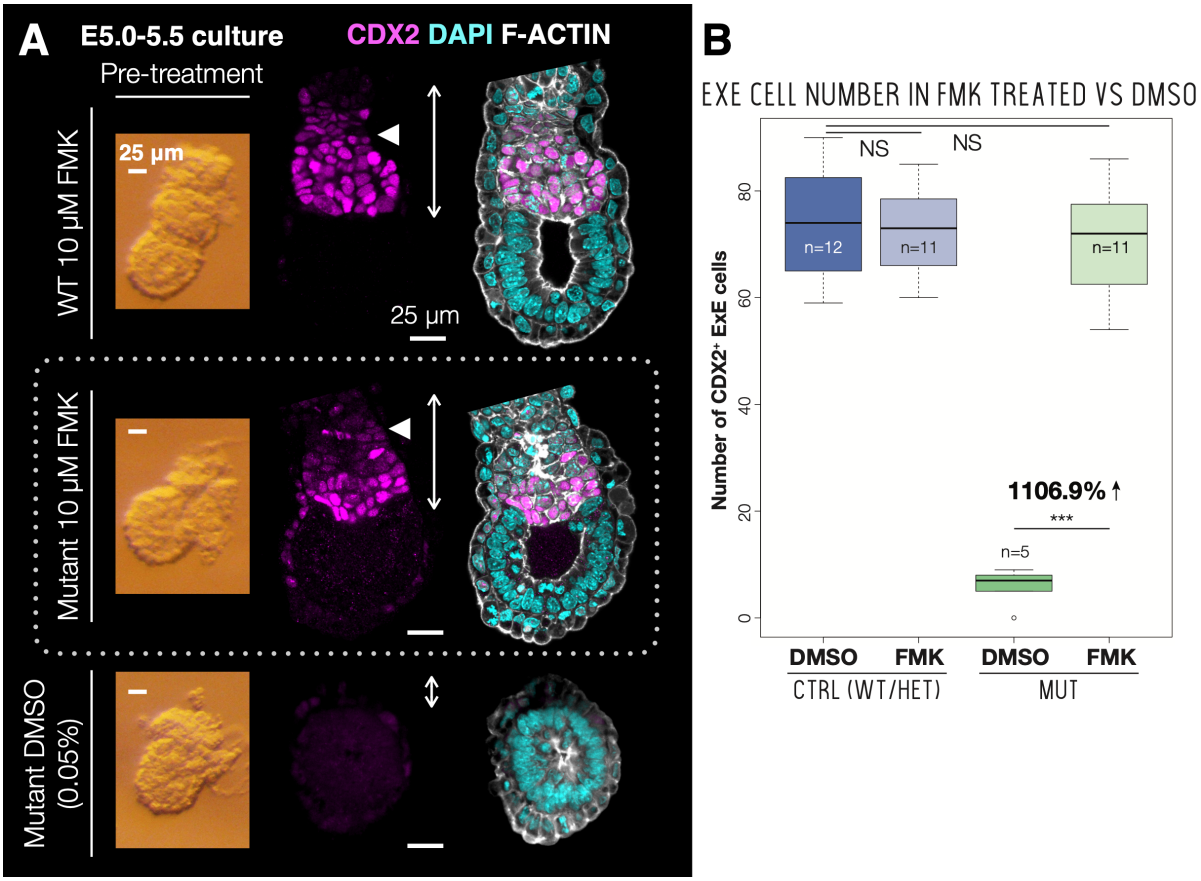


Figure 27. Mutant ExE phenotypic rescue by apoptosis inhibition from peri- to post-implantation
(A) Pre-treatment bright field image, and representative mid-sagittal sections of post-treatment embryos stained for CDX2 (ExE – magenta); F-ACTIN (gray); and DAPI (cyan). Conditions presented from top to bottom are: WT treated with 10 µM FMK, mutant treated with 10 µM FMK, and mutant treated with 0.05% DMSO control (which shows absence of ExE rescue). Arrowhead points at proximal ExE in WT versus mutant 10 µM FMK-treated. Scale bars: 25 µm.
(B) Quantification of ExE cell number in control versus mutant in FMK and DMSO-treated, showing recovery of ExE cells in FMK-treated mutants to near control levels. (Statistical tests: hierarchical ANOVA and Tukey; NS = not significant; *** $p < 0.001$).

3.3 Discussion

This Chapter provides detailed characterisation of the *Ets2*-mutant reduced ExE phenotype and uncovers its developmental origins. The severity of the ExE defect varied depending on genetic background with the mixed background (with *Hhex-GFP* reporter) giving a milder phenotype as compared to pure 129 upon tissue morphometric analysis. This could be due to compensatory mechanisms present in the mixed background. I conducted subsequent immunofluorescence and detailed tissue growth and death rate characterisation using the pure 129 *Ets2*-mutant strain to investigate the basis for its more severe ExE phenotype.

Georgiades and Rossant, 2006 did not detect pre-implantation *Ets2* expression by ISH, whereas I did by immunofluorescence. This could be due to differences in mRNA versus protein expression, or that ISH and immunofluorescence have different sensitivities. Georgiades and Rossant, 2006 also did not find expression of ExE markers in the reduced ExE of their type 1 mutants at post-implantation. Here I report that the reduced mutant ExE retains signature marker gene expression, shown by immunofluorescence of ExE marker genes CDX2, EOMES, and ESRRB. This suggests that, although diminished, the mutant ExE at post-implantation is normal in its identity. I quantitatively assessed the ExE tissue size reduction by analysis of tissue morphometrics. Both ExE length, as well as that normalised to epiblast length was markedly reduced in mutants at E5.5 and E6.5. The reduced mutant ExE, however, was sufficient for normal epiblast length and OCT4 marker gene expression. Epiblast width was decreased in mutants by E6.5, which may relate to A-P axis developmental defects in the *Ets2*-mutant, investigated in Chapter 4. The mutant epiblast at late post-implantation stages E6.5 and E7.0 was abnormal, lacking T-positive mesoderm induction. Together these data suggest that while the reduced mutant ExE appears sufficient for normal epiblast marker gene expression and morphometrics at early post-implantation, it may be insufficient for providing signals to support late post-implantation epiblast patterning.

I found the post-implantation mutant ExE to retain normal proliferation and apoptosis, showing that following its reduction, the mutant ExE has regular tissue growth and death rates. The small handful of mutant ExE cells found at E5.5 that grow normally to E6.5 may be independent of ETS2, which would suggest that the ExE is a heterogeneous population. The peri-implantation mutant ExE was also normal in proliferation. As the E5.0 mutant ExE was unperturbed in length morphometrics and cell number, then was rapidly reduced in size by E5.5, I tested whether increased peri-implantation apoptosis could account for the mutant ExE reduction. Indeed, I found an elevation of CC3 apoptotic index in ExE progenitors of the pTE

at E4.75, followed by upregulated Click-TUNEL index in the peri-implantation ExE at E5.0. These observations are in line with the progression of the apoptotic pathway where activation of executioner Caspase-3 precedes fragmentation of DNA (Taylor et al., 2008). To validate whether apoptosis is the causal mechanism for mutant ExE tissue disappearance, I cultured control and mutant embryos in the Pan-Caspase apoptosis inhibitor FMK, and was able to rescue the mutant ExE phenotype, i.e. recovery of ExE cell number close to WT levels. However, this was not a complete rescue, as CDX2 expression levels were not restored to fully mirror that of WT counterparts. This suggests a more nuanced mechanism at play, which apoptosis inhibition alone is insufficient to recapitulate.

Overall, I have shown that ETS2 is required during a short time window at peri-implantation to prevent apoptosis of a subset of ExE cells. These data suggest a heterogeneous ExE with at least 2 populations: one dependent on ETS2 for survival at peri-implantation and one that does not. The ETS2-dependent population highlights the anti-apoptotic effect that ETS2 can have, which may relate to its oncogenic role in the stroma of breast (Tynan et al., 2005), prostate (Grasso et al., 2012), and pancreatic (Pitarresi et al., 2016) cancers.

4. ETS2 regulates AVE proliferation, orientation of cell division, and cell density

4.1 Introduction

4.1.1 Maintenance of epithelial cell division rates and orientation

During *in vivo* tissue morphogenesis, epithelial cell divisions occur in a constrained manner to balance cell number homeostasis, proper epithelial architecture, and growth whilst preserving barrier function (Ragkousi and Gibson, 2014). Epithelial mitosis is coordinated with tissue wide maintenance of polarity. Cells that undergo planar cell divisions parallel to the basement membrane orient their mitotic spindles and establish new adherens junctions in a manner conformed to overall tissue organisation. The proper functioning of epithelia is dependent on cell polarity that involves formation of adherens junctions along the apicolateral membranes with neighbouring cells (Cavey and Lecuit, 2009). Polarisation in apicolateral domains feature differential localisation of polarity proteins and distinct lipid compartments (Tepass et al., 2001). Both polarity and intercellular adhesion are important for epithelial barrier function (Papusheva and Heisenberg, 2010).

Orientation of mitotic spindle

Epithelial cell divisions can be oriented in the planar or orthogonal axis; randomised oblique cell divisions are less frequent (Gillies and Cabernard, 2011). Programmed perpendicular cell divisions can be a means of achieving asymmetric segregation of fates to establish different cell types, for example in the developing neural cortex (Fietz et al., 2010), and stratifying embryonic skin (Lechler and Fuchs, 2005). Planar cell divisions result in daughter cells of equivalent fate and apical-basal polarity, and involves tightly coordinated establishment of *de novo* cell-cell adherent junctions to enable smooth integration within the

monolayer (Morin and Bellaïche, 2011). The cellular mechanism of mitotic spindle orientation involves molecular anchors Dynein-Dynactin that are positioned in a restricted fashion via Gnai3-LGN-NuMA complex in polarised cells (Peyre et al., 2011). Absence or mislocalisation of these factors lead to defects in spindle orientation that eventually result in loss of progenitors in a simple epithelium (Peyre et al., 2011). At a global scale, tension that results from cell growth and spreading can also influence orientation of cell division.

Therefore, during growth, epithelial cells have the ability to retain their distinct morphology, barrier function and adhesion to adjacent cells. Maintenance of epithelial integrity, order, and division rates and orientation relate to the origin of epithelial cancers where cells that over proliferate fail to integrate back in the epithelial monolayer (Ragkousi and Gibson, 2014).

4.1.2 Characterisation of proliferation in the VE and AVE at post-implantation stages

Proliferation is regulated in the VE epithelium for homeostatic and developmental purposes. Stuckey et al., 2011 conducted detailed *in vivo* quantification of proliferation at post-implantation across various stages of A-P axis development. Specifically, they quantified distribution of cells in S, G2, and M phases of the cell cycle using KI67, BrdU incorporation, and PHH3. They found that while growth of the VE epithelium as a whole is influenced by epiblast proliferation rates, no difference in proliferative index between the anterior versus posterior VE was detected throughout stages of AVE induction and anterior shift. These data suggest that differential proliferation within the VE is insufficient to account for repositioning of the AVE.

4.1.3 ETS2 regulation of epithelial proliferation in intestinal tumorigenesis

In cancer, ETS2 may act as either an oncogene or tumour suppressor, reviewed in 1.4.1 and 1.4.2, respectively. ETS2 has been reported to exert a tumour suppressive function in epithelial cells. Specifically, Munera et al., 2011 demonstrated a role for ETS2 in mouse intestinal epithelial tumorigenesis via tissue-specific conditional inactivation of *Ets2* in colon epithelia of young and adult mice. This perturbation led to an increased proportion of *Ets2*-null intestinal crypts due to elevated proliferation. These crypts additionally displayed increased rates of fission that resulted in large patches of *Ets2*-null intestinal epithelia, suggesting a competitive advantage. *Ets2*-deficient crypts also exhibited higher cell density. When presented with carcinogens, these crypts had elevated propensity for tumorigenesis. ETS2 regulation of intestinal tumorigenesis may be specific to tumours driven by WNT signalling (Munera et al., 2011). ETS2 appears to modulate the proliferative response to WNT signalling in the intestinal stem cell compartment marked by LGR5, which is also a WNT target (Garcia et al., 2009).

In this Chapter, I examined ETS2 function in another epithelial context, where it may be regulating cell growth and proliferation in the embryonic VE.

4.2 Results

4.2.1 ETS2 protein is expressed in the VE and AVE at post-implantation stages

To assess where ETS2 is expressed and may be transcriptionally active – i.e. display nuclear localisation – at post-implantation, I conducted whole mount immunofluorescent staining analysis. I used an antibody that recognises the DNA-binding domain of ETS2, where absence of signal in ETS2 immunostains of mutants validated the antibody's specificity (Chapter 3, Figure 3E). In the previous chapter, I reported presence of nuclear ETS2 expression in the pre-implantation mouse embryo starting as early as the morula stage, with increasing

expression in the CDX2-positive TE lineage upon progression through pre-implantation development (3.2.1). Here I report high levels of ETS2 protein expression in the ExE at post-implantation. However, when I increased the brightness of the ETS2 signal, I could detect its nuclear expression in the epiblast and throughout the VE (Figure 28). Notably, presence of nuclear ETS2 in the AVE raises the possibility that the mutant AVE phenotype could be due to a cell autonomous effect of ETS2. The observation that ETS2 is expressed in all lineages at post-implantation is supported by multiple scRNA-seq reports at E5.5 and E6.5 (Mohammed et al., 2017; Wen et al., 2017; Scialdone et al., 2016).

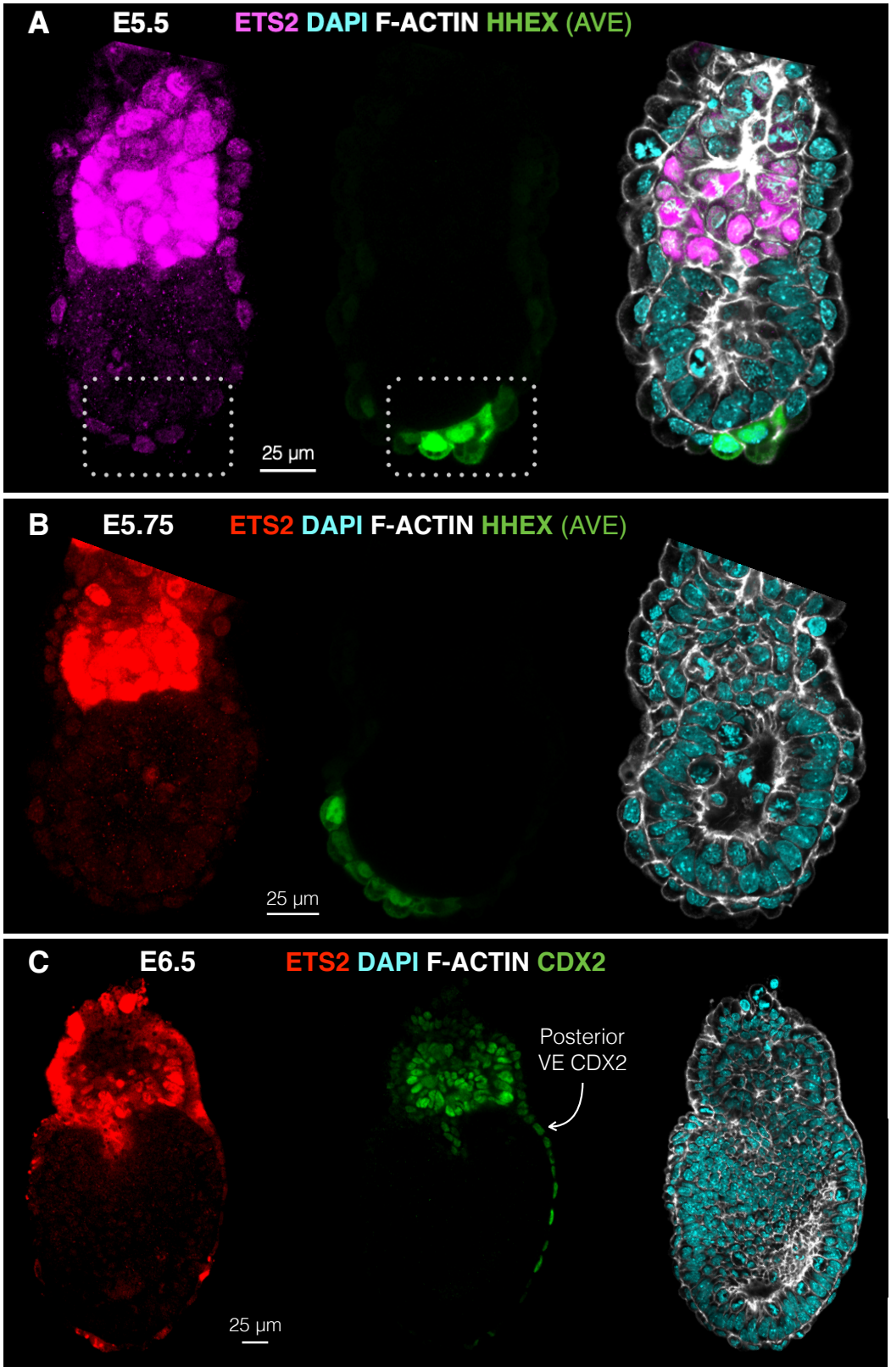


Figure 28. ETS2 expression at post-implantation stages (E5.5 to E6.5)
(A) Representative mid-sagittal sections of whole mount immunofluorescent staining at E5.5 of ETS2 (magenta) and HHEX-GFP (green – AVE), showing ETS2 nuclear expression in all lineages including the AVE (boxed).
(B) ETS2 (red) is expressed in the migrating AVE, and all cell lineages at E5.75.
(C) ETS2 (red) retains similar expression pattern at E6.5 as compared to early post-implantation. Here, CDX2 (green) marks the ExE, as well as displays nuclear expression in the posterior VE (arrow). F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

4.2.2 AVE migration is complete in mutants

Similar to Georgiades and Rossant, 2006, I observed a thickened AVE in the mutants by E6.5 (Figure 29). With confocal immunofluorescence, I could obtain details on the cellular and molecular nature of the thickening.

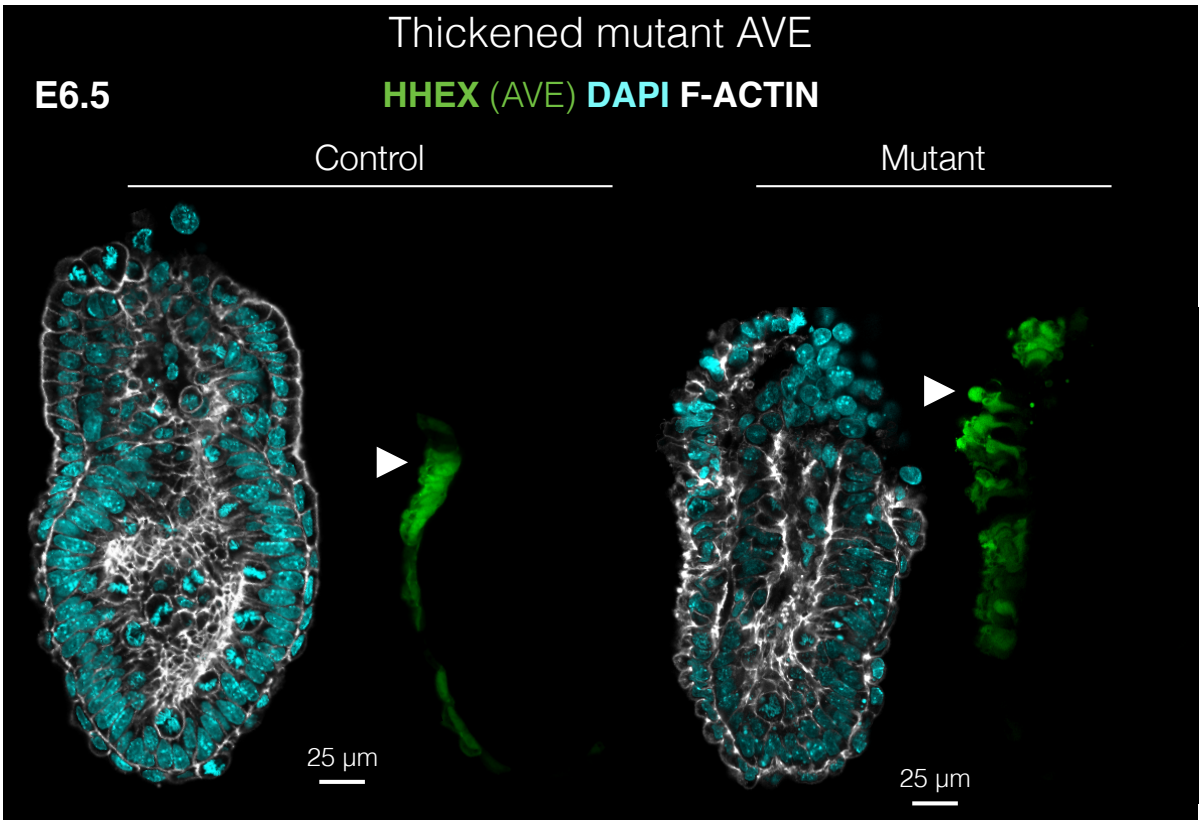


Figure 29. Thickened mutant AVE at E6.5
Representative mid-sagittal sections of whole mount immunofluorescent images at E6.5 of HHEX-GFP (green – AVE); F-ACTIN (gray); DAPI (cyan). The HHEX-GFP-positive AVE (arrowhead) is markedly thickened in mutant compared to control. Mutant AVE epithelium also appeared disorganised, visible with F-ACTIN. Scale bars: 25 µm.

Differing from Georgiades and Rossant, 2006 where they characterised type 1 mutants to have incomplete AVE anterior shift by E6.75, I found presence of complete AVE migration in *Ets2*^{db1}-mutants. The AVE, ubiquitously marked by AMOT (Shimono and Behringer, 2003), migrated up the entire surface of the epiblast long axis by E5.75; migration was also completed at E6.5 (Figure 30). This observation of stereotypic AVE anterior shift over the expected time window of E5.5 to E6.5 suggests that failure of T-positive-mesoderm induction (3.2.5) is not a result of defective AVE migration.

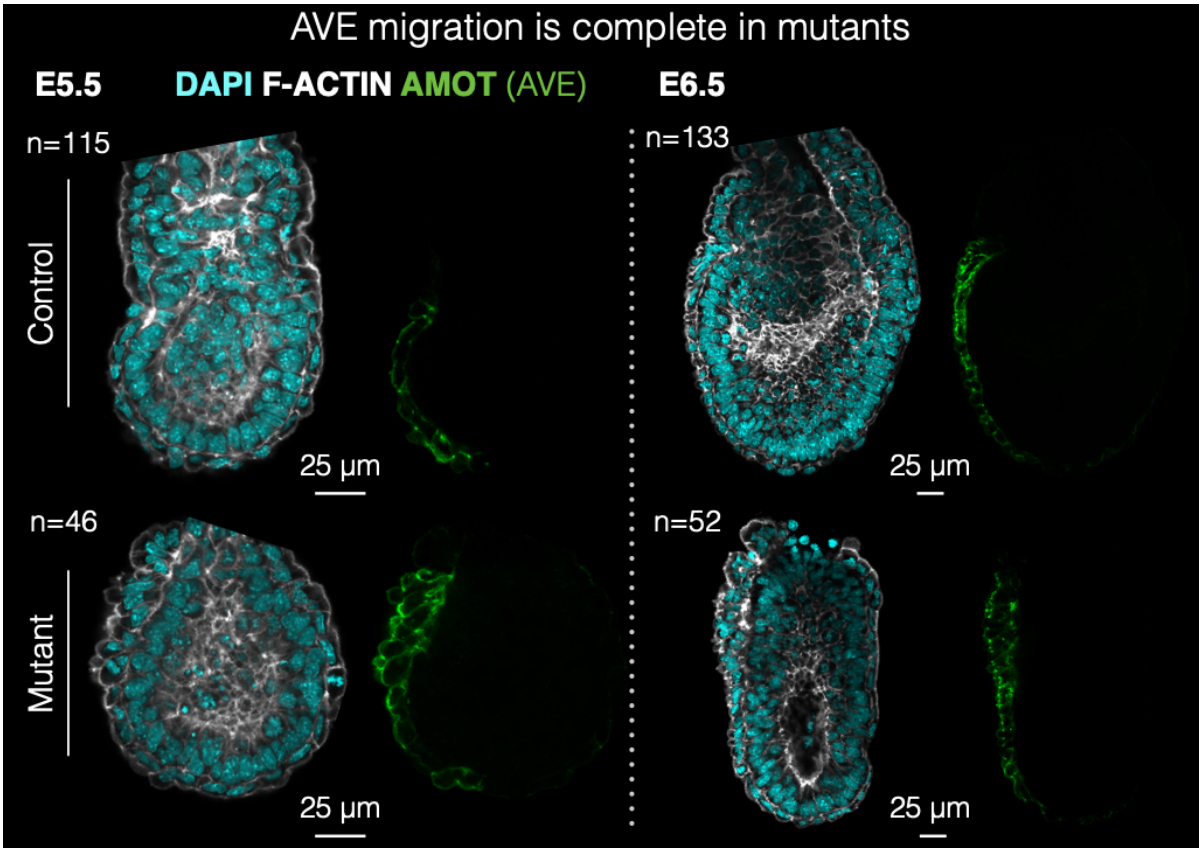


Figure 30. AVE migration is completed in mutants at late E5.5 and E6.5
Representative mid-sagittal sections of whole mount immunofluorescent images at E5.5 and E6.5 of AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). A later stage E5.5 embryo is shown where the AVE has just completed migration in both mutant and littermate control. AVE migration is also completed at E6.5. Scale bars: 25 µm.

4.2.3 Apical-lateral polarity is intact in mutants

To interrogate whether the thickened mutant AVE is characterised by epithelial disorganisation or loss of epithelial integrity, I examined polarity components by confocal immunofluorescence. I analysed: apical junctional proteins PAR6B (Figure 31, 32) and PAR3 (Figure 33, 34); tight junctional marker ZO-1 (Figure 35, 36); as well as lateral junctional markers SCB (Figure 37, 38) and ECAD (Figure 39, 40). I conducted these immunofluorescent characterisations at both early (E5.5) and late (E6.5) post-implantation stages in control versus mutant. All apical-lateral polarity components I examined were unperturbed in mutants in terms of localisation, however intensities may differ. I also stained for AMOT junctional component, which serves as an AVE marker (Shimono and Behringer, 2003), and found normal

expression in mutants at different phases of AVE migration ('distal', 'migrating', 'migrated' – Figure 41, 42, 43, respectively). For these immunofluorescent data, I present both sections and 3D opacity renderings. Sections enable visualisation of subcellular details in protein localisation, while 3D opacity renderings aid visualisation of the junctional meshwork in topological view. Together, these allow assessment of whether stereotypic polarity organisation is disrupted in mutants.

Although not a polarity marker, I also examined GATA6 to verify, in addition to AMOT, that the mutant retained stereotypic VE marker gene expression. Though GATA6 expression was normal in the mutant VE (Figure 44), I found presence of approximately 2-4 ectopic GATA6 cells in the epiblast cavity at E5.5. To trace the origins of this defect, I assessed GATA6 expression at peri-implantation stages E4.75 and E5.0. I found presence of approximately 2 ectopic GATA6-positive cells that displayed incomplete sorting to the PrE at E4.75 (Figure 45A). All control counterparts had completely sorted PrE by that stage. In normal development, incompletely sorted PrE cells in the late blastocyst undergo apoptosis (Plusa et al., 2008). These ectopic GATA6-positive cells at E4.75 appeared to subsequently end up sandwiched between the ExE and epiblast at E5.0 (around 4-6 cells; Figure 45B). The potential final fate of these mislocalised cells could be cell death by extrusion into the epiblast cavity at E5.5. Proof of this path for ectopic GATA6 cells in the *Ets2*-mutant would require live imaging with a *Gata6* reporter, which had not been conducted in this thesis.

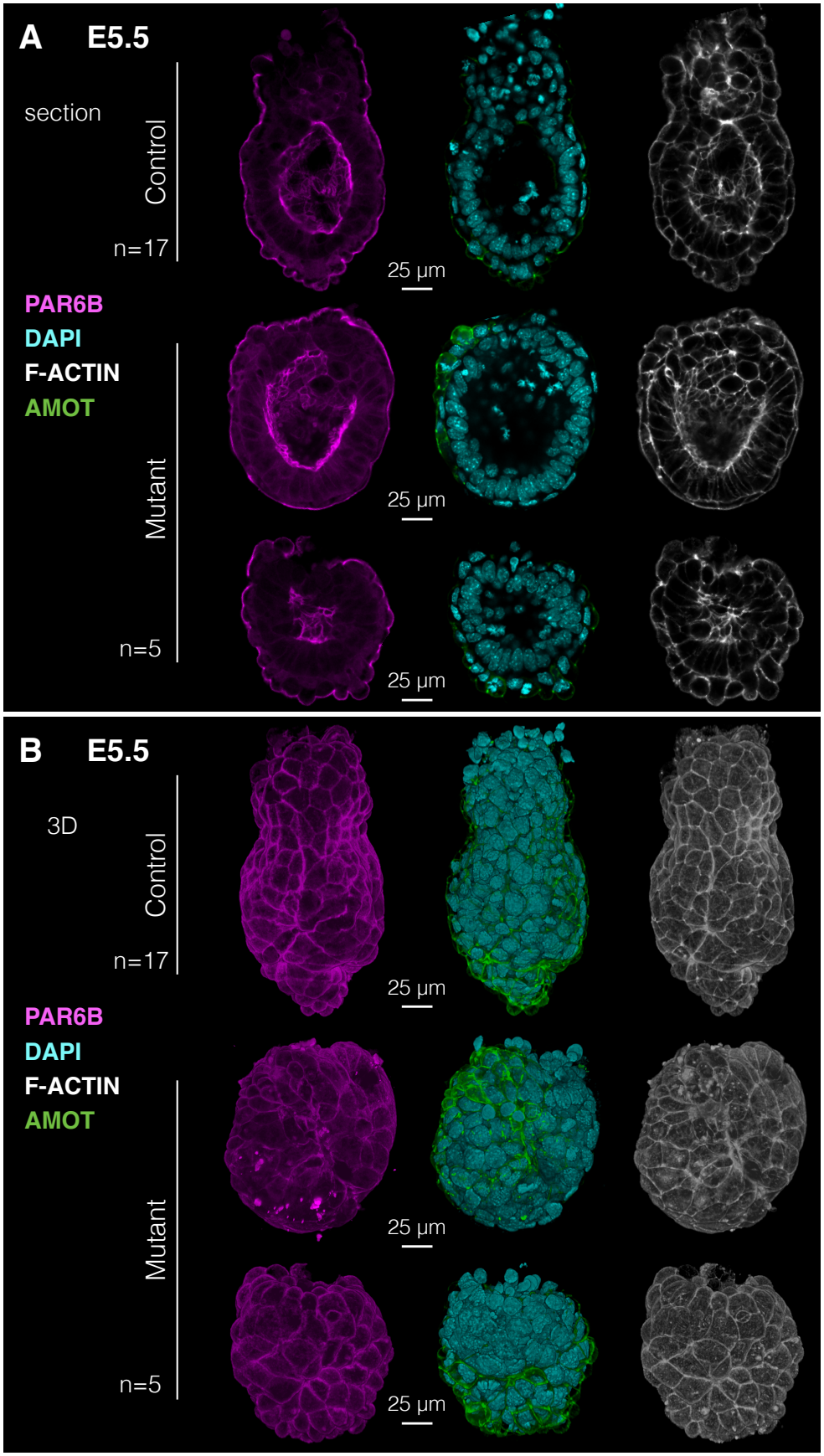


Figure 31. PAR6B apical polarity marker expression at E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of PAR6B (magenta) co-stained with AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

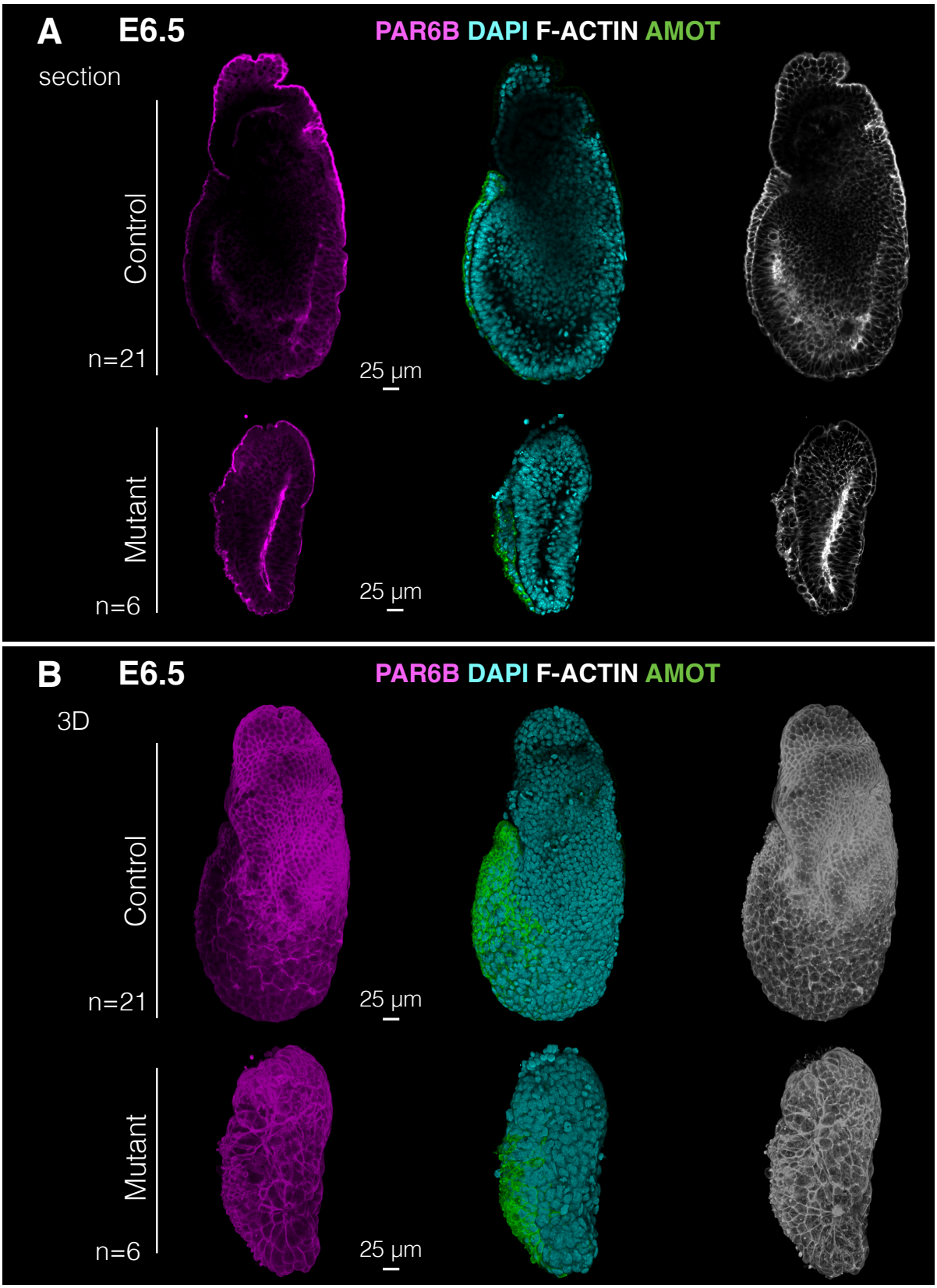


Figure 32. PAR6B apical polarity marker expression at E6.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of PAR6B (magenta) co-stained with AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). PAR6B retained apical membrane localisation in the exposed surface of the thickened mutant AVE and did not display ectopic expression. Scale bars: 25 µm.

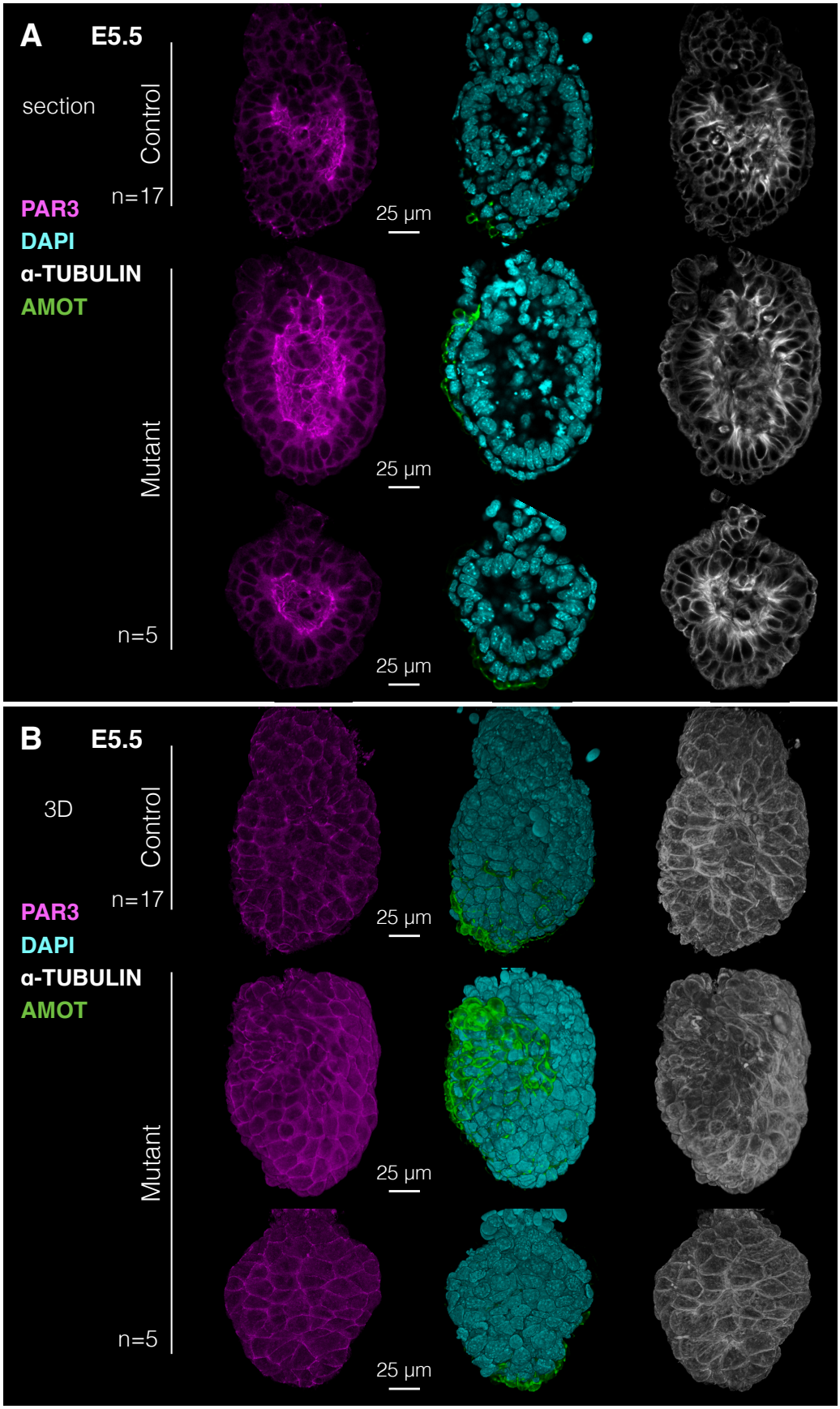


Figure 33. PAR3 apical junctional marker expression at E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of PAR3 (magenta) co-stained with AMOT (green – AVE); α-TUBULIN (gray); DAPI (cyan). PAR3 retained apical junctional localisation in the VE and AVE of mutants. Scale bars: 25 μm.

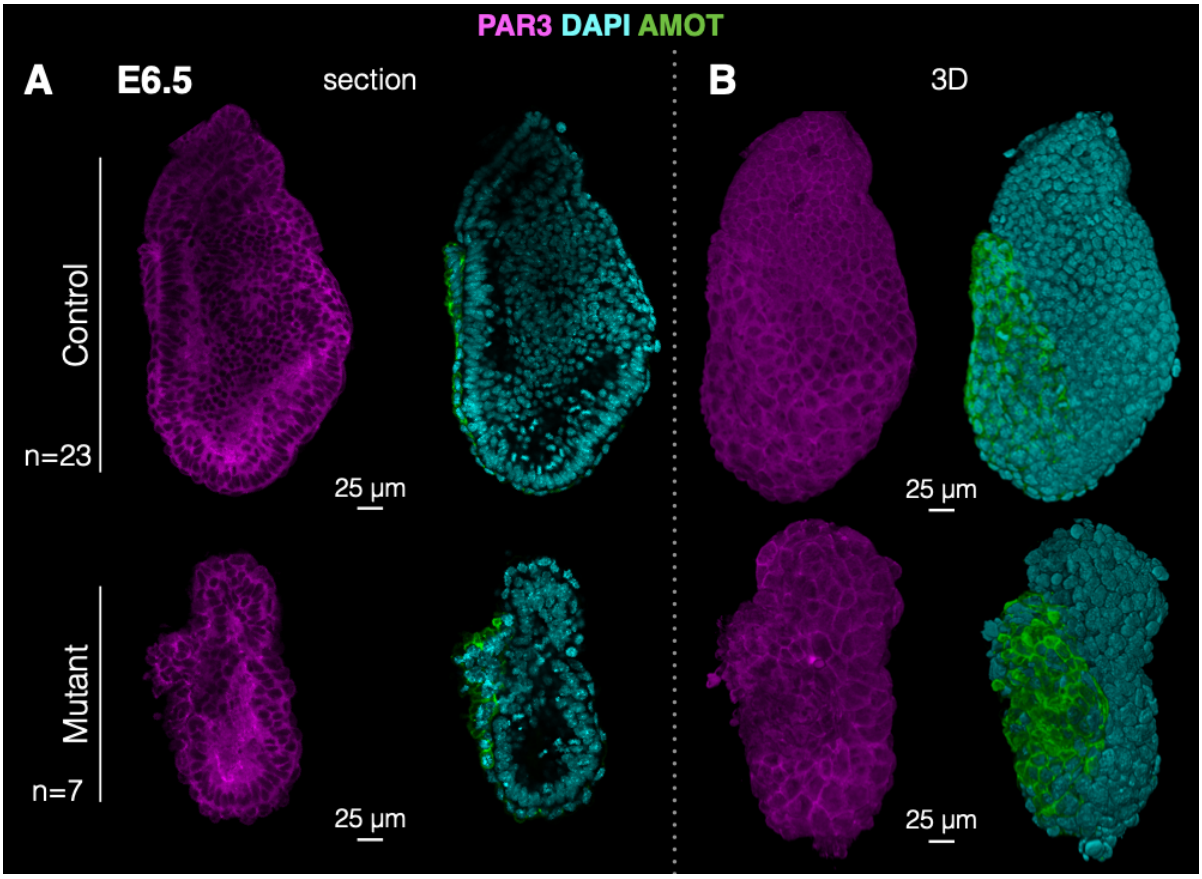


Figure 34. PAR3 apical junctional marker expression at E6.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of PAR3 (magenta) co-stained with AMOT (green – AVE); DAPI (cyan). PAR3 retained apical junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

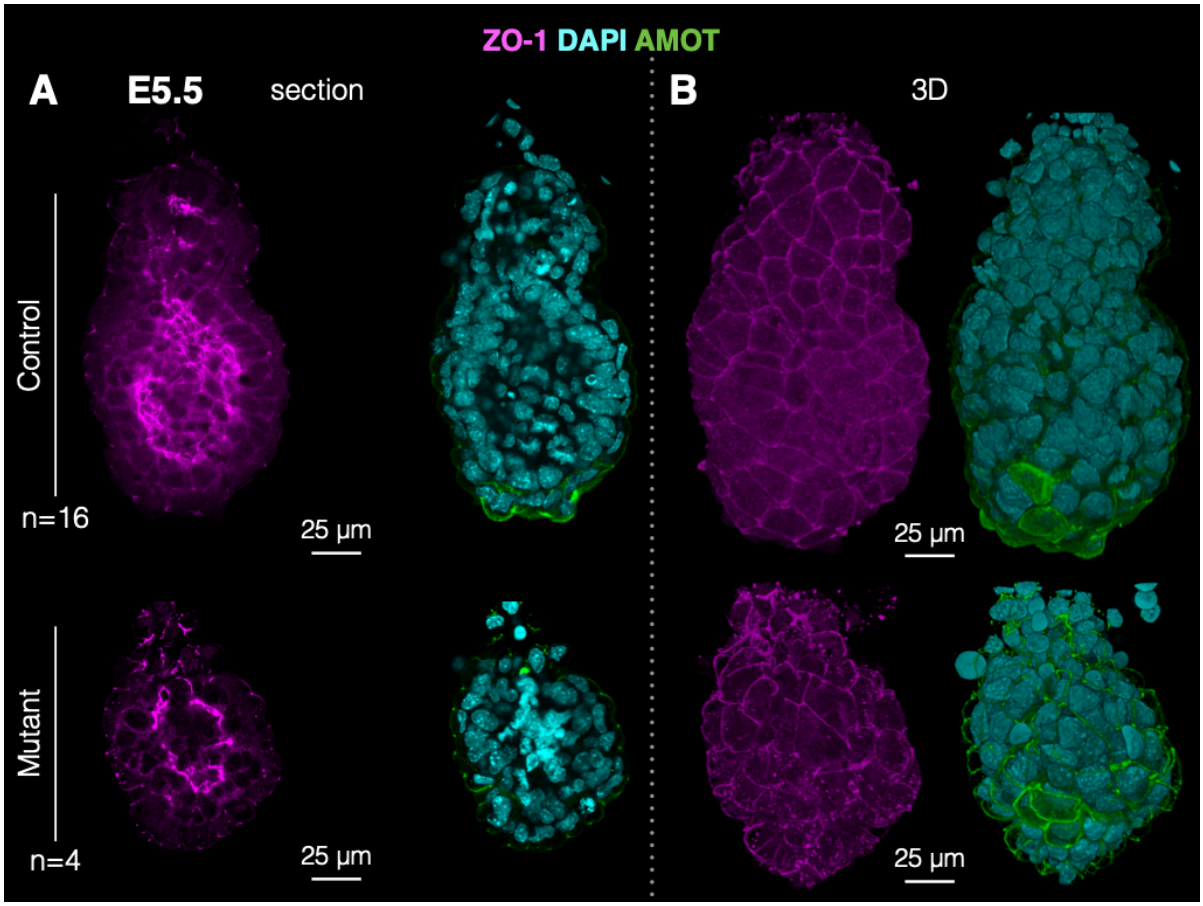


Figure 35. ZO-1 tight junctional marker expression at E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of ZO-1 (magenta) co-stained with AMOT (green – AVE); DAPI (cyan). ZO-1 retained tight junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

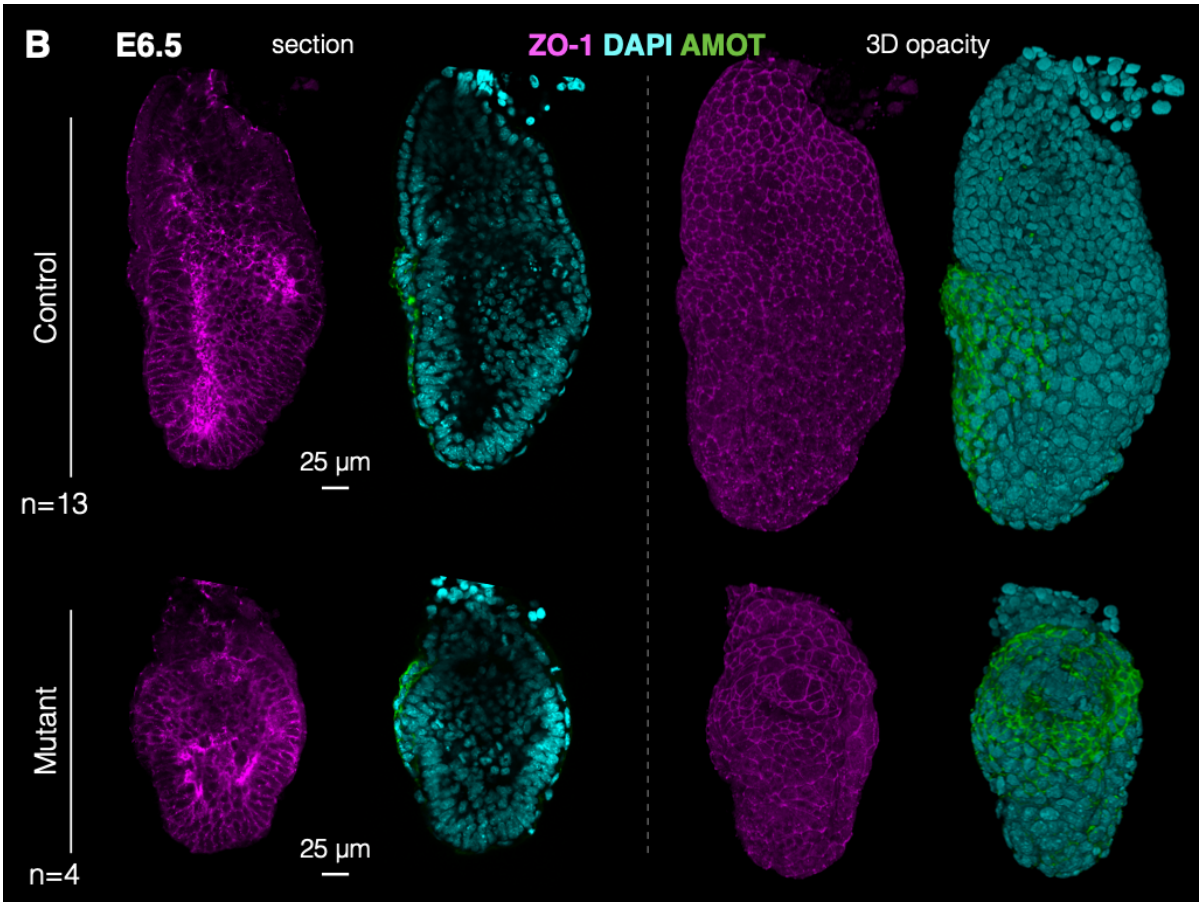


Figure 36. ZO-1 tight junctional marker expression at E6.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of ZO-1 (magenta) co-stained with AMOT (green – AVE); DAPI (cyan). ZO-1 retained tight junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

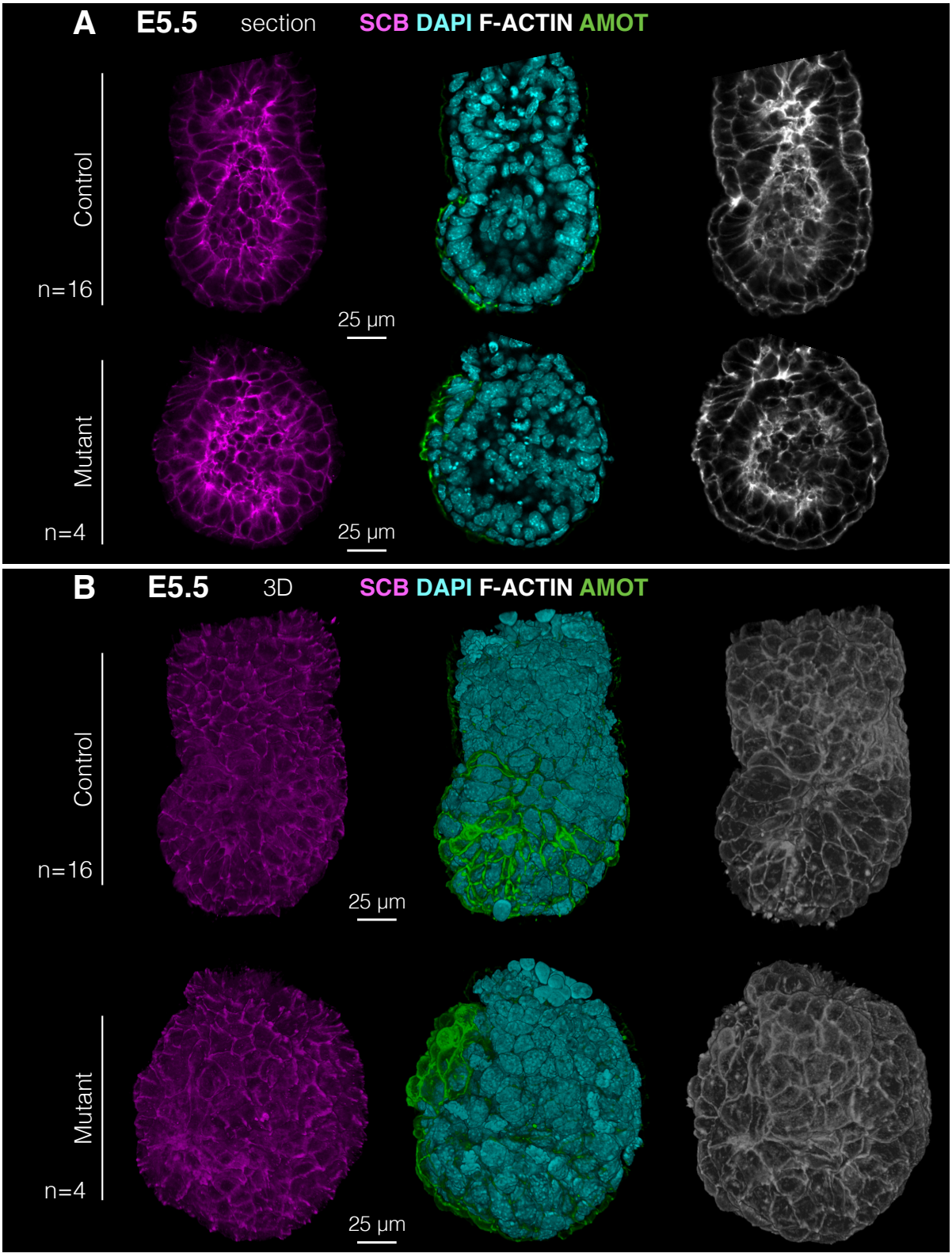


Figure 37. SCB lateral junctional marker expression at E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of SCB (magenta) co-stained with AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). SCB retained lateral junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

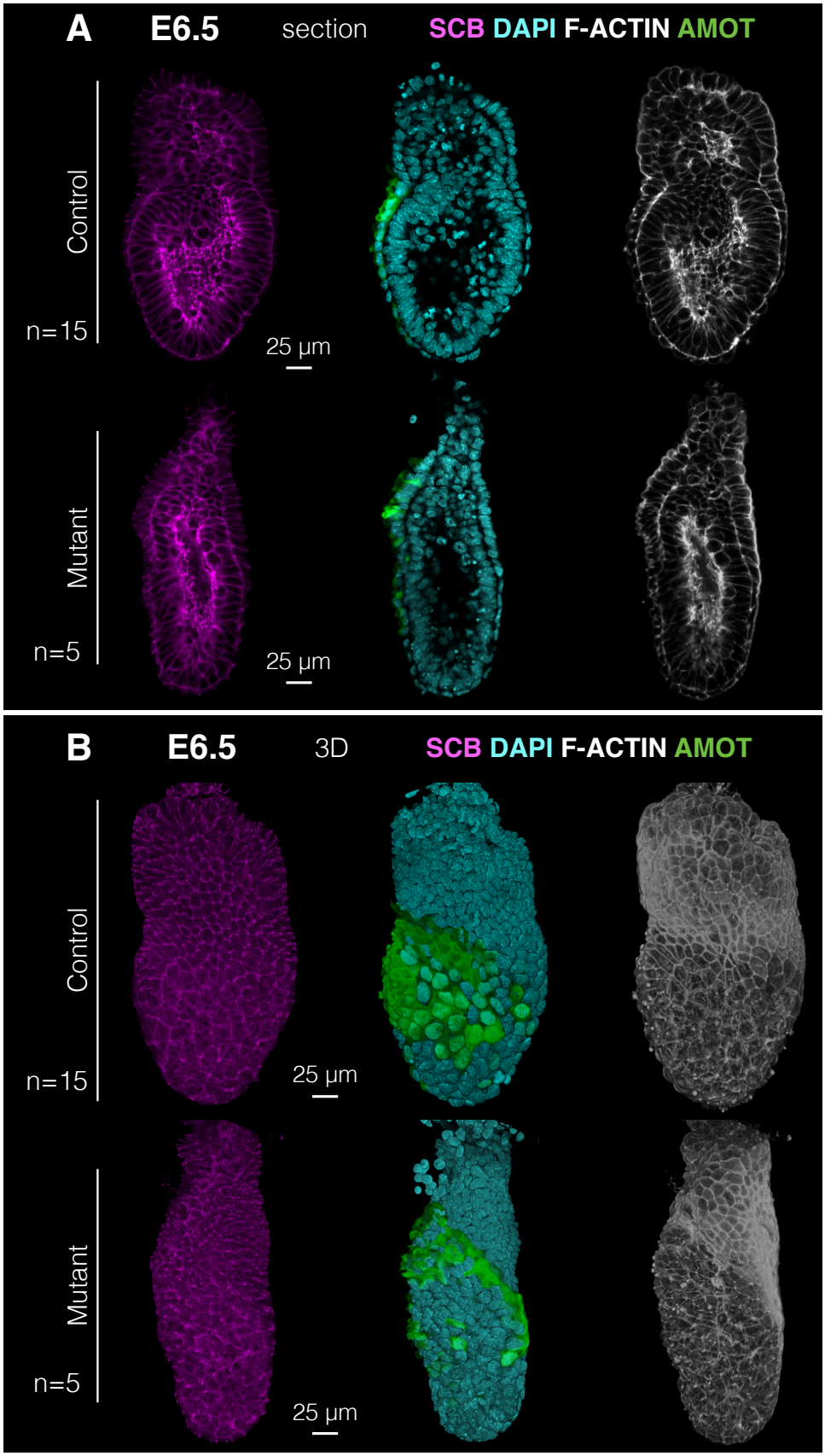


Figure 38. SCB lateral junctional marker expression at E6.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of SCB (magenta) co-stained with AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). SCB retained lateral junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

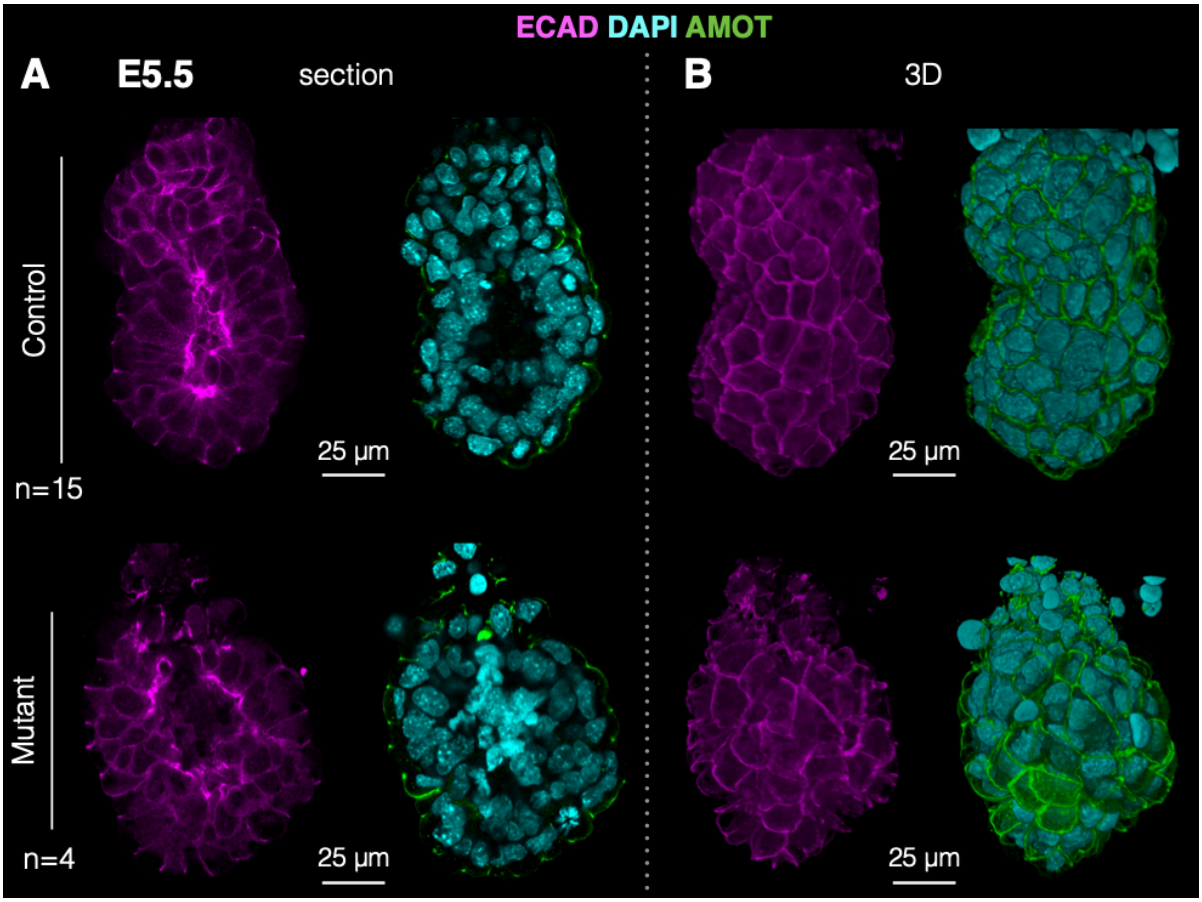


Figure 39. ECAD lateral junctional marker expression at E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of ECAD (magenta) co-stained with AMOT (green – AVE); DAPI (cyan). ECAD retained lateral junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

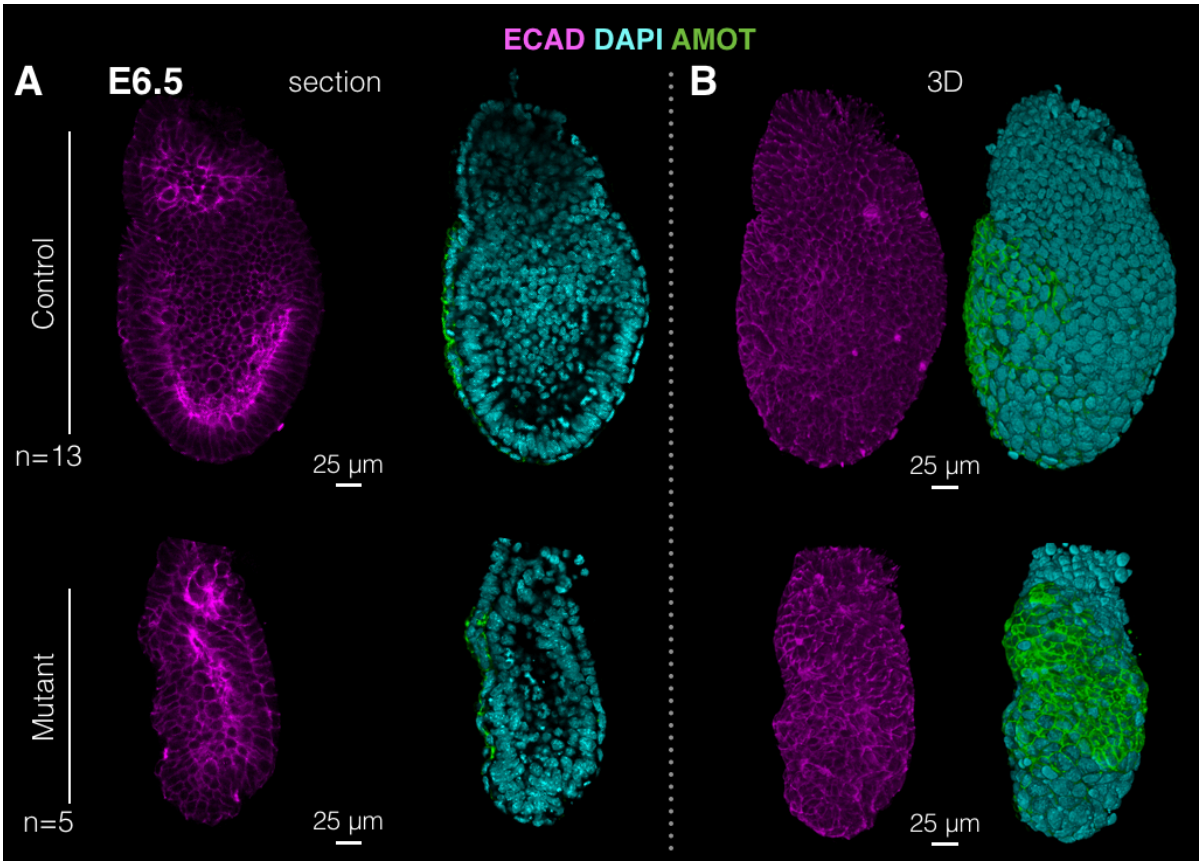


Figure 40. ECAD lateral junctional marker expression at E6.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of ECAD (magenta) co-stained with AMOT (green – AVE); DAPI (cyan). ECAD retained lateral junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

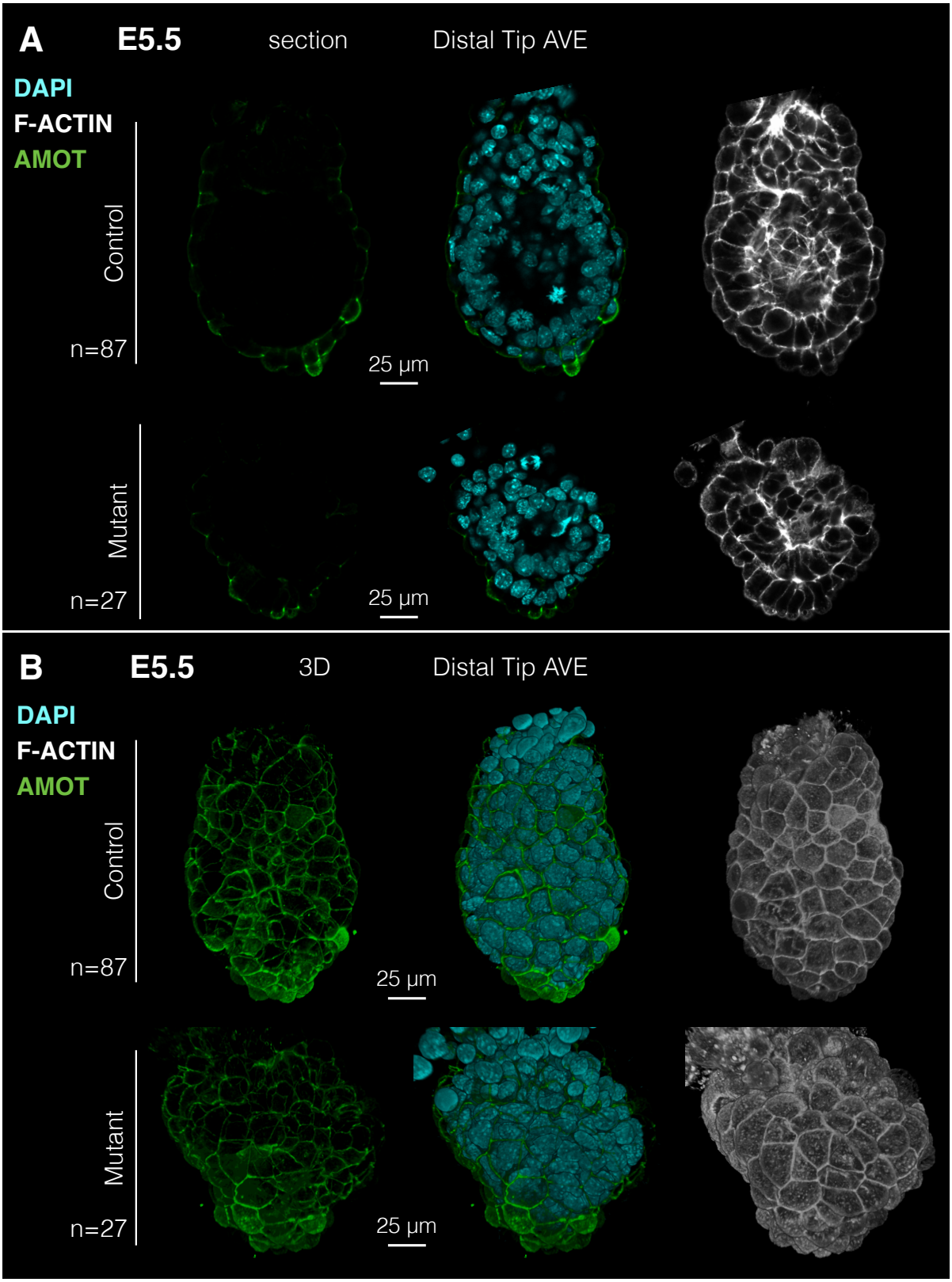


Figure 41. AMOT expression at distal tip AVE stage E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). AMOT retained normal junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

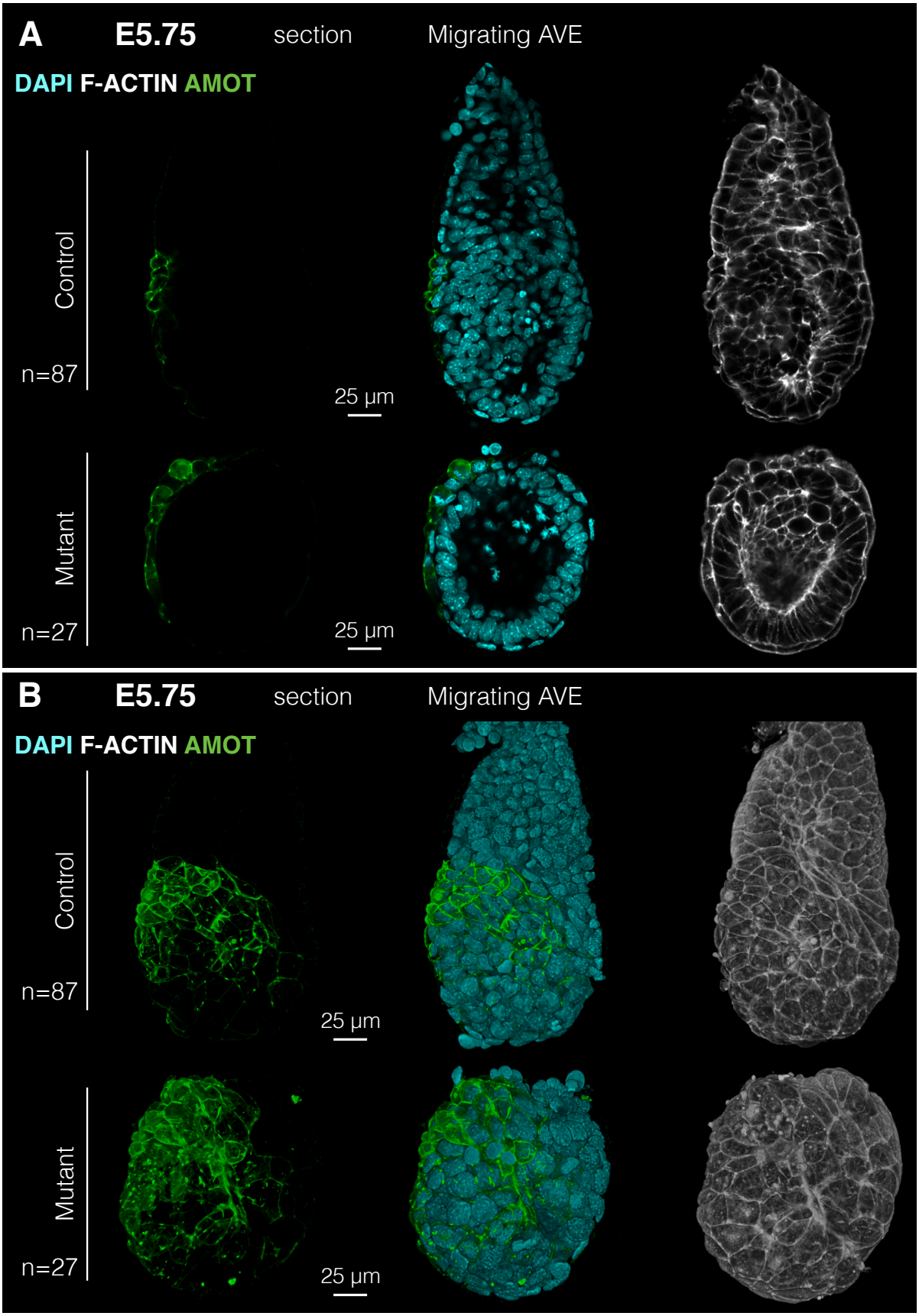


Figure 42. AMOT expression at migrating AVE stage E5.75
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.75 of AMOT (green – AVE); F-ACTIN (gray); DAPI (cyan). AMOT retained normal junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

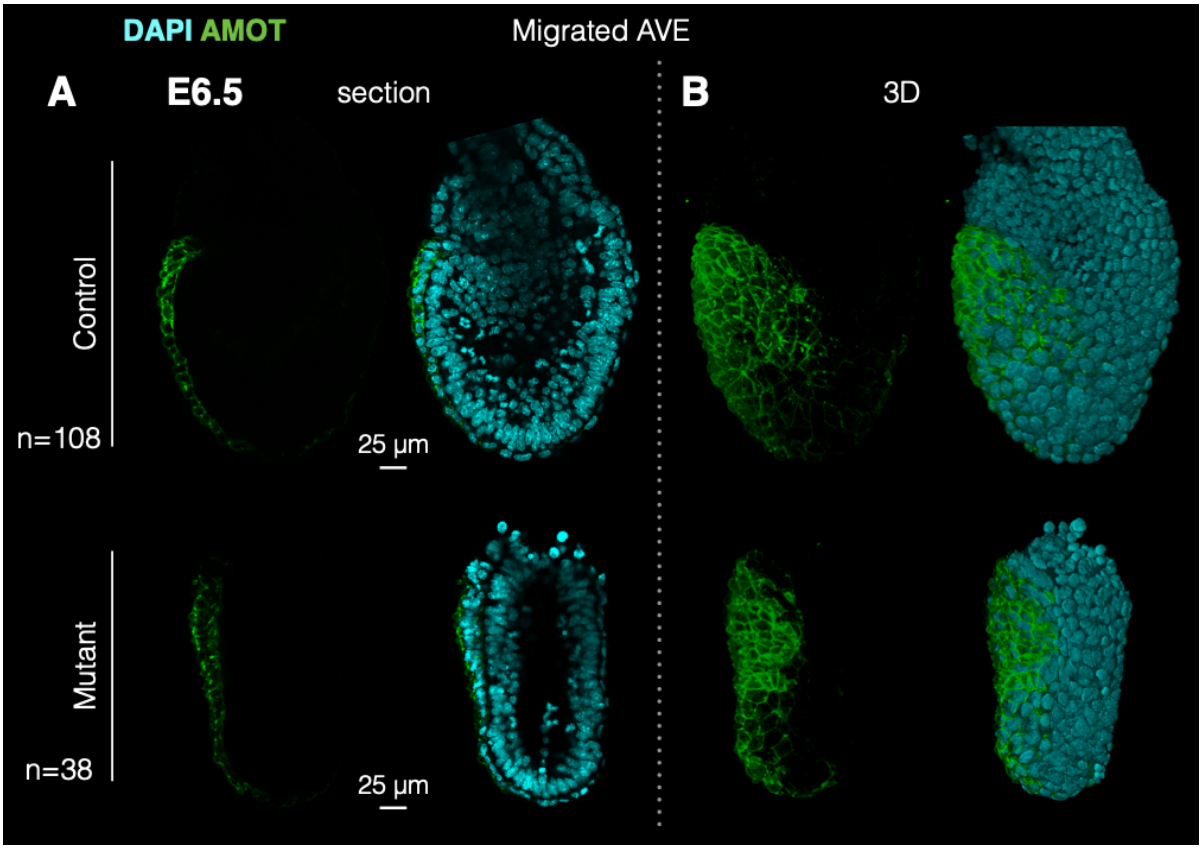


Figure 43. AMOT expression at migrated AVE stage E6.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of AMOT (green – AVE); DAPI (cyan). AMOT retained normal junctional localisation in the VE and AVE of mutants. Scale bars: 25 µm.

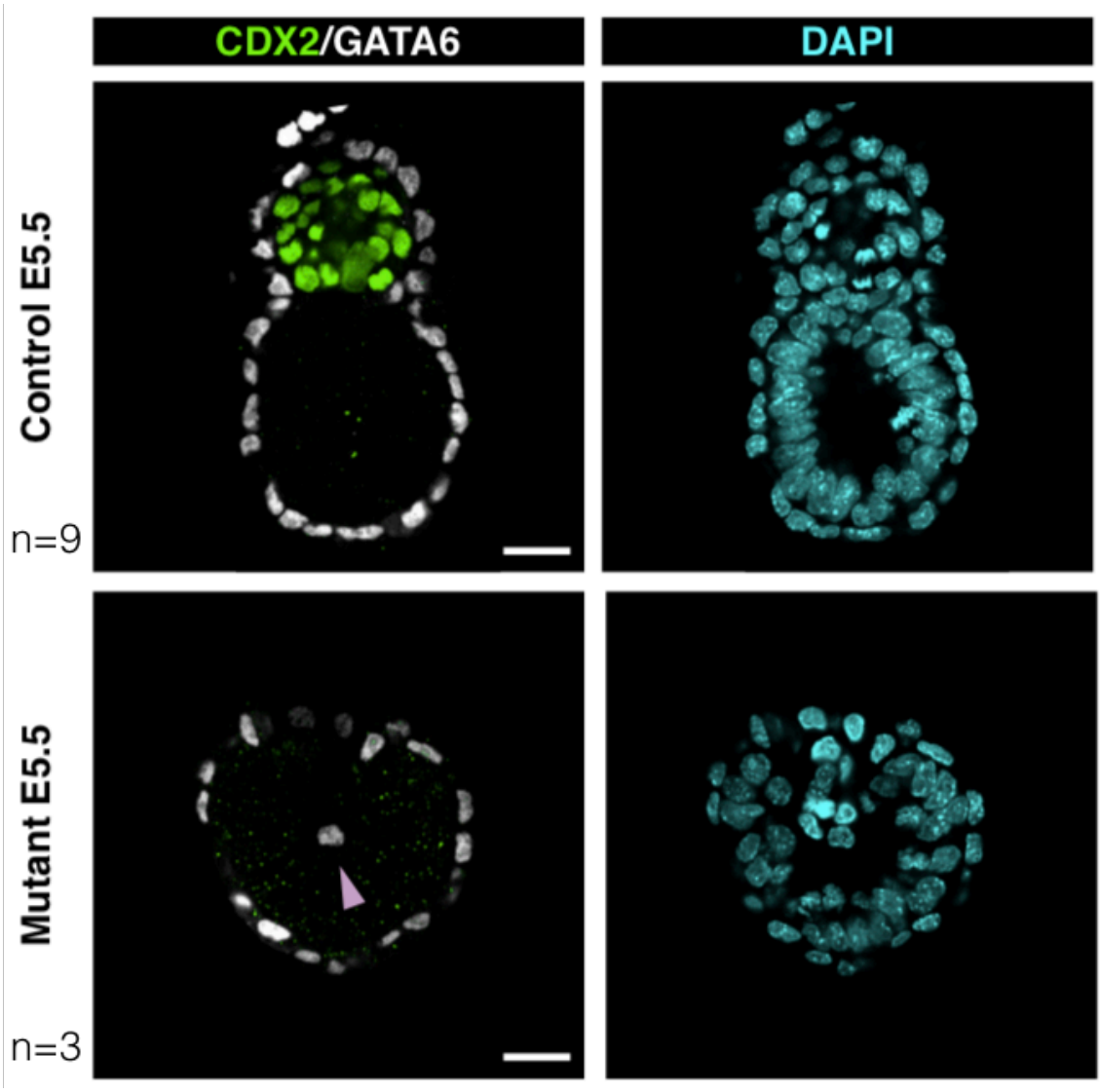


Figure 44. GATA6 VE marker expression at E5.5
Representative mid-sagittal sections and of whole mount immunofluorescence at E5.5 of GATA6 (gray – VE); CDX2 (green – ExE); DAPI (cyan). Normal GATA6 expression in the mutant VE. Presence of ~2-4 ectopic GATA6-positive cells (arrowhead) per embryo in the epiblast cavity at E5.5. Scale bars: 25 μm.

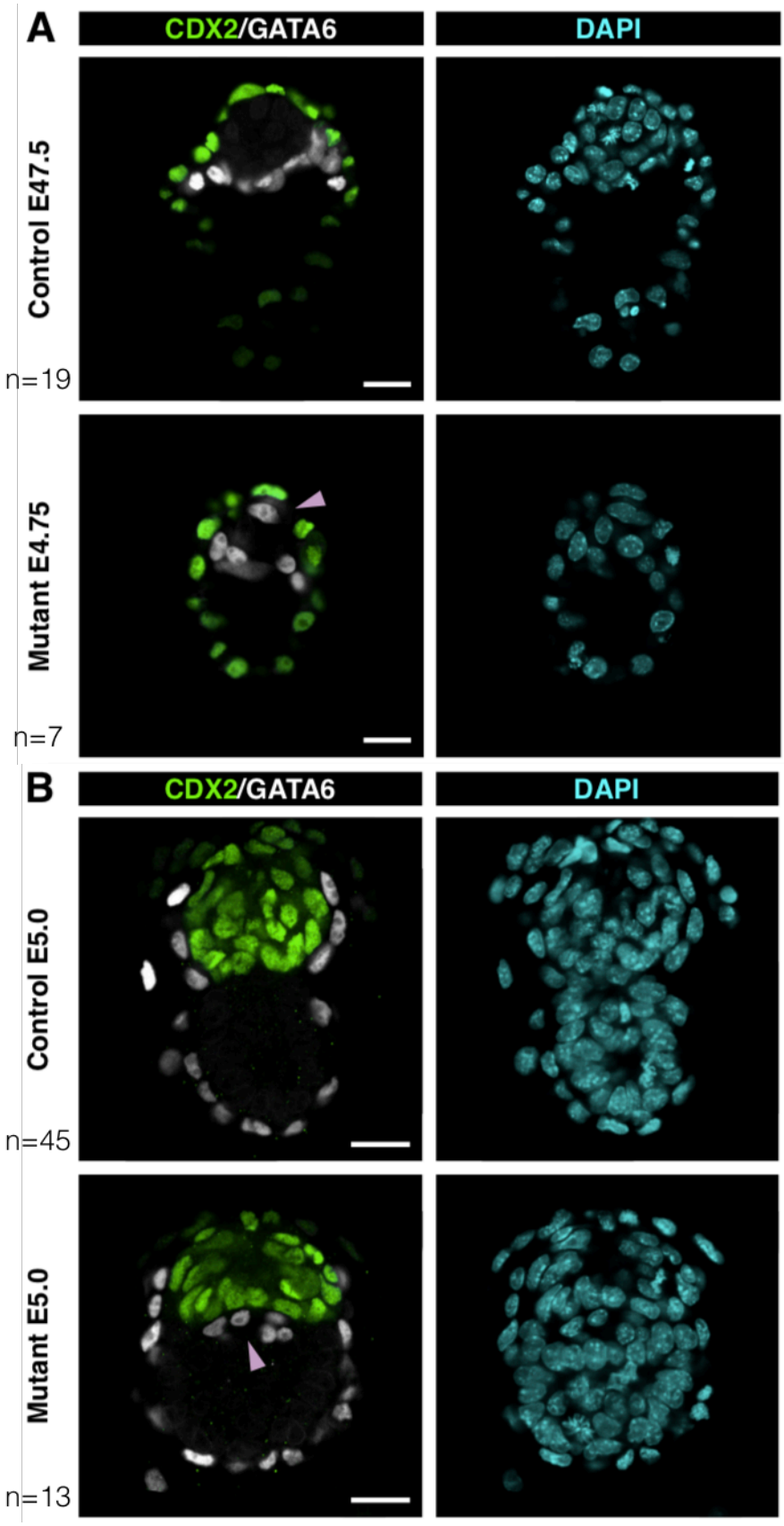


Figure 45. GATA6 PrE/VE marker expression at E4.75 and E5.0
Representative mid-sagittal sections and whole mount immunofluorescence at E4.75 (A) and E5.0 (B) of GATA6 (gray – VE); CDX2 (green – ExE); DAPI (cyan). (A) Incomplete PrE sorting of ~2 ectopic GATA6-positive cells (arrowhead) per embryo in mutants at E4.75. (B) ~4-6 ectopic GATA6-positive cells (arrowhead) per embryo sandwiched between the ExE and epiblast in mutants at E5.0. Scale bars: 25 μm.

I additionally characterised cytoskeletal components F-ACTIN (Figure 46) and phospho-Myosin Light Chain 2 (pMLC2; Figure 47), in control versus mutant to test whether these effectors of cellular dynamics were perturbed. F-ACTIN also aided in visualisation of cell morphology. Upon Imaris 3D outlining of F-ACTIN in control versus mutant AVE cells (at 0.8 μm z-intervals), I found F-ACTIN intensity – normalised to cell surface area – to be elevated in the thickened E6.5 mutant AVE (Figure 46B).

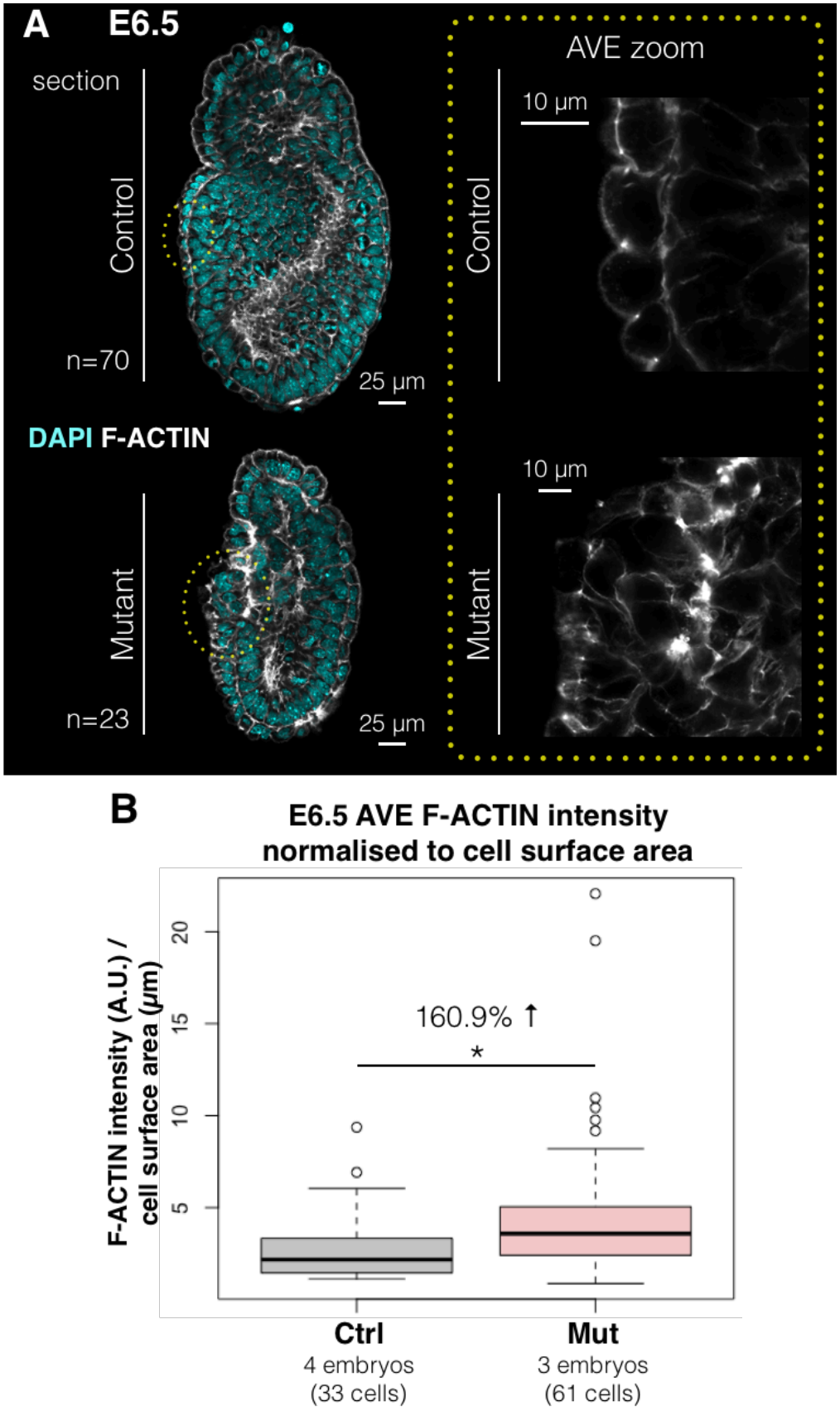


Figure 46. F-ACTIN and normalised intensity quantification at E6.5 to analyse thickened mutant AVE
(A) Representative mid-sagittal sections of whole mount staining of F-ACTIN at E6.5 in control versus mutant. AVE zoom of circled (yellow) region is shown. Scale bars: 25 µm, or 10 µm as indicated.
(B) Increased junctional F-ACTIN intensity normalised to cell surface area in the thickened E6.5 mutant AVE compared to control upon 3D Imaris outlining. (Statistical test: hierarchical ANOVA; * $p < 0.05$).

The actomyosin machinery is the ultimate effector of cellular dynamics where pMLC2 is a readout for engagement of the cytoskeleton and active contractility, required for cell division and motility (Fletcher and Mullins, 2010). I examined pMLC2 at E6.5 (Figure 47) to gain insight into whether the thickened AVE in mutants have altered frequency and localisation of active actomyosin contractility. I found pMLC2 to appear as punctate accumulations along different regions of the cell membrane, where mutants displayed increased junctional pMLC2 compared to control (Figure 48).

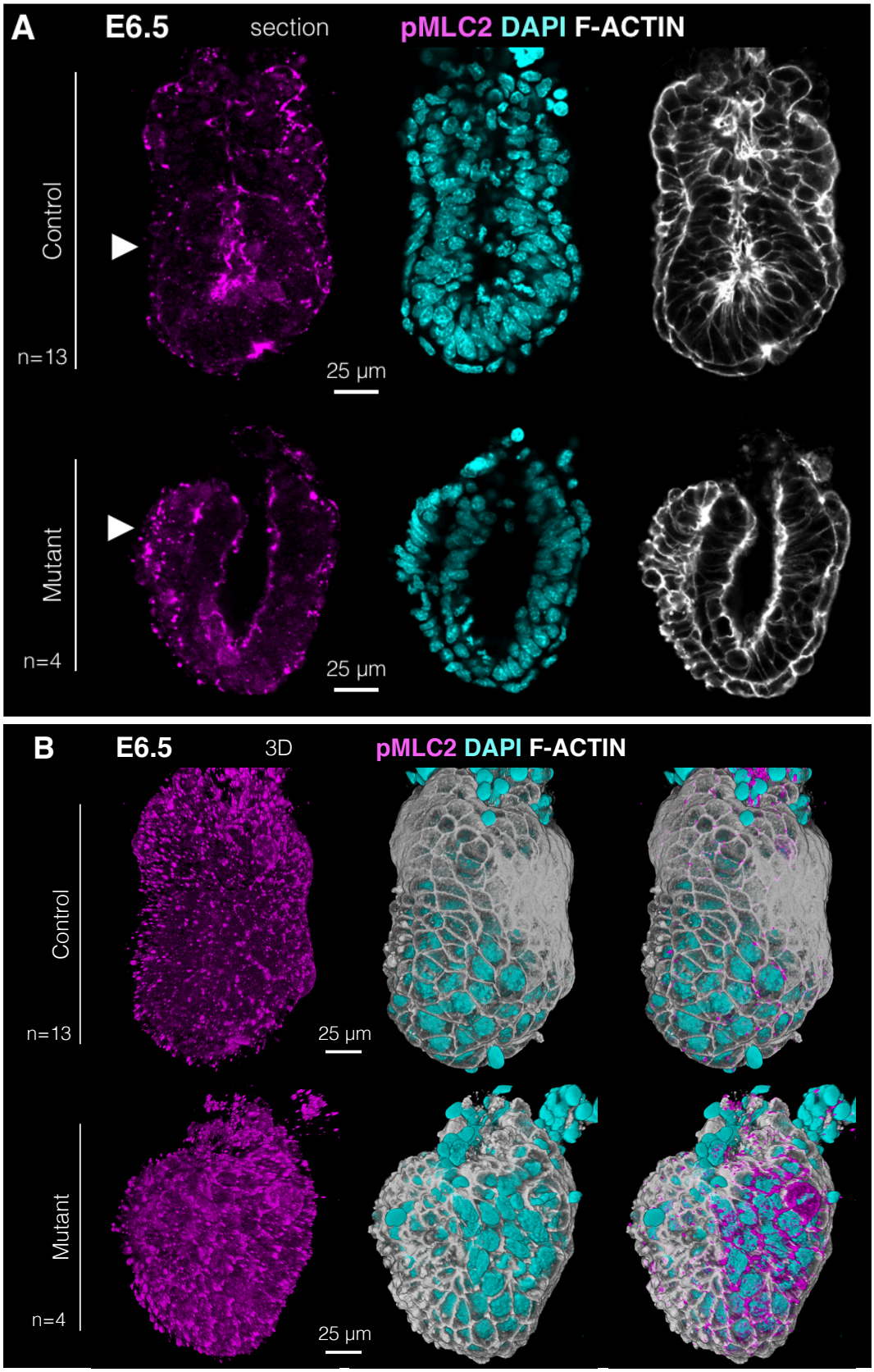


Figure 47. E6.5 immunofluorescent detection of active actomyosin contractility using pMLC2
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E6.5 of pMLC2 (magenta); F-ACTIN (gray); DAPI (cyan). Increased frequency of pMLC2-positive membrane accumulations in mutant AVE compared to control (clearer in section view; arrowhead). Scale bars: 25 µm.

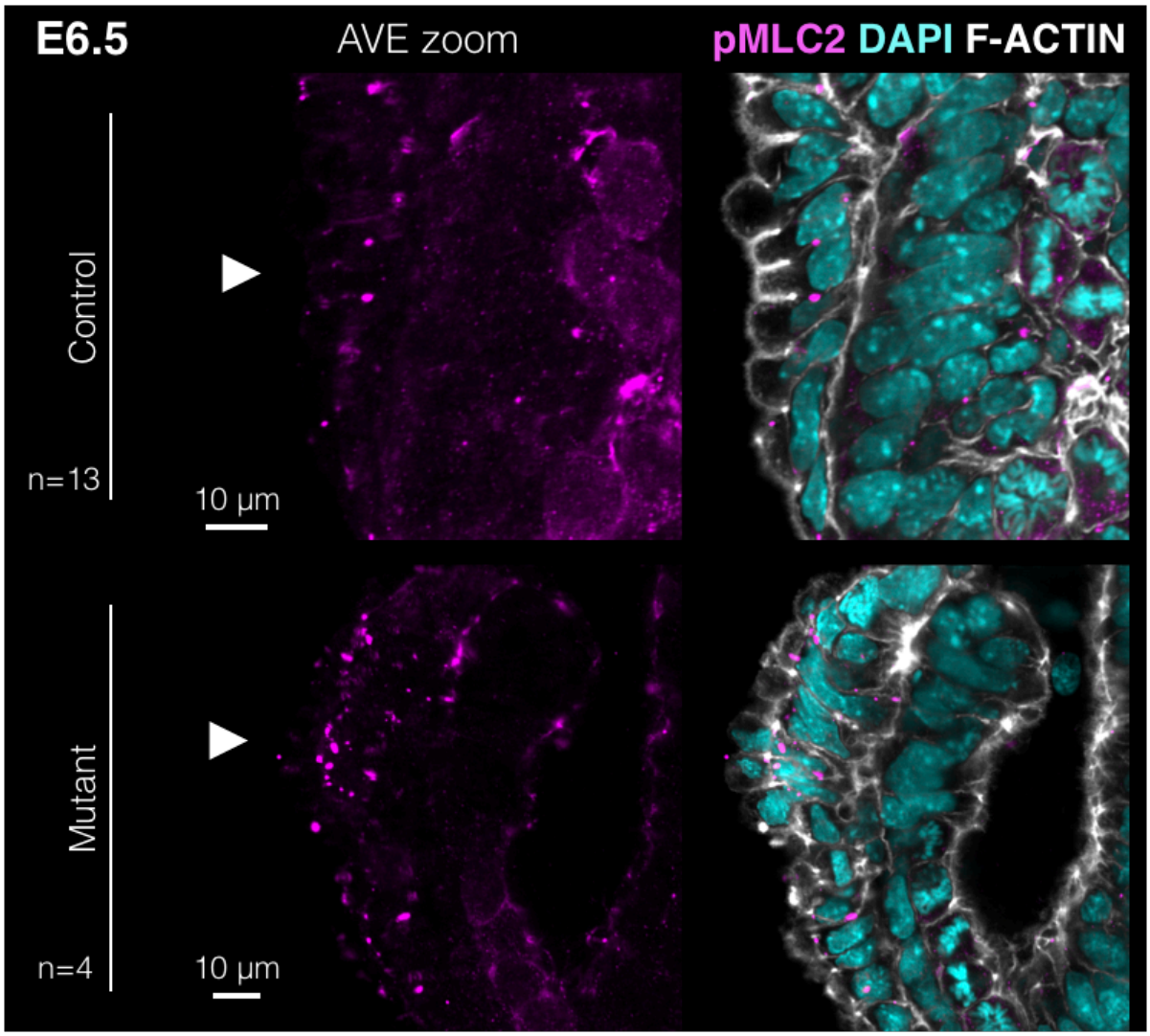


Figure 48. E6.5 AVE zoom of pMLC2 stain in control versus mutant
Representative mid-sagittal sections of AVE zoom from the same embryos presented in Figure 47, stained for pMLC2 (magenta); F-ACTIN (gray); DAPI (cyan). Increased frequency of pMLC2-positive membrane accumulations in mutant AVE compared to control (arrowhead). Scale bars: 10 µm.

4.2.4 LAMININ- and COLLAGEN IV-positive basement membrane is thickened beneath mutant AVE cells

As apical-lateral polarity was unperturbed, I next checked whether the basement membrane was abnormal in mutants. I assessed basement membrane by immunofluorescent staining analysis of LAMININ (Figure 49, 50) and COLLAGEN IV (Figure 52) in control versus mutant. I found the basement membrane beneath mutant AVE cells to be drastically thickened by E6.5 (Figure 50), whereas there were no differences at E5.5 (LAMININ stain only; Figure 49). The LAMININ-positive basement membrane is normally already higher in

the anterior, beneath AVE cells, compared to posterior by late post-implantation (seen in E6.5 control example – Figure 50; Figure 51). Therefore, the thickened basement membrane found beneath E6.5 mutant AVE cells could be a more severe manifestation of a pre-existing upregulation of basal LAMININ in the anterior. In mutants, the degree of basement membrane thickness was not consistent throughout the AVE (i.e. some regions had thicker basement membrane than others within an embryo), which suggests heterogeneity within the mutant AVE that in turn differentially influences its extracellular environment.

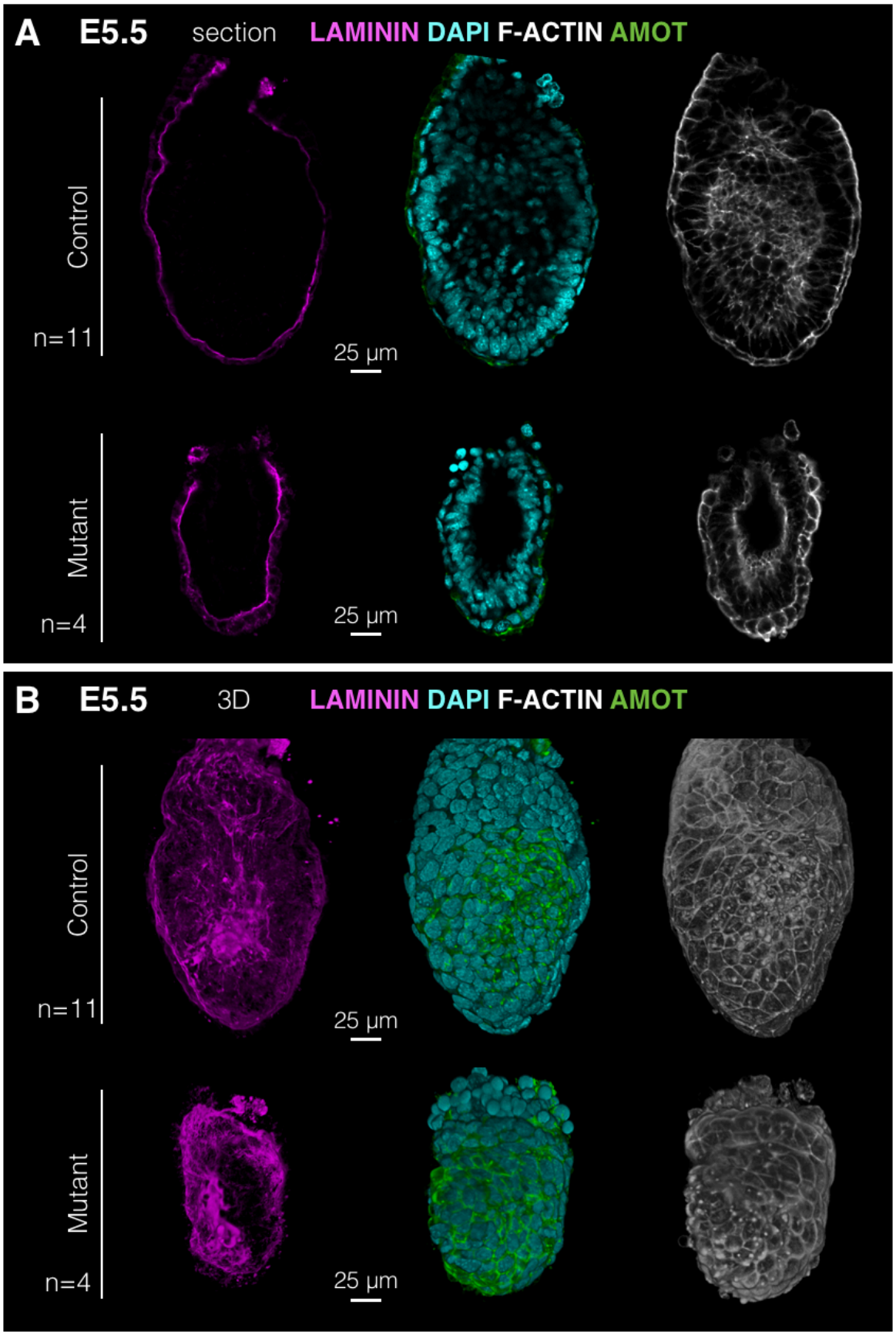


Figure 49. LAMININ basement membrane marker expression at E5.5
Representative mid-sagittal sections (A) and 3D opacity renderings (B) of whole mount immunofluorescence at E5.5 of LAMINN (magenta) co-stained with AMOT (green – AVE); DAPI (cyan). LAMININ retained basement membrane localisation in the VE and AVE of mutants. Scale bars: 25 µm.

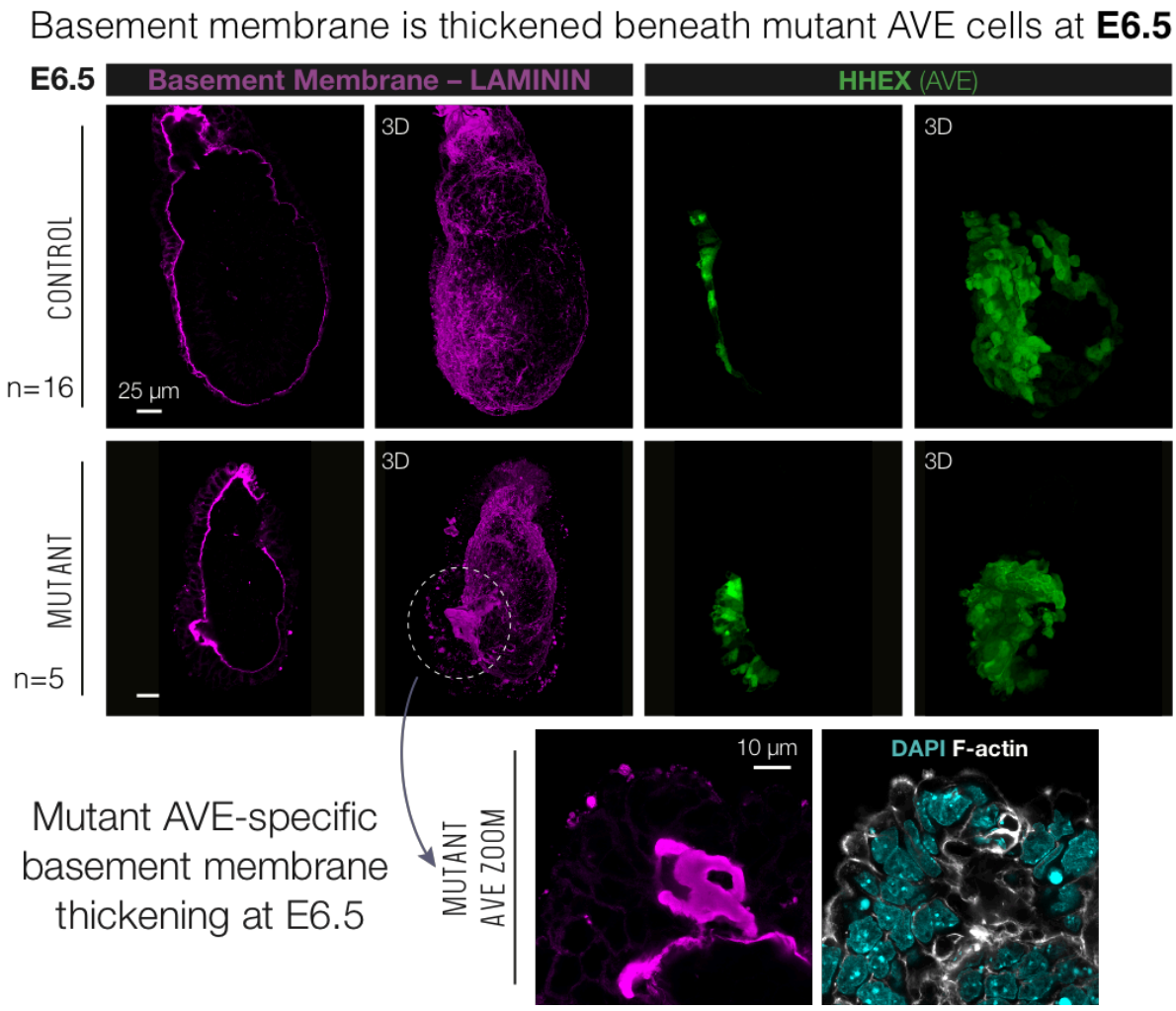


Figure 50. LAMININ basement membrane marker expression at E6.5
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescence at E6.5 of LAMINN (magenta); HHEX-GFP (green – AVE); F-ACTIN (gray); DAPI (cyan). LAMININ retained basement membrane localisation in the VE and AVE of mutants, but was thickened beneath mutant AVE cells (shown as AVE zoom in mutant). Scale bars: 25 µm, or 10 µm as indicated.

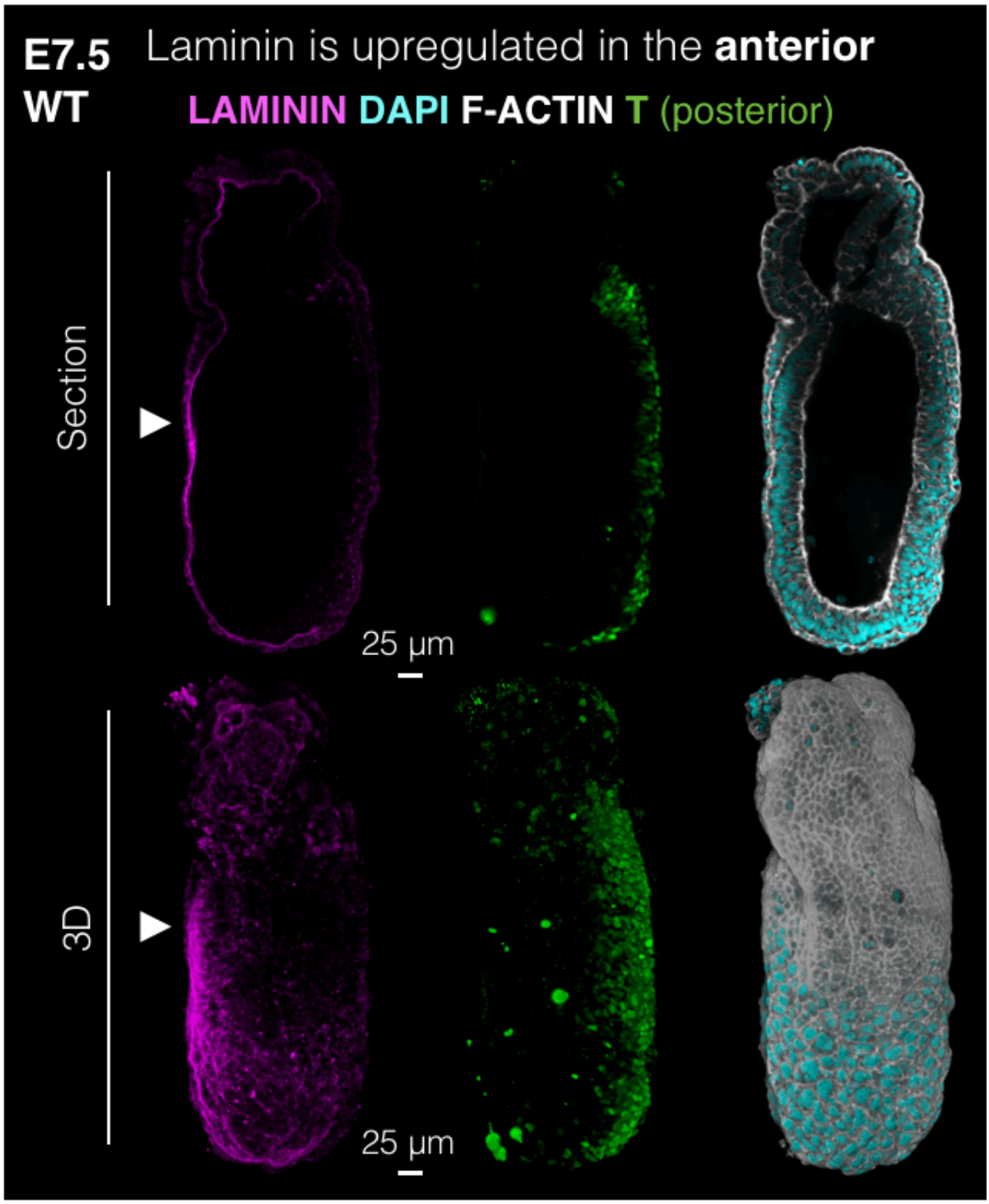


Figure 51. LAMININ expression in E7.5 WT showing upregulation in the anterior
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescence at E7.5 of LAMININ (magenta) co-stained with T (green – posterior marker); F-ACTIN; DAPI (cyan). LAMININ was visibly upregulated in the anterior (arrowhead) of WT embryos at E7.5, a stage where A-P axis has been established and gastrulation is underway in the posterior. Scale bars: 25 µm.

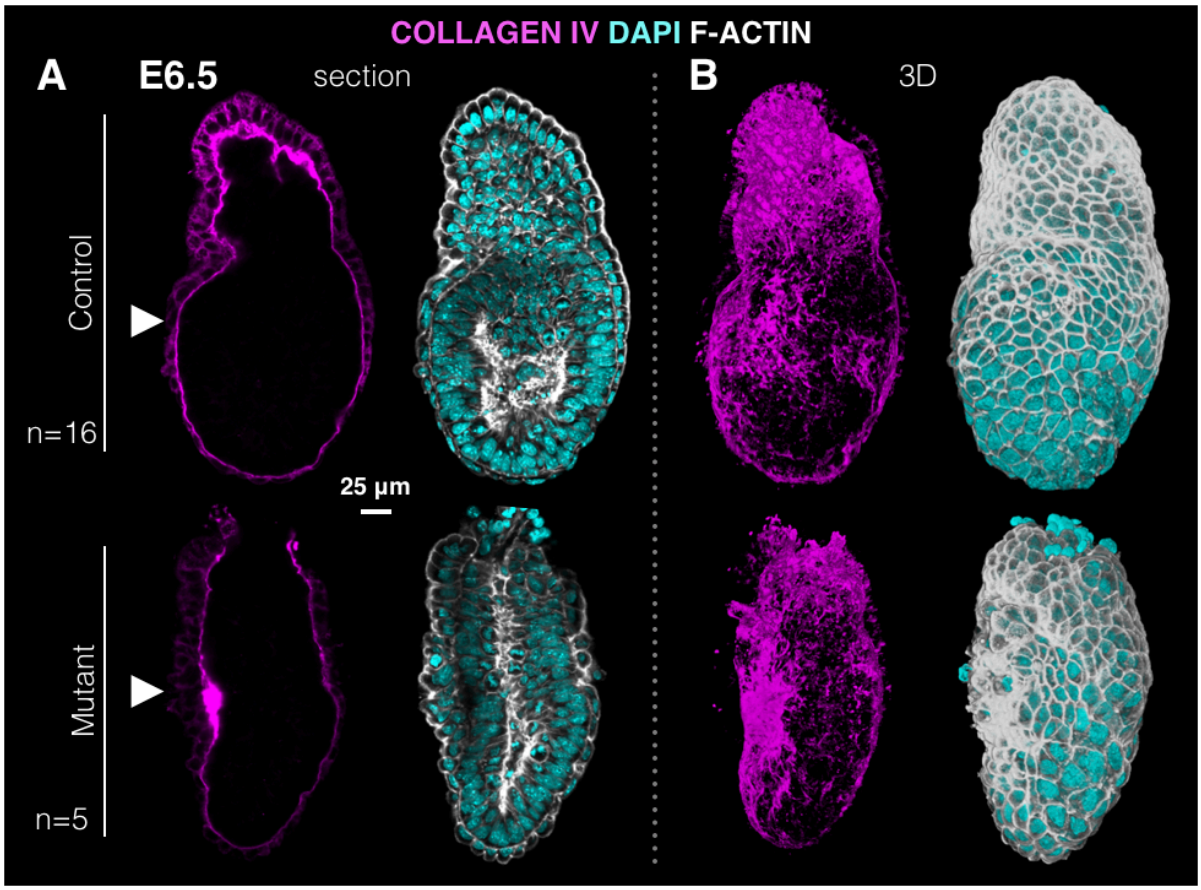


Figure 52. COLLAGEN IV basement membrane marker expression at E6.5
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescence at E6.5 of COLLAGEN IV (magenta); F-ACTIN (gray); DAPI (cyan). COLLAGEN IV retained basement membrane localisation in the VE and AVE of mutants, but was thickened beneath mutant AVE cells compared to control (arrowhead). Scale bars: 25 µm.

In summary, thorough characterisation of apical-basolateral polarity in mutants demonstrated preservation of intact polarity protein localisation throughout the VE, as well as in the thickened mutant AVE. However, F-ACTIN and pMLC2 were found to be upregulated in the E6.5 mutant AVE; and LAMININ- and COLLAGEN-IV-positive basement membrane were visibly thickened beneath E6.5 mutant AVE cells.

4.2.5 Mutant AVE epithelial stratification and increased cell number by E6.5

To assess whether the thickened mutant AVE retained simple epithelial organisation, i.e. with all cells attached to the basement membrane, I conducted Imaris 3D cell outlining of the AVE using high resolution Airy scan images of F-ACTIN. I found the mutant AVE thickening to be characterised by epithelial stratification (Figure 53).

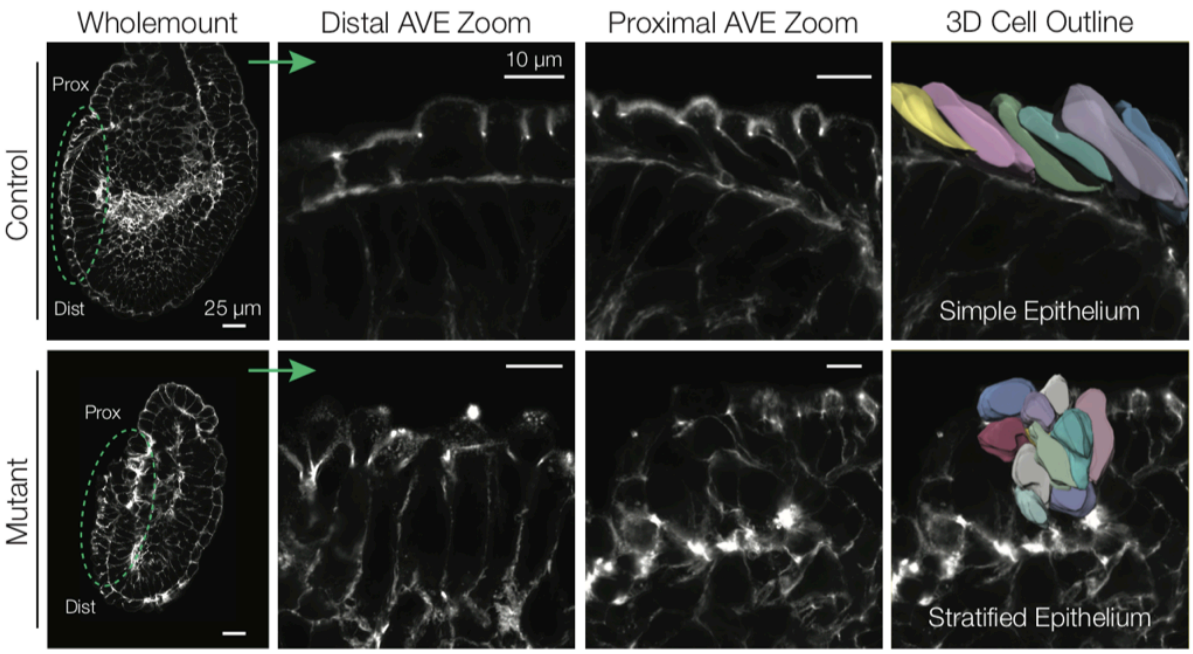


Figure 53. Thickened mutant AVE is characterised by epithelial stratification
Representative mid-sagittal sections of control versus mutant embryos at E6.5 with AVE zoom. Distal (dist) mutant AVE is excessively columnar compared to control, whereas proximal (prox) mutant AVE is more severely disorganised. 3D Imaris outlining showed presence of mutant AVE cells detached from the basement membrane indicating a bona fide multilayered epithelium, as opposed to pseudostratified (with a stratified appearance). AVE cells in controls remained as a monolayered simple epithelium.

Imaris outlining provided additional metrics on the geometric differences of the mutant AVE, which revealed significant reduction in cell volume and cell surface area (Figure 54), while there was a significant increase in normalised F-ACTIN intensity (Figure 46B).

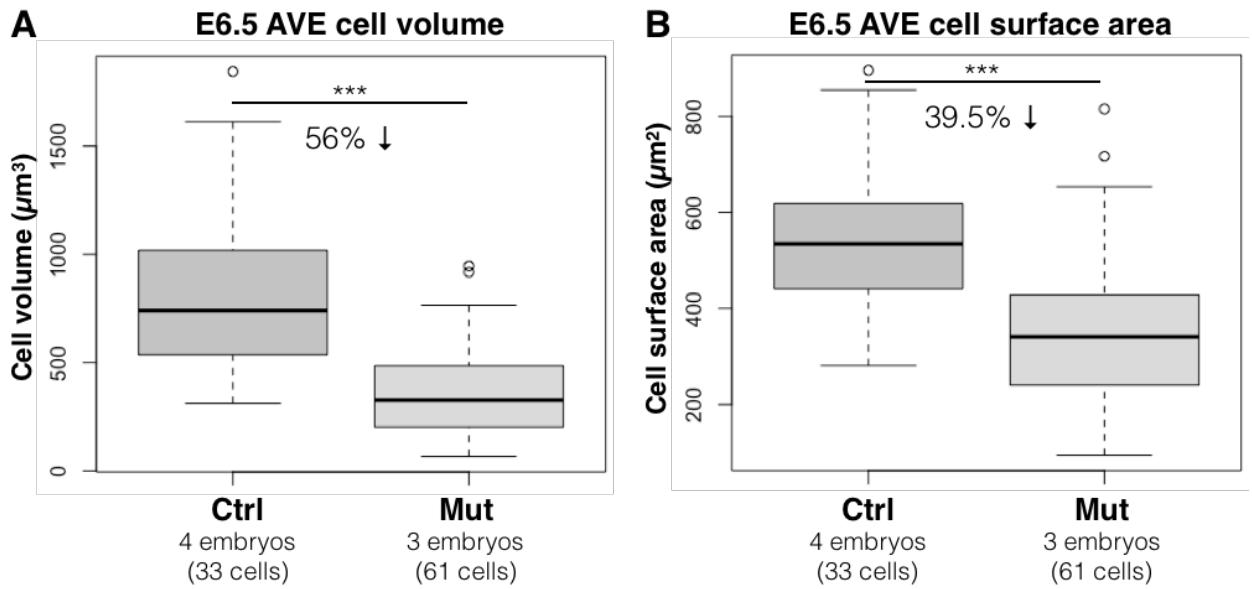


Figure 54. E6.5 mutant AVE cells have decreased volume and surface area
Significant reduction in E6.5 mutant AVE cell volume (A) and surface area (B). (Statistical test: hierarchical ANOVA; *** $p < 0.001$).

To investigate whether the mutant AVE thickening was due to increased cell number, or whether the same number of AVE cells were being tightly packed together in a smaller surface area, I quantified the number of AVE cells using AMOT as a ubiquitous AVE marker. I found normal AVE cell number at E5.5, a stage when the AVE thickening phenotype is not yet evident. By E6.5, I detected a significant 125.5% increase in mutant AVE cell number concomitant with the overt thickening phenotype (Figure 55).

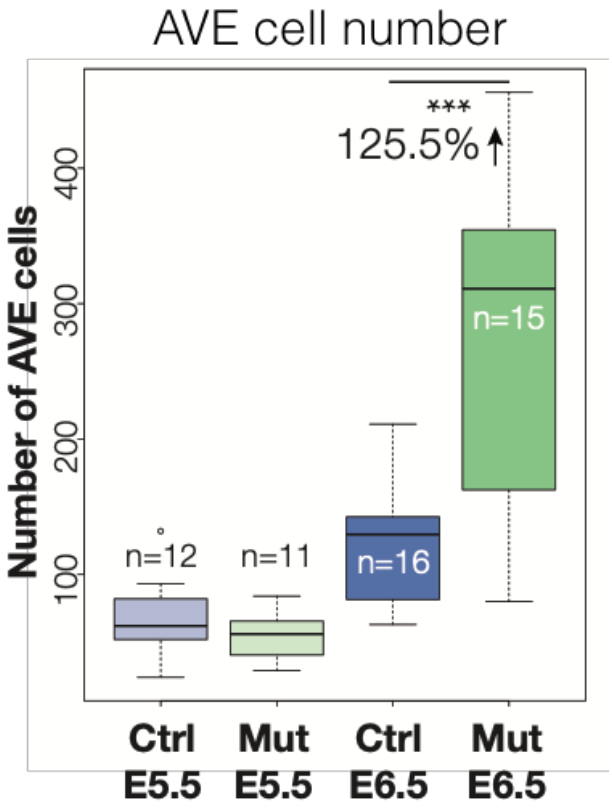


Figure 55. Increased mutant AVE cell number by E6.5
AVE cell number in control versus mutant at E5.5 and E6.5. No significant difference between control and mutant AVE cell number at E5.5, but a significant elevation in mutant AVE cell number at E6.5. AMOT was used to mark AVE cells. Each embryo analysed was co-stained with DAPI and F-ACTIN to ascertain individual nuclei and cell boundaries, respectively. (Statistical test: hierarchical ANOVA. *** $p < 0.001$).

4.2.6 Increased mutant AVE proliferation at E5.5

To test whether the increase in AVE cell number by E6.5 is due to increased proliferation at an earlier stage, I quantified the proliferative index at E5.5 using PHH3 to mark cells in all active phases of mitosis (Kim et al., 2017), as well as ALPHA-TUBULIN co-stained with DAPI to visualise dividing mitotic nuclei. Again, I used AMOT as the AVE marker. The proliferative index was defined by the number of PHH3, or mitotic AVE cells divided by total number of AVE cells per stage. I additionally quantified proliferative index at E5.0 and E6.5 stages to pinpoint the time window during which growth rates of AVE cells are affected in the absence of ETS2. At E5.0 when the AMOT marker is evenly junctional in distribution throughout the Epi-VE, I considered the distal half of the embryonic portion as the 'distal VE'. I found no difference in E5.0 distal VE cell number (Figure 56A) and proliferative index

between mutant versus control (Figure 56B). At E5.5, I detected a significant increase in mutant AVE proliferation (Figure 57). By E6.5, mutant AVE proliferation subsided and was not significantly different from control. These data demonstrate that the mutant AVE undergoes a burst of proliferation at the specific time frame of early post-implantation stage E5.5.

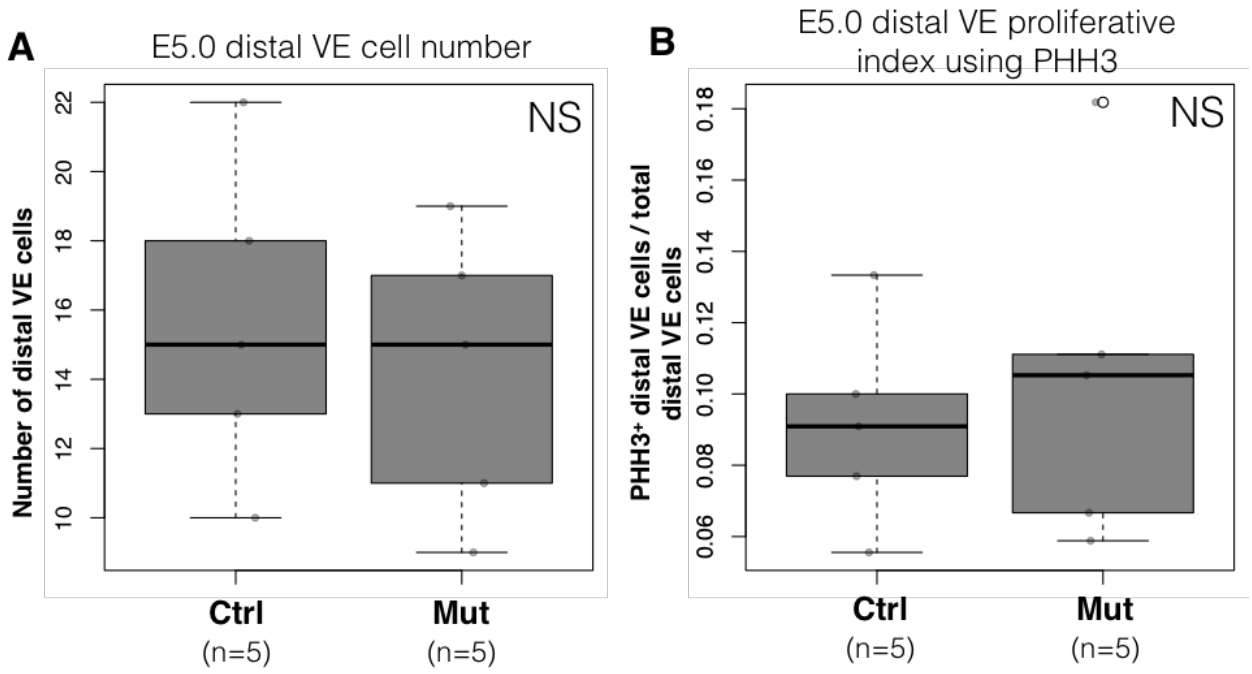


Figure 56. No significant difference in proliferation in the E5.0 mutant AVE
Distal VE cell number (A) and PHH3 proliferative index (B) at E5.0. No significant differences were detected between mutant versus control. (Statistical test: hierarchical ANOVA. NS = not significant).

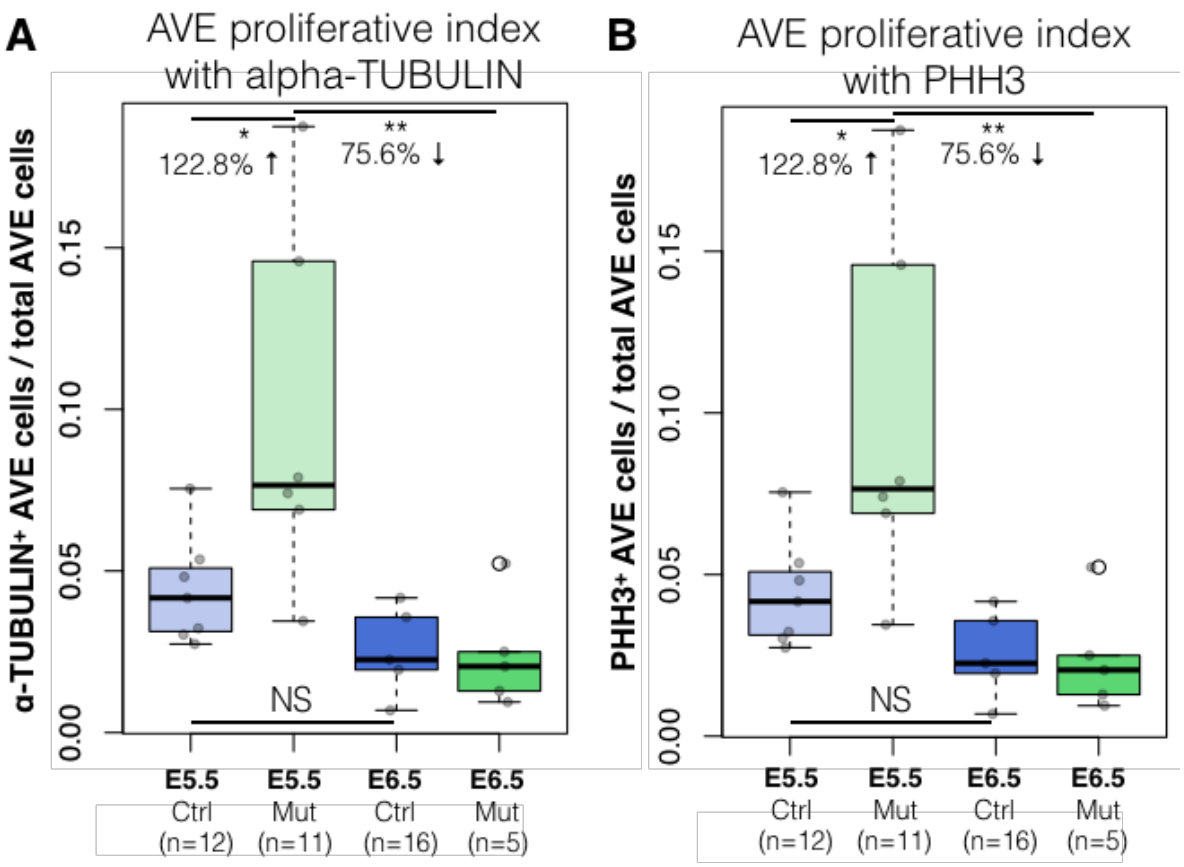


Figure 57. Increased proliferation in the mutant AVE at E5.5
Quantification of proliferative index in AMOT-positive AVE cells using (A) ALPHA-TUBULIN to mark mitotic spindles for visualisation of dividing cells, and (B) PHH3 to mark cells in all active phases of mitosis. Proliferative index was taken as the number of proliferating AVE cells divided by total number of AVE cells. Both methods of quantification showed similar results: a significant increase in mutant AVE proliferation compared to control at E5.5. This was accompanied by a significant decrease in mutant AVE proliferation at E6.5 compared to that in the mutant at E5.5, suggesting an intact sensing mechanisms in the mutant to detect and respond to over-proliferation. There is no significant decrease in AVE proliferation in control E6.5 compared to control E5.5 AVE. (Statistical tests: hierarchical ANOVA and Tukey. NS = not significant; * $p < 0.05$; ** $p < 0.01$).

4.2.7 Increased proportion of perpendicular cell division in the mutant at E6.5

Thus far, I have found that the E6.5 thickened mutant AVE is characterised by epithelial stratification and increased cell number. This phenotype is preceded by elevated mutant AVE proliferation at E5.5. I next asked whether misorientation of cell division angle additionally contributes to mutant AVE epithelial stratification. To do so, I quantified orientation of cell division in control versus mutant AVE at E5.5 and E6.5. I used ALPHA-TUBULIN to visualise mitotic spindles and DAPI to detect mitotic nuclei by morphology; co-stained with AMOT to mark AVE cells in whole mount immunostains (Figure 58A). I resliced the embryo for each

dividing cell under consideration such that the opposite poles of the mitotic spindle were in the same 2D plane. This allowed me to conduct angle measurements in 2D, where I drew a line tangent to the basement membrane and another line across either centrosome poles of the dividing cell, then measuring the angle created by these two lines using Fiji (detailed in 2.9). While I found no difference in mutant AVE orientation of cell division at E5.5 (Figure 58B), I detected a significant increase in perpendicular cell divisions by E6.5 (Figure 58C). Significance calculation was conducted both by comparing overall cell divisions (boxplot; hierarchical ANOVA), as well as with circular statistics (addition of Fisher's exact test) in the form of a rose plot where the angles of division were divided into bins of biological relevance: 0-30° = planar, 30-60° = oblique, 60-90° = perpendicular (Luxenberg et al., 2010). In controls, in line with normal maintenance of simple epithelia, cell divisions were mostly planar with a small proportion of oblique and no perpendicular. In mutants at E6.5, there was a 'bimodal' distribution of approximately half planar and half orthogonal, similar to the stratifying mouse embryonic skin at E16.5-E17.5 (Williams et al., 2014). There were no significant differences in the distribution of cell division orientation between proximal versus distal mutant AVE (Figure 59). Overall, in addition to the increase in mutant AVE proliferation, there was also an increase in the proportion of perpendicular cell divisions. This, however, occurred at a later stage than the burst in mutant AVE proliferation at E5.5, suggesting that misorientation of cell division angle could be a secondary effect.

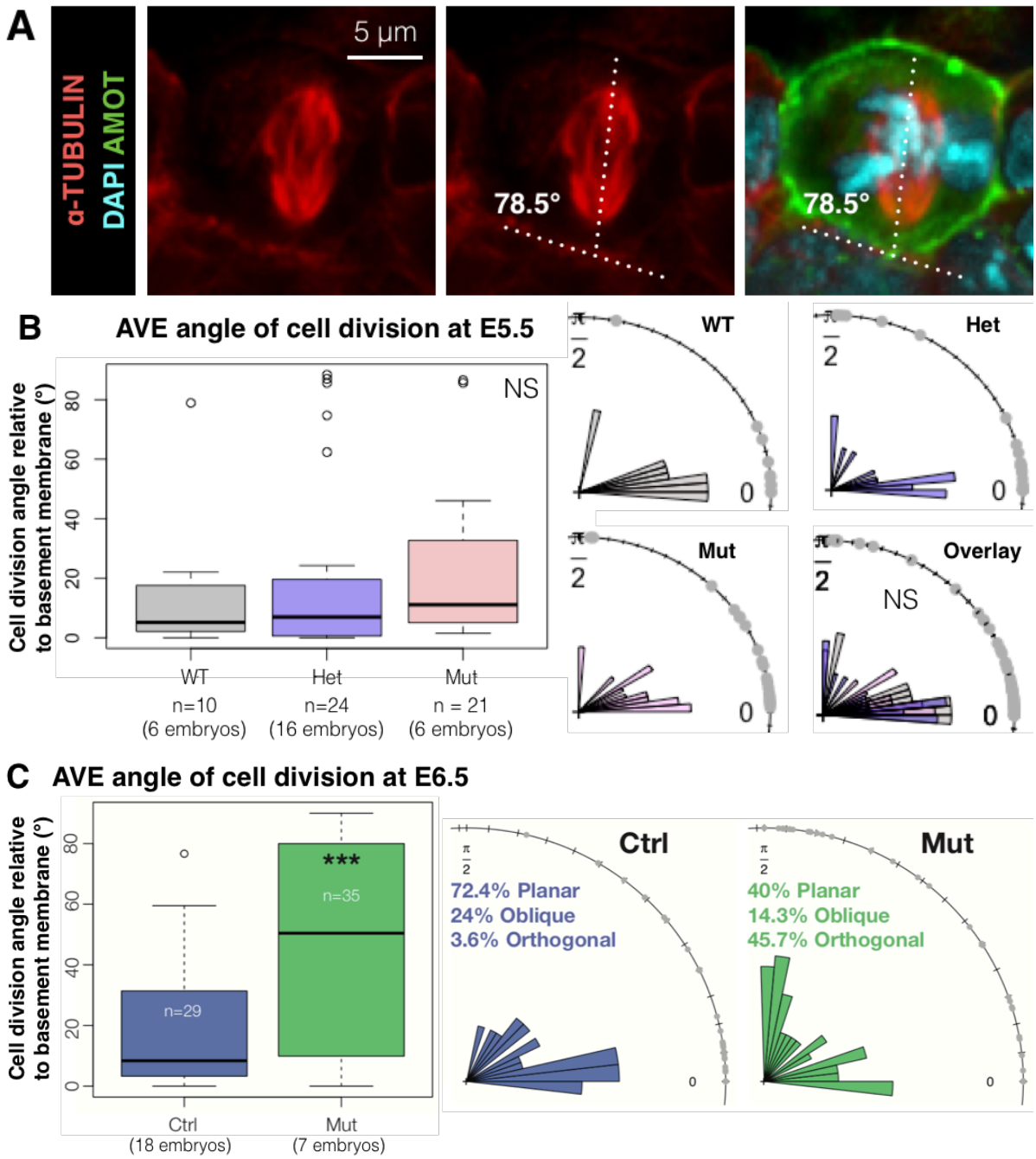


Figure 58. Increased proportion of perpendicular cell divisions in mutant AVE cells at E6.5
(A) Measurements for orientation of cell division in the AVE (AMOT-positive – green) was conducted using ALPHA-TUBULIN (red) in sections to visualise the mitotic spindle. The angle measurement was made using the intersection of a line tangent to the basement membrane and another line that connects the two poles of the mitotic spindle. DAPI (cyan). Scale bar: 5 µm.
(B) Orientation of cell division at E5.5 was not significantly different between control and mutant.
(C) Significantly increased proportion of perpendicular cell divisions in mutant AVE cells at E6.5. (Statistical tests: hierarchical ANOVA and Fisher's exact test. NS = not significant; *** $p < 0.001$).

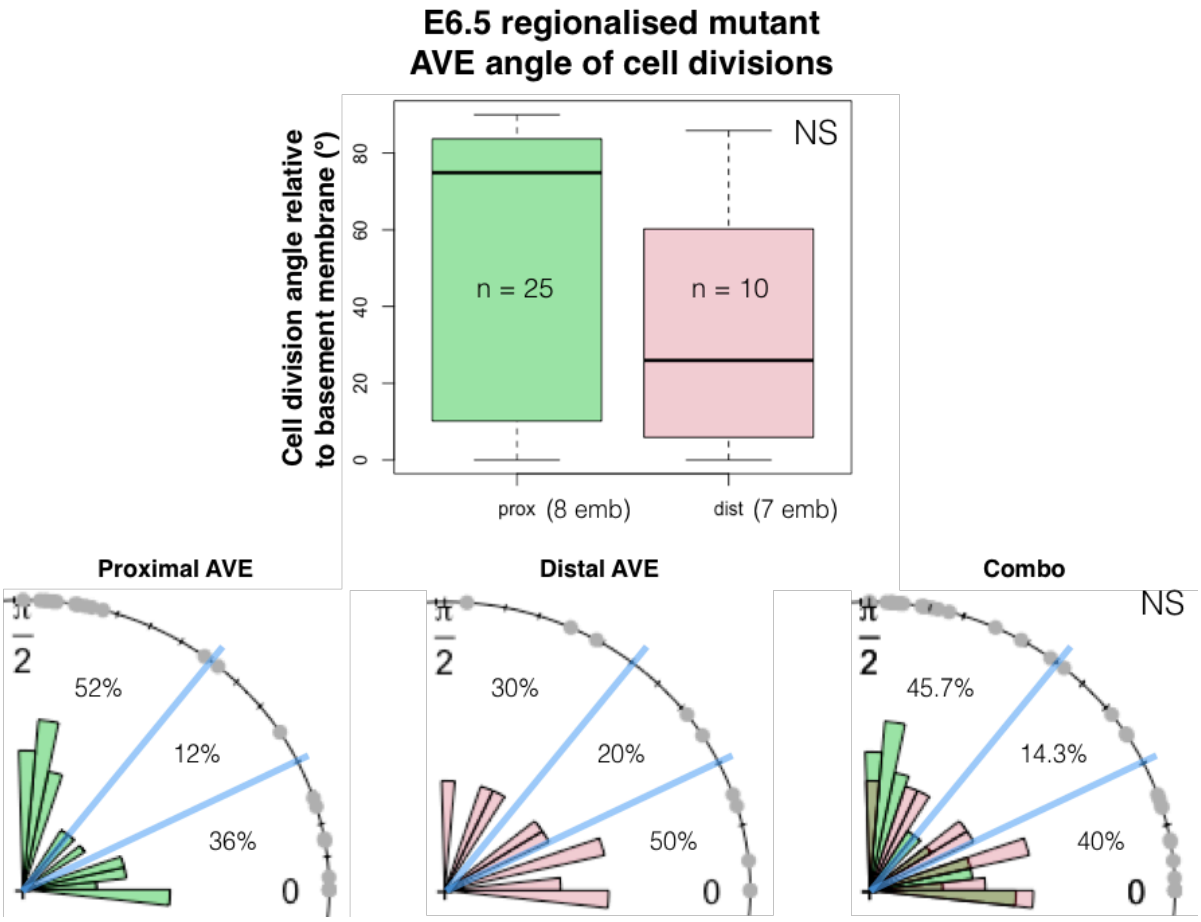


Figure 59. Regionalised mutant AVE angle of cell division at E6.5
Proximal versus distal mutant AVE at E6.5 do not significantly differ in distribution of cell division orientation. (Statistical tests: hierarchical ANOVA and Fisher's exact test. NS = not significant).

4.2.8 Increased mutant AVE cell crowding by E6.5

To further quantitatively describe the mutant AVE, I asked whether there was increased cell density, similar to that observed in *Ets2*-deficient intestinal epithelia (Munera et al., 2011). I quantified cell crowding in a regionalised manner where proximal versus distal AVE was segregated by dividing the distance along the epiblast on which the AVE migrates by half. Cell crowding was measured by taking the number of nuclei in 3 representative central sections spaced 10 µm apart, divided by the distance those nuclei span (detailed in 2.9). This has been used as a measure of cell crowding and packing to characterise pseudostratified epithelia, which has the developmental purpose of fitting more nuclei over a given surface area compared to simple epithelia (Lee and Norden, 2013; Figure 60A). At E5.5, there was no significant

difference in mutant AVE cell density (Figure 60B), whereas by E6.5, there was a significant increase in mutant AVE cell density in both proximal and distal domains (Figure 60C). This observation is in line with the burst of mutant AVE proliferation at E5.5 to result in more cells, as well as epithelial stratification, which would naturally fit an increased number of cells over a defined surface area. Furthermore, AVE migration, which is completed by E6.5 could contribute to increased cell density, i.e. presence of a cell migration end point could mean that trailing cells pile up on leading cells that have reached the endpoint. As the Hippo signalling pathway is known to regulate cell density and proliferation (Dupont, 2016), I examined the expression pattern and cellular localisation of Hippo signalling transcriptional effector YAP. YAP has been shown to display nuclear localisation throughout the VE with the exception of the AVE where YAP is excluded from the nucleus (Dr. Matthew Stower; Srinivas Group, unpublished data). I found no difference in YAP expression between control versus mutant at E5.5 (Figure 61A) and E6.5 (Figure 61B), suggesting that a YAP-independent mechanism regulates mutant AVE proliferation and cell density.

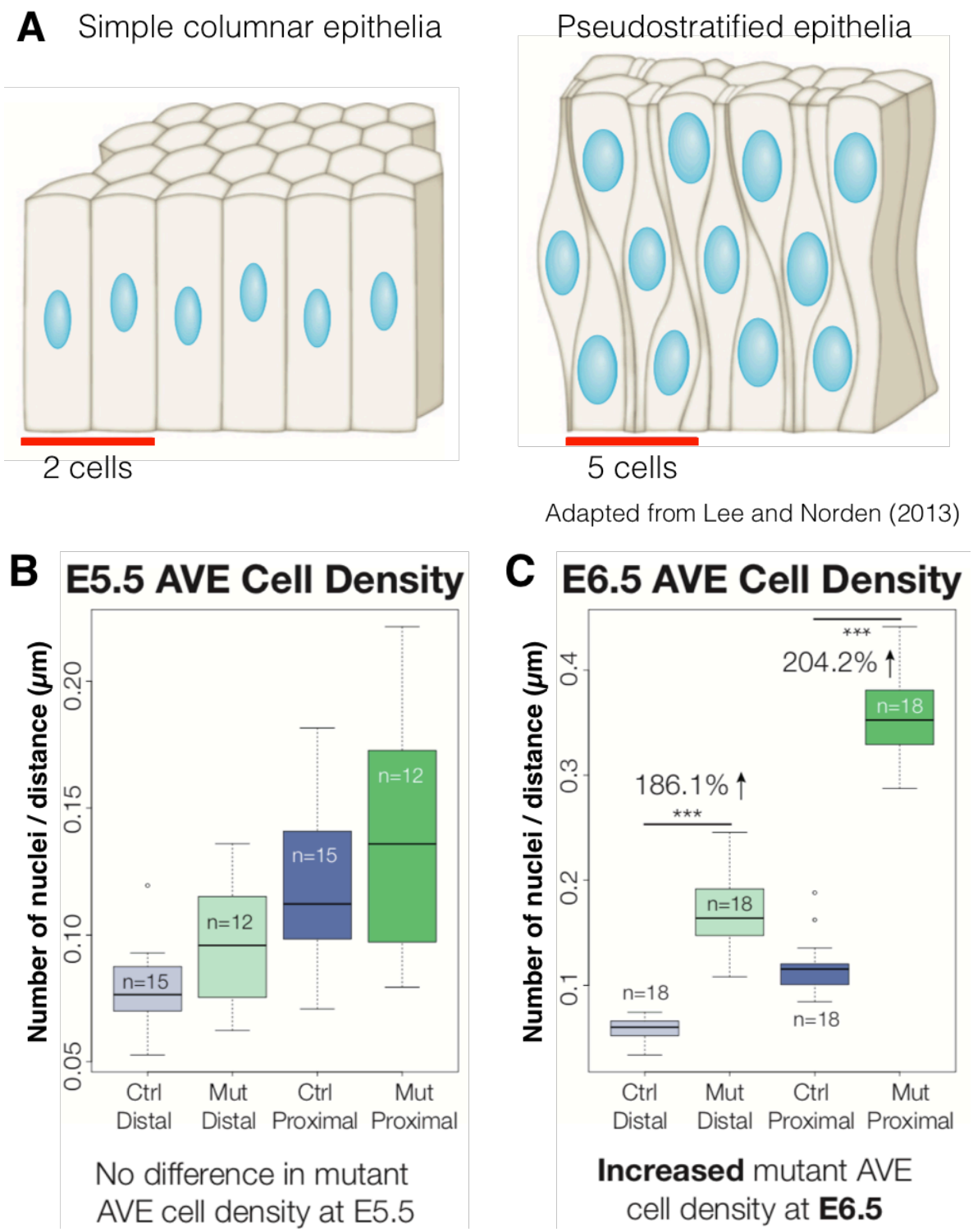


Figure 60. Regionalised AVE cell density quantification in control versus mutant at E5.5 and E6.5
Cell crowding was taken as the number of nuclei per distance along the epiblast (µm). (A) Schematic of more nuclei packed per unit length (shown as a red line for reference) in pseudostratified epithelia compared to monolayered (adapted from Lee and Norden, 2013).
(B) No significant difference in cell crowding in distal and proximal AVE at E5.5.
(C) Significant increase in mutant AVE cell crowding at E6.5 in both distal and proximal AVE. (Statistical tests: hierarchical ANOVA; *** $p < 0.001$).

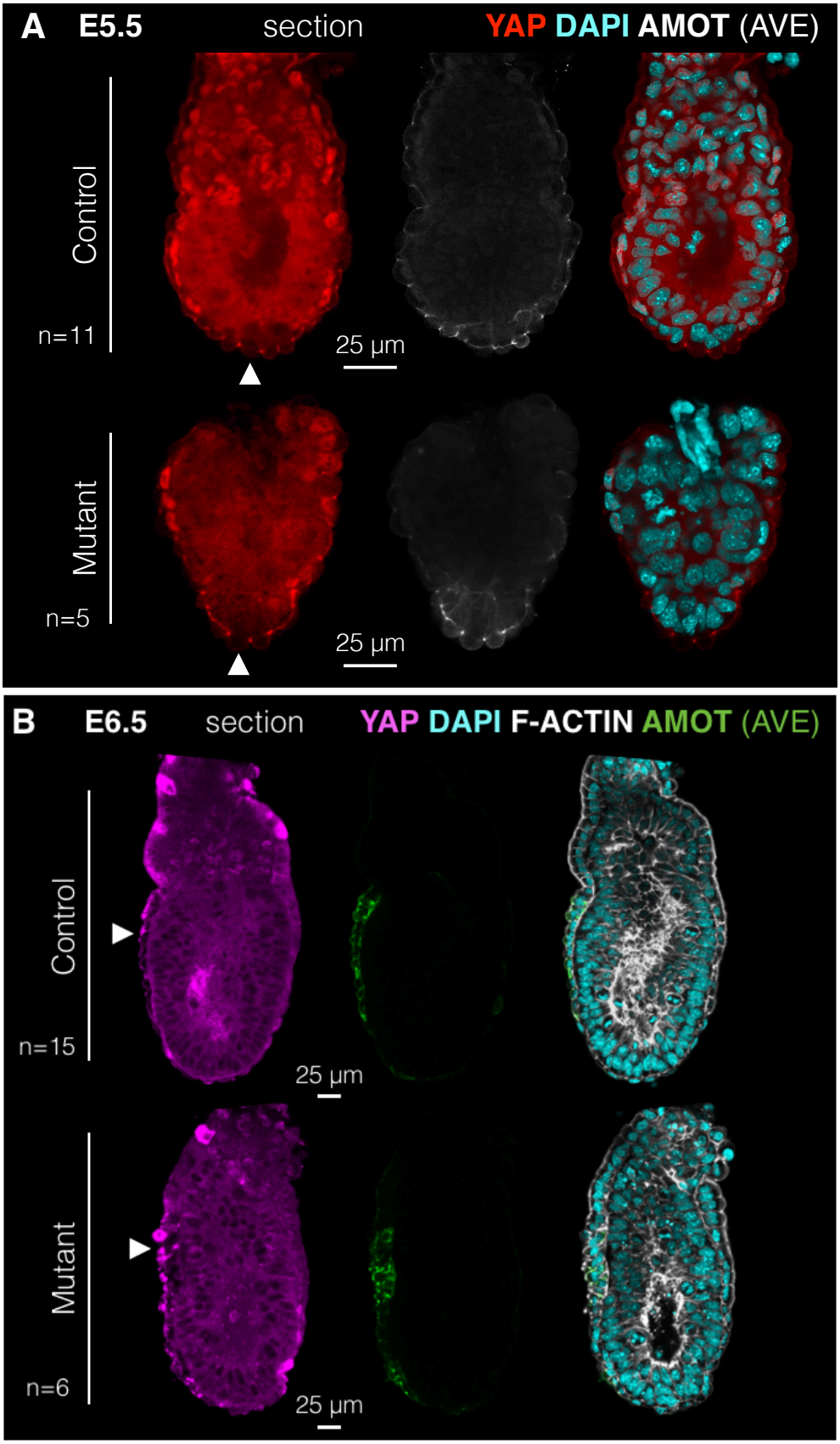


Figure 61. YAP is excluded from the nucleus of AVE cells in both control and mutant at post-implantation
Representative mid-sagittal sections at E5.5 (A) and E6.5 (B) of whole mount immunofluorescence of YAP (E5.5 red; E6.5 magenta) co-stained with AMOT (E5.5 gray; E6.5 green – AVE); F-ACTIN (gray); DAPI (cyan). YAP is excluded from the nucleus of AVE cells at E5.5 and E6.5 (arrowhead) in both control and mutants. Scale bars: 25 µm

4.2.9 Mutant AVE phenotype is consistent regardless of genetic background

All AVE phenotypic characterisations conducted had been verified using both the pure (129S6/SvEvTac, referred to as '129') and mixed (with *Hhex-GFP*; mix of CBA/J, C57BL/6, and 129) genetic backgrounds where I found no background-dependent differences. All AVE-related quantifications presented thus far had been conducted using the pure 129 background for consistency. To quantitatively compare the mutant AVE in pure versus mixed backgrounds, I quantified cell crowding and found no significant difference in the extent of crowding in the mutant AVE at E6.5 between the different backgrounds. Both backgrounds displayed significant elevation in mutant AVE cell crowding (Figure 62). Therefore, contrary to the ExE phenotype that demonstrated variable penetrance, the AVE phenotype was found to be unchanged in severity amongst the two genetic backgrounds tested, suggesting that the underlying biology for the AVE phenotype is not sensitive to strain differences.

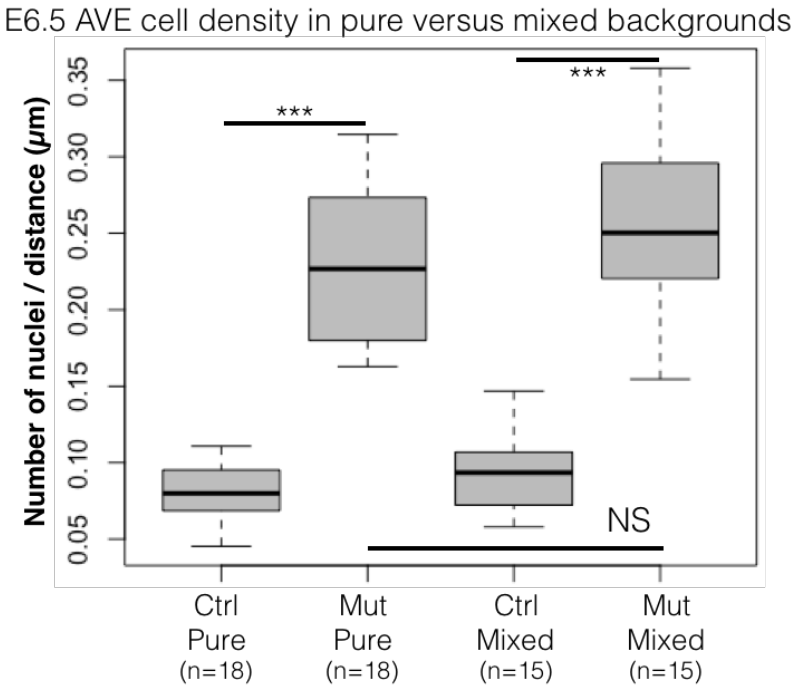


Figure 62. Thickened mutant AVE phenotype is consistent in both pure and mixed genetic backgrounds
Quantification of AVE cell crowding at E6.5 in control versus mutants in pure and mixed genetic backgrounds. Both backgrounds displayed a significant increase in mutant AVE cell crowding. (Statistical test: hierarchical ANOVA. NS = not significant; *** $p < 0.001$).

4.2.10 Mutant AVE phenocopy upon treatment with two independent inhibitors of known ETS2 upstream activator MEK

Having extensively characterised the mutant AVE phenotype, I next asked whether I could recapitulate the defect by inhibiting known upstream signals of ETS2 activation. ETS2 becomes activated by phosphorylation at its T72 residue situated within its signalling domain, which triggers its nuclear translocation and enables its DNA binding capability (Foulds et al., 2004; detailed in 1.2.2). Normally, ETS2 becomes activated downstream of ERK phosphorylation, which is activated by MEK (Figure 63A). I therefore used two MEK inhibitors that have validated efficacy in other contexts: PD0325901 (henceforth referred to as PD03), one of the i's in 2i for ESC culture (Ying et al., 2008); and U0126 tested in cardiac differentiation (Razy-Krajka et al., 2018). I cultured WT HHEX-GFP-positive embryos, where the AVE is marked by GFP expression, from E5.5 to E6.5 (the time window of mutant AVE defect progression) with either of these inhibitors to test whether I could recapitulate the mutant AVE defect (Figure 63B). PD03 and U0126 work in different ways to block MEK phosphorylation of ERK: PD03 is a non-ATP-competitive MEK1/2 inhibitor (0.5 μM-1 μM working concentration), and U0126 inhibits the kinase activity of MEK1/2 (10-50 μM working concentration).

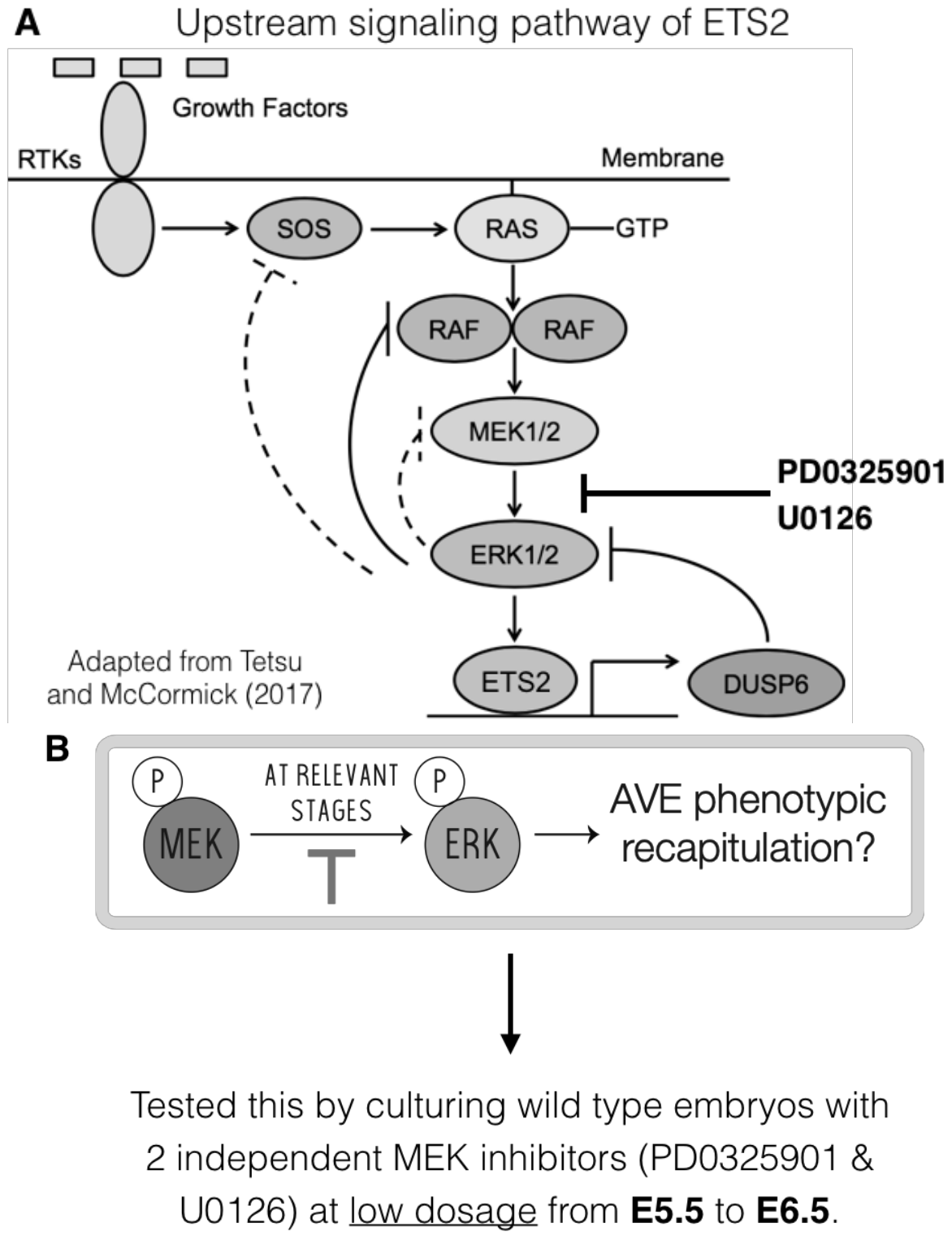


Figure 63. Inhibition of known upstream signalling cascade of ETS2 in WT embryos
(A) Schematic of MAPK signalling cascade that leads to ETS2 activation, showing presence of negative feedback loops, as well as the level at which MEK inhibitors target (adapted from Tetsu and McCormick, 2017).
(B) Schematic showing experimental design to inhibit MEK activation of ERK. 2 independent MEK inhibitors were used. Embryos were cultured from E5.5 to E6.5, which is the time window of mutant AVE phenotypic progression, to test whether inhibiting known upstream signals of ETS2 in WT could lead to mutant AVE defect recapitulation.

I was able to recapitulate the mutant AVE thickening phenotype using both inhibitors at the lower end of their recommended concentration: 1 μM for PD03 (Figure 64) and 10 μM for U0126 (Figure 65). I present an AVE zoom of the U0126 MEK inhibitor treated example from Figure 65 to highlight the severity of the AVE thickening defect recapitulation (Figure 66). I additionally found that I could titrate the defect, since 5 μM of U0126 led to a less severe AVE phenotype, i.e. less of a thickening (Figure 67), as well as a significantly milder effect on epiblast width morphometrics (Figure 68). A significant reduction of epiblast A-P width has been characterised in mutants at E6.5 (3.2.2).

To visualise the effect of MEK inhibition on ERK phosphorylation and, in turn, ETS2 activation, I conducted immunofluorescent staining analyses of pERK1/2 and ETS2 in both MEK inhibitor treated conditions and DMSO controls. I found ETS2 expression to be retained in the ExE of MEK inhibitor treated embryos (Figure 64B; PD03 example), suggesting that perturbation of ETS2 expression in the ExE is not required to result in an AVE thickening phenotype. However, the pERK1/2 signal in the ExE was diminished from low dose MEK inhibitor treatment (Figure 65B), demonstrating the requirement for a more tissue-specific approach to interrogate whether there is an AVE cell autonomous effect of MAPK/ETS2 signalling.

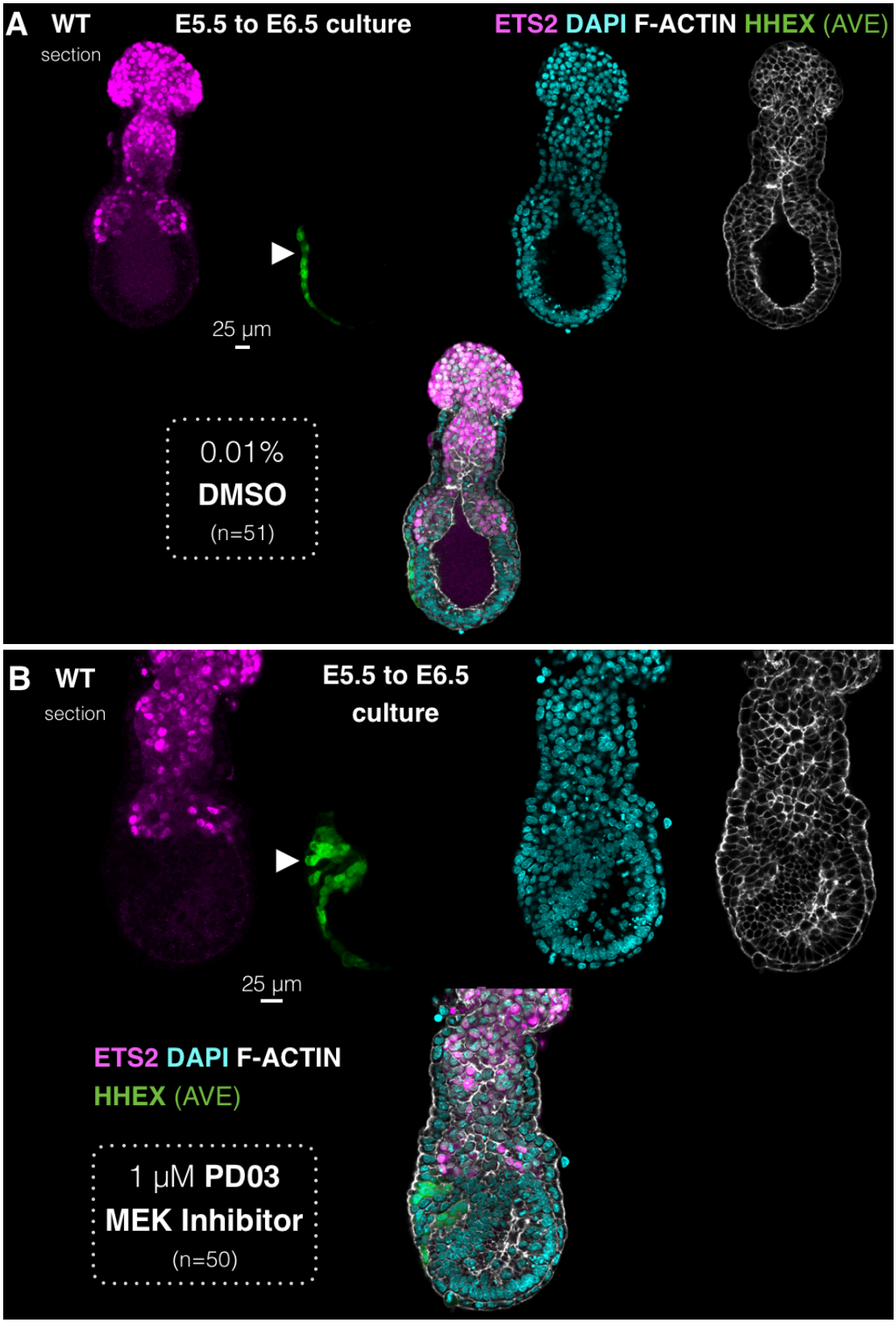


Figure 64. E5.5 to E6.5 PD03 (1 µM) MEK inhibitor culture of WT embryos resulted in mutant AVE phenotypic recapitulation

Representative mid-sagittal sections of whole mount immunofluorescence of 1 µM PD03-treated (A) versus 0.01% DMSO control (B). Inhibitor treatment led to recapitulation of the thickened mutant AVE phenotype in treated and not in DMSO control (arrowhead). ETS2 (magenta) ExE signal is retained in 1 µM treated (comparable to WT), suggesting that the AVE thickening phenotype may be independent of the ExE. HHEX-GFP (green – AVE); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

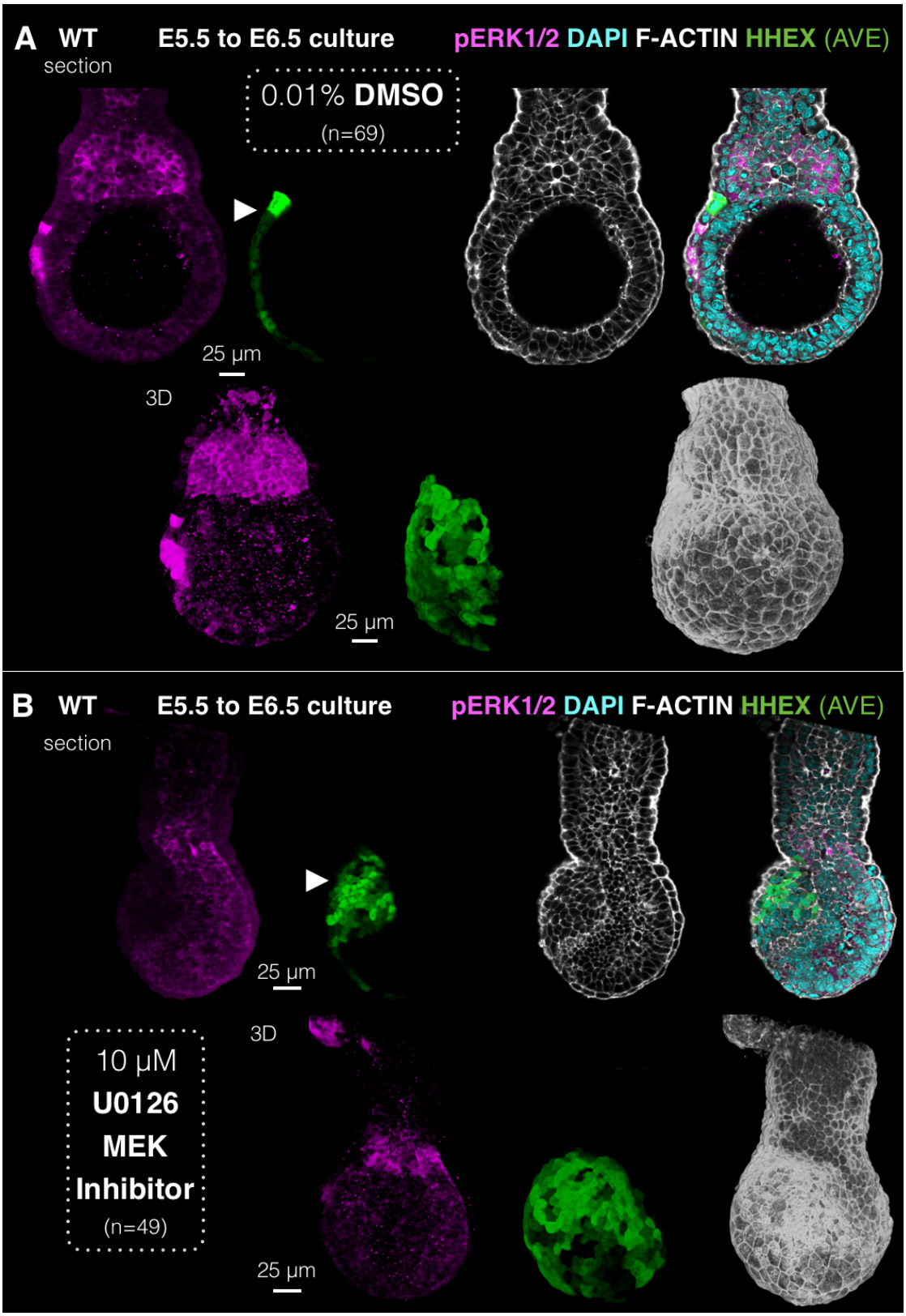


Figure 65. E5.5 to E6.5 U0126 (10 µM) MEK inhibitor culture of WT embryos resulted in mutant AVE phenotypic recapitulation

Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescence of 10 µM U0126-treated (A) versus 0.01% DMSO control (B). Inhibitor treatment led to recapitulation of the thickened mutant AVE phenotype in treated and not in DMSO control (arrowhead). ExE pERK1/2 (magenta) signal appeared diminished in the treated condition compared to control, demonstrating that whole embryo culture in low dose MEK inhibitor does affect the ExE to a certain degree. HHEX-GFP (green – AVE); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

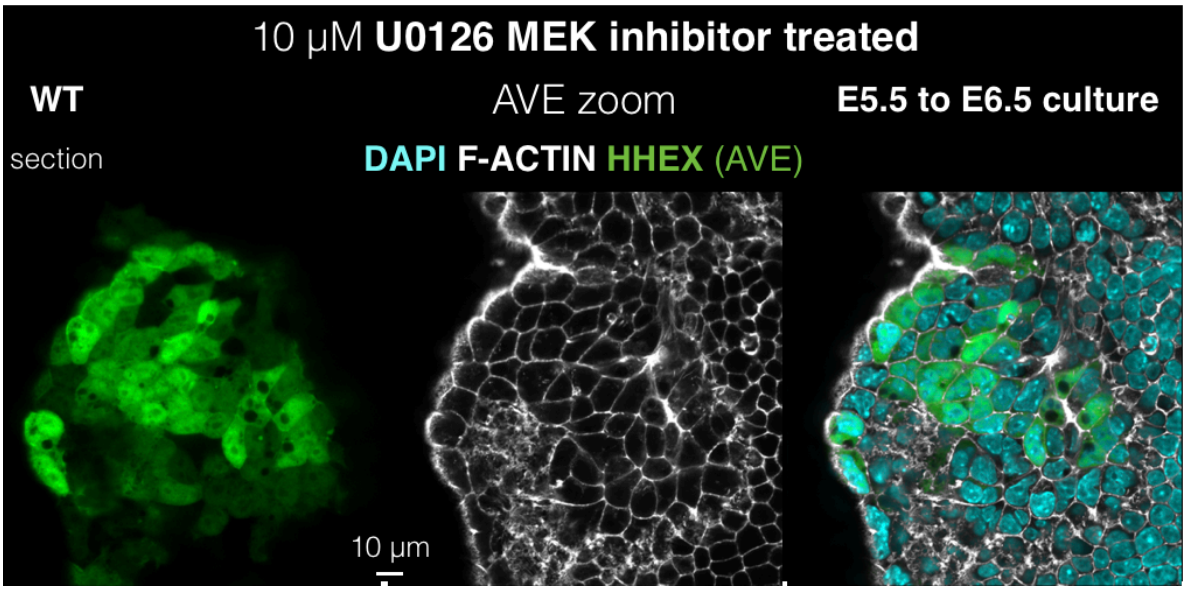


Figure 66. AVE zoom for U0126 (10 μM) MEK inhibitor treated highlighting the severity of the AVE thickening defect recapitulation
Representative mid-sagittal sections of zoomed in AVE for MEK inhibitor treated example presented in Figure 65. HHEX-GFP (green – AVE); F-ACTIN (gray); DAPI (cyan). Scale bar: 10 μm.

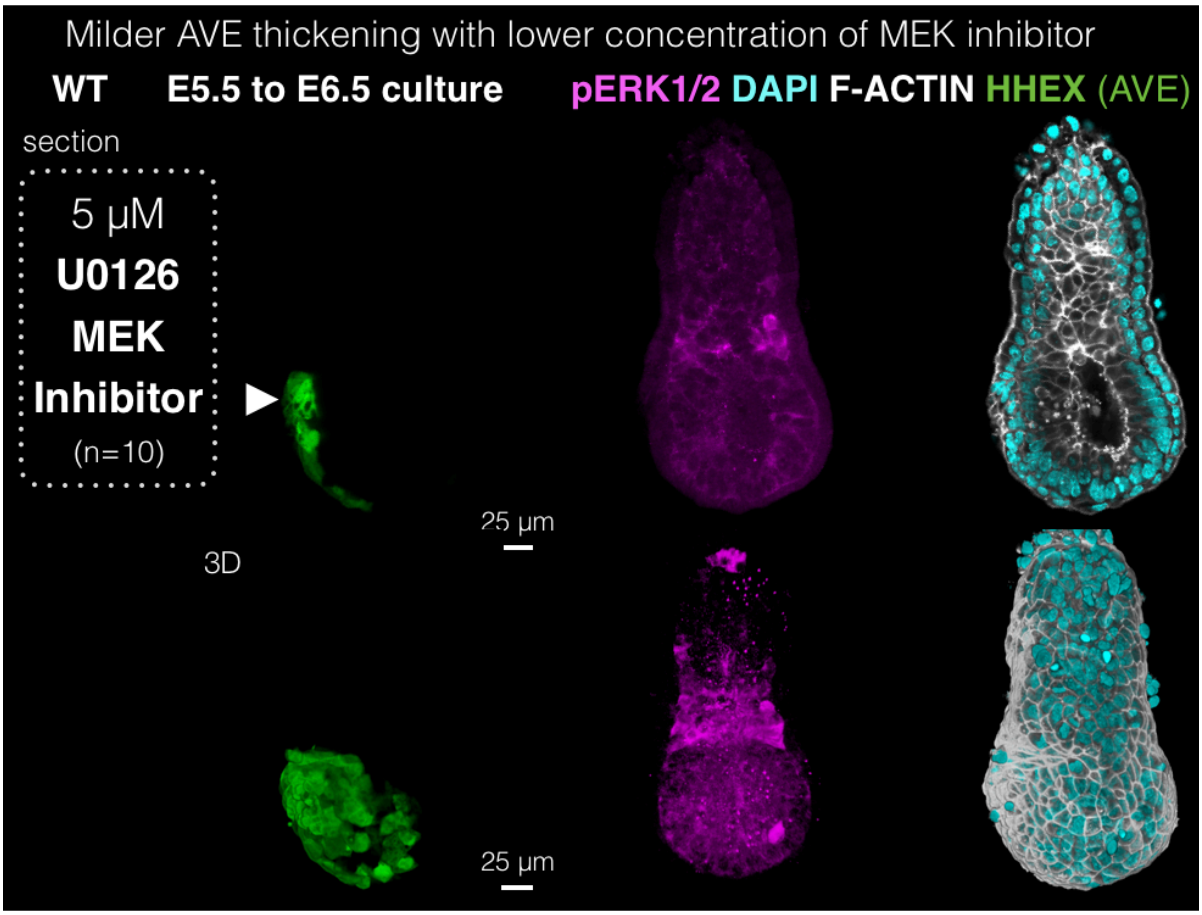


Figure 67. MEK inhibitor-mediated mutant AVE defect recapitulation in WT embryos can be titrated
Representative mid-sagittal sections of whole mount immunofluorescence and 3D opacity renderings of 5 μM U0126-treated WT where the mutant AVE defect recapitulation (arrowhead) is less severe compared to 10 μM U0126 treated (Figure 65, 66). ExE pERK1/2 (magenta) also appeared diminished to a less severe degree by qualitative assessment; HHEX-GFP (green – AVE); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 μm.

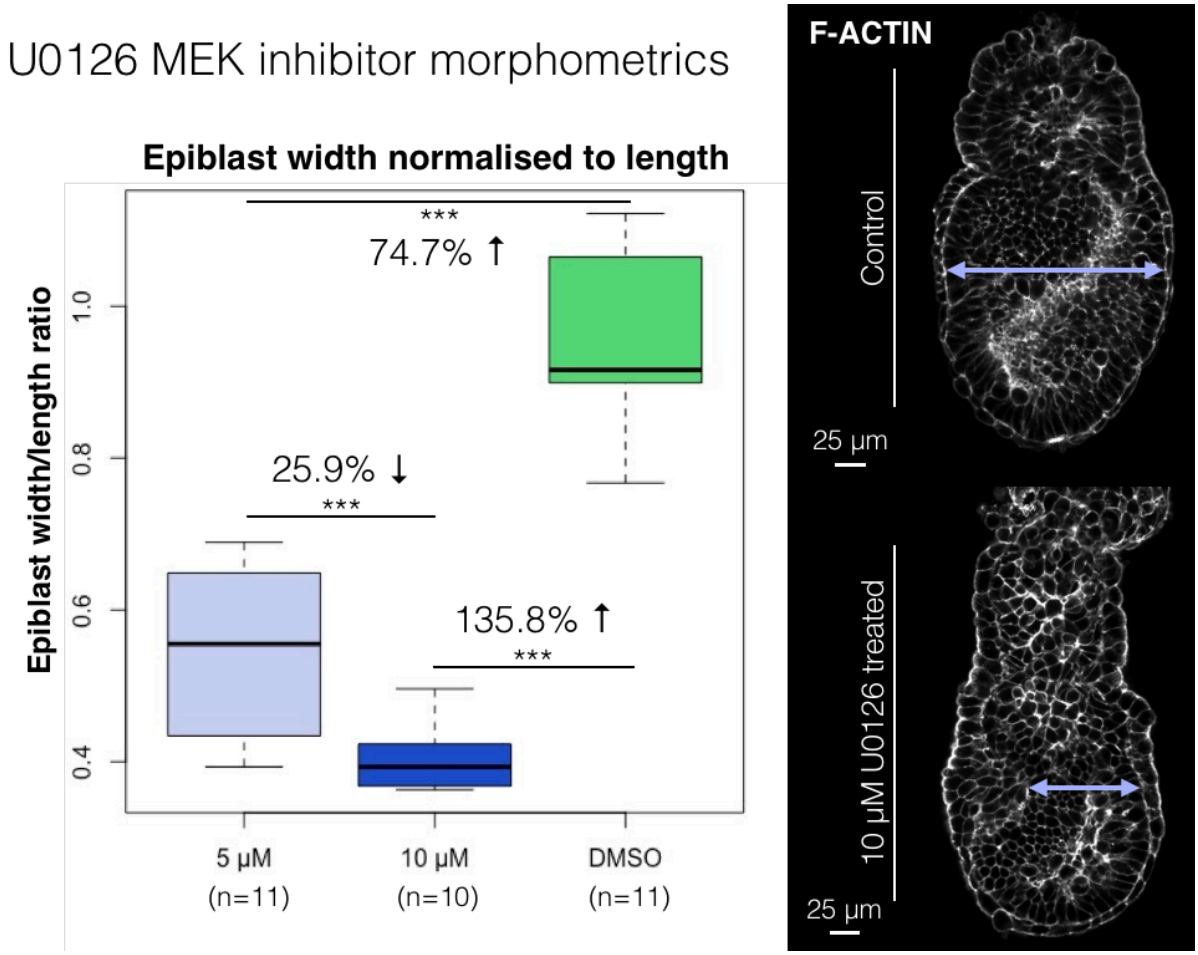


Figure 68. Quantification of epiblast width morphometrics normalised to length in U0126 MEK inhibitor treated conditions compared to DMSO control

U0126 MEK inhibitor treated embryos have a more severe A-P epiblast width reduction in 10 µM-treated compared to 5 µM-treated in line with the defect being milder with a lower inhibitor concentration. A-P width measurement shown in F-ACTIN (gray) staining of DMSO control versus MEK inhibitor treated, where the latter has a narrower A-P width (double-headed arrow). (Statistical tests: hierarchical ANOVA. *** $p < 0.001$). Scale bars: 25 µm.

To further analyse the extent to which MEK inhibition recapitulated the molecular features of the *Ets2*-mutant phenotype, I conducted immunofluorescent staining analyses of: 1. GATA6 (Figure 70) to see if the AVE/VE retained stereotypic marker gene expression as seen in mutants (Figure 69); 2. PAR6B (Figure 69) to test whether the thickened AVE in MEK inhibitor-treated embryos retained polarity like in mutants (Figure 32); and 3. LAMININ (Figure 71) to check whether the thickened basement membrane phenotype found in mutants (Figure 50) can be recapitulated with inhibitor culture. In all cases, MEK inhibitor treatment mimicked what was observed in mutants, apart from the LAMININ-positive basement membrane, which was thickened in inhibitor-treated conditions, but displayed different organisation (Figure 71). Overall, these data suggest that MEK inhibition in WT embryos from E5.5 to E6.5 can result in mutant AVE phenotypic recapitulation, but not to a complete degree.

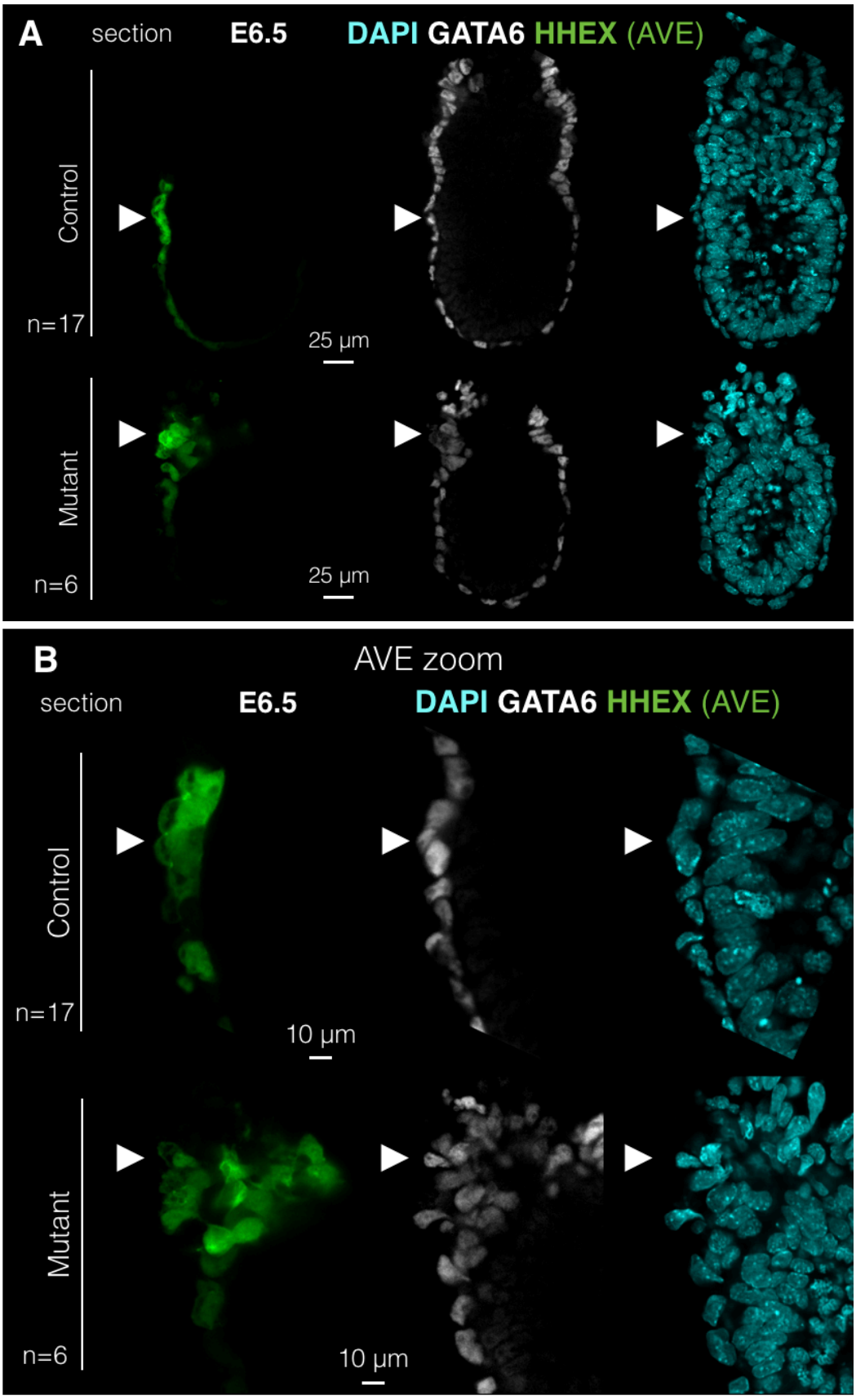


Figure 69. GATA6 expression at E6.5 in mutant versus control
Representative mid-sagittal sections of E6.5 (A) whole mount immunofluorescence and AVE zoom (B) of GATA6 (gray); HHEX-GFP (green – AVE); DAPI (cyan). Arrowheads point at AVE. GATA6 nuclear expression is present throughout thickened mutant AVE. Scale bars: 25 µm, or 10 µm as indicated.

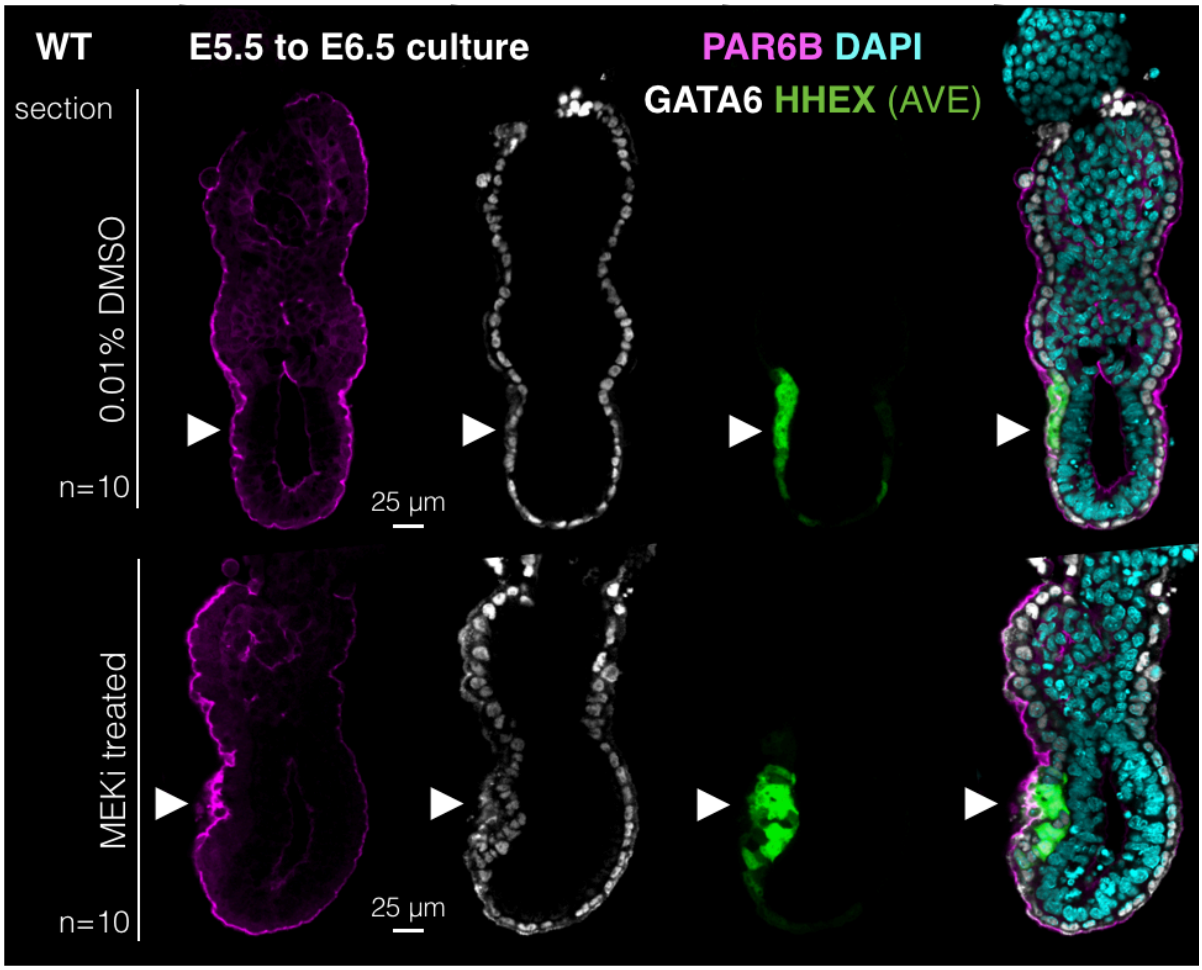


Figure 70. MEK inhibitor-mediated recapitulation of mutant AVE PAR6B and GATA6 expression
Representative mid-sagittal sections of whole mount immunofluorescence in DMSO control versus MEK inhibitor (MEKi) treated embryos (verified with both PD03 and U0126), stained for PAR6B (magenta) and GATA6 (gray); HHEX-GFP (green – AVE); DAPI (cyan). Arrowheads point at AVE. Retention of PAR6B expression in the apical domain as seen in mutants, and preservation of GATA6 nuclear signal throughout the thickened AVE in MEKi treated. Scale bars: 25 µm.

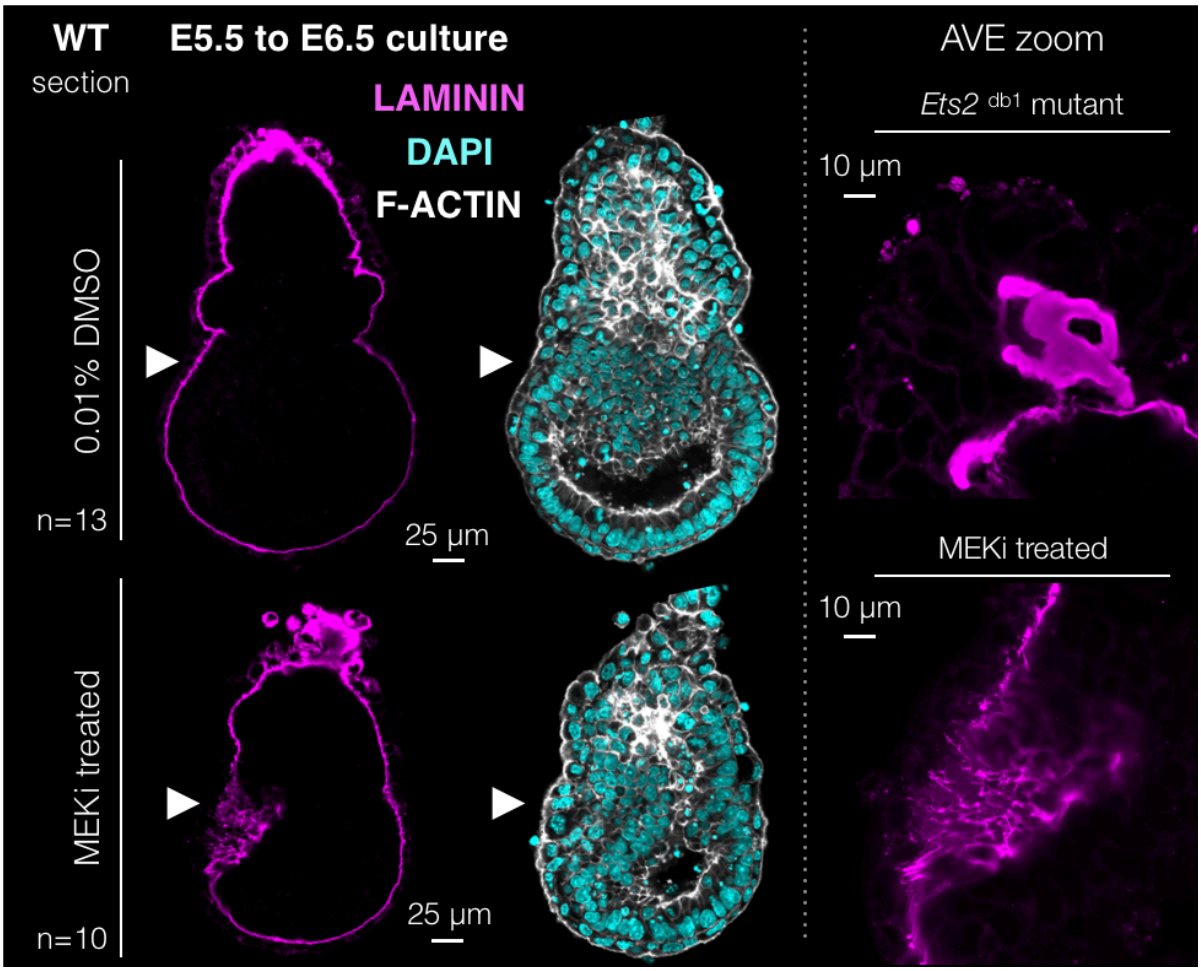


Figure 71. MEK inhibitor-mediated partial recapitulation of mutant AVE LAMININ thickening phenotype
Representative mid-sagittal sections of whole mount immunofluorescence in DMSO control versus MEK inhibitor treated embryos (verified with both PD03 and U0126) stained for LAMININ (magenta); F-ACTIN (gray); DAPI (cyan). Arrowheads point at AVE. LAMININ is thickened beneath AVE in MEKi treated embryos; however, its organisation differs from what is observed in the mutant AVE – shown as AVE zoom – suggesting that other factors may contribute to the LAMININ basement membrane thickening phenotype in mutants. Scale bars: 25 µm, 10 µm as indicated.

4.2.11 MEK-inhibitor-mediated mutant AVE phenocopy is still present upon ExE removal prior to culture

To gauge whether the MEK-inhibitor-treated AVE defect could still arise without the ExE, I removed the ExE prior to inhibitor culture treatment, then utilised the same conditions as other MEK WT cultures (1 µM PD03, or 10 µM U0126, culturing from E5.5 to E6.5). Absence of ExE during this time frame and under these conditions still led to an AVE thickening phenotype, but to a less severe degree (Figure 72). This suggests that the AVE thickening can still arise in the absence of signals from the ExE, but that the mutant ExE *in*

vivo may enhance the severity of the defect. However, these are artificial conditions. To better assess whether the mutant AVE phenotype is cell autonomous, *in vivo* genetics via conditional gene ablation or misexpression is needed to only alter the AVE or VE.

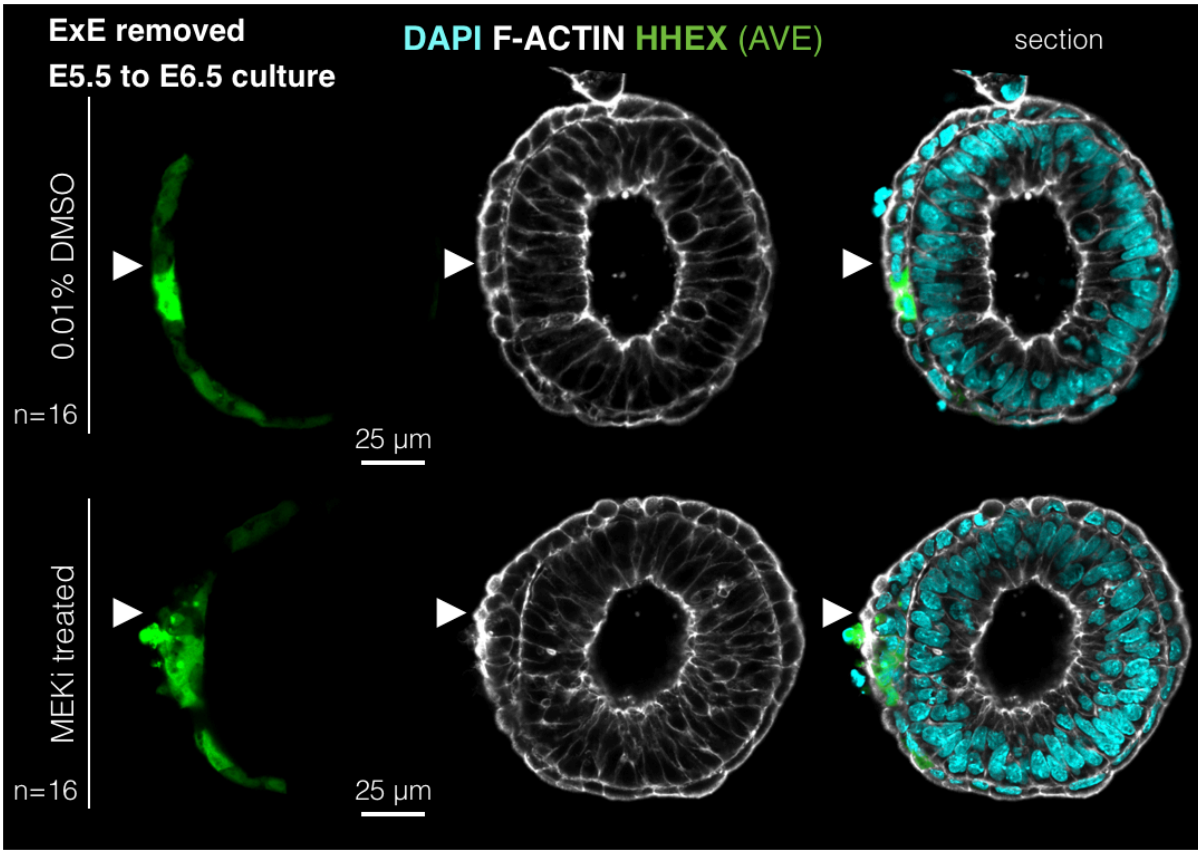


Figure 72. WT embryos with ExE removed then cultured in MEK inhibitor from E5.5 to E6.5 still develop AVE thickening
Representative mid-sagittal sections of MEK inhibitor treated versus DMSO control with ExE removed prior to culture (E5.5 to E6.5). AVE thickening was still observed in the absence of ExE during inhibitor culture treatment, suggesting that the defect could be autonomous to the embryonic portion. However, the phenotype is less severe, indicating a potential contribution from the ExE. Arrowheads point at AVE. HHEX-GFP (green); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

4.2.12 Additional evidence to support the independence of mutant ExE and AVE phenotypes

To further demonstrate the potential independence of the mutant ExE and AVE phenotypes, I employed FMK-mediated apoptosis inhibition from E5.0 to E5.5 to rescue the mutant ExE phenotype (3.2.10); then chased away the inhibitor by culturing the ExE-rescued

mutants in regular media from E5.5 to E6.5 to see whether the AVE phenotype would still develop. Normal embryo morphology and AVE migration was observed in WT/het controls treated with DMSO or FMK (data not shown). Upon culture of mutant embryos with ExE rescue up to E6.5, I found retention of the mutant AVE thickening defect (Figure 73), supporting the mutual independence of the ExE and AVE phenotypes.

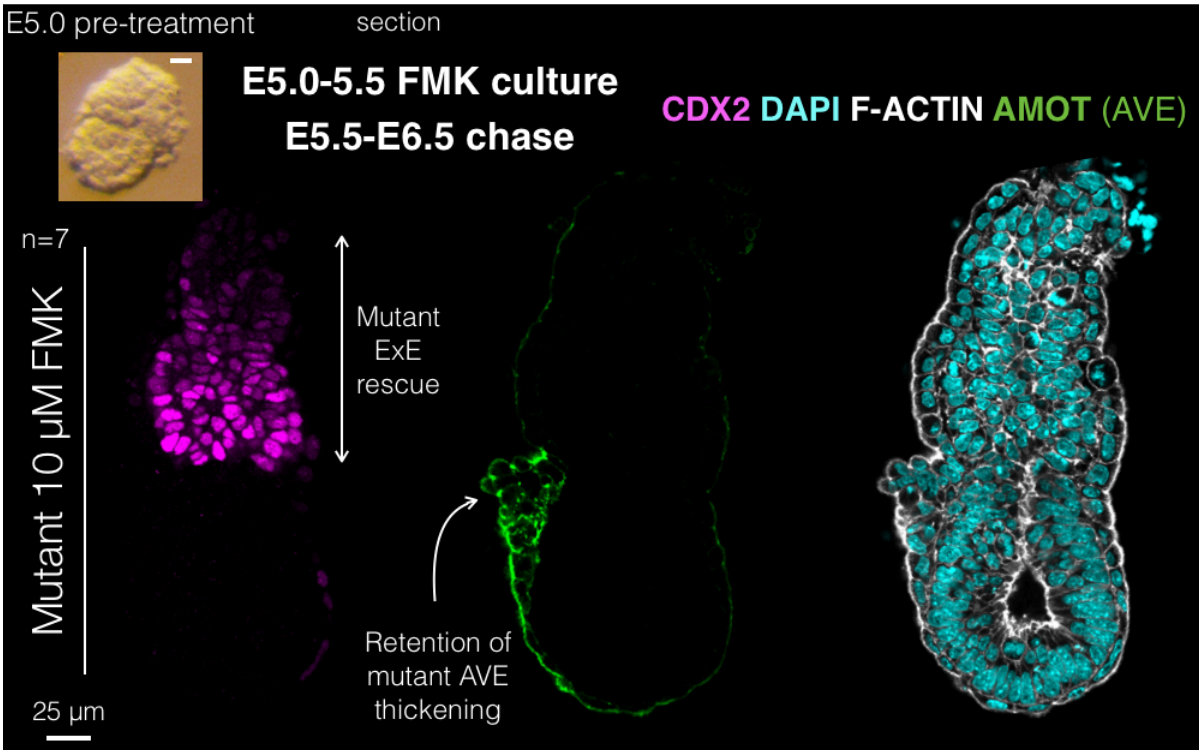


Figure 73. Mutant ExE and AVE defects may be independent, as rescue of ExE defect by FMK-mediated apoptosis inhibition still results in mutant AVE thickening by E6.5
Representative mid-sagittal sections of FMK-treated (10 µM) mutant cultured from E5.0 to E5.5, then in regular media from E5.5 to E6.5, showing a recovered ExE (marked by CDX2 – magenta); thickened AVE (AMOT – green); F-ACTIN (gray); DAPI (cyan). Scale bar: 25 µm.

4.3 Discussion

Georgiades and Rossant, 2006 reported an AVE thickening phenotype in the *Ets2*-mutant. As this was an opposite effect on tissue growth from the reduced ExE defect, I aimed to characterise the origins and cellular bases for the AVE phenotype to better understand how ETS2 influences AVE tissue size.

ETS2 expression and patterning defects in the mutant

ETS2 expression pattern at early and late post-implantation stages was previously assessed via ISH by Georgiades and Rossant, 2006, and Donnison et al., 2015. In both reports, *Ets2* was found to be highly expressed in the ExE, but absent from the VE. Here I found, using immunofluorescence, presence of nuclear ETS2 throughout the VE and in the epiblast, in addition to high levels of ExE expression. This observation is supported by recent scRNA-seq reports (Mohammed et al., 2017; Wen et al., 2017). The different results attained from ISH versus confocal immunofluorescence and scRNA-seq could be attributed to the latter techniques being more sensitive.

Using AMOT as a ubiquitous marker of the AVE, I found AVE migration to be complete at late E5.5, as well as by E6.5. This is in contrast to Georgiades and Rossant, 2006 where AVE migration had not been completed even by E6.75. I validated the observation of complete AVE migration in mutants with another AVE marker HHEX-GFP. However, despite the completion of the AVE anterior shift in mutants, mesoderm marked by T was not found to be induced in mutants by E6.5 and E7.0. This could be due to deficient signalling from the ExE known to be required for *T* induction (Arnold et al., 2008), or a defect within the epiblast.

Mutant polarity characterisation

To investigate whether polarity is intact in the thickened mutant AVE, I examined a number of polarity components expressed at different regions along the apical-basolateral aspects of epithelia. I used PAR6B to mark the apical exposed surface; PAR3 to visualise apical junctions; ZO-1 for tight junctions; SCB and ECAD for lateral junctions; AMOT AVE marker present in apical-lateral junctions; and LAMININ and COLLAGEN IV to mark the basement membrane. Apical and lateral polarity components were all intact in localisation in the thickened mutant AVE in a manner that is expected in a stratified epithelium. Specifically, this

involves the most superficial layer of cells retaining expression of PAR6B, while the other layers express apical and lateral junctional markers; and the basement membrane is restricted to the most basal layer (Williams et al., 2014; Ragkousi and Gibson, 2014). The basement membrane was thickened beneath the mutant AVE, detected using both LAMININ and COLLAGEN IV basement membrane markers. Given the known roles of YAP in sensing and responding to differences in the basement membrane (Dupont et al., 2016), I analysed YAP expression in the mutant AVE and found retention of its stereotypic localisation, suggesting that Hippo/YAP signalling is intact in mutants. The ECM thickening I found is also in line with downregulation of ECM-remodelling enzymes *MMP3/9/13* observed in *Ets2*-mutants (Yamamoto et al., 1998). However, I did not characterise the expression of these matrix remodelling enzymes *in vivo*. Antibody stain, or hybridisation chain reaction (HCR) ISH (Choi et al., 2018) could be used to test via co-stain whether E6.5 AVE basement membrane thickening occurs concomitant with downregulation of matrix remodelling enzymes, to correlate these two effects *in situ*.

As actomyosin contractility is the ultimate effector of cellular dynamics including cell division (Fletcher and Mullins, 2010), I analysed F-ACTIN and active pMLC2 in the mutant AVE. I found elevated junctional F-ACTIN intensity normalised to cell surface area upon Imaris outlining. There were also visibly increased junctional pMLC2 puncta along apical and lateral membranes in mutant AVE cells. Together, these observations suggest increased actomyosin-mediated cellular dynamics in mutants. This could be verified in the future by live imaging in mutants using the *Lifeact* (live actin) reporter mouse line (Riedl et al., 2010).

Mutant AVE phenotypic characterisation

To unequivocally assess whether the mutant AVE thickening is characterised by epithelial stratification, I conducted 3D Imaris cell outlining using high resolution F-ACTIN

Airy scan imaging of the AVE. Imaris outlining showed that the mutant AVE was indeed stratified, whereas control AVE cells remained as a monolayered epithelium. Imaris additionally demonstrated that cell volume and surface area were decreased in the mutant AVE, which may contribute to increased density of cells over a given surface area. To investigate whether the mutant AVE stratification is due to more AVE cells, or possibly, the same number of AVE cells packed in a smaller surface area, I quantified AVE cell number at E5.5 and E6.5. I found a significant increase in mutant AVE cell number by E6.5 when the AVE thickening phenotype is present. There was no significant increase in AVE cell number at E5.0 and E5.5, in line with absence of AVE thickening at these stages.

To interrogate the cause of the stratified mutant AVE that has increased cell number, I quantified proliferative index in the mutant AVE versus control at E5.0, E5.5, and E6.5. Mutant AVE proliferative index was significantly elevated at E5.5, but was unchanged from control at E5.0 and E6.5. This suggests a narrow time window during which ETS2 regulates AVE proliferation. The decrease in proliferation at E6.5 could potentially be due to cells sensing increased crowding. Upon quantification of cell crowding, there was significantly upregulated cell density in both the proximal and distal regions in mutants by E6.5. Overall, these data demonstrate a short developmental time frame of mutant AVE defect progression, characterised by increased proliferation at E5.5, followed by epithelial stratification, increased cell number, and cell crowding by E6.5.

I further investigated whether misorientation of the mitotic spindle could contribute to mutant AVE stratification. I assessed this by using ALPHA-TUBULIN to visualise mitotic spindles co-stained with DAPI to detect mitotic nuclei in AMOT-positive AVE cells. I found orientation of mutant AVE cell division to be mostly planar at E5.5, followed by a significant increase in the proportion of perpendicular cell divisions by E6.5. Interestingly the quantitative distribution of cell division orientation in the mutant AVE at E6.5 is reminiscent of the

'bimodal' distribution seen in the stratifying embryonic skin at E16.5-E17.5 (Williams et al., 2014) – approximately 50% planar and 50% orthogonal cell divisions with very few oblique. Misorientation of the mitotic spindle was detected at E6.5, which occurs after the initial burst in mutant AVE proliferation at E5.5, suggesting that the spindle defect could be secondary to the proliferative defect that may be driving the *Ets2*-mutant AVE phenotype.

MEK inhibitor mediated mutant AVE phenotypic recapitulation

Having characterised the nature and origins of the mutant AVE phenotype in detail, I next asked whether I could recapitulate the phenotype by inhibiting the known upstream MAPK signalling cascade of ETS2 in WT embryos. To do so, I utilised well-characterised MEK inhibitors: PD03 (used in ESC maintenance) and U0126. As MAPK signalling is ubiquitous and active in the entire embryo, not just the VE/AVE, I aimed to arrive at the lowest dosage that would give an AVE-specific effect by morphological assessment. I cultured WT embryos in MEK inhibitor during the time window of mutant AVE defect progression: E5.5 to E6.5. Indeed, upon treatment with the lowest recommended dosage of either MEK inhibitor (PD03 – 1 μM; U0126 – 10 μM), the thickened AVE phenotype could be recapitulated in WT embryos without morphologically affecting the ExE. Throughout the MEKi studies, I used HHEX-GFP fluorescent reporter to visualise AVE cells. To characterise the molecular consequences of MEKi treatment, I stained cultured embryos with pERK1/2 (which should be downregulated in the VE from inhibition of its upstream activator MEK), and ETS2 (which should have decreased nuclear expression in the VE upon blocking its upstream cascade that triggers its nuclear localisation). I also stained for other molecular readouts that I have analysed in the *Ets2*-mutant, including PAR6B, GATA6 and LAMININ to assess the extent of phenotypic recapitulation in greater detail. I was able to titrate the MEKi phenocopy, as demonstrated by using half the recommended lowest dosage of U0126 where I detected a

milder AVE thickening defect, as well as a milder effect of the AVE thickening on the epiblast. At the lowest dosage of MEKi that resulted in an AVE phenotype, I found the ExE ETS2 signal to be retained, suggesting an independence of the ExE on the mutant AVE defect. However, pERK1/2 in the ExE was reduced in the treated condition, highlighting that these inhibitor studies are not AVE-specific, and that the AVE defect recapitulation could involve signals from the ExE. Overall, I was able to recapitulate, with MEK inhibition in WT, most molecular features associated with the mutant AVE thickening phenotype, including: 1. GATA6 expression throughout the thickened AVE at E6.5; 2. localisation of PAR6B in the most apical layer of the thickened epithelium; and 3. Increased LAMININ beneath the thickened AVE epithelium, although the LAMININ organisation in MEKi-treated embryos did not fully mimic the LAMININ distribution seen beneath the mutant AVE.

To begin to assess whether the mutant AVE defect may be independent of the ExE, I removed the ExE, then cultured the embryonic portion of WT embryos in MEKi from E5.5 to E6.5. In the absence of ExE, the AVE thickening could be recapitulated to a less severe degree. This suggests that although there may be contribution from the ExE in enhancing the AVE thickening defect, the ExE may not be essential for the AVE thickening phenotype. However, this approach is crude to address the question of whether the AVE thickening defect is cell autonomous. One weakness is that the epiblast is still present. Whether the AVE thickening phenotype is cell autonomous can hypothetically be addressed by culturing mutant versus control AVE cells *in vitro*. However, an efficient method to derive and culture AVE cells alone *in vitro* has not yet been reported in the literature. The *in vitro* embryoid body system for generating VE-like cells in the outer layer (Evans, 2011) still retains ESCs in the centre, and therefore does not remove the possibility that the defect is due to some signalling contribution from the epiblast. To demonstrate in a robust fashion whether the AVE defect is cell autonomous, tissue-specific *in vivo* genetics is required.

To further test whether mutant ExE and AVE phenotypes may be independent, I rescued the mutant ExE defect by Pan-Caspase inhibition from E5.0 to E5.5, then chased away the inhibitor, culturing the same embryos to E6.5. I found retention of the thickened AVE phenotype in mutants with rescued ExE. This suggests that signalling from the reduced ExE is not required for the mutant AVE defect to arise.

Interestingly, absence of *Ets2* throughout the VE seems to only affect the AVE. There is precedence of ETS2 playing a distinct role within an epithelial subpopulation in the context of *Ets2*-deletion in the intestinal epithelium, where only the intestinal crypt base (containing the stem cell population) displayed increased proliferation (Munera et al., 2011). Both of these observations could be due to presence of different transcriptional co-factors of ETS2, mediating varied transcriptional responses in different epithelial compartments.

Overall this chapter demonstrates that the effect of ETS2 on tissue size is opposite in the AVE as compared to the ExE. I have characterised the AVE phenotype in-depth and the time frame of its origins and progression. I have also provided evidence that the known upstream MAPK signalling cascade for ETS2 activation is relevant in the AVE via MEKi-mediated mutant AVE phenotypic recapitulation in WT. However, the transcriptional effectors downstream of ETS2 in the AVE remain unknown.

5. ScRNA-seq to elucidate downstream effectors of ETS2 transcriptional activity

5.1 Introduction

Chapters 3 and 4 characterised the origins and progression of the *Ets2*-mutant ExE and AVE phenotypes, highlighting the opposing effects of transcription factor ETS2 on tissue size. I showed that the ExE and AVE defects have different embryonic origins and involve separate mechanisms. In summary, I found the reduced mutant ExE phenotype to initiate at peri-implantation due to elevated apoptosis, which I could rescue by Pan-Caspase-mediated apoptosis inhibition. The diminished ExE that remained in mutants at post-implantation retained normal tissue growth and death rates, as well as signature marker gene expression (CDX2, EOMES, ESRRB), suggesting that ETS2 is not required for this particular ExE 'subpopulation'. For the mutant AVE, I demonstrated that its stereotypic anterior shift is normal; however, the tissue becomes abnormally thickened. This mutant AVE thickening was characterised by epithelial stratification, increased cell number, and increased cell crowding by E6.5. Overt AVE thickening was preceded by elevated proliferation at E5.5. Increased proportion of mutant AVE perpendicular cell divisions additionally contributed to its epithelial stratification. While thickened and stratified, the mutant AVE retained stereotypic localisation of apical-basolateral polarity components. Blocking known upstream MAPK signalling cascade of ETS2 via MEK inhibitor pharmacological culture from E5.5 to E6.5 – using either PD03 or U0126 at low dosage – resulted in mutant AVE phenocopy. However, the downstream effectors of ETS2 transcriptional activity in the ExE and AVE are unknown.

5.1.1 Rationale for use of scRNA-seq to investigate the *Ets2*-mutant

This Chapter investigates the transcriptional responses elicited by ETS2 in the ExE and AVE using scRNA-seq. Because ETS2 has tissue-specific effects, bulk sequencing (e.g. by RNA-seq) of the entire embryo would be insufficient to resolve context-dependent roles of ETS2. Furthermore, the early embryonic origins of *Ets2*-mutant phenotypes involve low cell numbers, which would present a challenge to obtain sufficient material for bulk RNA-seq. In addition, the AVE is known to be heterogeneous, expressing different combinations of its signature marker genes at different intensities within the population (Hoshino et al., 2015). Studying the AVE at single-cell resolution could help resolve these heterogeneities and better interrogate biological anomalies in mutants. Therefore, scRNA-seq was selected as the preferred approach to examine the downstream transcriptional response elicited by ETS2 in a tissue-specific fashion in *Ets2*-mutants versus WT.

5.1.2 ScRNA-seq biology and technology

Smart-seq2

ScRNA-seq technology is a high throughput method to quantify gene expression in individual cells and uncover heterogeneity within cell populations. Smart-seq2 was used for all single-cell sequencing studies in this thesis. The technique offers quantitative, reproducible results using small amounts of purified RNA from single cells (Ramsköld et al., 2012). The advantage of Smart-seq2 is the possibility to recover full-length cDNA via the MMLV reverse transcriptase. This enables recovery of exons for each transcript, as well as splice variants. The disadvantage of the technique is that it does not allow for analysis of polyA-negative RNA, as it preferentially selects polyA-positive RNA; and mRNA strand specificity is also lost (Picelli et al., 2014).

Biology of scRNA-seq

Differences amongst individual cells within a population can manifest in important functional changes. Compared to bulk RNA-seq, scRNA-seq provides additional details on gene regulatory networks (GRNs), transcriptional dynamics, and composition of changing tissues with respect to different cell populations (Kolodziejczyk et al., 2015). ScRNA-seq technology has been used when there are limited amounts of biological tissue available; for example, from the early embryo to elucidate global trends of gene expression in detail and to track lineage progression changes (Tang et al., 2009; Deng et al. 2014). On the other hand, this high throughput approach also enables parallel processing and profiling on the order of 100,000 cells, which in turn allows for discovery of heterogeneity within cell populations and detection of rare populations in complex organ systems (Fan et al., 2015). Together, scRNA-seq offers the possibility to interrogate new questions, such as frequency and dynamics of cellular transition states, stochasticity of gene expression, and extensive GRN establishment. However, bioinformatics challenges remain in arriving at the most suitable method to interpret and extract inferences from scRNA-seq datasets (Stegle et al., 2015).

ScRNA-seq methodology

ScRNA-seq involves: 1. Single cell collection and lysis; 2. Poly[T] primed reverse transcription to select for mRNA and produce cDNA; 3. Amplification of cDNA by *in vitro* transcription or PCR approaches; and 4. Use of amplified cDNA for sequencing library preparation.

Single cell collection

Single cells must be collected rapidly to preserve cell viability and in a way to ensure that individual cells are indeed captured. This step presents one of the main challenges for single-cell sequencing and involves different protocols depending on the biological system of

interest. To obtain single cells from sparse samples, such as the early embryo, micro-manipulation approaches may be employed (Yan et al., 2013). Specifically, single cells can be recovered using microcapillaries guided by the microscope. Though low throughput, this approach allows for verification of single cell recovery by visual inspection. Laser capture microdissection (LCM) can be employed for tissues embedded in sections whereby a laser is used to outline individual cells of interest that can then be excised from microscopic sections (Frumkin et al., 2008). This technique can preserve spatial information within a tissue and allow recovery of particular cell types of interest if their morphologies are distinct (Keays et al., 2005). Higher throughput methods involve dissociating a tissue into single cell suspension in a suitable buffer and collecting the entire tissue sample as a whole. Depending on the tissue, dissociation may be challenging and the enzymatic treatment for single cell disaggregation could be harmful to the extent of causing artefacts in cellular transcriptomes. Therefore, the mildest enzymatic method sufficient to dissociate cells into single cell suspension is preferred. Cells suspended in single cell lysis buffer can subsequently be sorted into individual wells by microfluidics (e.g. Fluidigm C1 or microdroplet approaches), or fluorescence activated cell sorting (FACS). FACS is a rapid and scalable approach if single cells can be precisely and repeatedly sorted into the centre of each well to ensure that the cell is immediately and completely immersed in lysis buffer. FACS also provides the added advantage of experimental design flexibility such that desired populations may be purified using fluorescence. In addition, FACS has the option of index sorting, whereby cell size, cell granularity, and fluorescence levels of proteins associated with each well may be recorded (Hayashi et al., 2010). In summary, manual cell collection, LCM, microfluidics, or FACS-mediated cell collection may be employed depending on the sample of interest and the biological question.

Computational approaches for assessing sample quality and technical noise from Smart-seq2 data

Appropriate interpretation of biological insights arising from scRNA-seq data depends on bioinformatics approaches, which are actively being developed (Stegle et al., 2015).

Quality control

Quality control is essential for removing low quality libraries – which could be due to technical variability upon cell collection, library preparation, and sequencing – prior to further analysis. Following read mapping to the appropriate reference genome, a number of bioinformatics approaches can be used to verify whether there are issues with the sequencing library, and identify cells that have RNA degradation. Some of these measures include checking: the number of genes that are expressed and proportion of mapped reads to mitochondrial genes (Islam et al., 2014); number of mapped regions and the proportion of mapped reads to endogenous genes (Shalek et al., 2014); and levels of housekeeping genes (Treutlein et al., 2014). Both biological and technical factors can contribute to variable gene expression amongst different cells within a population. Identifying biological variability involves being able to approximate technical variability in a quantitative and accurate manner. To account for technical noise, external RNA spike-ins, such as the ERCC RNA spin-in mix, can be incorporated at equal and known amounts in the lysis buffer to assess with the cell lysate (Jiang et al., 2011).

Quantitatively defining biological noise and identifying cell-to-cell variability in gene expression

Following quantification and identification of technical noise, biological heterogeneity across cells for each gene can then be assessed. One approach to do so is to measure the statistical variance of counts, whereby the observed total variance can be segregated into technical and biological contributions using RNA spike-ins (Brennecke et al., 2013). The

biological variance can be further subdivided to account for variation from cell cycle, or other pertinent biological process expected to have underlying activity in the system (Buettner et al., 2015). ScRNA-seq offers the predominant advantage of being able to identify cellular subpopulations that arise from heterogeneous populations. Having detected different clusters in an unbiased manner, they can then be annotated using markers of expected, known cell types. For clustering, a number of approaches can be employed, for example, using: all genes expressed, tissue-specific markers (subset of genes), or highly variable genes (HVGs) (Jaitin et al., 2014). Selection of the ideal clustering algorithm should take into consideration the 'separability' and 'compactness' of clusters that will influence the final number of clusters, which can vary depending on the algorithm selected (Treutlein et al., 2014). Known signature marker gene expression for a cell population can be used to validate the identity of the resultant clusters (Jaitin et al., 2014) (Figure 74). Whether transcriptional differences from scRNA-seq analysis represent functionally distinct cellular populations necessitate experimental validation (Patel et al., 2014). A main computational challenge that remains is the ability to clearly separate technical noise from biologically-driven heterogeneity (Pijuan-Sala et al., 2018).

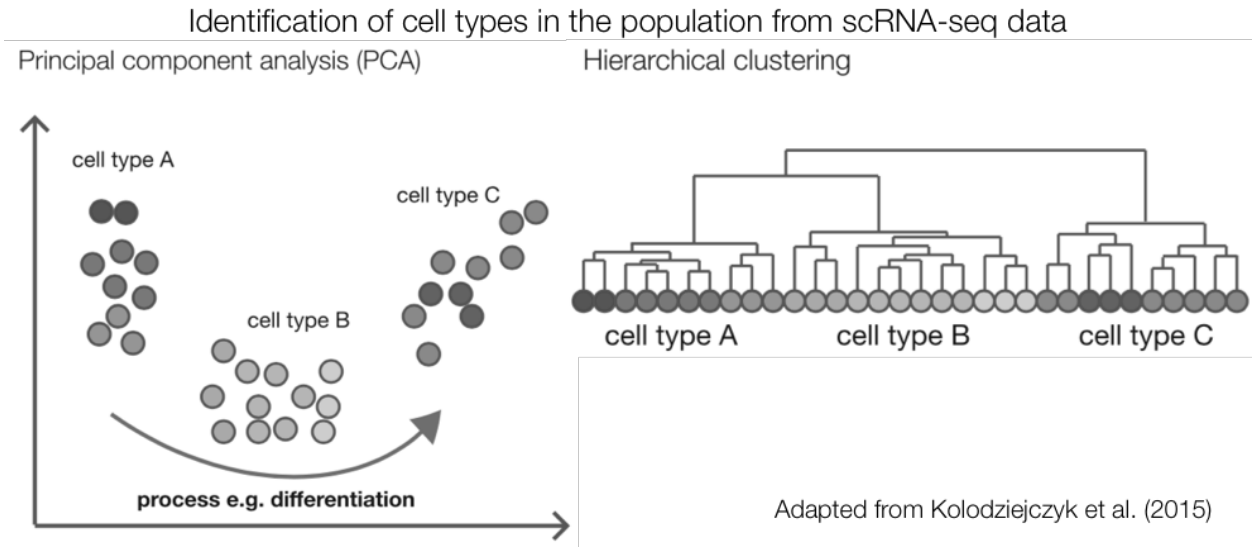


Figure 74. Identification and characterisation of cell types in a heterogeneous population
This can be done by principal component analysis (PCA), t-SNE (not shown), or hierarchical clustering (Kolodziejczyk et al., 2015).

Trajectory analysis can additionally be employed to elucidate genes and pathways that regulate transitions in cell fate and underlying GRNs driving cell fate decisions. For example, pseudotime quantitative analysis using the Monocle algorithm has been developed to quantify progression of biological processes like cellular differentiation (Trapnell et al., 2014).

Analysis of differentially expressed genes (DEGs) can further help assess differences between cell types, or when comparing control to mutant or treatment conditions. Computation of DEGs from scRNA-seq data requires different algorithms from those used for bulk-seq (Stegle et al., 2015). As there are no replicates per individual cell, single cell studies run the risk of having batch effects, and therefore accounting for technical noise is critical. Overall, scRNA-seq involves a distinct set of computational approaches from bulk-seq in order to detect biological significance.

5.1.3 Bioinformatics for analysis of misregulated genes and pathways

Multiple gene ontology (GO) algorithms were employed to identify recurrently enriched biological processes and signalling pathways. Although these different algorithms vary in methodology for calculation of statistical significance, they are mostly similar in the source of their reference dataset for genes annotated to a biological phenomenon. GOs are constructed from a framework of structural vocabularies (Ashburner et al., 2000). GO categories include: biological process, molecular function, and cellular component. Biological process relates to the resultant effect of biological pathways that lead to a physical or chemical change in the cell. Molecular function highlights the biochemical effect of a gene, but does not include information on the cellular localisation of the gene's functional activity. Cellular component specifies the cellular localisation where a gene product exerts its function in relation to eukaryotic cellular compartments. The aim for GO is to categorise and comprehensively describe biological entities, as well as statistically evaluate the relationships amongst these

entities, where each entity is clearly defined. This methodology therefore serves as a structural guideline for organisation and assessment of query gene lists and novel biological datasets. The data can be tailored by adjusting statistical parameters of the GO search, i.e. selecting a specific cutoff. Overall, GO enables translation of large biological datasets into biologically-relevant hypotheses to be tested by experimentation. GO annotations are organised as nodes in the context of a network. The corollary from this organisation is that the ancestry of these nodes is known and preserved for any given annotation. GO databases are constantly updated to reflect current knowledge (Ashburner et al., 2000). The remainder of this section presents brief background on the different GO methodologies used in this thesis.

Gene set enrichment analysis (GSEA) and Reactome

GSEA aids in the interpretation of whole genome expression profile data based on known reference datasets that pertain to biological function, as well as curated cancer gene lists. The significant advantage of this approach is the ability to evaluate gene expression data in relation to functional gene sets attained from published knowledge of biochemical signalling pathways in normal and disease conditions (Subramanian et al., 2005). Drawing from the database stored in GSEA, the Reactome pathway method (Fabregat et al., 2018) is a knowledge-based approach that combines information on molecular pathways – including metabolism, DNA replication, cellular trafficking, signal transduction, etc. – in the context of a defined, organised network map. This system may serve both as a reference source for known cellular processes, as well as a platform to investigate novel relations between gene expression profiles through its ability to link proteins to molecular functions.

PANTHER classification for gene function analysis of large datasets

Protein annotation through evolutionary relationship (PANTHER) is a classification

system for GO, biological function, and signalling pathways that incorporates the relevant statistical methods for analysing large genomic datasets. In addition to annotating gene families to GO terms, PANTHER also organises genes into subfamilies and functional classes, such that genes within subfamilies that would otherwise appear individual are shown to share related functions where applicable (Mi et al., 2013). The PANTHER software outputs the following information from DEG lists: 1. Background and sample frequency, where background frequency is the number of genes associated with a GO term throughout the background set, and sample frequency is that for the input list; 2. Over and under-representation of a GO term, represented by '+' and '-', respectively; 3. P-value – the likelihood for genes in the input list to occur in relation to the gene list of the reference background in terms of GO annotation; and 4. List of significant GO terms and their enrichment with respect to background and sample frequencies.

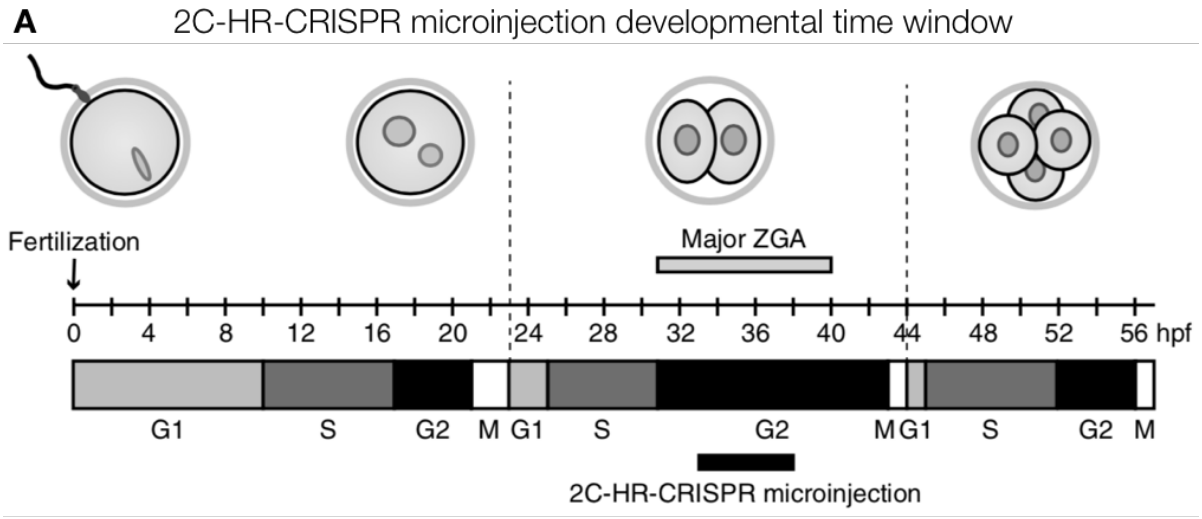
Overall, GO enables enrichment analysis of query gene sets to elucidate significant over- or under-represented biological processes and signalling pathways. This aids in interpretation of DEGs from comparative sequencing analysis, and proposes biological questions to focus on by experimentation.

Following GO analysis for the most enriched biological processes and signalling pathways, associated top most significant DEGs require *in vivo* validation by qPCR, (HCR)-ISH, or antibody staining. QPCR may be conducted on purified cell populations to directly verify whether a specific cell type displays differential gene expression in line with scRNA-seq results. ISH and antibody staining provide *in vivo* spatial information to assess whether candidate genes are differentially expressed in the expected cell types. However, both of these techniques are not quantitative; rather, they may inform presence or absence of expression, and whether expression domains are diminished or expanded in control versus mutant.

5.1.4 Two-cell homologous recombination (2C-HR)-CRISPR to validate scRNA-seq findings *in vivo*

Genetic *in vivo* validation of scRNA-seq results can be conducted via misexpression of top upregulated DEGs under a tissue-specific promoter. To do so, 2C-HR-CRISPR can be employed to efficiently knock-in large insertion constructs by gene targeting to a specific genomic locus. This technology involves microinjecting CRISPR components into two-cell stage mouse embryos, a unique developmental time window with a long G2 phase, which features open chromatin that increases the likelihood for HR to occur (Figure 75). With the further addition of biotin and streptavidin tagging of donor DNA template and Cas9, respectively, Gu et al., 2018 could achieve a CRISPR knock-in targeting efficiency of up to 95% for *Gata6*. However, this is dependent on the endogenous locus to be targeted.

CRISPR-Cas9 mediated genome editing has been advantageous over lengthy and laborious genome editing in ESCs where precise editing occurs at low frequencies in embryos at the zygotic stage (Carroll, 2014). The CRISPR-Cas9 system is a significantly more efficient approach for gene editing in zygotes to generate: point mutations, small insertions, and indels, mostly utilising nonhomologous-end-joining and homology-directed-repair mechanisms (Yang et al., 2013). The bottleneck had been the inability to insert genomic fragments larger than 2kb by HR (Quadros et al., 2017). With 2C-HR-CRISPR, Gu et al., 2018 demonstrated that reporter alleles of up to 4.5kb could be inserted. Presence of targeting was verified by PCR genotyping of individual embryos, or through conducting immunofluorescence of targeted embryos to detect protein upregulation of the knocked-in gene. The efficiency for targeting was found to be invariant upon comparing outbred versus inbred (C57Bl/6J) mouse genetic backgrounds (Gu et al., 2018).



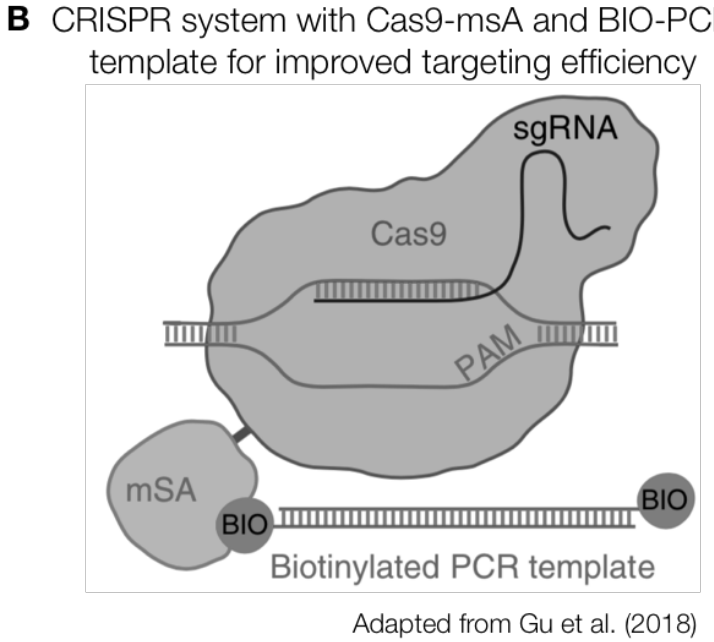


Figure 75. Mouse developmental time window and components for 2C-HR-CRISPR microinjection
(A) Diagram illustrating the timing of microinjection for 2C-HR-CRISPR in early pre-implantation development – i.e. during the long G2 phase at the 2-cell stage, and at the time of major zygotic genome activation (ZGA). This time window increases the likelihood for homologous directed repair to occur. Hpf = hours postfertilisation.
(B) Diagram of Cas9-mSA bound to BIO-PCR template in the 2C-HR-CRISPR system. PAM = protospacer-adjacent motif (Adapted from Gu et al., 2018).

This Chapter utilises Smart-seq2 scRNA-seq technology to interrogate downstream effectors of ETS2 transcriptional activity; GO analysis to select candidate pathways and genes to interrogate; and attempts 2C-HR-CRISPR technology for *in vivo*, tissue-specific genetic validation of scRNA-seq results.

5.2 Results

5.2.1 ScRNA-seq of pure 129 background E5.5 control versus mutant – mostly epiblast dataset

I first selected E5.5 for scRNA-seq analysis of mutants, as this was the initiation stage of the mutant AVE defect. I initially attempted scRNA-seq of the entire embryo at E5.5 with the aim of recovering all cell populations. I conducted all E5.5 dissections at 1pm, using M2 and standard protocol. I dissociated entire E5.5 embryos into single cells by the following method (details are presented in Methods 2.8; however, note that this method is not effective for recovering ExE and VE cells): 6 minute TrypLE treatment at 37°C, keeping each embryo separate (100 µl TrypLE per embryo), followed by brief, gentle mouthpipeting (2-3 times) guided by light microscopy. The cells in TrypLE were then transferred back into an Eppendorf, gently tapped 2-3 times, then neutralised with equal volume heat inactivated FBS. Control and mutant embryos collected from the pure 129 background were determined by phenotype, assessed based on the degree of ExE, as the ExE phenotype is already present by this stage (Figure 76). However, WT versus het could not be segregated, and phenotyping did not allow for secondary, unbiased validation of the correct genotype. Once dissociated, single cells from individual embryos were collected by FACS (gating strategy and settings detailed in Methods 2.8) into lysis buffer, followed by storage in -80°C prior to Smart-seq2 conducted at the Wellcome Sanger Institute. Single cell collection work flow and number of embryos and cells collected are presented in Figure 76.

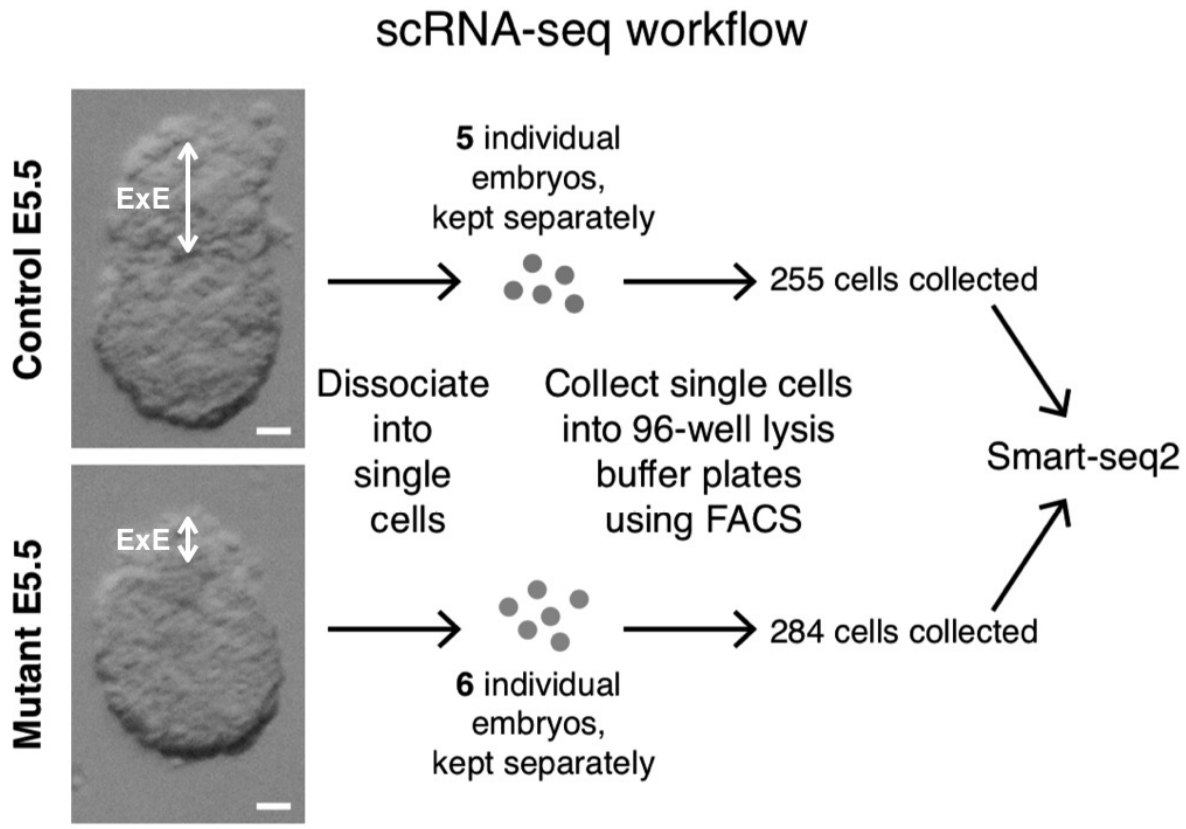


Figure 76. Workflow of single cell collection from phenotyped E5.5 control versus mutant embryos
Embryos were dissected at E5.5 (1pm). Mutants were identified based on the reduced ExE phenotype. No genotyping was conducted. Embryos were disaggregated individually and kept separate with a total of 5 controls and 6 mutant embryos obtained. Single cells were collected into 96-well lysis buffer plates by FACS with a total of 255 control and 284 mutant cells collected. This was followed by Smart-seq2. Brightfield images of control and mutant embryos are shown, and the relatively differences between control versus mutant ExE length are highlighted (double-headed arrow). Scale bars: 25 μm.

Mutant versus control embryos could not be segregated based on their single cell sequences, as the *Ets2* ^{db1}-mutant is a neo-fusion that retains the 3' UTR, and Smart-seq2 is conducted from the 3' end (Picelli et al., 2014). Because the 3' UTR of *Ets2* is very long (1911 nucleotides), neo was not detectable and therefore could not be used to segregate control versus mutant (Figure 77).

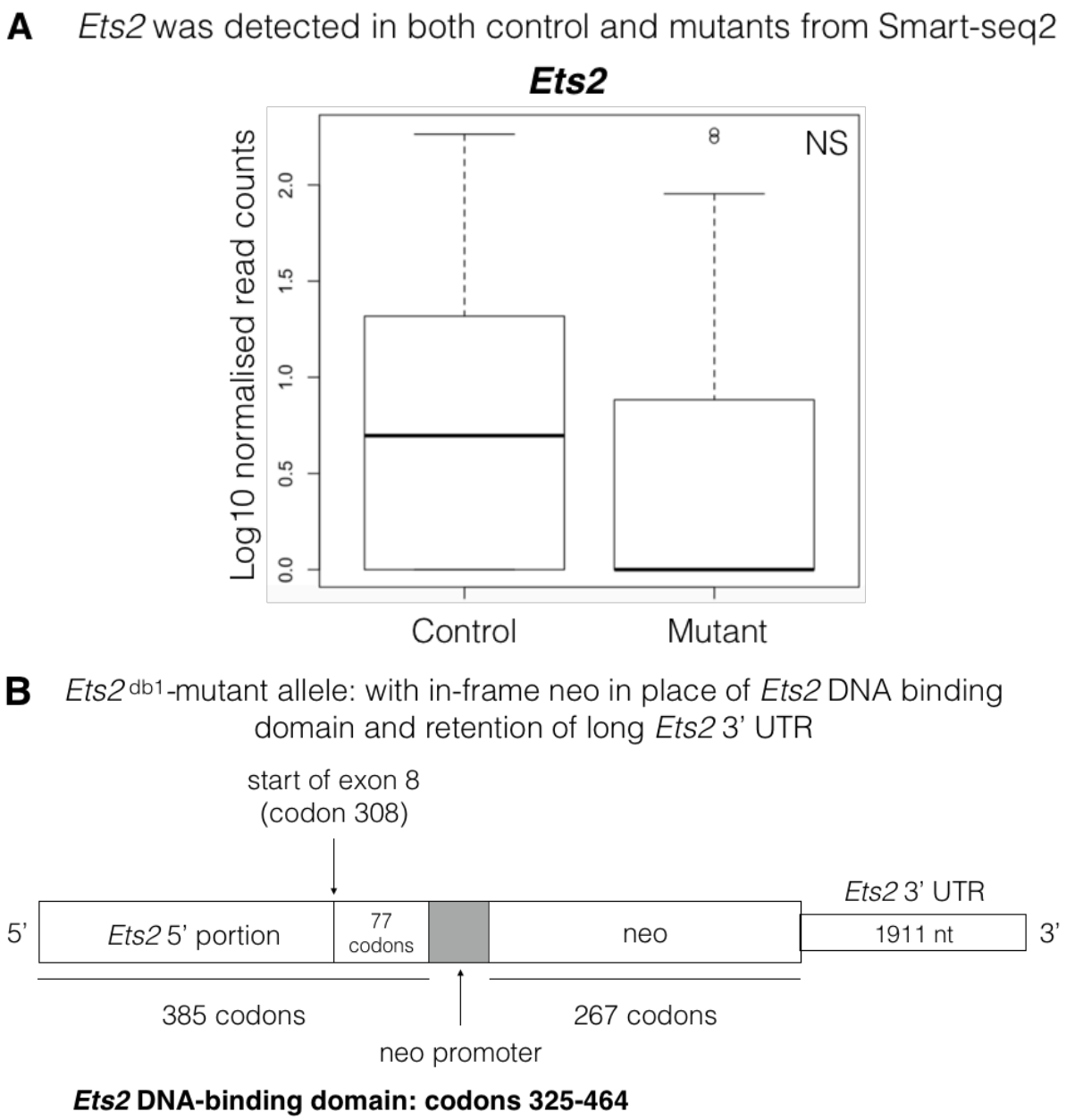


Figure 77. Detection of *Ets2* in both control and mutants by Smart-seq2 due to retention of the long *Ets2* 3' UTR in the mutant allele
(A) Control and mutant *Ets2* expression do not significantly differ from Smart-seq2 conducted at E5.5 (Statistical test: two-way ANOVA. NS = not significant).
(B) The presence of *Ets2* from scRNA-seq in mutants is in line with the mutant allele having an in-frame deletion of the DNA-binding domain where the long 3' UTR (1911 nt) of *Ets2* is retained. As sequencing with Smart-seq2 is conducted from the 3' end of the cDNA, the technology would not be able to distinguish between control and mutant.

All analyses of scRNA-seq data presented in this Chapter were conducted by Dr. Maria Solovey and Dr. Antonio Scialdone, apart from GO, which was conducted by myself. Following quality control (QC) assessment, the final recovery of E5.5 control versus mutant cells from the pure 129 background was 75% of total cells collected (198/255 cells for control;

206/284 cells for mutant). Unbiased t-SNE clustering analysis revealed no clear separation between control versus mutant (Figure 78B). Assignment of cell fate based on known marker genes demonstrated that the majority of cells collected were epiblast with very few AVE/VE and ExE cells (Figure 78A). This dataset was therefore only of sufficient statistical power to analyse the epiblast for DEGs with conventional techniques. While there were not enough AVE/VE cells for any statistical approaches, there had been just enough ExE cells for adjusted analysis with sufficient statistical strength.

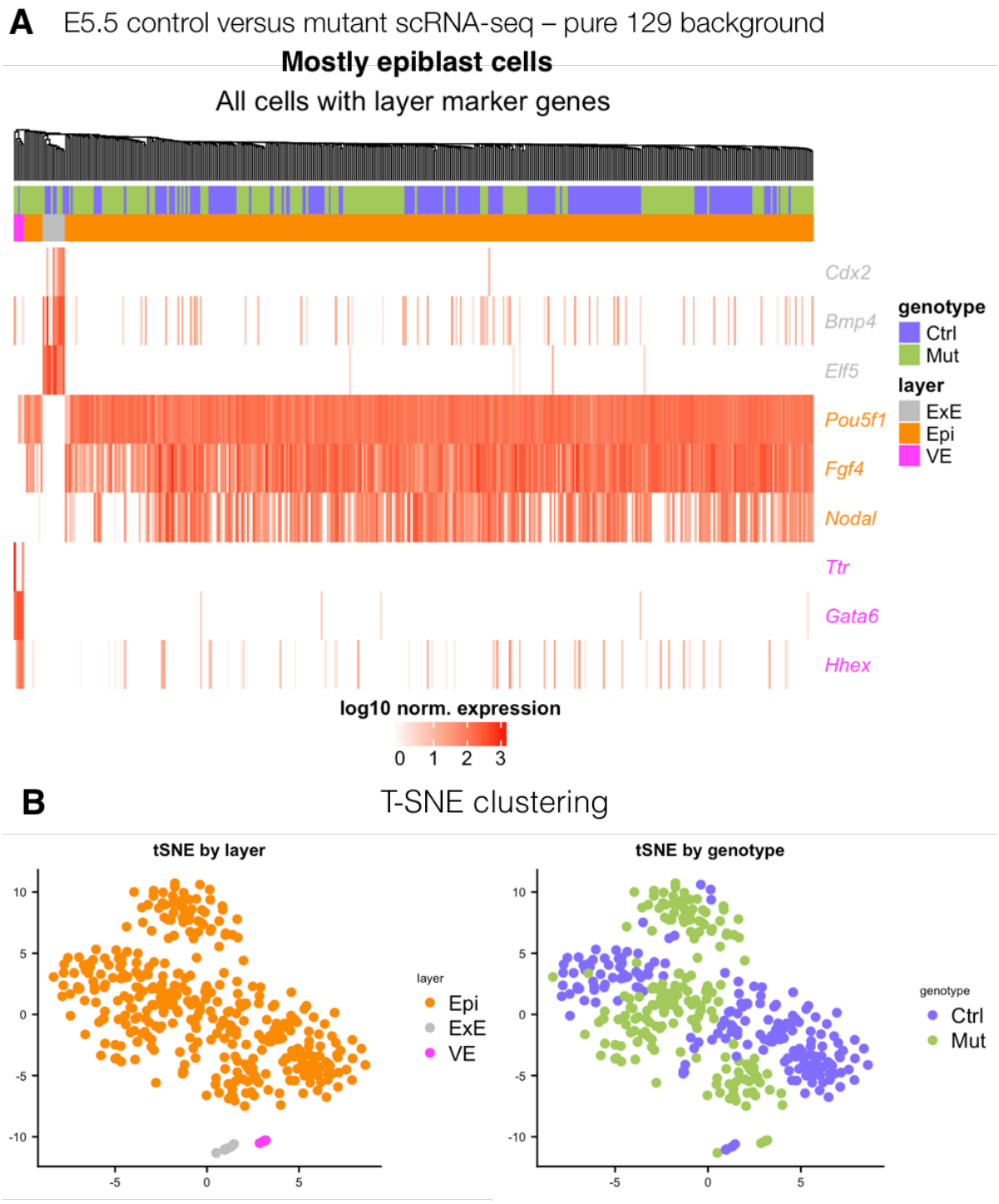


Figure 78. E5.5 pure 129 background scRNA-seq – heatmap clustering based on tissue-specific marker genes and t-SNE

(A) Heatmap clustering based on E5.5 layer marker genes from Mohammed et al., 2017. Control and mutants do not segregate. Most cells are epiblast ('Epi').

(B) T-SNE clustering of all cells by layer markers, again showing presence of mostly epiblast cells and very few ExE and AVE/VE cells. T-SNE clustering by genotype is also presented where control and mutant do not cluster separately.

To check where our E5.5 epiblast scRNA-seq data cluster with respect to single-cell epiblast data reported in the literature, we plotted our dataset with that from Mohammed et al.,

2017 (Figure 79). This previous report aimed at uncovering the development progression and regulatory mechanisms of the murine epiblast from when it first forms in the pre-implantation blastocyst to a stage just prior to gastrulation. To do so, the authors conducted scRNA-seq analysis of the wild type epiblast at E3.5, E4.5, E5.5, and E6.5. The staging criteria they used for E5.5 was embryos lacking overt thickening of the AVE, i.e. pre-AVE migration, which is an early E5.5 stage. We found that our E5.5 epiblast data fell between their 'early E5.5' and E6.5 epiblast cells, suggesting that our staging criteria were similar.

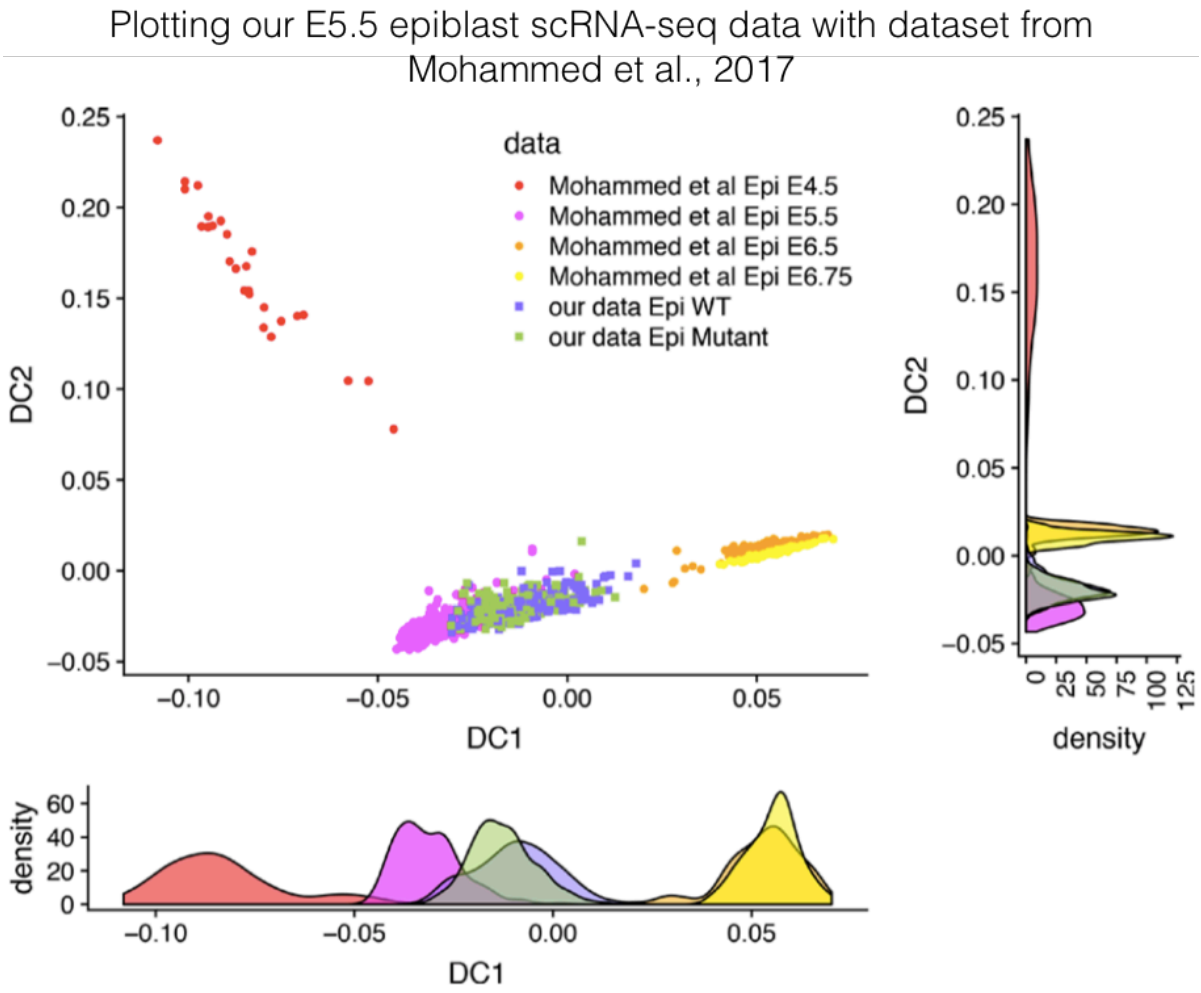


Figure 79. Diffusion map to show the embryo stages and cell types recovered in relation to reference dataset from Mohammed et al., 2017

Diffusion map technique to obtain low-dimensional representation of published scRNA-seq data of epiblast cells from E4.5 to E6.75 (circles; Mohammed et al., 2017), which we projected our epiblast data onto (squares). The scatter plot shows the first two diffusion components DC1 and DC2, while the plots at the bottom and on the right-hand side display the density of points corresponding to cells along DC1 and DC2. Both our WT and mutant data overlap to the greatest extent with previously published E5.5 epiblast.

5.2.2 E5.5 mutant epiblast transcriptome is broadly unaltered

As the ExE is known to signal reciprocally with the epiblast (Ang and Constam, 2004), I interrogated whether the reduced mutant ExE (presented in 3.2.3) had an effect on the mutant epiblast transcriptome. I previously showed that the mutant epiblast retained core signature marker expression OCT4 (3.2.4), and was normal in length morphometrics at both E5.5 and E6.5 (3.2.2). Here, scRNA-seq independently demonstrated that the mutant epiblast was broadly unaltered. This was demonstrated by absence of separate clustering amongst mutant versus control cells upon heatmap and t-SNE analysis (Figure 78), suggesting that mutant epiblast cells have not transdifferentiated into another cell type. However, scRNA-seq is unable to detect defects in epiblast patterning, as the technique does not preserve spatial information. An additional computational method was next used to verify whether the broad transcriptome of the mutant epiblast is indeed comparable to control. This involved cluster stability analysis, i.e. scoring the robustness of global transcriptional differences based on 2 cluster quality indices: Pearson gamma and average silhouette width (Figure 80). The value of these indices was compared to 2 control cases: 1. Positive control: dataset where global transcriptional differences were known to exist (epiblast transcriptional profiles from E5.5 versus E6.5; data from Mohammed et al., 2017); and 2. Negative control: dataset where no difference between groups was expected, which was our epiblast dataset with genotype labels randomly shuffled 100 times. Values of both indices calculated for comparison of WT versus mutant demonstrated high similarity to the negative control, suggesting that the transcriptome of epiblast cells at E5.5 amongst the different genotypes could not be separated into 2 stable clusters. This further validated the absence of global transcriptional differences between control versus mutant epiblast at E5.5.

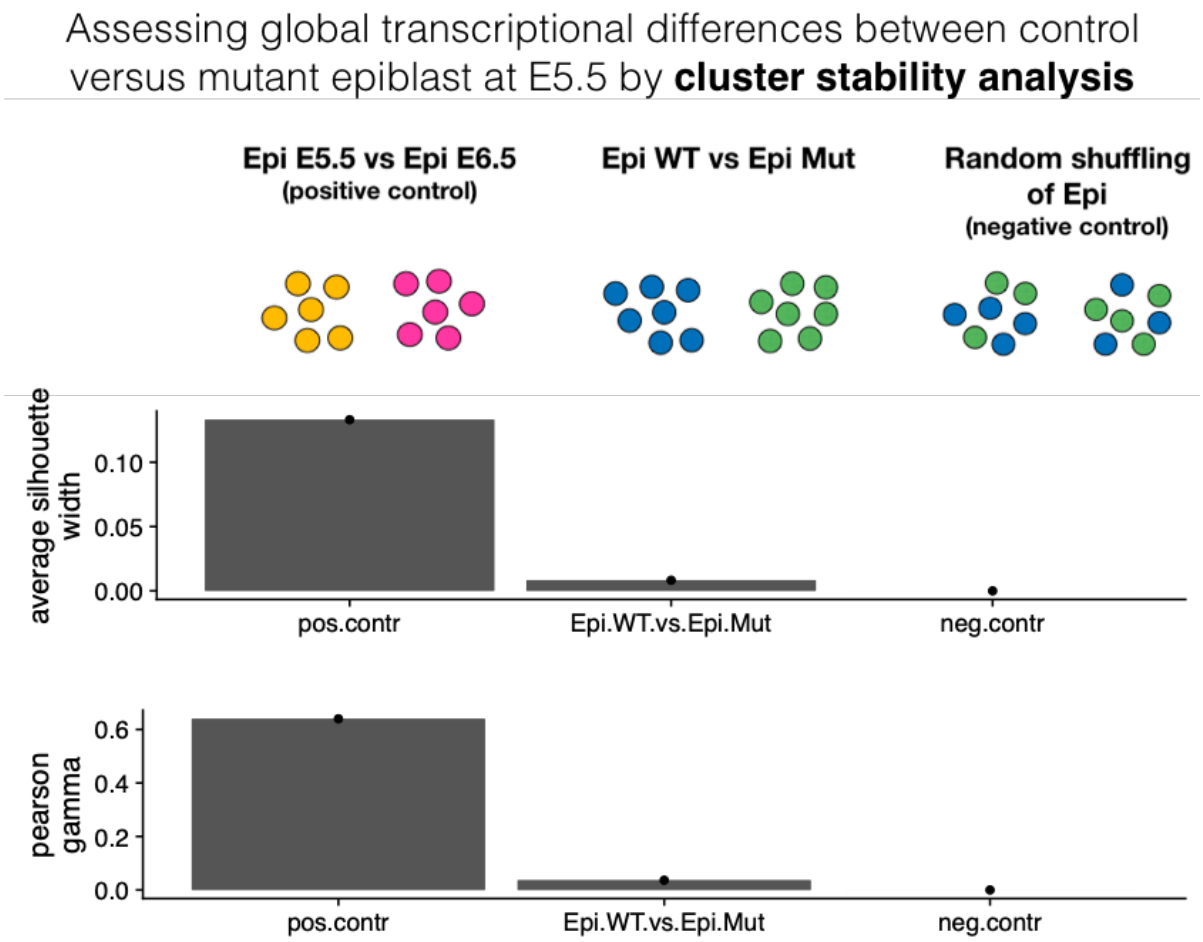


Figure 80. Cluster stability analysis additionally demonstrated absence of differences between control versus mutant epiblast at E5.5

E5.5 control versus mutant epiblast was next assessed for DEGs. First, upon evaluation of individual epiblast signature marker genes (obtained from Boroviak et al., 2015), we found *Nanog* to be significantly downregulated, and *Pou3f1* and *Sox3* to be significantly upregulated (Figure 81). In total, there were only 6 significantly upregulated (Table 13) and 6 significantly downregulated (Table 14) candidates in the epiblast, which included several genes from the ID family of transcriptional negative regulators (*Id1/2/3*) (Hong et al., 2011), and mesoderm marker *Msx1* (Drukker et al., 2012). The latter could foreshadow the mesoderm induction defect seen at E6.5 in the *Ets2*-mutant (presented in 3.2.5). However, absence of spatial information from this dataset means that it is difficult to ascertain whether the downregulation of *Msx1* is indeed expressed in the right location to contribute to mesoderm defects at E6.5.

The next step would be to validate these DEGs *in vivo*. Bolded DEGs in Tables 13 and 14 indicate cases where the antibody is available in the Srinivas laboratory. *In situ* hybridisation chain reaction (HCR) (Choi et al., 2018), which requires specific complementary probes to endogenous mRNA, may be employed for candidates where a suitable antibody is not available.

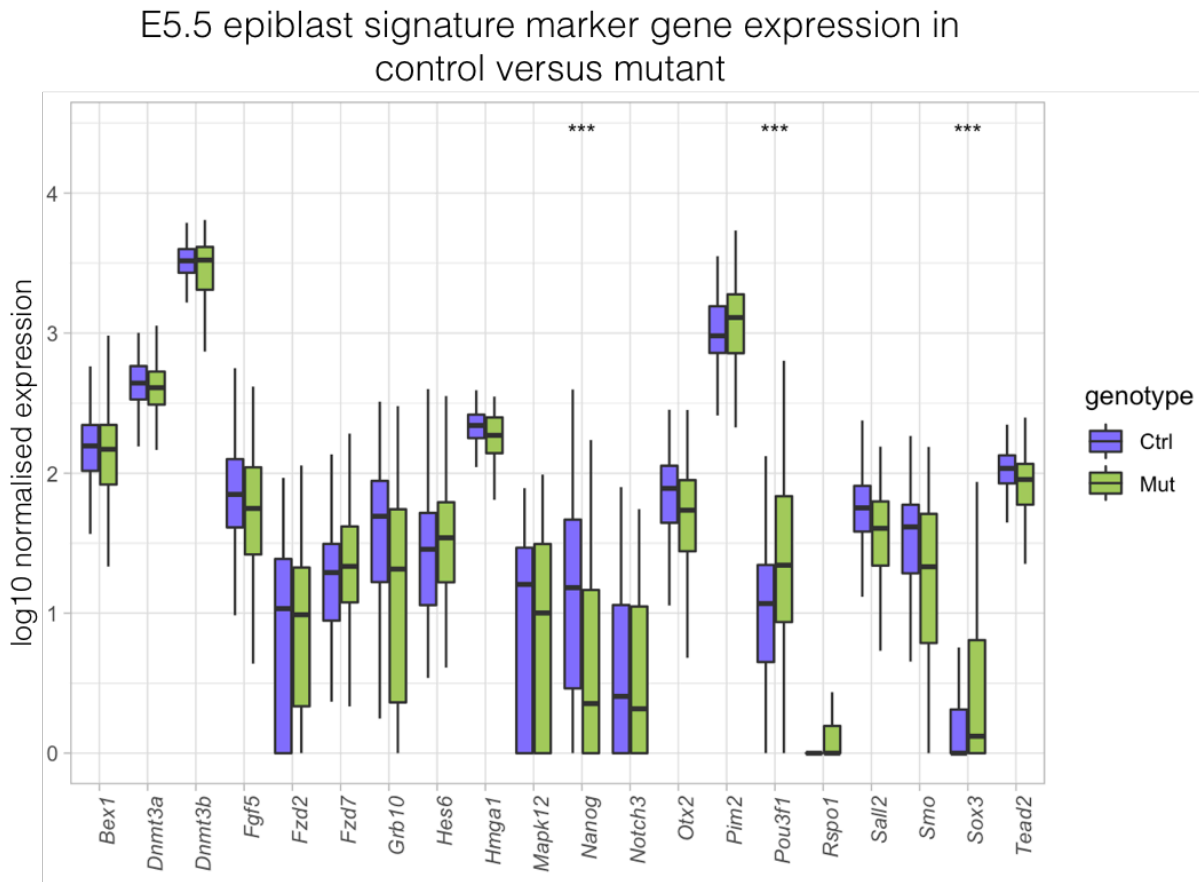


Figure 81. E5.5 control versus mutant epiblast signature marker gene expression
Boxplot showing the log10 normalised expression levels of signature marker genes for the epiblast (obtained from Boroviak et al., 2015) in our control versus mutant E5.5 epiblast scRNA-seq dataset. (Statistical test: two-way ANOVA. *** $p < 0.001$).

Table 13: **Upregulated** DE genes in the E5.5 mutant epiblast

Gene	Fold change	p-adj	Gene information
Basp1	3.02	7.49E-10	**Brain Abundant Membrane Attached Signal Protein 1:** Membrane bound protein with several transient phosphorylation sites and conserved PEST motifs.
Xist	2.60	1.22E-05	**X Inactive Specific Transcript (Non-Protein Coding):** Involved in initiation and spread of X-inactivation, an early developmental process in mammalian females that transcriptionally silences one of the pair of X chromosomes, thereby providing dosage equivalence between males and females.
Pou3f1	2.60	1.26E-12	**POU Class 3 Homeobox 1 (or Oct6):** Transcription factor that binds to the octamer motif (5'-ATTTGCAT-3'). Involved in early embryogenesis and neurogenesis.
Arf6	2.18	1.17E-08	**ADP Ribosylation Factor 6:** Member of the ARF gene family, which is part of the RAS superfamily. Localised to plasma membrane, and regulates vesicular trafficking, remodelling of membrane lipids, and signaling that results in actin remodelling.
Smarcd2	2.13	2.74E-05	**SWI/SNF-Related Matrix-Associated Actin-Dependent Regulator Of Chromatin Subfamily D Member 2:** Nuclear localisation; displays helicase and ATPase activities, and regulate transcription of certain genes by altering the chromatin structure around those genes.
Bag1	2.10	5.38E-15	**BCL2 Associated Athanogene 1:** Binds to BCL2 and enhances the anti-apoptotic effects of BCL2, linking between growth factor receptors and anti-apoptotic mechanisms.

Table 14: Top **downregulated** DE genes in the E5.5 mutant epiblast

Gene	Fold change	p-adj	Gene information
Id1	18.31	4.06E-27	**Inhibitor Of DNA Binding 1, HLH Protein:** Transcriptional repressor that can form heterodimers with members of the basic HLH family of transcription factors. Role in cell growth, senescence, and differentiation.
Id2	8.16	1.78E-25	**Inhibitor Of DNA Binding 2, HLH Protein:** Transcriptional repressor; role in negatively regulating cell differentiation.
Msx1	3.68	3.30E-06	**Msh Homeobox 1:** Member of the muscle segment homeobox family; transcriptional repressor during embryogenesis via interactions with components of the core transcription complex and other homeoproteins. Role in limb-pattern formation, craniofacial development, and tumor growth inhibition.
Id3	2.98	1.84E-11	**Inhibitor Of DNA Binding 3, HLH Protein:** Transcriptional repressor; role in cellular processes, including growth, senescence, differentiation, apoptosis, angiogenesis, and neoplastic transformation.
Nanog	2.52	5.48E-06	**Homeobox Transcription Factor Nanog:** DNA binding homeobox transcription factor involved in embryonic stem (ES) cell proliferation, self-renewal, and pluripotency. Can inhibit ES cell differentiation and can also repress its own expression in differentiating cells.
Smad7	2.18	0.000398792	**SMAD Family Member 7:** Antagonist of signaling by TGF-beta receptor type-1 (TGFBR1) superfamily members and activin signaling. Nuclear protein that binds E3 ubiquitin ligase SMURF2. Upon binding, the complex translocates to the cytoplasm where it interacts with TGFBR1 and leads to receptor degradation. Also a downstream effector of TGFBR1 signaling.

Table 13 and 14. Significantly up- and downregulated DEGs in the E5.5 mutant epiblast
Fold change (FC) and adjusted p value (p-adj; also known as 'false discovery rate') are shown. Antibodies are available in the Srinivas laboratory for genes in bold.

To further investigate whether the E5.5 mutant epiblast transcriptome may foreshadow the later stage mesoderm defect, we overlaid this dataset with genes known to positively

correlate with *T* (mesoderm marker) expression (obtained from Scialdone et al., 2016). We found 8 genes positively correlated with *T* to be significantly downregulated in the mutant epiblast (Figure 82). These included *Id1*, *Msx1*, and *Nanog.* The next step is to validate the expression patterns of these candidates *in vivo* by ISH or antibody stain. Together, these data show that although the mutant epiblast is broadly normal in its transcriptome at E5.5 – i.e. to the extent where it does not cluster separately from control by t-SNE – a few DEGs exist. These DEGs include those that positively correlate with mesoderm and therefore could relate to the E6.5 *T*-induction failure seen in mutants.

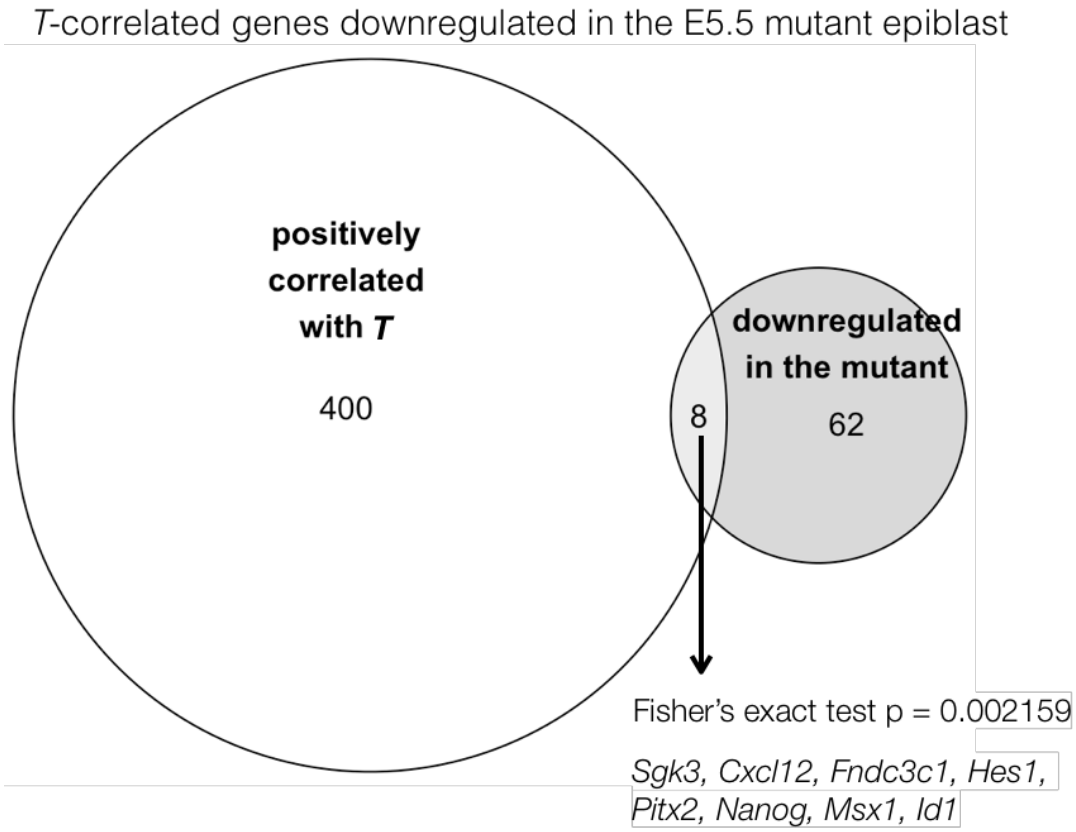


Figure 82. Identification of mesoderm (*T*)-correlated genes that are also downregulated in the E5.5 mutant epiblast
Euler diagram of the overlap between genes downregulated in the E5.5 mutant epiblast and genes positively correlated with *T* (i.e. upregulated in the primitive streak) in epiblast cells from E6.5 mouse epiblast (Scialdone et al., 2016). The 8 genes in the overlap are listed below. Note that the 62 downregulated candidates were defined by p-adj < 0.1 only. Therefore, there were more candidates here than the 6 downregulated candidates presented in Table 2 with p-adj < 0.1 and FC > 2. (Statistical test: Fisher's exact).

5.2.3 Mutant ExE at E5.5 retains signature marker expression by scRNA-seq

In 3.2.3, I showed by immunofluorescence that the reduced mutant ExE retains signature marker expression CDX2, EOMES, and ESRRB. To further validate whether the E5.5 mutant ExE retains its stereotypic character, we examined expression of 17 signature marker genes (references in Appendix A) from our control versus mutant E5.5 ExE scRNA-seq data. We found no significant difference in mutants in the expression of any of these markers, represented in heatmap and boxplot (Figure 83). Therefore, scRNA-seq independently showed that the reduced mutant ExE present at E5.5 is likely normal in its identity.

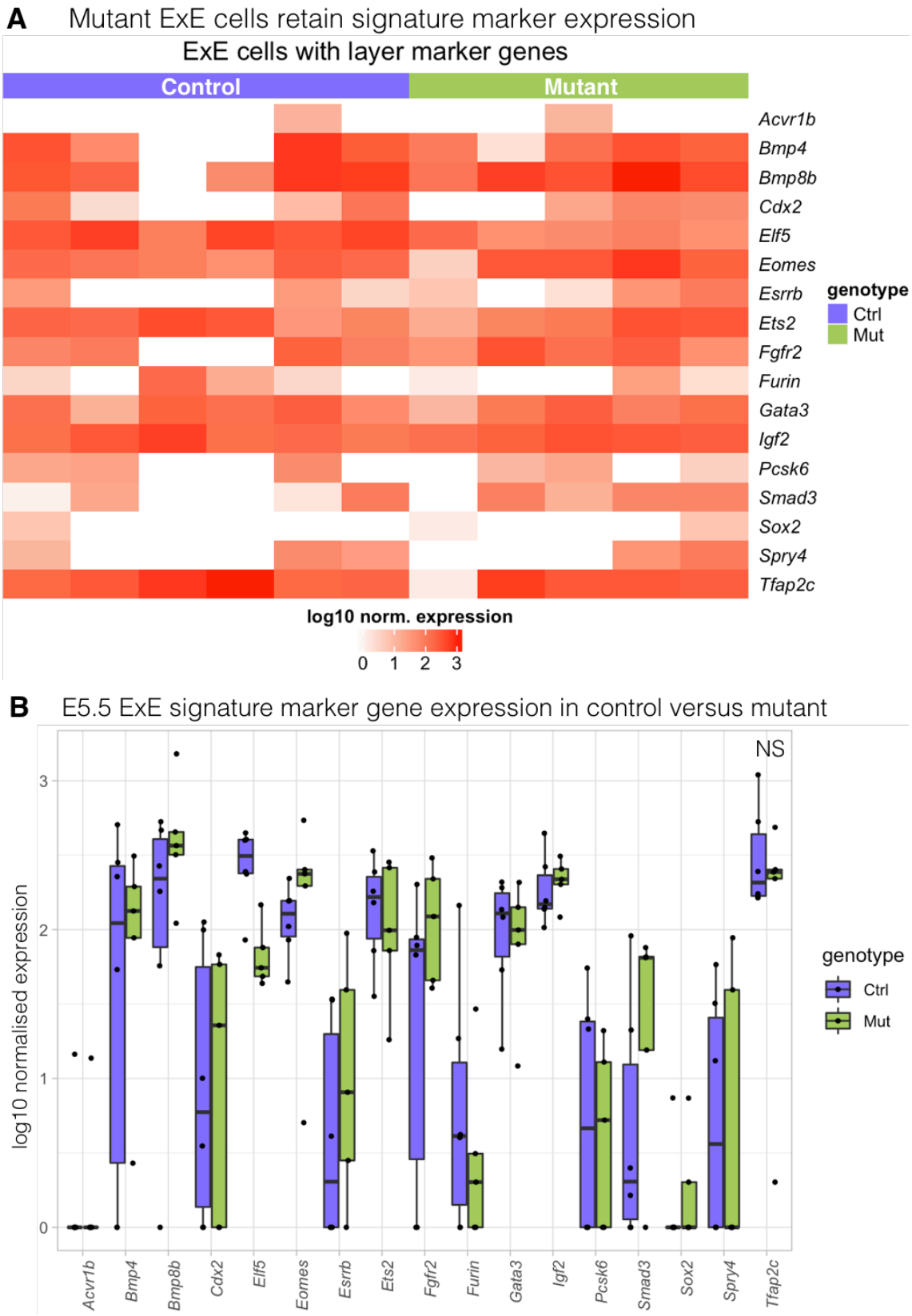


Figure 83. Signature marker gene expression of E5.5 control versus mutant scRNA-seq data, showing no significant difference in mutants for all markers

(A) Heatmap (B) Boxplot showing normalised 1og10 expression. (Statistical test: two-way ANOVA. NS = not significant). References for these ExE markers are listed in Appendix A.

5.2.4 E5.5 mutant ExE expresses normal proliferation and apoptotic pathway regulators by scRNA-seq

Chapter 3.2.6 and 3.2.7 showed that the reduced mutant ExE at post-implantation retains normal rates of proliferation and apoptosis. To verify this finding, we analysed GO terms associated with proliferation and apoptosis in our E5.5 control versus mutant ExE scRNA-seq data. Because the ExE was reduced in the mutant at E5.5, we examined GO terms: 'negative regulation of proliferation' (GO_0008285; Figure 84) and 'positive regulation of apoptosis' (GO_0043065; Figure 85). By both heatmap and SinaPlot (Sidiropoulos et al., 2018) representation, we observed no significant difference between control versus mutant E5.5 ExE for either of these GO terms. This suggests that at least at a transcriptional level, the reduced post-implantation mutant ExE has normal tissue growth and death pathway regulation. Together, these data support that the few ExE cells remaining in mutants at post-implantation may be independent of ETS2 for retention of ExE character, and normal growth and survival.

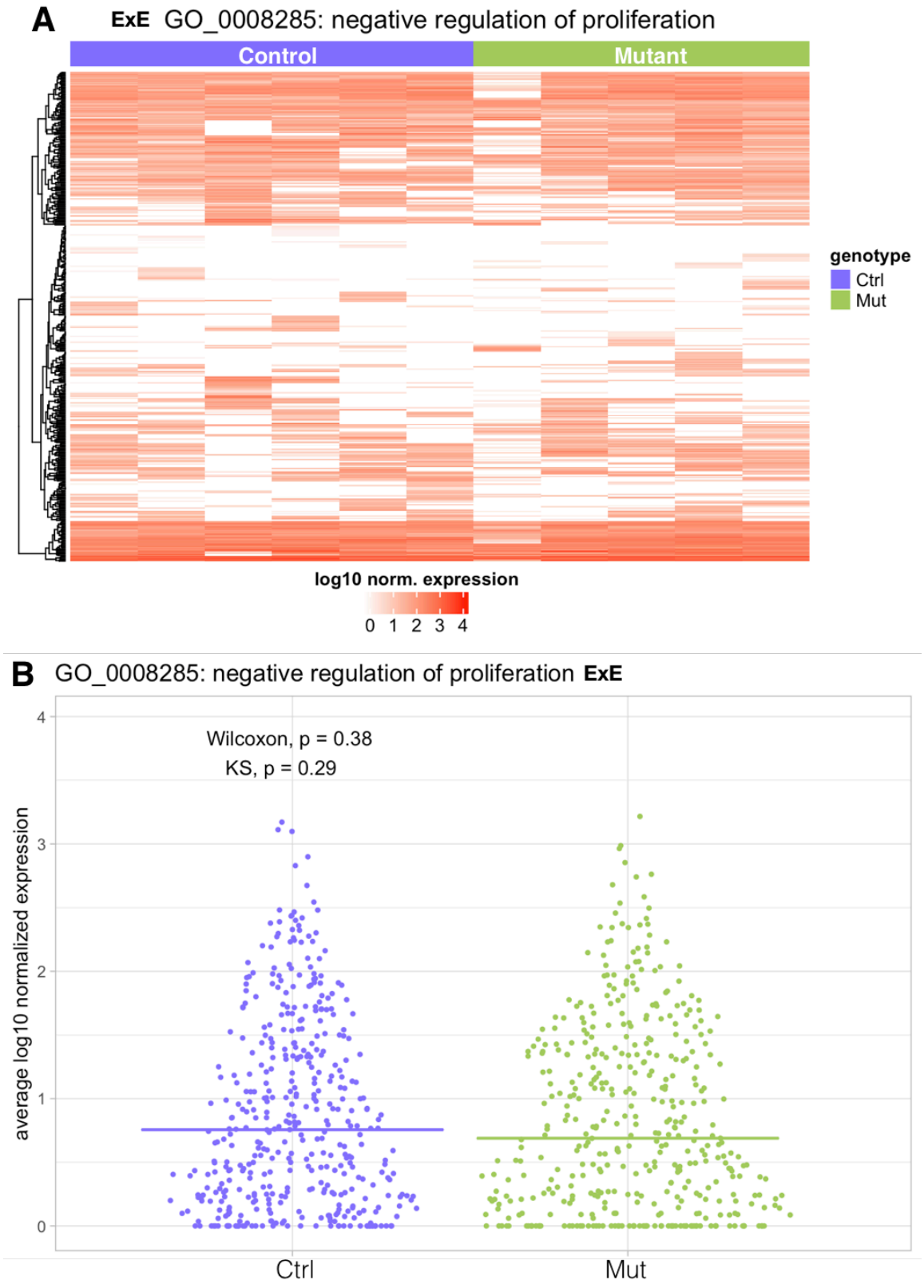


Figure 84. GO_0008285 'negative regulation of proliferation' is unchanged between control versus mutant ExE scRNA-seq data
(A) Heatmap (B) SinaPlot (Statistical tests: Student's t-test and Kolmogorov-Smirnov test for data distribution).

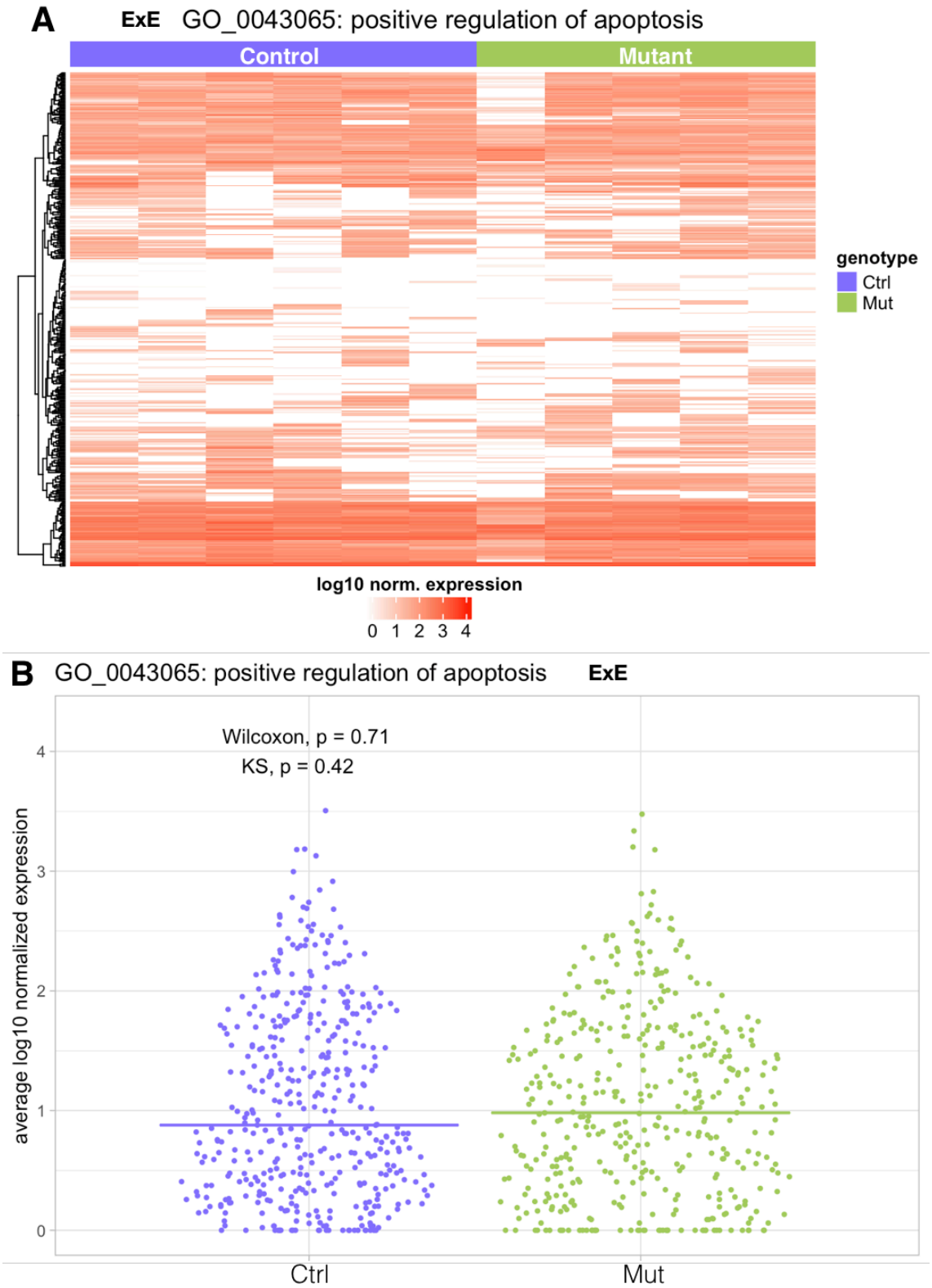


Figure 85. GO_0043065 'positive regulation of apoptosis' is unchanged between control versus mutant ExE scRNA-seq data

(A) Heatmap (B) SinaPlot. (Statistical tests: Student's t-test and Kolmogorov-Smirnov test for data distribution).

In summary, scRNA-seq of the pure 129 background in control versus mutant at E5.5 recovered mostly epiblast cells and a few ExE cells. These data were analysed to reveal that the mutant epiblast transcriptome was broadly unaltered; however, there were a few DEGs, including pluripotency factors, and genes associated with *T*-positive mesoderm that could foreshadow later stage epiblast defects see in the *Ets2*-mutant (3.2.5; Polydorou et al., 2013). The reduced mutant ExE at E5.5 retains core signature marker gene expression and normal tissue proliferation and apoptosis, supported by scRNA-seq GO analysis. This suggests that the diminished mutant ExE found at post-implantation is largely normal and possibly independent of ETS2 for growth and survival. However, this single-cell dataset did not include sufficient AVE cells to investigate the transcriptional response mediated by ETS2 that results in the mutant AVE phenotype. Therefore, I next employed a targeted approach to purify for the AVE population.

5.2.5 ScRNA-seq of HHEX-GFP-positive AVE in control versus mutant

Since the single cell sample preparation approach in 5.2.1 led to insufficient recovery of AVE cells, I devised a new method to enrich for the AVE population. To investigate the transcriptional basis for the mutant AVE phenotype, I conducted scRNA-seq at different time points of its defect progression using HHEX-GFP to purify for AVE cells. The specific stages I selected were: 1. E5.5 (dissected at 1pm) – when there is increased proliferation, but no overt thickening in the mutant AVE; and 2. E6.25 (dissected at 7am) – when the thickening defect is present and progressing to become more severe. Since E6.5 is when the mutant AVE has evident stratification, cell crowding, and increased cell number, I did not conduct scRNA-seq at that stage, as the transcriptional differences would likely be secondary effects by that point. E5.5 and E6.25 were therefore deemed to be suitable stages for tracking the initiation and progression of the mutant AVE phenotype. Sample preparation and FACS-mediated single cell

collection for HHEX-GFP-positive AVE cells are detailed in Methods 2.8. Dissociation conditions were identical to those used in 5.2.1; however, TrypLE treatment time was shortened to 3.5 minutes for E5.5 and 4.5 minutes for E6.25. These timings enabled complete single cell dissociation of the VE layer while the epiblast and ExE appeared intact by visual inspection under the microscope. Instead of phenotyping, I genotyped the Reichert's membrane by fast PCR prior to FACS-mediated single cell collection. The following is a summary workflow of mutant and control AVE scRNA-seq (Figure 86).

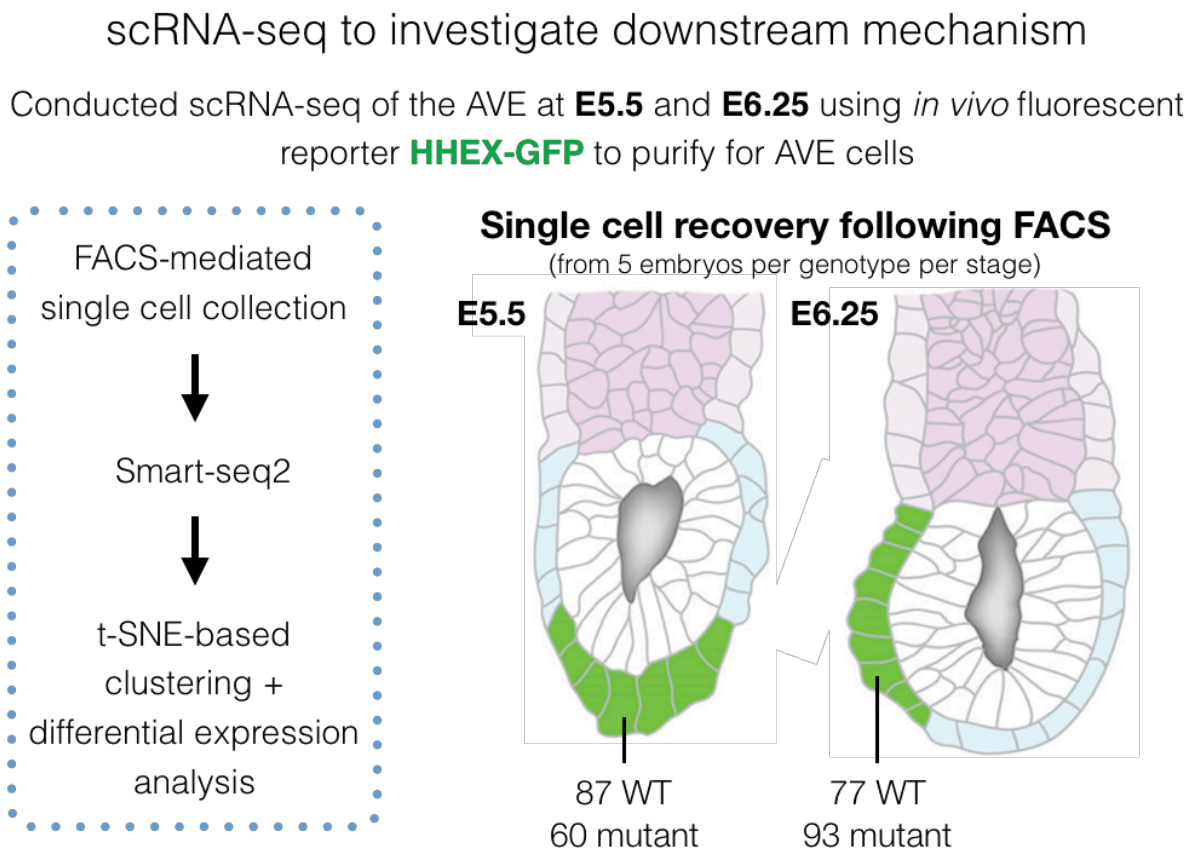


Figure 86. Workflow of scRNA-seq collection to purify for mutant and control AVE cells at E5.5 and E6.25 using HHEX-GFP fluorescent reporter

Single cell recovery following quality control analysis is shown in Table 15. The dropout rate was higher at E5.5 as compared to E6.25 potentially due to: 1. E5.5 AVE cells being more sensitive to dissociation conditions; or 2. Logistical consideration that there was a longer period of time between cell collection and sequencing for E5.5 cells (during which time the cells were stored in lysis buffer at -80°C), possibly leading to RNA degradation.

HHEX-GFP-positive AVE single cell recovery

Stage	Number of cells collected	Single cell recovery after QC	Dropout rate
E5.5 (from 5 embryos per genotype)	WT = 87 Mut = 60	WT = 43 Mut = 27	WT = 50.6% Mut = 55%
E6.25 (from 5 embryos per genotype)	WT = 77 Mut = 93	WT = 67 Mut = 89	WT = 13% Mut = 4.3%

Table 15. HHEX-GFP-positive AVE single cell recovery from FACS, and after Smart-seq2 QC analysis

Heatmap analysis of cell layer markers at post-implantation and t-SNE based clustering (Figure 87) revealed that most cells recovered were indeed AVE cells, and that mutants do not cluster separately from control. This was applicable for both E5.5 and E6.25 mutant AVE scRNA-seq data. This demonstrates that the mutant AVE retains its stereotypic character and is broadly normal in its transcriptome, similar to the E5.5 mutant ExE and epiblast. The observation was further validated by cluster stability analysis, which had been employed in 5.2.2 to examine E5.5 mutant versus control epiblast. Both cluster quality indices of 'pearson gamma' and 'average silhouette width' did not significantly differ in the mutant AVE (Figure 88) compared to known positive and negative controls (again similar to those used for 5.2.2). This independently showed that the mutant AVE at E5.5 and E6.25 is normal in its overall transcriptional profile.

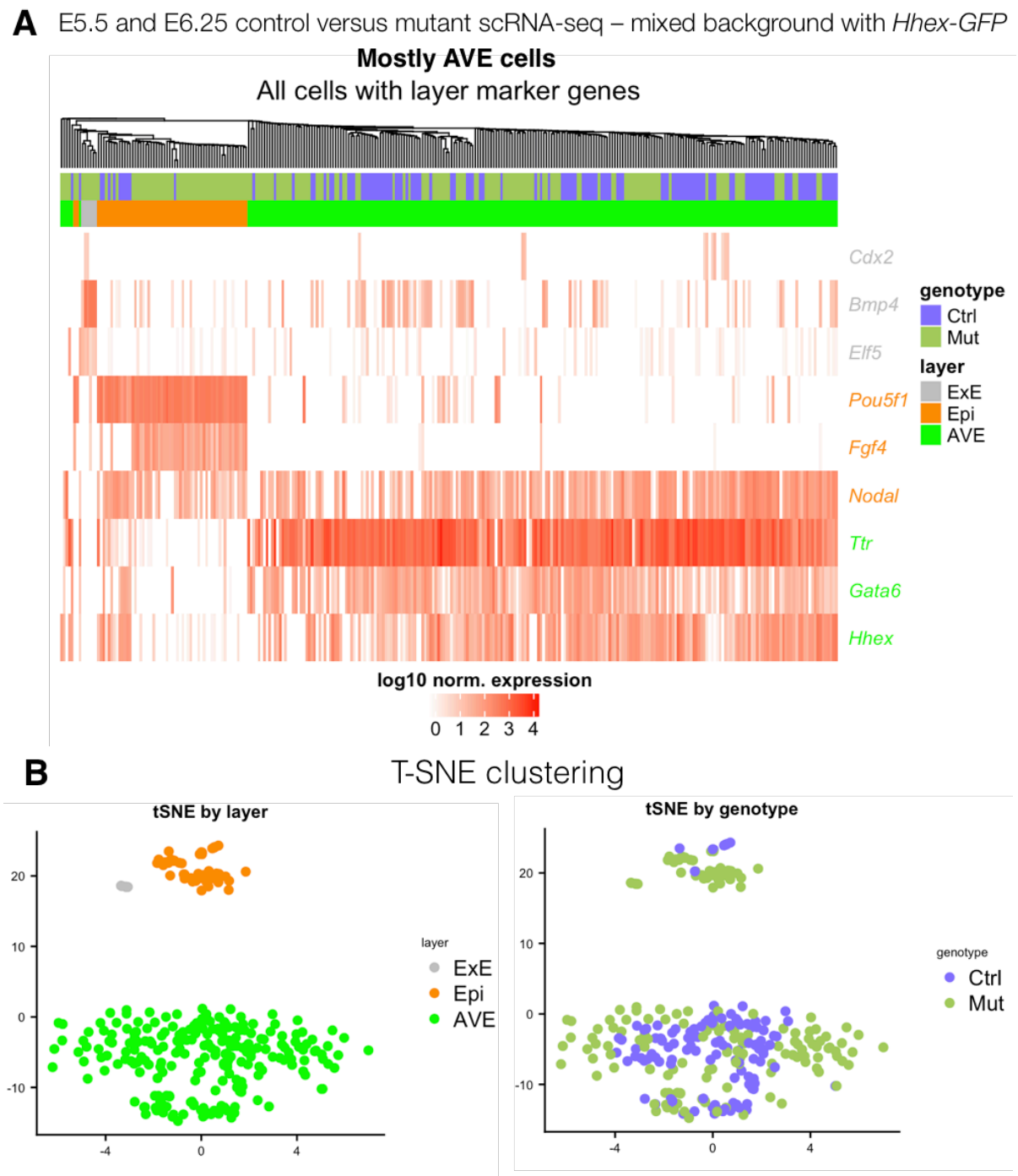


Figure 87. E5.5 and E6.25 HHEX-GFP-positive (mixed background) scRNA-seq – heatmap clustering based on tissue-specific marker genes and t-SNE

(A) Heatmap clustering based on E5.5 layer marker genes from Mohammed et al., 2017. Control and mutants do not segregate. Most cells are AVE.

(B) T-SNE clustering of all cells by layer markers, again showing presence of mostly AVE cells and fewer ExE and epiblast cells. T-SNE clustering by genotype is also presented where control and mutant do not cluster separately.

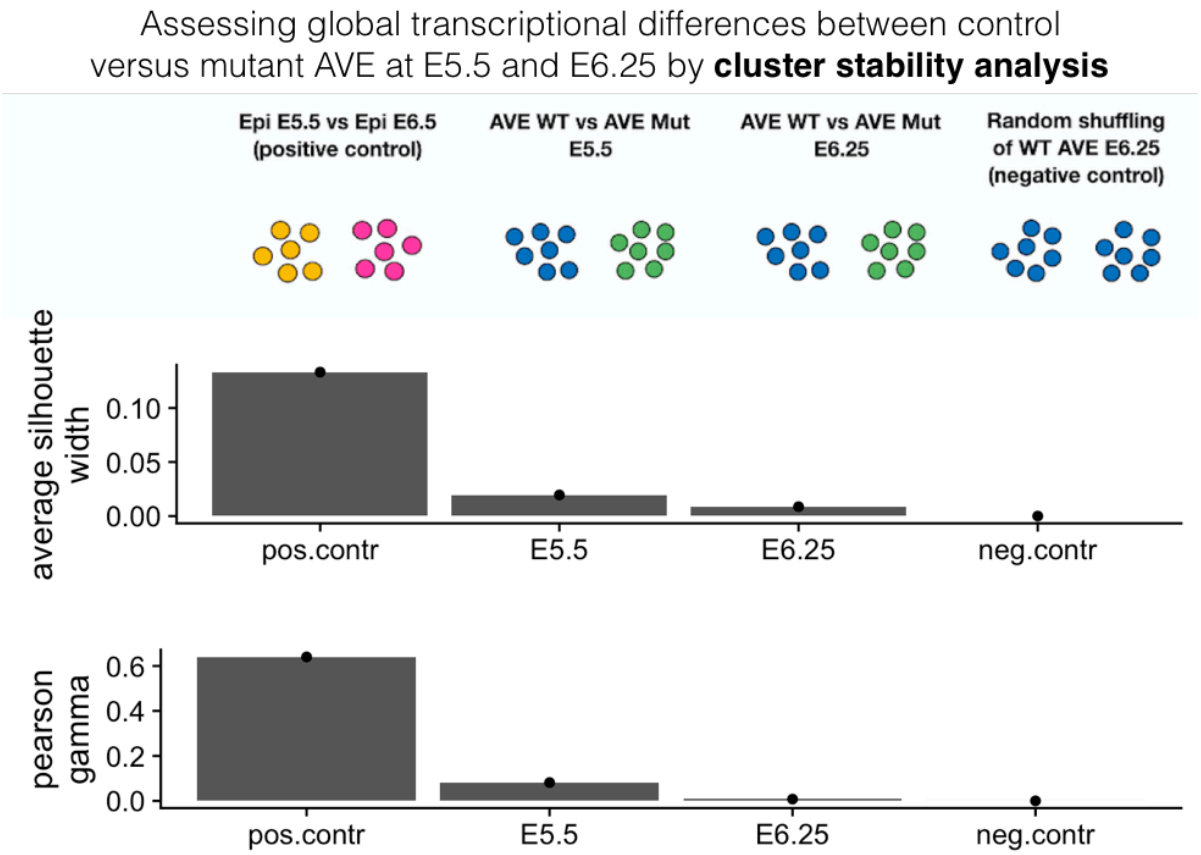


Figure 88. Cluster stability analysis additionally demonstrated absence of differences between control versus mutant AVE at E5.5 and E6.25

Statistical comparison (by two-way ANOVA) of known signature AVE markers revealed no significant differences in expression levels between control versus mutant at E5.5 (Figure 89A). However, there was a significant downregulation in *Cer1* and *Lefty1* expression in mutants by E6.25 (Figure 89B). Presence of normal expression levels of other AVE markers in mutants at E6.25 suggests that the decrease in *Cer1* and *Lefty1* expression could indicate fewer cells of a particular AVE subpopulation in mutants, in line with the AVE being a heterogeneous population (Hoshino et al., 2015).

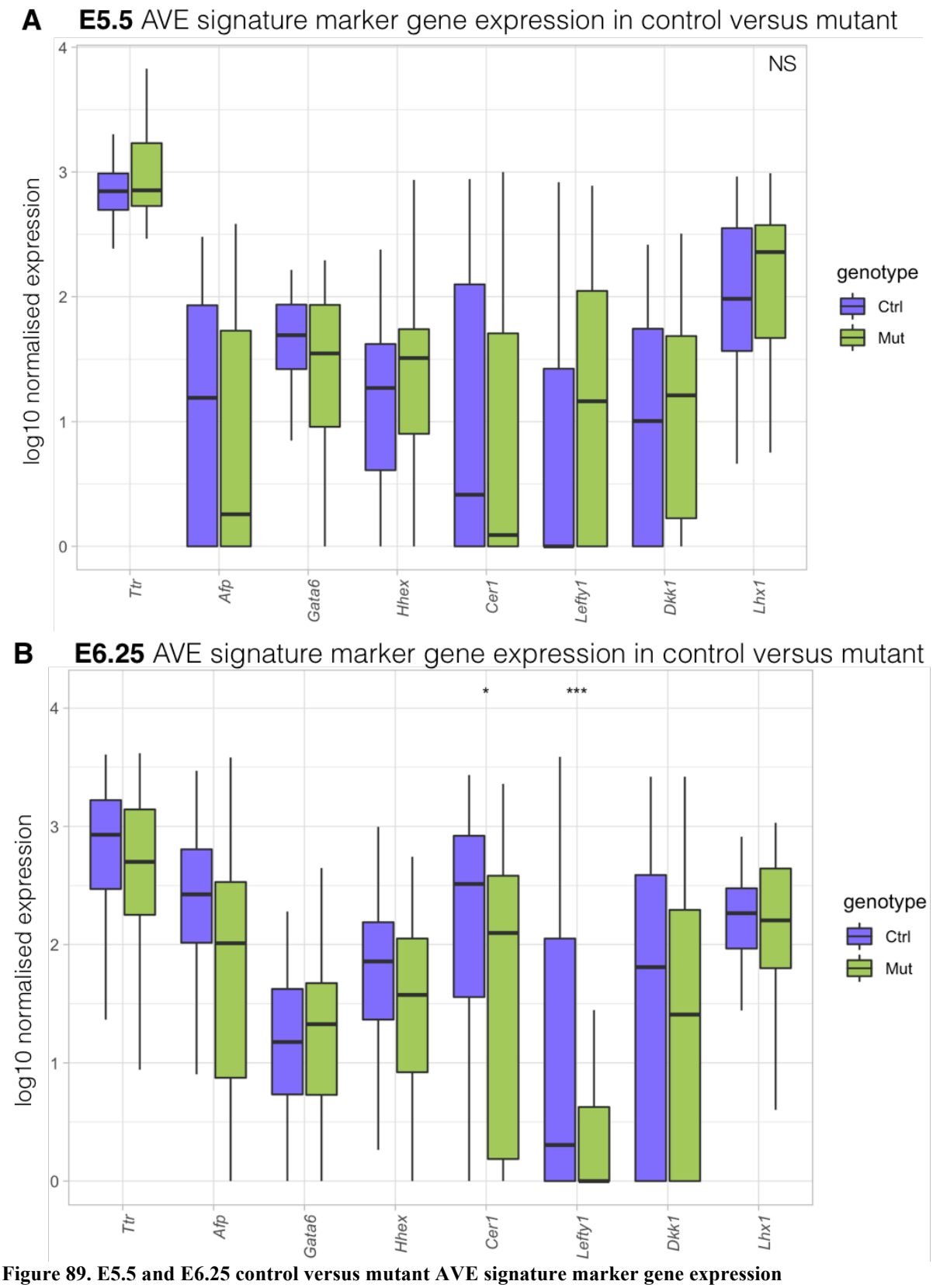


Figure 89. E5.5 and E6.25 control versus mutant AVE signature marker gene expression
Boxplot showing the log10 normalised expression levels of marker genes for epiblast at E5.5 (obtained from Scialdone et al., 2016) for control versus mutant AVE cells. (Statistical test: two-way ANOVA. NS = not significant; * $p < 0.05$; *** $p < 0.001$).

In summary, mutant AVE cells have stereotypic tissue-specific marker expression (*Lefty1, Cer1, Hhex, Lhx1, Dkk1*), and simultaneously express VE markers (*Gata6, Afp, Ttr*), further verifying that mutant AVE cells retain normal character. This has been confirmed *in vivo* where mutant AVE cells express tissue-specific markers HHEX-GFP and AMOT by genetic fluorescent reporter and antibody stain, respectively (4.2.2). Together, these observations decrease the possibility of asymmetric cell division, i.e. increased proportion of perpendicular cell divisions in the mutant AVE (4.2.7), to result in changes in cell fate in the superficial layers of the stratified epithelium. While asymmetric cell divisions have been shown in early neurodevelopment to be a means to achieve cellular diversification (Lancaster and Knoblich, 2012), this mechanism does not appear to be applicable to the mutant AVE.

5.2.6 GO analysis of biological processes and signalling pathways misregulated in the mutant AVE

I next conducted GO analysis on entire statistically significant DEG lists in the mutant AVE, defined by fold change (FC) greater than 2 and false discovery rate (FDR; or p-adj) less than 0.1. With these cutoff parameters, there were a total of 22 DE genes at E5.5 and 910 at E6.25, with only 12 genes that overlapped between the two stages (Figure 90).

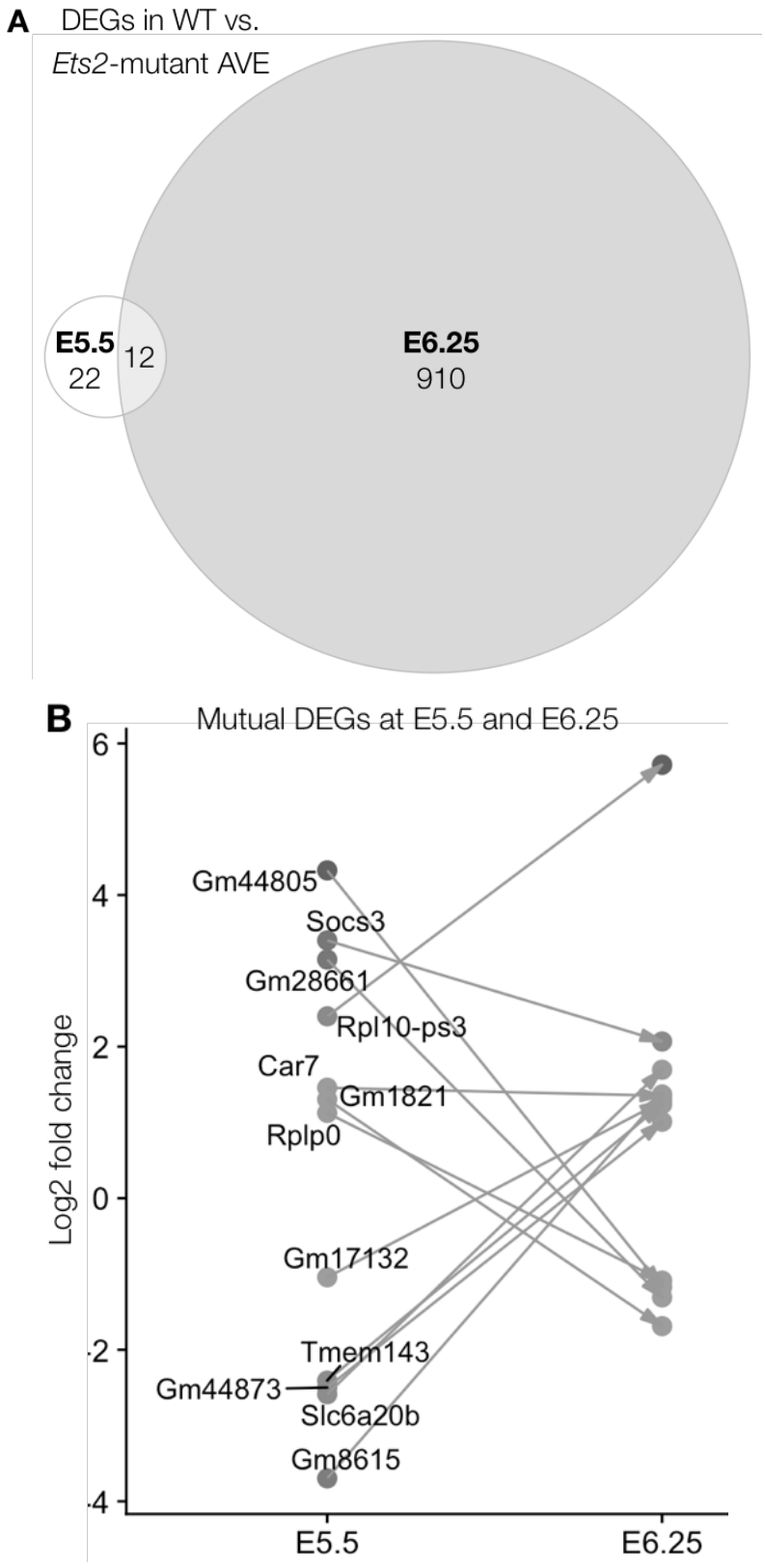


Figure 90. Number of DEGs in the mutant AVE at E5.5 and E6.25 with FC > 2 and p-adj < 0.1
(A) Venn diagram showing the relative proportion of mutant DEGs at E5.5 and E6.25
(B) 12 DEGs that overlap in the E5.5 and E6.25 mutant AVE.

To investigate genes and pathways enriched in mutant AVE DEG lists, I conducted GO analysis using several algorithms to assess top misregulated biological processes. These include: Princeton GO::TermFinder (Boyle et al., 2004), Panther (Mi et al., 2013), Reactome (Fabregat et al., 2018), and g:Profiler (Reimand et al., 2016). Recurrent genes and pathways from each software verified the significance of top enriched pathways dysregulated in mutants.

For the E5.5 mutant AVE, there were too few genes in the DE list for any statistically significant GO terms to arise. In the E6.25 mutant AVE, the enriched biological processes in order of most to least significant were: metabolism, cell cycle, positive regulation of gene expression, chromatin organisation, histone modification, cytoskeletal organisation, protein ubiquitination, cell projection, microtubule-based process, maintenance of cell number, and stem cell population maintenance (Figure 92). Amongst the signalling pathways, the most significantly enriched was WNT (Figure 91 – showing PANTHER and Reactome GO results as an example). The WNT pathway was more enriched than the second most significant signalling pathway TGFβ. Although other biological processes (e.g. metabolism, cell cycle) were more enriched than the WNT pathway (Figure 92), I chose to focus on WNT, as the pathway may be more directly targeted by experimentation.

A **WNT enrichment**

PANTHER Classification System

Mus musculus (REF) #	#	expected	Fold Enrichment	+/-	raw P value
positive regulation of canonical Wnt signaling pathway	93	39	21.05	1.85	+	2.17E-03
↳positive regulation of Wnt signaling pathway	124	55	28.07	1.96	+	6.42E-05
↳regulation of Wnt signaling pathway	280	117	63.38	1.85	+	7.33E-08
↳regulation of signal transduction	2650	837	599.83	1.40	+	1.57E-18
↳regulation of cell communication	3018	922	683.12	1.35	+	3.48E-17
↳regulation of signaling	3043	931	688.78	1.35	+	1.73E-17
↳regulation of response to stimulus	3709	1038	839.53	1.24	+	5.73E-11
↳positive regulation of signal transduction	1441	437	326.17	1.34	+	6.54E-08
↳positive regulation of cell communication	1623	483	367.37	1.31	+	8.38E-08
↳positive regulation of signaling	1630	486	368.95	1.32	+	5.53E-08
↳regulation of canonical Wnt signaling pathway	213	84	48.21	1.74	+	3.14E-05
Wnt signaling pathway	244	94	55.23	1.70	+	2.11E-05
↳cell-cell signaling by wnt	244	94	55.23	1.70	+	2.11E-05
↳cell surface receptor signaling pathway involved in cell-cell signaling	291	104	65.87	1.58	+	1.08E-04

B

reactome

WNT enrichment

Pathway name	Entities found
Mitotic Metaphase and Anaphase	59
Formation of tubulin folding intermediates by CCT/TriC	2
G2/M Transition	60
Synthesis of bile acids and bile salts via 27-hydroxycholesterol	3
Retinoid metabolism and transport	15
Incretin synthesis, secretion, and inactivation	6
Mitochondrial tRNA aminoacylation	6
Viral mRNA Translation	26
Synthesis, secretion, and inactivation of Glucagon-like Peptide-1 (GLP-1)	5
Chromosome Maintenance	23
Integrin cell surface interactions	17
Organic cation transport	4
RA biosynthesis pathway	4
Signaling by WNT	104
HSF1-dependent transactivation	9
RHO GTPases activate PKNs	15
Synthesis of Ketone Bodies	1
p53-Dependent G1 DNA Damage Response	12
p53-Dependent G1/S DNA damage checkpoint	12
MAPK6/MAPK4 signaling	23

Figure 91. PANTHER and Reactome GO analyses of E6.25 mutant AVE recurrently show significant enrichment of the WNT pathway

WNT pathway enrichment found upon analysis of entire mutant AVE DE list using (A) Panther GO (Mi et al., 2013), (B) Reactome (Fabregat et al., 2018).

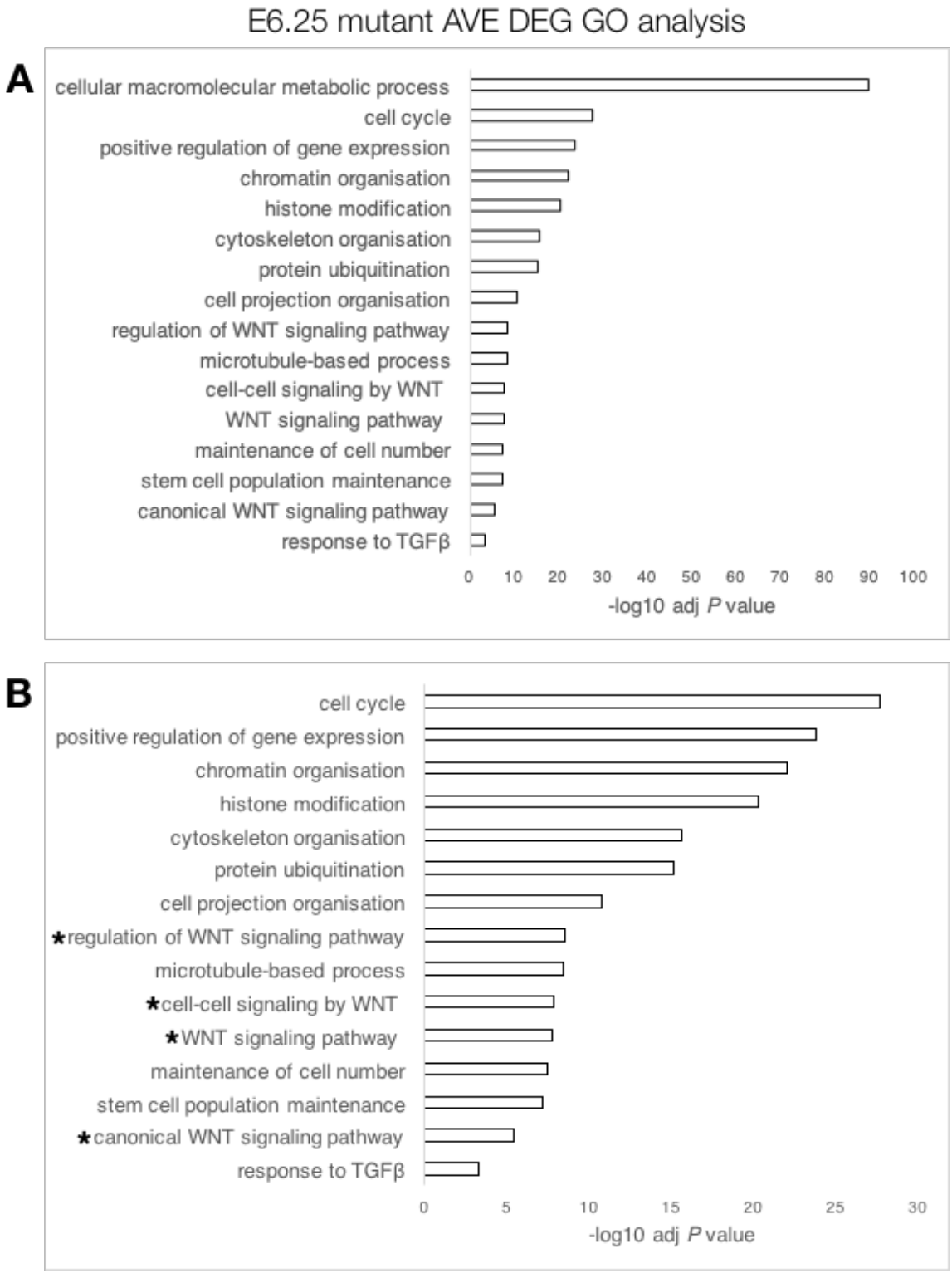


Figure 92. Top enriched GO terms in E6.25 mutant AVE DEG list
(A) Including metabolic process.
(B) Excluding metabolic process for clearer visualisation of the relative differences amongst other GO terms. WNT is the most enriched signalling pathway, where several WNT-related terms (asterisk) were all more significant than the second most enriched pathway TGFβ.

5.2.7 Top misregulated genes in the mutant AVE

Top DEGs in the mutant AVE (defined by statistics) were next assessed with a focus on the WNT pathway. At E5.5, the entire DEG list was examined, as there were only a small number of significantly misregulated candidates with FC > 2 and p-adj < 0.1 (Table 16, 17 – these lists were ordered by descending FC). 11/34 upregulated candidates and 5/13 downregulated candidates were part of the WNT/β-CATENIN pathway. The top upregulated candidate in the mutant AVE at E5.5 that was not a pseudogene was *Moesin* (Msn; Table 16).

E6.25 mutant AVE top upregulated DEGs with p-adj < 0.001 (stringent cutoff) and FC > 2 were next analysed in relation to WNT. There was a total of 124 genes that met these criteria (Table 18, 19) when I excluded non-coding genes, pseudogenes and Rik genes (Table 20), which were difficult to interpret and classify to a biological process. By E6.25, I hypothesised that the effects of WNT/β-CATENIN pathway activation should be evident in the mutant AVE, shown by significant upregulation of WNT pathway members. Upon annotation of the top 124 upregulated candidates, I found 43 WNT upstream signalling mediators and 23 WNT downstream effectors. In addition, 26 out of the top 124 candidates, which were not annotated to the WNT pathway, have been reported as oncogenes (GSEA; Subramanian et al., 2005). Upregulation of oncogenes fits with the observation that the mutant AVE becomes overproliferative and transforms from a simple to stratified epithelium (Muroyama and Lechler, 2012). Therefore, both GO analysis of the entire DE list, as well as inspection of the top 124 most statistically significant candidates in the E6.25 mutant AVE demonstrated that WNT/β-CATENIN is the most enriched signalling pathway misregulated in the mutant AVE. This pathway then became the focus for subsequent *in vivo* validation.

E5.5 mutant AVE **upregulated** candidates

Rank	Gene	Name	Fold change	p-adj (FDR)	WNT pathway?
1	ENSMUSG00000108443	predicted gene	134.9575937	3.76E-05	
2	ENSMUSG00000086922	processed pseudogene	127.5250633	0.053552098	
3	ENSMUSG00000084416	pseudogene	121.594611	0.053552098	
4	Msn	Moesin	80.85408378	0.041666997	yes
5	Slc5a5	Solute Carrier Family 5 (Sodium/Iodide Cotransporter), Member 5	43.09341642	0.041666997	
6	Rac3	Rac Family Small GTPase 3	38.90336811	0.053552098	WNT/PCP
7	Fsd1	Fibronectin Type III And SPRY Domain Containing 1	36.64189458	0.053552098	
8	Tdgf1	Teratocarcinoma-Derived Growth Factor 1	28.52039329	0.07781677	yes
9	Slc38a1	Solute Carrier Family 38 Member 1	26.91557979	0.067791805	yes
10	Sult4a1	Sulfotransferase Family 4A Member 1	26.06300849	0.07781677	
11	Fos	Fos Proto-Oncogene, AP-1 Transcription Factor Subunit	25.3358638	0.053552098	yes - target
12	Irak4	Interleukin 1 Receptor Associated Kinase 4	19.53660836	0.07781677	WNT/Calcium
13	Socs3	Suppressor Of Cytokine Signaling 3	16.48406084	0.003572224	
14	Hist1h2ag	Histone Cluster 1 H2A Family Member G	12.81037993	0.062945223	
15	ENSMUSG00000097808		8.312962707	0.059164915	
16	ENSMUSG00000038155		8.312962707	0.059164915	
17	Hist1h2ae	Histone Cluster 1 H2A Family Member E	6.19694299	0.075259202	
18	Hist1h2ap		4.954145181	0.001217898	
19	ENSMUSG00000087299		3.978577403	0.080090527	
20	Dctpp1	DCTP Pyrophosphatase 1	3.655269037	0.062945223	yes - target
21	ENSMUSG00000044285		3.266570907	0.07781677	
22	Lsm7	LSM7 Homolog, U6 Small Nuclear RNA And MRNA Degradation Associated	3.045434705	0.053552098	
23	Rplp0	Ribosomal Protein Lateral Stalk Subunit P0	2.920069803	0.008973357	
24	Car7	carbonic anhydrase 7	2.897235553	0.001217898	
25	Cited1	Cbp/P300 Interacting Transactivator With Glu/Asp Rich Carboxy-Terminal Domain 1	2.782191503	0.097206686	yes
26	Fau	FAU, Ubiquitin Like And Ribosomal Protein S30 Fusion	2.768307308	0.075469116	
27	Atox1	Antioxidant 1 Copper Chaperone	2.624457965	0.053552098	
28	Arpc1b	Actin Related Protein 2/3 Complex Subunit 1B	2.590449501	0.099043662	
29	Clic1	Chloride Intracellular Channel 1	2.492737506	0.080090527	
30	Atp5g1	ATP Synthase, H+ Transporting, Mitochondrial Fo Complex, Subunit C1	2.484847793	0.07983249	yes
31	Pga5	Pepsinogen 5, Group I (Pepsinogen A)	2.431694569	0.060242874	yes - target
32	Cox6c	Cytochrome C Oxidase Subunit 6C	2.402575703	0.07781677	yes - target
33	Sparc	Secreted Protein Acidic And Cysteine Rich	2.356397785	0.046810458	yes
34	Spink1	Serine Peptidase Inhibitor, Kazal Type 1	2.268726865	0.052161147	yes

Table 16. Significantly upregulated DEGs in the E5.5 mutant AVE
FC > 2, p-adj < 0.1. Relation to WNT pathway is indicated where applicable.

E5.5 mutant AVE **downregulated** candidates

Rank	Gene	Name	Downregulated fold change	p-adj (FDR)	WNT pathway?
1	Immunoglobulin Superfamily Member 5	Igsf5	78.65946751	0.031265675	
2	RAS Guanyl Releasing Protein 3	Rasgrp3	54.28631112	0.034347038	yes
3	Serine/Threonine Kinase 10	Stk10	49.08901873	0.067791805	yes
4	Proprotein Convertase Subtilisin/Kexin Type 4	Pcsk4	47.55809427	0.067791805	
5	SMAD Family Member 6	Smad6	47.4938918	0.053552098	yes
6	Small VCP Interacting Protein	Svip	44.91042561	0.053552098	
7	Calmodulin-Lysine N-Methyltransferase	Camkmt	27.74597175	0.067791805	
8	Slc6a20b solute carrier family 6 (neurotransmitter transporter), member 20B	Slc6a20b	22.40933333	0.046810458	
9	EF-Hand Calcium Binding Domain 11	Efcab11	16.07713865	0.095046679	
10	TEF, PAR BZIP Transcription Factor	Tef	11.54404257	0.05590773	
11	Transmembrane Protein 143	Tmem143	9.768933125	0.024801678	
12	Inhibitor Of DNA Binding 3, HLH Protein	Id3	7.135379329	0.003572224	yes
13	Inhibitor Of DNA Binding 1, HLH Protein	Id1	2.985922692	0.041666997	yes

Table 17. Significantly downregulated DEGs in the E5.5 mutant AVE
FC > 2, p-adj < 0.1. Relation to WNT pathway is indicated where applicable.

E6.25 mutant AVE **upregulated** candidates (top 62)

Rank	Gene	Name	Fold Change	p-adj (FDR)	WNT pathway? Oncogene?
1	Zfp936	zinc finger protein 936	9.138783902	0.000298751	
2	ENSMUSG00000110974	JunD Proto-Oncogene, AP-1 Transcription Factor Subunit	7.890876761	1.22E-11	WNT
3	ENSMUSG00000071076	Transcription factor jun-D	7.890876761	1.22E-11	WNT downstream
4	Acvr1b	Activin A Receptor Type 1B	5.486755277	0.000127637	WNT downstream
5	Mmab	Methylmalonic Aciduria (Cobalamin Deficiency) CblB Type	5.383055469	0.000883062	
6	Slc25a23	Solute Carrier Family 25 Member 23	5.276474236	0.000211891	
7	Chmp4b	Charged Multivesicular Body Protein 4B	5.034300616	4.46E-07	WNT upstream
8	Ldlr	low density lipoprotein receptor	4.924129366	1.10E-08	WNT upstream
9	Bhlhe41	Basic Helix-Loop-Helix Family Member E41	4.915684093	2.09E-34	oncogene
10	Vegfa	Vascular Endothelial Growth Factor A	4.792133257	2.15E-07	WNT downstream
11	Gfpt1	Glutamine--Fructose-6-Phosphate Transaminase 1	4.718872892	5.56E-08	oncogene
12	Fbxw7	F-Box And WD Repeat Domain Containing 7	4.573471577	0.000236775	WNT upstream
13	Pura	Purine Rich Element Binding Protein A	4.185949174	0.000515605	oncogene
14	Scd1	Stearoyl-CoA Desaturase	4.062400783	2.13E-06	WNT upstream
15	Gas2l3	Growth Arrest Specific 2 Like 3	3.918907829	0.000573369	WNT downstream
16	Pank3	Pantothenate Kinase 3	3.906092549	2.77E-06	
17	Tle4	Transducin Like Enhancer Of Split 4	3.8396984	0.000694421	WNT upstream
18	Rictor	RPTOR Independent Companion Of MTOR Complex 2	3.829598131	2.59E-06	WNT upstream
19	Zfp36	ZFP36 Ring Finger Protein	3.65670491	0.000200118	WNT upstream
20	Ttyh3	Tweety Family Member 3	3.588670798	2.84E-06	WNT upstream
21	Naa30	N(Alpha)-Acetyltransferase 30, NatC Catalytic Subunit	3.515439082	2.63E-05	oncogene
22	H19	H19, Imprinted Maternally Expressed Transcript (Non-Protein Coding)	3.501252868	1.50E-08	WNT upstream
23	Xiap	X-Linked Inhibitor Of Apoptosis	3.411060068	7.16E-10	WNT upstream
24	Lgr4	Leucine Rich Repeat Containing G Protein-Coupled Receptor 4	3.374944221	6.78E-07	WNT upstream
25	Slc6a20b	Sodium- and chloride-dependent transporter XTRP3B	3.331741099	0.000108691	WNT downstream
26	Ubqln2	Ubiquilin 2	3.314262481	3.11E-09	WNT upstream
27	Ppp1r9a	Protein Phosphatase 1 Regulatory Subunit 9A	3.303132023	0.000836868	WNT upstream
28	Mef2d	Myocyte Enhancer Factor 2D	3.284824155	0.000808127	WNT upstream
29	Smim1	Small Integral Membrane Protein 1	3.188806471	0.000890099	
30	Ldb1	LIM Domain Binding 1	3.180085728	1.94E-05	WNT upstream
31	Dhcr24	24-Dehydrocholesterol Reductase	3.157688082	1.14E-08	WNT downstream
32	Zfc3h1	Zinc Finger C3H1-Type Containing	3.152111769	0.000564165	
33	Igf2	Insulin Like Growth Factor 2	3.151617127	2.24E-05	WNT upstream
34	Meg3	Maternally Expressed 3	3.139654151	1.92E-11	WNT upstream
35	Fam193b	Family With Sequence Similarity 193 Member B	3.137970088	0.000433777	
36	Ap1s3	Adaptor Related Protein Complex 1 Sigma 3 Subunit	3.130838747	0.000119258	
37	Msi2	Musashi RNA Binding Protein 2	3.117706234	2.91E-08	WNT upstream
38	Ak9	Adenylate Kinase 9	3.101061546	0.000170302	oncogene
39	Rfx7	Regulatory Factor X7	3.077919806	1.21E-05	
40	Dock10	dedicator of cytokinesis 10	3.038009798	1.01E-20	WNT upstream
41	Casc3	Cancer Susceptibility 3	3.004243634	2.14E-05	WNT upstream
42	Olfr1307	olfactory receptor 1307	2.992864869	9.39E-07	
43	Plxna4	Plexin A4	2.982733164	2.14E-05	oncogene
44	Cyp51	Cytochrome P450 Family 51 Subfamily A Member 1	2.939912513	0.000429018	WNT downstream
45	Vmn1r218	vomeronasal 1 receptor 218	2.933579928	0.000242996	
46	Tmem189	Transmembrane Protein 189	2.889435583	0.000279449	
47	Slfn5	Schlafen Family Member 5	2.886859057	1.36E-06	WNT upstream
48	Setd1b	SET Domain Containing 1B	2.876348897	2.64E-05	WNT downstream
49	Ptms	Parathymosin	2.708176011	6.09E-07	
50	Dynll2	Dynein Light Chain LC8-Type 2	2.691203634	7.84E-05	
51	Ccni	Cyclin I	2.690917299	1.70E-07	oncogene
52	Naf1	Nuclear Assembly Factor 1 Ribonucleoprotein	2.689443877	0.000375338	
53	Kmt2c	Lysine Methyltransferase 2C	2.68760756	0.000411184	WNT upstream
54	Tmem259	Transmembrane Protein 259	2.677099412	0.0002646	
55	Agap3	ArfGAP With GTPase Domain, Ankyrin Repeat And PH Domain 3	2.668834004	0.000899095	oncogene
56	Kcnq1ot1	KCNQ1 Opposite Strand/Antisense Transcript 1 (Non-Protein Coding)	2.651658864	5.30E-09	WNT downstream
57	Tpcn1	Two Pore Segment Channel 1	2.62028192	9.25E-05	WNT upstream
58	Cux2	Cut Like Homeobox 2	2.618335731	4.68E-05	WNT upstream
59	Brd1	Bromodomain Containing 1	2.613736975	0.000116243	oncogene
60	Tnfrsf14	TNF Receptor Superfamily Member 14	2.57741037	9.50E-10	WNT downstream
61	Trpm2	Transient Receptor Potential Cation Channel Subfamily M Member 2	2.571489747	1.65E-05	WNT downstream
62	Vmn1r81	vomeronasal 1 receptor 81	2.568063643	0.000355796	

Table 18. Top 62 significantly upregulated DEGs in the E6.25 mutant AVE
FC > 2, p-adj < 0.001; excluding non-coding genes, pseudogenes, Rik genes; ranked by descending FC. Relation to WNT pathway and oncogenes are indicated where applicable.

E6.25 mutant AVE **upregulated** candidates (top 63-124)

Rank	Gene	Name	Fold Change	p-adj (FDR)	WNT pathway? Oncogene?
63	Ppp3cb	Protein Phosphatase 3 Catalytic Subunit Beta	2.552692449	0.00032167	WNT upstream
64	Malat1	Metastasis Associated Lung Adenocarcinoma Transcript 1, non-protein coding	2.547666923	5.29E-11	WNT upstream
65	Bdp1	B Double Prime 1, Subunit Of RNA Polymerase III Transcription	2.53062631	9.11E-07	
66	Msmo1	Methylsterol Monooxygenase 1	2.527120451	2.57E-05	
67	Gigyf1	GRB10 Interacting GYF Protein 1	2.523979215	0.000391732	oncogene
68	Mirg	miRNA containing gene	2.522643953	3.91E-05	WNT upstream
69	Atad2b	ATPase Family, AAA Domain Containing 2B	2.513991554	0.000463389	oncogene
70	Scd2	Acyl-CoA desaturase 2	2.509894558	1.41E-09	WNT upstream
71	Ccdc85c	Coiled-Coil Domain Containing 85C	2.508210387	1.31E-05	WNT upstream
72	Spry2	Sprouty RTK Signaling Antagonist 2	2.505283793	0.000249285	WNT downstream
73	Lipe	Lipase E, Hormone Sensitive Type	2.493835674	5.56E-08	WNT downstream
74	Adam10	ADAM Metallopeptidase Domain 10	2.473592451	0.000299305	WNT downstream
75	Ckap4	Cytoskeleton Associated Protein 4	2.472535859	1.13E-05	WNT upstream
76	Bri3	Brain Protein I3	2.461906852	8.24E-08	WNT downstream
77	Noct	Nocturnin	2.45995845	1.95E-06	oncogene
78	Rock2	Rho Associated Coiled-Coil Containing Protein Kinase 2	2.449381263	3.05E-05	oncogene
79	Hnrnpa0	Heterogeneous Nuclear Ribonucleoprotein A0	2.445408799	5.30E-08	oncogene
80	Pdcd4	Programmed Cell Death 4	2.420990788	0.000127637	WNT downstream
81	Dazap1	DAZ Associated Protein 1	2.414445035	1.39E-05	oncogene
82	Tead1	TEA Domain Transcription Factor 1	2.407556676	0.000134627	WNT upstream
83	Bmpr2	Bone Morphogenetic Protein Receptor Type 2	2.405842129	0.000559031	WNT upstream
84	Dyrk1a	Dual Specificity Tyrosine Phosphorylation Regulated Kinase 1A	2.387856238	0.000573027	WNT upstream
85	Col23a1	Collagen Type XXIII Alpha 1 Chain	2.35941513	0.000920421	oncogene
86	Eva1c	Eva-1 Homolog C	2.346149647	2.59E-06	WNT downstream
87	Acnat1	acyl-coenzyme A amino acid N-acyltransferase 1	2.33531924	9.29E-05	
88	Cnot6	CCR4-NOT Transcription Complex Subunit 6	2.333087896	0.000147142	
89	Zscan26	Zinc Finger And SCAN Domain Containing 26	2.330831313	7.93E-05	oncogene
90	Ppp1r14b	Protein Phosphatase 1 Regulatory Inhibitor Subunit 14B	2.328458777	9.79E-06	WNT upstream
91	Ppip5k1	Diphosphoinositol Pentakisphosphate Kinase 1	2.324157682	4.73E-05	oncogene
92	Dot1l	DOT1 Like Histone Lysine Methyltransferase	2.316516304	0.000155923	WNT upstream
93	Sag	S-Antigen Visual Arrestin	2.310785479	8.94E-05	WNT upstream
94	Ireb2	Iron Responsive Element Binding Protein 2	2.291527041	0.000234232	
95	Klf9	Kruppel Like Factor 9	2.281892259	8.40E-05	WNT upstream
96	Enah	ENAH, Actin Regulator	2.281096493	7.41E-06	oncogene
97	Cbfb	Core-Binding Factor Beta Subunit	2.239333272	0.000176385	oncogene
98	Suz12	SUZ12 Polycomb Repressive Complex 2 Subunit	2.237990719	2.92E-06	WNT upstream
99	Ubald2	UBA Like Domain Containing 2	2.234194355	1.36E-06	
100	Irak1	Interleukin 1 Receptor Associated Kinase 1	2.233993719	0.000309122	
101	Fgd4	FYVE, RhoGEF And PH Domain Containing 4	2.214741961	0.000343448	WNT downstream
102	Prr15	Proline Rich 15	2.212552305	0.000746307	
103	Pkn2	Protein Kinase N2	2.205561522	2.62E-05	WNT upstream
104	Selenot	Selenoprotein T	2.203367128	0.000178652	oncogene
105	Csmd2	CUB And Sushi Multiple Domains 2	2.202282553	0.000662524	
106	Arid4b	AT-Rich Interaction Domain 4B	2.194806686	0.000309122	oncogene
107	Rtn4	Reticulon 4	2.183808804	2.72E-05	oncogene
108	Mtch1	Mitochondrial Carrier 1	2.180333594	0.000262063	
109	Rian	Maternally Expressed 8, Small Nucleolar RNA Host Gene	2.179816575	2.46E-07	oncogene
110	Lamc1	Laminin Subunit Gamma 1	2.172810916	0.000193933	WNT downstream
111	Igf2r	Insulin Like Growth Factor 2 Receptor	2.168913528	0.000193933	WNT upstream
112	Mllt10	Histone Lysine Methyltransferase DOT1L Cofactor	2.165664	0.000421654	WNT upstream
113	Olfr46	Olfactory receptor D12	2.163033132	0.000398845	
114	Brwd3	Bromodomain And WD Repeat Domain Containing 3	2.162519053	6.18E-05	WNT downstream
115	Akirin2		2.140502108	0.0009837	oncogene
116	Tpd52	Tumor Protein D52	2.10712517	3.14E-05	WNT downstream
117	Ythdc1	YTH Domain Containing 1	2.089464673	3.05E-05	
118	Uhrf2	Ubiquitin Like With PHD And Ring Finger Domains 2	2.076769183	0.00072325	WNT upstream
119	Acin1	Apoptotic Chromatin Condensation Inducer 1	2.043207488	9.11E-07	WNT downstream
120	Olfr1271	olfactory receptor 1271	2.027582073	0.00087744	
121	Otud4	OTU Deubiquitinase 4	2.019765527	7.64E-07	WNT upstream
122	Nova1	NOVA Alternative Splicing Regulator 1	2.017603285	0.000355796	oncogene
123	Nktr	Natural Killer Cell Triggering Receptor	2.008013382	7.49E-05	
124	Pbrm1	Polybromo 1	2.006721381	0.000137439	

Table 19. Top 63-124 significantly upregulated DEGs in the E6.25 mutant AVE
FC > 2, p-adj < 0.001; excluding non-coding genes, pseudogenes, Rik genes; ranked by descending FC. Relation to WNT pathway and oncogenes are indicated where applicable.

E6.25 mutant AVE – genes excluded (1-49)

Rank	Gene	Fold Change	p-adj (FDR)
1	Gm42418	19.99329629	2.09E-34
2	mt-Rnr1	4.078625753	1.37E-26
3	BX936338.1	17.40127486	1.67E-25
4	Gm12896	17.40127486	1.67E-25
5	Gm12895	17.40127486	1.67E-25
6	mt-Nd2	3.347887035	7.20E-24
7	Gm23935	8.370691703	6.09E-20
8	Gm24245	8.370691703	6.09E-20
9	Gm24270	8.370691703	6.09E-20
10	Lars2	5.117471497	2.39E-19
11	Gm45665	22.57631696	4.80E-16
12	mt-Nd6	3.299086561	7.30E-16
13	Gm47164	3.188415521	1.62E-15
14	mt-Ts1	2.573209144	1.98E-15
15	Mir6236	5.138653058	4.94E-14
16	A430010J10Rik	3.56471109	1.34E-12
17	Gm42535	2.182638063	3.45E-10
18	Gm17132	2.491524608	8.25E-09
19	Gm32219	11.45668101	1.08E-08
20	Gm26917	2.325796884	2.61E-08
21	Gm38246	7.522872778	5.68E-08
22	mt-Rnr2	1.752919666	6.82E-08
23	Gm15564	4.861857016	7.27E-08
24	AC122774.1	6.796658125	7.62E-08
25	Gm7324	5.313193693	8.41E-08
26	Gm43652	2.652102623	8.41E-08
27	AC131033.1	13.18988345	1.71E-07
28	Gm42566	4.101235131	1.73E-07
29	AC152063.6	2.179816575	2.46E-07
30	Gm37485	3.047826916	3.15E-07
31	Gm42519	3.250030501	4.41E-07
32	Gm4034	2.288111773	4.61E-07
33	Gm42047	4.420128667	5.41E-07
34	Gm36955	6.391652187	5.49E-07
35	Gm11225	2.269534001	7.96E-07
36	Gm10093	20.04587581	1.04E-06
37	Gm48789	2.699009585	1.43E-06
38	AC158986.1	2.445325931	1.64E-06
39	Gm31182	5.886970963	1.67E-06
40	Gm48119	3.777859475	1.67E-06
41	Gm43461	3.515740521	2.52E-06
42	Gm42826	3.991287051	2.80E-06
43	Gm37954	3.646390264	2.91E-06
44	Gm26718	26.45847337	4.24E-06
45	Gm15066	2.472195425	6.64E-06
46	Gm29666	2.345882	7.82E-06
47	2310022B05Rik	4.581272863	7.95E-06
48	Gm20721	15.59907792	9.61E-06
49	Gm44487	2.545875522	1.14E-05

E6.25 mutant AVE – genes excluded (50-98)

Rank	Gene	Fold Change	p-adj (FDR)
50	Gm26909	2.861391911	1.23E-05
51	Gm47598	7.106050037	1.31E-05
52	Gm44228	2.008439915	1.51E-05
53	Rps12l1	1.706587855	1.54E-05
54	Gm48600	2.63007481	1.61E-05
55	Gm43775	2.597381387	1.61E-05
56	Gm37607	2.539801832	1.61E-05
57	Gm15832	11.53974964	2.58E-05
58	Gm47551	2.099282506	2.77E-05
59	Gm13075	2.542343518	3.05E-05
60	mt-Nd3	1.974900019	4.14E-05
61	1810037I17Rik	1.62478044	4.79E-05
62	Gm43980	2.168422496	4.83E-05
63	Gm30948	4.145444018	4.87E-05
64	Gm26905	1.876142032	6.03E-05
65	Gm10076	1.539145358	6.21E-05
66	4930546K05Rik	2.412633475	6.80E-05
67	mt-Nd4	1.347193535	7.54E-05
68	Gm38099	3.241316517	7.71E-05
69	Gm47542	1.960344329	9.16E-05
70	mt-Tq	2.554387989	0.000103047
71	Gm17275	15.04486276	0.000111331
72	Gm10275	1.724552619	0.000122443
73	Gm11373	4.005165494	0.000136286
74	Gm37242	3.508910575	0.000140512
75	Gm42577	3.004087137	0.000147142
76	Gm17494	2.067349418	0.000195856
77	Gm45792	3.741153853	0.000213172
78	Gm12940	2.173646339	0.00022013
79	Gm37750	2.063695336	0.000257381
80	2700016F22Rik	16.91745793	0.000282254
81	Gm43336	6.715316929	0.000316523
82	Gm29455	7.021117808	0.000318281
83	Gm11131	2.404045785	0.000330544
84	Gm37357	2.935881984	0.000351769
85	Gm44639	2.233691102	0.000354886
86	Gm47652	2.628870806	0.000356008
87	Gm14324	2.297187872	0.000379391
88	Gm21056	3.356228924	0.000385205
89	Gm28791	4.902974745	0.000399013
90	AC165106.3	2.82149408	0.000485797
91	Gm44044	2.252361721	0.000535593
92	Gm2961	2.411290626	0.000556945
93	2810425M01Rik	6.728180184	0.000718582
94	1810013D15Rik	3.585440743	0.00074761
95	Gm9754	24.17125741	0.000787136
96	D130062J10Rik	2.529666422	0.000817122
97	Gm47626	3.161424751	0.0008254
98	mt-Tc	2.294178929	0.000909572

Table 20. Top significantly upregulated DEGs in the E6.25 mutant AVE excluded from Table 18, 19
FC > 2, p-adj < 0.001; excluded, as they are: non-coding genes, pseudogenes, Rik genes; ranked by descending FC.

We next represented these top WNT-annotated candidates in the E6.25 mutant AVE in volcano plot and boxplot format to aid visualisation of candidates with high fold change and low p-adj value. The boxplot showing log10 normalised expression on the y-axis helped

identify genes with low expression levels in control. These graphs were segregated into ‘WNT downstream’ (Figure 93), ‘WNT upstream’ (Figure 94), and ‘oncogene’ (Figure 95).

Overall, assessment of top DEGs in the mutant AVE at E5.5 and E6.25 supported WNT enrichment found from GO analysis. I next examined some of these top candidates and the WNT/β-CATENIN pathway *in vivo*.

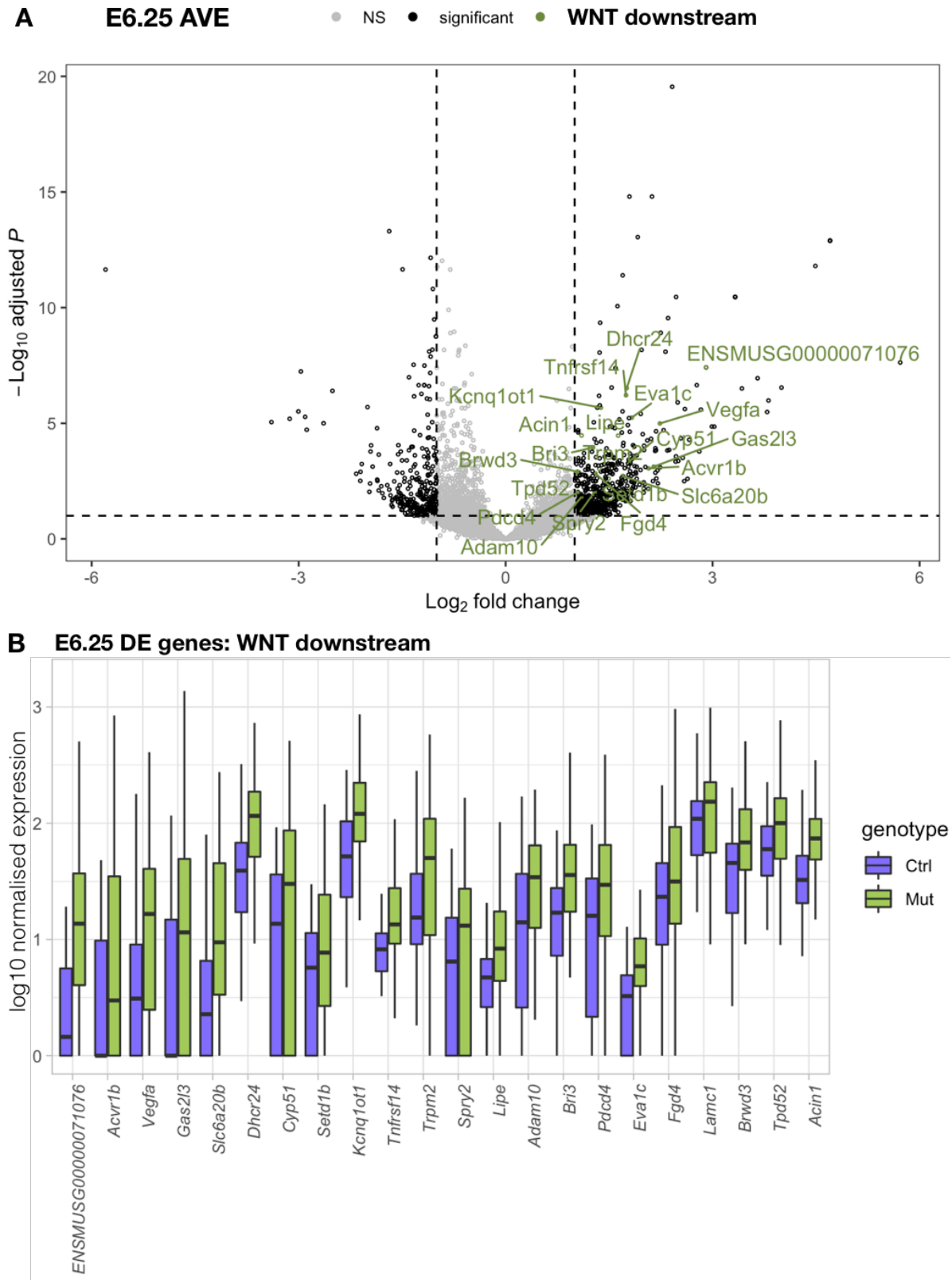


Figure 93. E6.25 mutant AVE upregulated DEGs that are WNT downstream components
(A) Volcano plot (B) boxplot representation

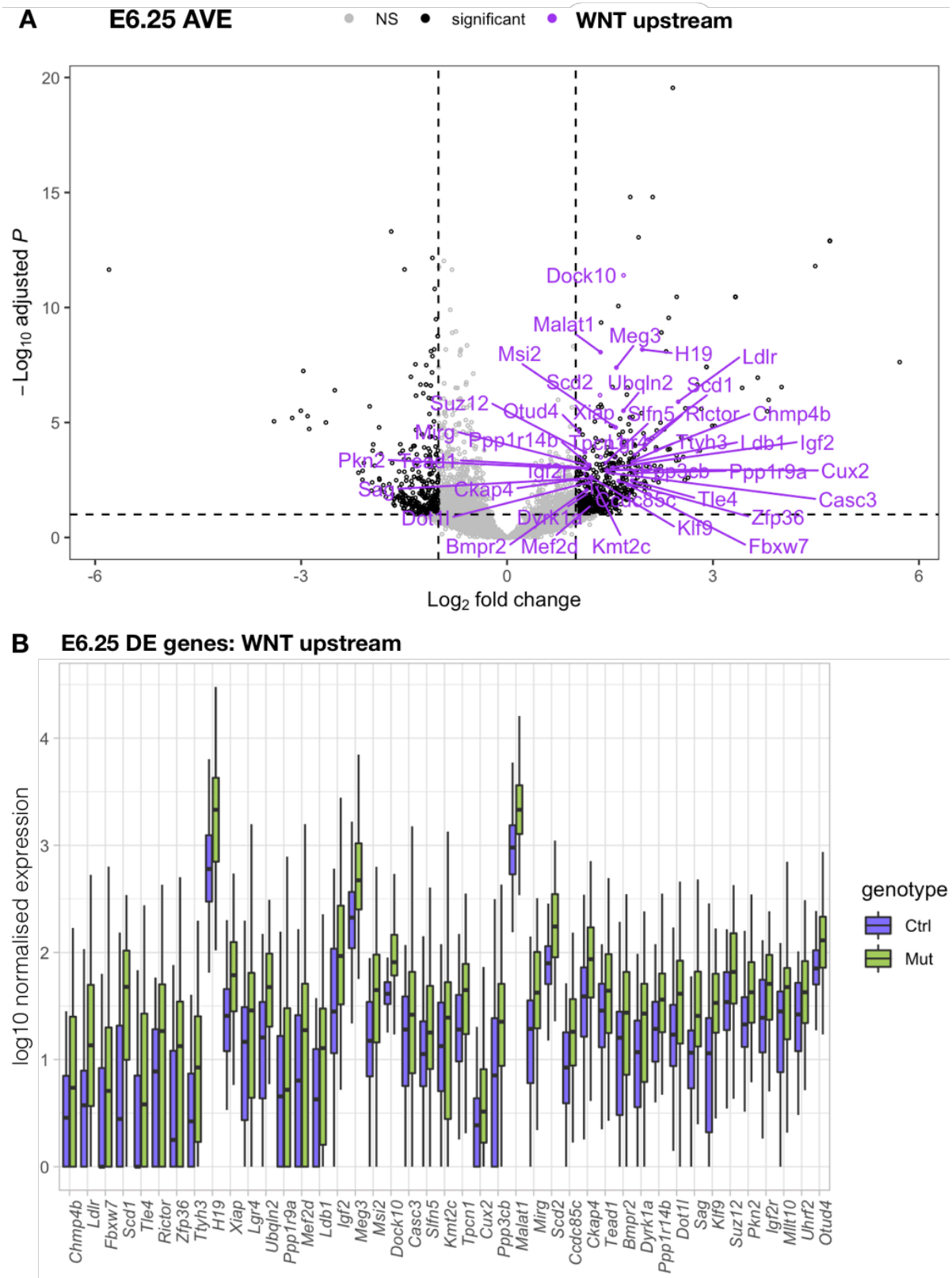


Figure 94. E6.25 mutant AVE upregulated DEGs that are WNT upstream components
(A) Volcano plot (B) boxplot representation

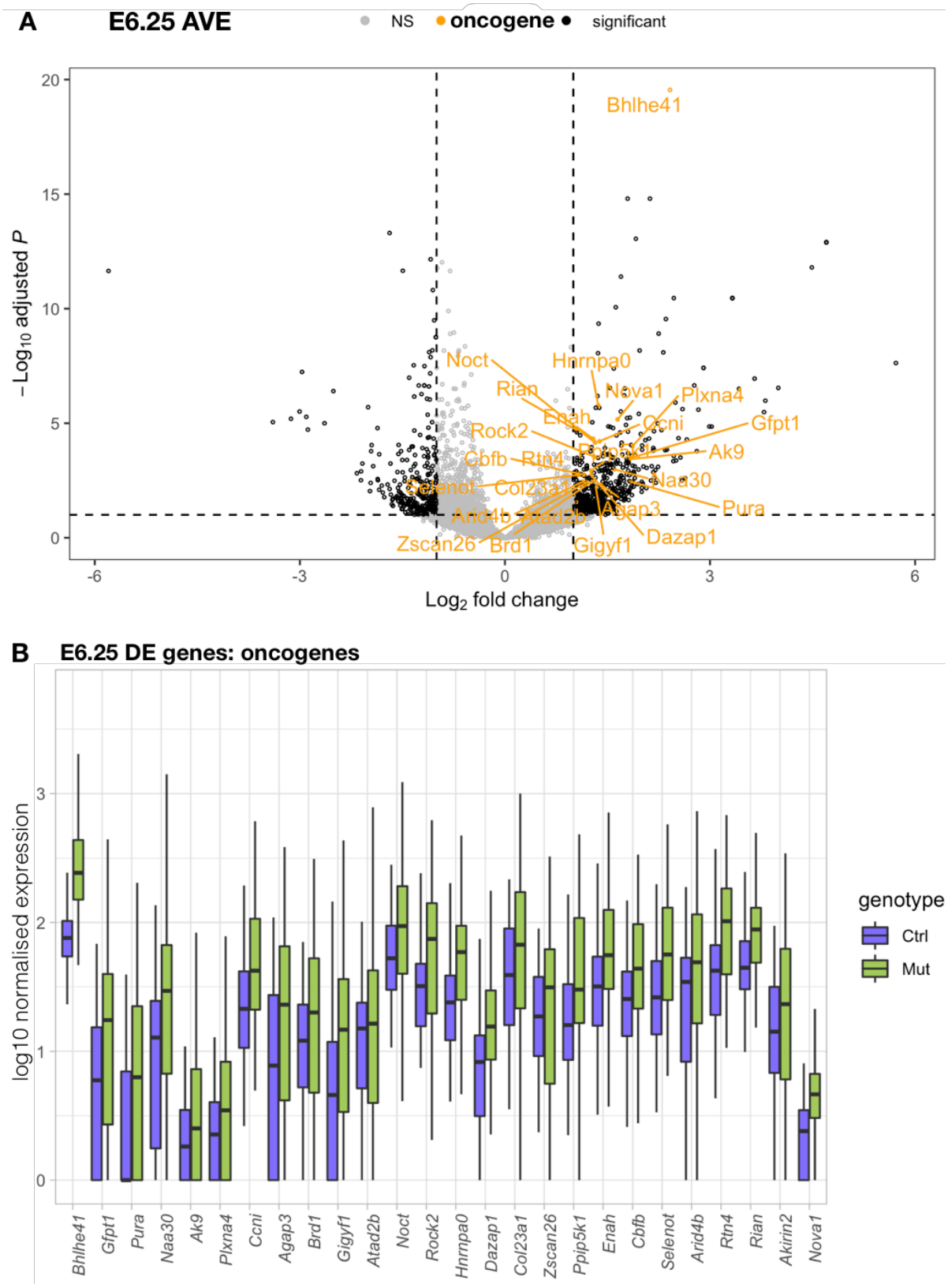


Figure 95. E6.25 mutant AVE upregulated DEGs that are oncogenes
(A) Volcano plot (B) boxplot representation

5.2.8 Validation of scRNA-seq WNT pathway enrichment by antibody stain and pharmacological intervention

To validate the results from scRNA-seq bioinformatics and GO analyses *in vivo*, I employed both antibody stain and pharmacological intervention. Antibody stain is limited to the availability of a suitable and specific antibody. I first tested MOESIN (top upregulated candidate in the mutant AVE at E5.5), and found expanded nuclear expression pattern of MOESIN in the mutant by E6.0 (Figure 96). I selected E6.0 for mutant versus control MOESIN antibody stain to allow time for transcripts to manifest as protein expression. MOESIN has been shown to act upstream of WNT/β-catenin signalling to trigger its pathway activation and induce proliferation (Zhu et al., 2012). MOESIN-induced proliferation has been demonstrated to be associated with its nuclear localisation (Batchelor et al., 2004). Furthermore, *Moesin* promoter analysis using Mulan Dcode software (Ovcharenko et al., 2005) revealed presence of 5 multi-conserved (amongst mouse and human) ETS2 binding sites (Figure 97), raising the possibility of *Msn* being a direct downstream target of ETS2 transcriptional activity.

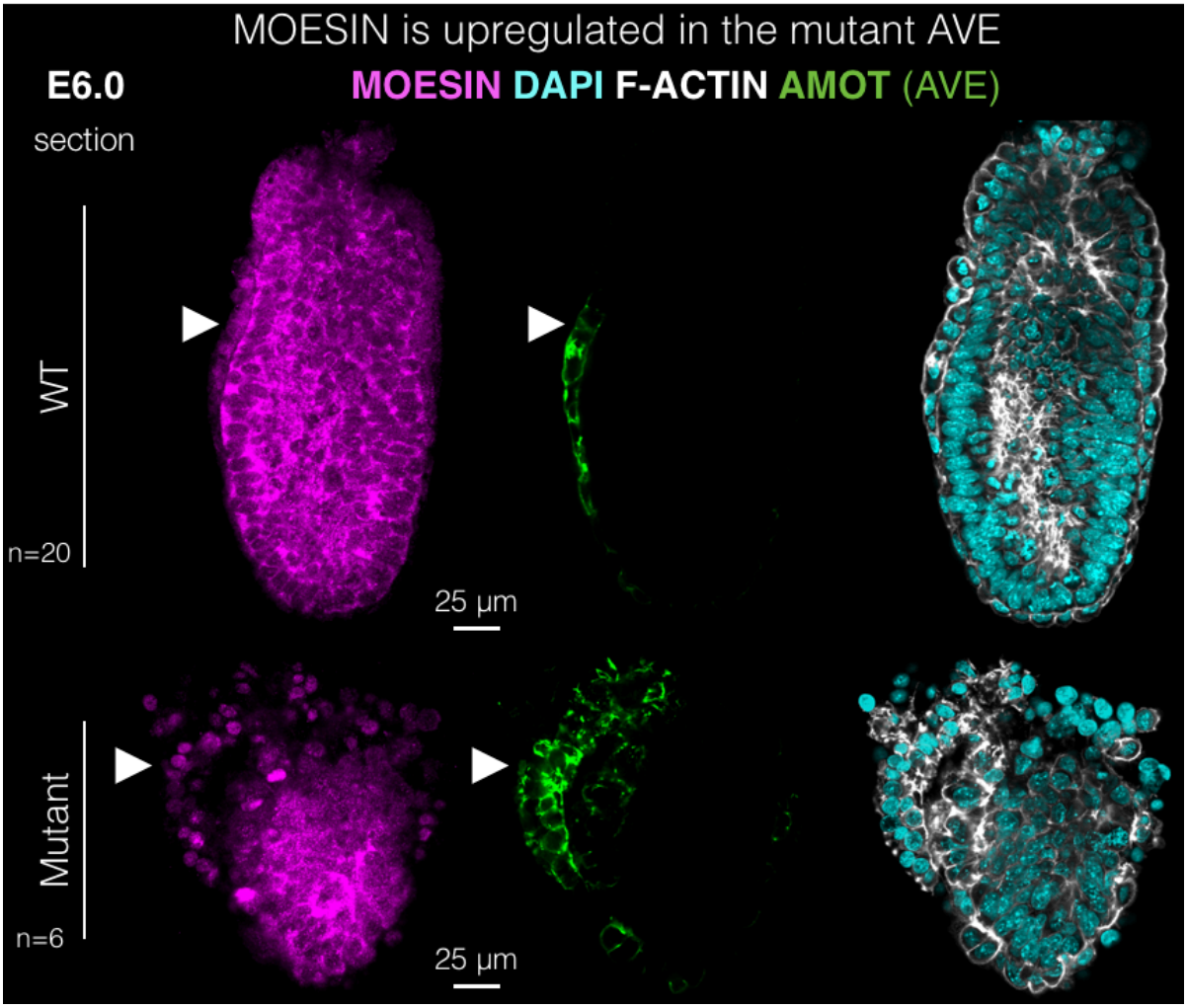


Figure 96. Validation of MOESIN upregulation in the mutant AVE at E6.0
Validation of E5.5 scRNA-seq top upregulated candidate in the mutant AVE. Because the upregulation in RNA expression was found at E5.5, the antibody stain to validate the scRNA-seq result was conducted at a slightly later stage to increase the chances of capturing upregulated protein expression.
Representative mid-sagittal sections of whole mount immunofluorescence of WT and mutant embryos at E6.0 stained for MSN (magenta), AMOT (green – AVE; arrowhead), F-ACTIN (gray); DAPI (cyan). Presence of expanded nuclear expression pattern of MSN in the mutant AVE compared to WT. Scale bars: 25 µm.

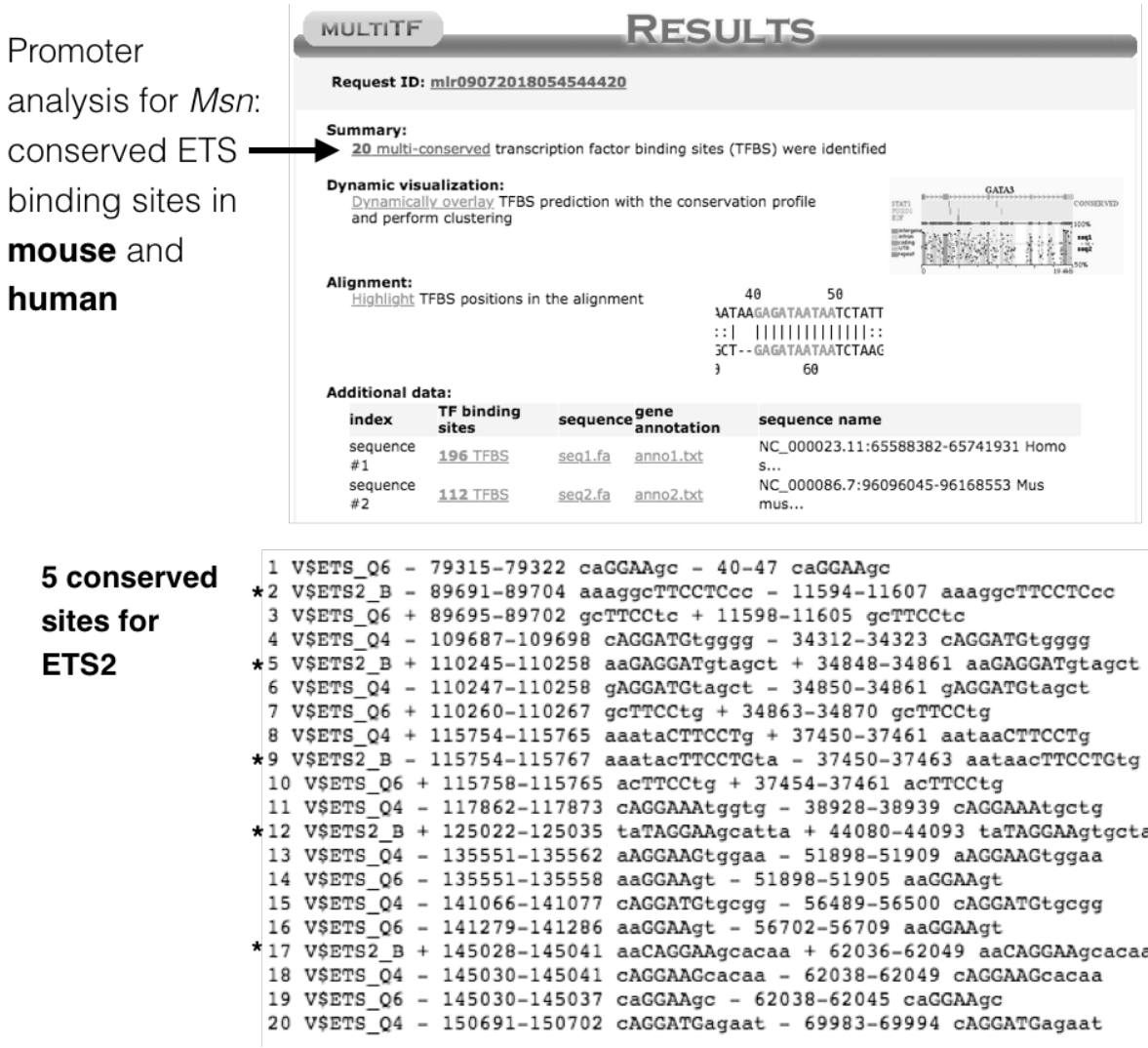


Figure 97. *Msn* promoter analysis to search for conserved ETS binding sites
Promoter analysis for conserved ETS binding sites amongst mouse and humans (using Mulan Dcode) revealed 20 multi-conserved sites for the ETS family of transcription factors, of which 5 are specific to ETS2 (noted with asterisk). This suggests that it is plausible for *Msn* to be a direct downstream target of ETS2.

Pharmacological activation of the WNT pathway in WT

Another approach to validate the upregulation of the WNT pathway *in vivo* is by pharmacological intervention. I first attempted to recapitulate the mutant AVE phenotype in WT embryos, as there are relatively more embryos available in WT compared to mutant, enabling thorough assessment on whether the inhibitor is having an effect. Similar to the MEK inhibitor study (4.2.10), I cultured WT embryos with HHEX-GFP fluorescent reporter of the AVE from E5.5 to E6.5 (time window of mutant AVE defect progression) with a WNT agonist. Coincidentally, the WNT agonist I used was CHIR (99021), a GSK3 inhibitor, which is the

other ‘i' of the ‘2i’ cocktail used to maintain ESCs (Ying et al., 2008). I found that inhibition of GSK3, with the equivalent CHIR concentration as that used in ESC culture (3 μM), in WT embryos from E5.5 to E6.5 was able to recapitulate the thickened mutant AVE phenotype (Figure 98). At this concentration, ETS2 expression was comparable in treated versus DMSO control. In addition, I was able to titrate the AVE thickening defect, whereby a higher concentration of CHIR (5 μM) led to a more severe AVE thickening (Figure 99). Co-staining with active β-CATENIN in 5 μM treated condition demonstrated its upregulation compared to control, verifying effectiveness of CHIR treatment. Upon comparison of ETS2 expression in CHIR-treated versus MEKi-treated, I found absence of nuclear ETS2 in the AVE of the MEKi condition, but presence in the CHIR condition (Figure 100). This observation could be in line with the model where MAPK/MEK activation is upstream of ETS2, which in turn activates WNT/β-CATENIN signalling to trigger AVE proliferation. While blocking MEK signals for ETS2 nuclear translocation is expected to abrogate nuclear ETS2 expression in the AVE, WNT/β-CATENIN pathway activation should not affect ETS2 nuclear localisation directly, which had been observed (Figure 100). Therefore, MEKi and CHIR-treatment of WT from E5.5 to E6.5 both recapitulate the mutant AVE phenotype with expected ETS2 expression patterns.

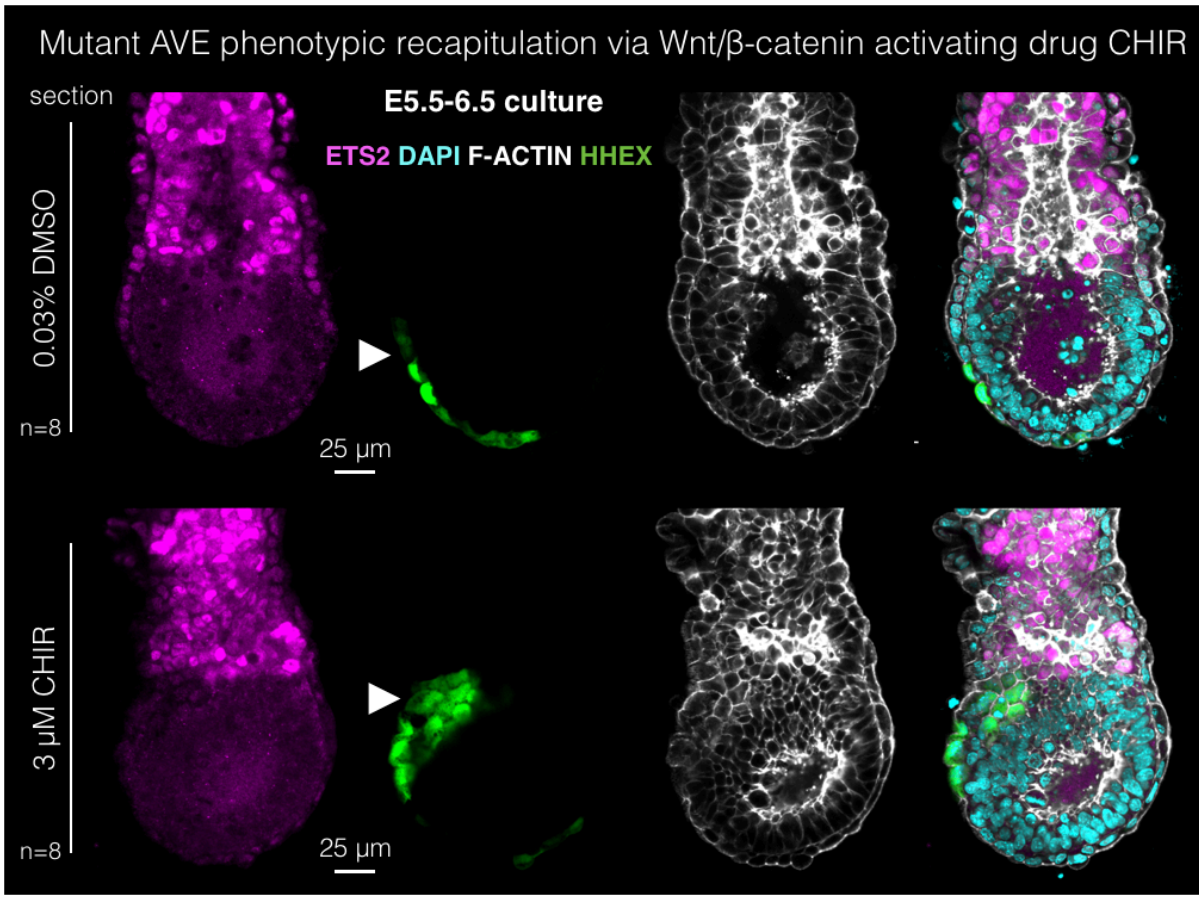


Figure 98. Mutant AVE phenotypic recapitulation upon culturing WT embryos in WNT agonist CHIR (3 µM) from E5.5 to E6.5, stained for ETS2

Representative mid-sagittal sections of whole mount immunofluorescence of DMSO control (0.03%) versus CHIR treated (3 µM) stained for ETS2 (magenta); HHEX-GFP (green – AVE; arrowhead); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

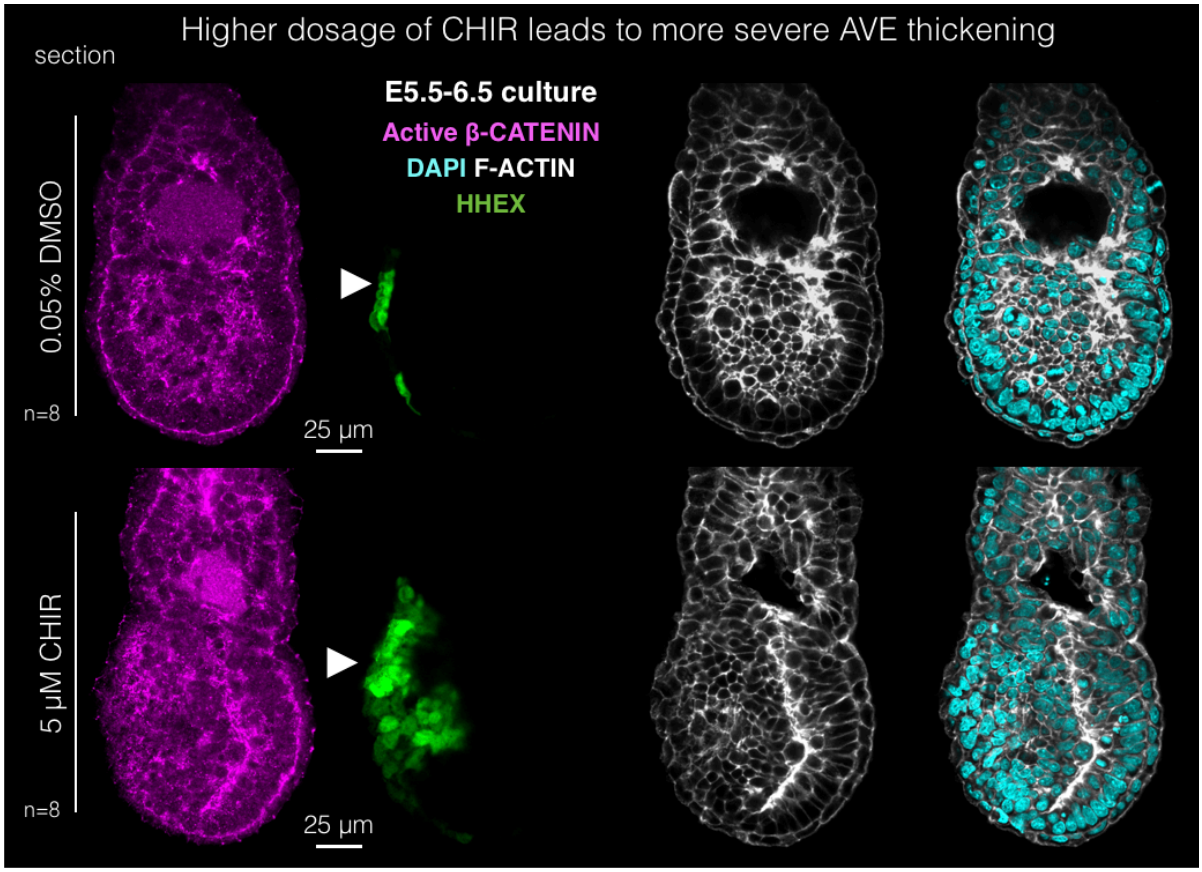


Figure 99. Mutant AVE phenotypic recapitulation with CHIR can be titrated whereby a higher dosage of CHIR (5 μM) led to a visibly more severe AVE thickening phenotype

Representative mid-sagittal sections of whole mount immunofluorescence of DMSO control (0.05%) versus CHIR treated 5 μM) stained for active β-CATENIN (magenta); HHEX-GFP (green – AVE; arrowhead); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 μm.

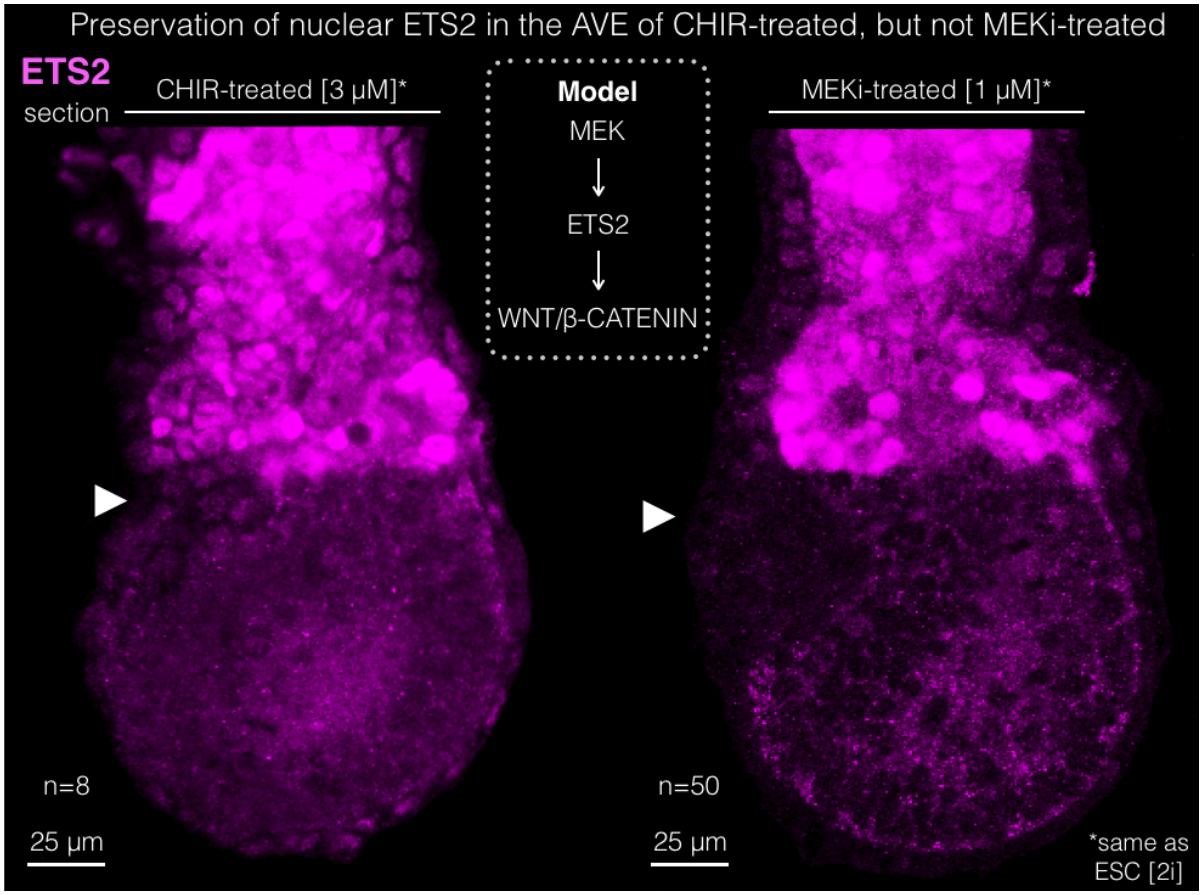


Figure 100. Comparison of ETS2 expression in the AVE of CHIR (3 µM) versus MEKi (1 µM) treated WT embryos from E5.5 to E6.5
Decreased nuclear ETS2 (magenta) in the AVE of MEKi-treated compared to CHIR-treated. Scale bars: 25 µm.

I next conducted immunofluorescent analysis to assess the degree to which CHIR culture recapitulated molecular features of the mutant AVE phenotype, similar to what I had done for MEKi analysis (4.2.10). I stained for GATA6 to check whether the thickened AVE from CHIR culture retains normal GATA6 expression, and found presence of nuclear GATA6 throughout the AVE/VE of CHIR-treated embryos (Figure 101). I also analysed cytoskeletal and polarity components by staining for F-ACTIN to visualise cell outline and more clearly see potential epithelial stratification; total β-CATENIN (lateral polarity component); and LAMININ (basement membrane marker). I found lateral polarity component total β-CATENIN to retain stereotypic localisation in the CHIR-treated thickened AVE (Figure 101). Although the WNT/β-CATENIN pathway is being perturbed, the stable junctional pool of β-CATENIN appears unaffected, which is expected since this pool is commonly not involved in

transcription (Clevers, 2006). Contrary to *Ets2*-mutants, the LAMININ-positive basement membrane beneath AVE cells did not become thickened in CHIR-treated WT embryos (Figure 102). In mutants, the basement membrane thickening is observed at E6.25-E6.5, which follows the burst of proliferation at E5.5. It is possible that the WNT pathway may be specifically influencing proliferation of the AVE at around E5.5, and that other cell biological phenotypes seen in the *Ets2*-mutant, such as thickened basement membrane and increased perpendicular cell divisions are secondary effects, occurring via other functions of ETS2 that are WNT-independent.

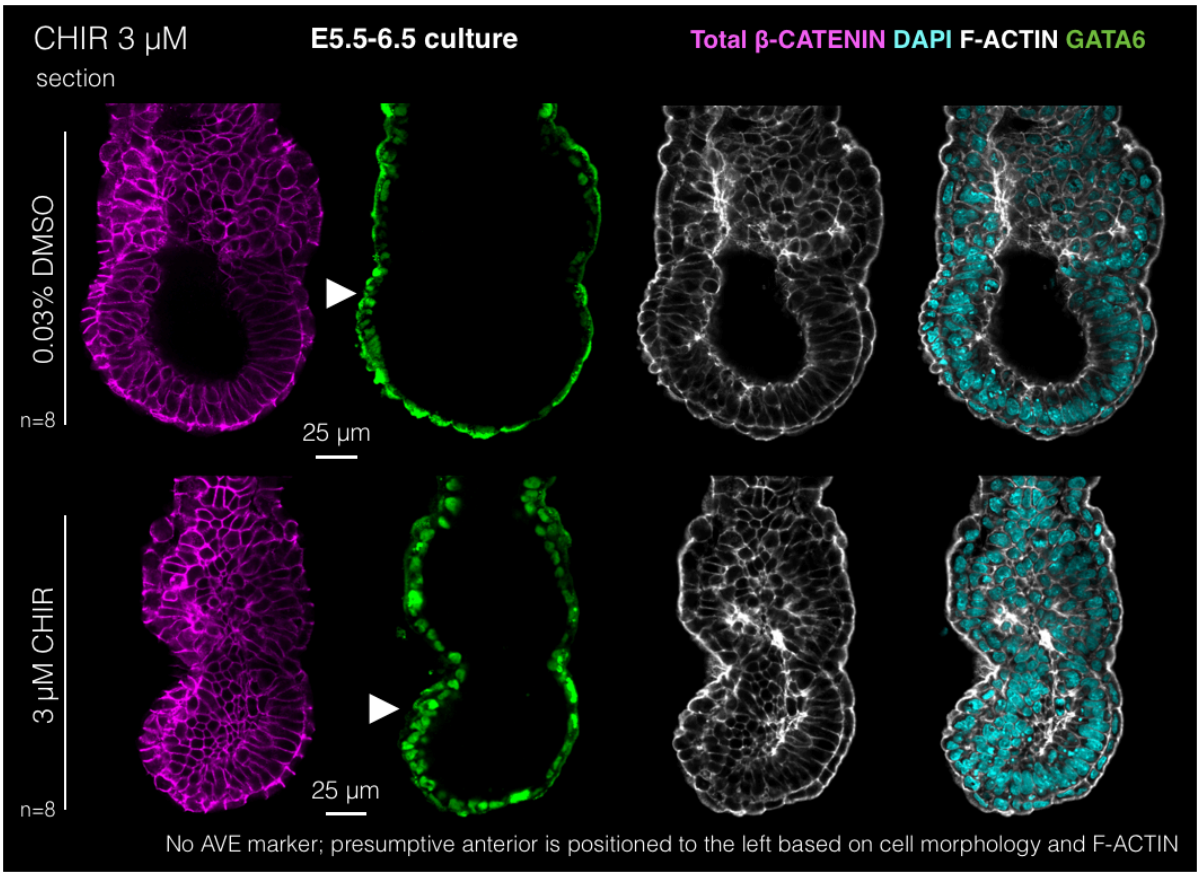


Figure 101. Mutant AVE phenotypic recapitulation upon culturing WT embryos in WNT agonist CHIR (3 µM) from E5.5 to E6.5, stained for total β-CATENIN and GATA6

Representative mid-sagittal sections of whole mount immunofluorescence of DMSO control (0.03%) versus CHIR treated (3 µM) stained for total β-CATENIN (magenta); GATA6 (green; AVE – arrowhead); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

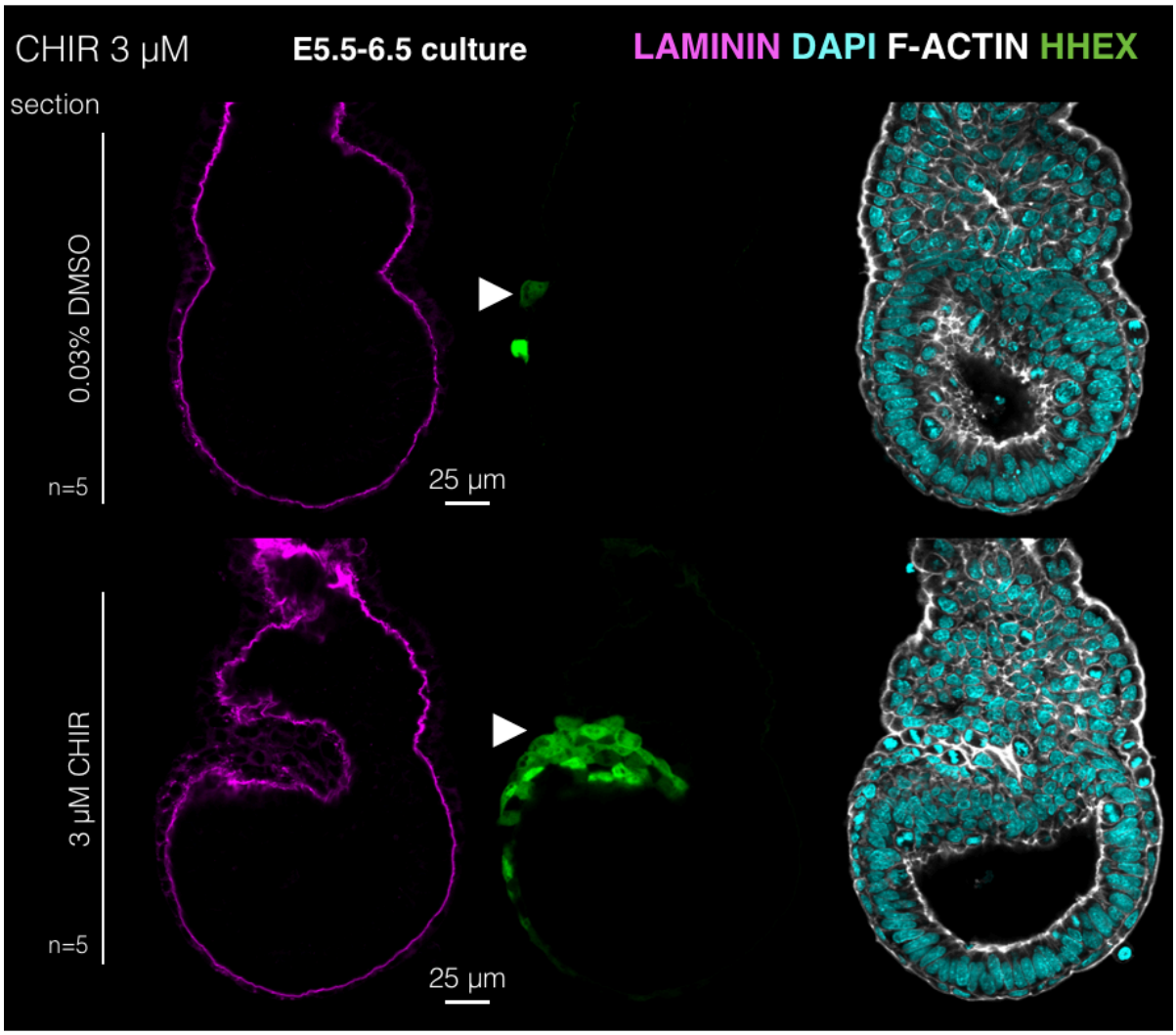


Figure 102. Mutant AVE phenotypic recapitulation upon culturing WT embryos in WNT agonist CHIR (3 µM) from E5.5 to E6.5, stained for LAMININ
Representative mid-sagittal sections of whole mount immunofluorescence of DMSO control (0.03%) versus CHIR treated (3 µM) stained for LAMININ (magenta); HHEX-GFP (green – AVE; arrowhead); F-ACTIN (gray); DAPI (cyan). Scale bars: 25 µm.

In summary, this section validated top E5.5 scRNA-seq upregulated candidate MOESIN by antibody stain in mutants. MOESIN is known to act upstream of WNT/β-CATENIN pathway activation and is agonistic of WNT-mediated proliferation. MOESIN may play a role in upregulating the canonical WNT pathway in the mutant AVE, which in turn may upregulate several WNT agonists seen by scRNA-seq at E6.25. Involvement of the WNT/β-CATENIN pathway in the mutant AVE phenotype was supported by phenotypic recapitulation in WT embryos cultured in WNT agonist CHIR from E5.5 to E6.5.

5.2.9 Validation of top misregulated genes from mutant AVE scRNA-seq by *in vivo* genetics using 2C-HR-CRISPR

Validation of E5.5 mutant AVE scRNA-seq top upregulated candidate MOESIN by antibody stain suggests WNT involvement in regulation of AVE cell number *in vivo*, as MOESIN can promote WNT-mediated proliferation (Zhu et al., 2012). Furthermore, the ability to recapitulate the mutant AVE phenotype in WT embryos with a WNT agonist supports an ETS2-WNT axis of signalling in the AVE to regulate its tissue size. However, inhibitor studies are not tissue specific. While the lowest concentrations that recapitulated the AVE phenotype were used (i.e. did not affect gross embryo morphology and ETS2 expression in the ExE), the possibility of contribution from non-AVE/VE tissues cannot be excluded. To demonstrate the role of the WNT pathway *in vivo* in a tissue-specific fashion, I attempted 2C-HR-CRISPR to misexpress top misregulated WNT agonists in the VE. A VE-specific promoter was selected, as ETS2 is expressed throughout the VE, but only the AVE cells appear to be sensitive to the absence of ETS2. This suggests that ETS2 transcriptional co-factors differ in the AVE versus the surrounding VE. Furthermore, only AVE cells were responsive to MEK inhibitor and CHIR treatments while the rest of the VE, which was similarly exposed to the pharmacological agents, was unperturbed. In addition, the AVE is a heterogeneous population making it difficult to select a promoter that will target the entire AVE population. The VE-specific promoter selected was *Ttr*, as it becomes activated at peri-implantation stage E5.0 (shown by Dr. Matthew Stower; Srinivas group, unpublished data). Coincidentally, E5.0 is a suitable time to misexpress a WNT agonist in the VE, as the stage is just prior to the mutant proliferation defect at E5.5. In addition, homozygous null *Ttr*-mutants were found to be viable (Episkopou et al., 1993). Therefore, knocking a gene into the *Ttr* endogenous locus is unlikely to have an adverse effect on normal embryogenesis because of absent TTR.

2C-HR-CRISPR was selected as an efficient approach to knock-in a large construct by gene targeting. Specifically, I planned to knock-in: the entire coding region of a top misregulated WNT agonist, tagged with a fluorescent reporter and other elements to improve the transcriptional and translational efficiency of the targeted allele. The technique involves injecting CRISPR components at the 2-cell stage, as well as the streptavidin-biotin tagged Cas9-PCR system to increase the likelihood of CRISPR components to interact with each other *in situ* (Gu et al., 2018).

Construct design for candidate WNT agonist misexpression by 2C-HR-CRISPR

The design for the biotinylated PCR donor construct for *Msn* was identical to that utilised for other candidates tested/to be tested. Dr. Satish A. Jayaram of the Srinivas laboratory generated a *Ttr* targeting construct consisting of 5' and 3' *Ttr* homology arms (HA) that can be specifically knocked-in to the *Ttr* genomic locus, positioned directly after *Ttr* endogenous promoter (Figure 103). Preceding the insertion of the construct of interest, e.g. *Msn*, is *Td-Tomato* fluorescent reporter to label where the insertion construct is expressed, and 2A peptide to improve the efficiency of co-expression of multiple genes. 2A peptides are of viral original and 18-22 amino acids in length. They enable cleavage of polypeptides in eukaryotic translation via 'ribosome skipping', which results in relatively high levels of subsequent protein expression as compared to other methods for simultaneous expression of multiple genes (Trichas et al., 2008; Liu et al., 2017). 2A peptides have the added advantage of being small in size, which decreases the chances of functional disruption of co-expressed genes.

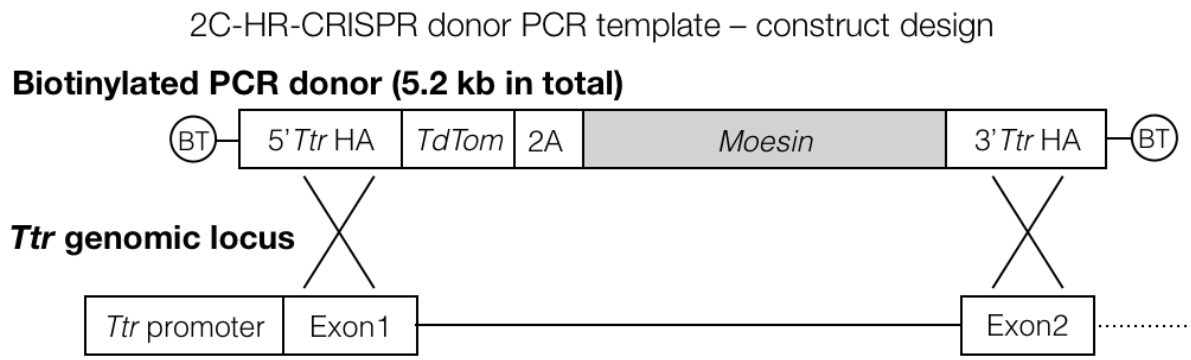


Figure 103. 2C-HR-CRISPR donor PCR template construct design with *Moesin* (5.2 kb total length)
Biotinylated (BT) donor DNA consisting of 5' and 3' *Ttr* homology arms (HA) for knock-in insertion downstream of the *Ttr* endogenous promoter, which would obliterate endogenous TTR function. *Td-Tomato* (*TdTom*) reporter and 2A peptide are added upstream of *Msn* to serve as a fluorescent tag and increase translation efficiency, respectively.

The gene of interest was obtained by creation of a cDNA library. I made this library using E9.0 embryos (Figure 104), which is a stage where top candidates from scRNA-seq are expressed at high levels, verified using scRNA-seq reference dataset from Pijuan-Sala et al., 2019. Following reverse transcription by polyA priming, as opposed to random priming, using MMLV reverse transcriptase, I was able to obtain full length cDNA libraries. I confirmed this by the ability to amplify entire cDNA's upon designing primers for genes of interest, e.g. *Msn, Xiap*.

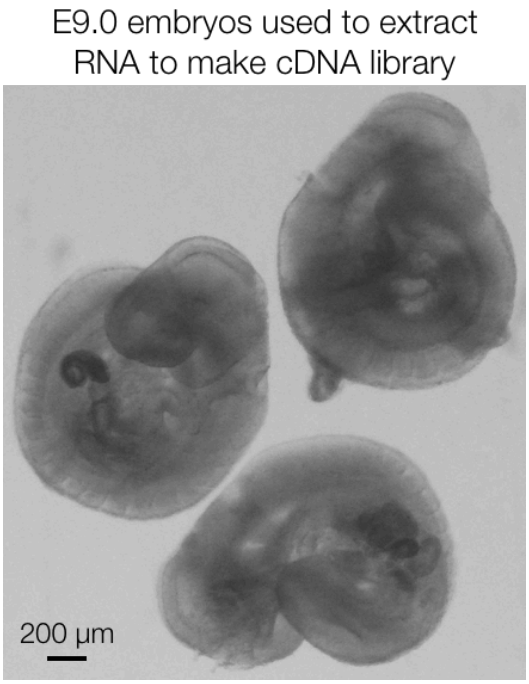


Figure 104. E9.0 embryos utilised to extract RNA to make cDNA library
3 E9.0 embryos were pooled together for RNA extraction to make cDNA. Brightfield image of embryos used is shown. Stage E9.0 was confirmed to express top candidates from mutant AVE scRNA-seq analysis confirmed with scRNA-seq reference dataset at that stage (Pijuan-Sala et al., 2019).

PCR for the gene of interest was conducted with HA's for the *Ttr* construct from Dr. Jayaram, such that the gene of interest could be cloned into the *Ttr* targeting construct. Following cloning of the gene of interest into the *Ttr* backbone, specifically downstream of *Td-Tomato* and 2A, a biotinylated PCR was conducted. The final product was a linearised *Ttr* donor backbone and inserted gene of interest with biotin on either ends, in accordance with Gu et al., 2018. The biotinylated PCR donor construct for injection was validated to be precise both by restriction enzyme digest for appropriate band lengths, and by Sanger sequencing. The final purified biotinylated donor PCR product was 20 µl at 20 ng/µl. Equal volume guide RNA was prepared and added by Dr. Jayaram. This then served as the solution for injection into 2-cell stage embryos followed by embryo transfer into pseudopregnant females, all conducted by Dr. Jonathan Godwin. Genotyping for positive integration was conducted in embryos dissected out at E6.5 using a forward primer that sits in the *Ttr* genomic locus and a reverse primer that recognises the integrated *Td-Tomato* (Figure 105).

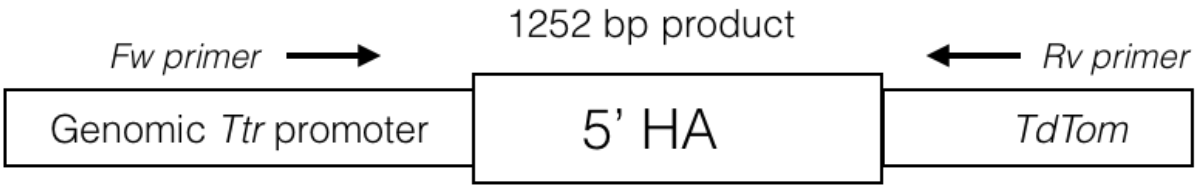


Figure 105. Genotyping for positive integration of 2C-HR-CRISPR targeting construct into the *Ttr* endogenous locus
Forward (Fw) primer sits in *Ttr* genomic promoter and reverse (Rv) primer sits in *Td-Tomato*, which is part of the targeting construct.

2C-HR-CRISPR attempt for *Msn*

Upon 2C-HR-CRISPR-mediated gene targeting of *Msn* into the *Ttr* locus and dissection at E6.5, an AVE thickening was observed in a fraction of the embryos. The AVE was visualised by co-staining with F-ACTIN and AMOT (Figure 106). Td-Tomato fluorescence, however, was not detected in any of the cases, potentially due to weak signal, or error in protein folding. Genotyping for positive integration was conducted by Dr. Satish A. Jayaram, so that it was blinded to phenotype. This, however, revealed that only 2/34 (5.88%) embryos recovered had positive integration by genotyping, which is low compared to the efficiencies attained in Gu et al., 2018. One embryo with genotyped positive integration is shown in Figure 106. The other had no overt AVE thickening phenotype. Of the 34 embryos, 13 had an abnormal appearance, i.e. thickening to varying degrees. Alternative PCR genotyping approaches could be considered to test whether the AVE thickening appearance I observed from 2C-HR-CRISPR *Msn* injections was random, or due to *Msn* misexpression in the VE.

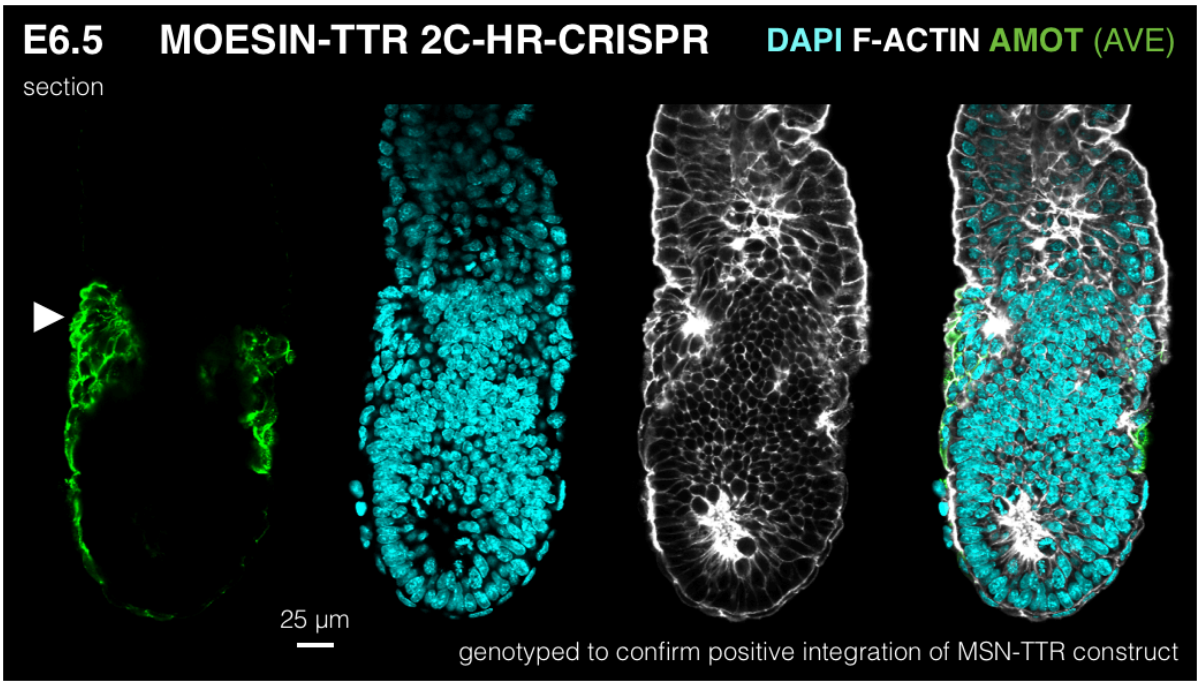


Figure 106. Potential mutant AVE phenotypic recapitulation in WT embryos upon 2C-HR-CRISPR-mediated knock-in of *Msn* driven by VE-specific promoter TTR, stained for AMOT
Thickened mutant AVE phenotype (arrow) may be recapitulated using 2C-HR-CRISPR to misexpress *Msn* in the VE. However, this experiment requires more numbers and more robust genotyping.
Representative mid-sagittal sections of whole mount immunofluorescence staining are shown marking AVE with AMOT (green); F-ACTIN (gray); DAPI (cyan). Scale bar: 25 µm.

2C-HR-CRISPR attempt for *Xiap*

E6.25 top upregulated candidate *Xiap*, nuclear agonist of WNT/β-CATENIN, was also assessed by 2C-HR-CRISPR. This candidate was, however, not validated by antibody stain in mutants. As *Xiap* is only upregulated at E6.25, when the mutant AVE phenotype is already present, its increased expression is less likely to be causal of the initiation of mutant AVE overproliferation, which occurs at E5.5. Nonetheless, because E6.25 is a stage during which the mutant AVE defect is progressing toward increased severity, *Xiap* was still considered for 2C-HR-CRISPR. Here I report no striking mutant AVE thickening phenotype upon *Xiap* misexpression when evaluating injected embryos at E6.75. However, a subtle defect could still be present (Figure 108), and could potentially be detected by quantification of AVE cell number and cell crowding. Since only 2/16 (12.5%) of *Xiap-Ttr* microinjected embryos had genotyped positive integration (Figure 107), increasing numbers could be helpful to attain

sufficient statistical strength to quantitatively determine whether *Xiap* misexpresison in the VE results in mutant AVE phenotypic recapitulation.

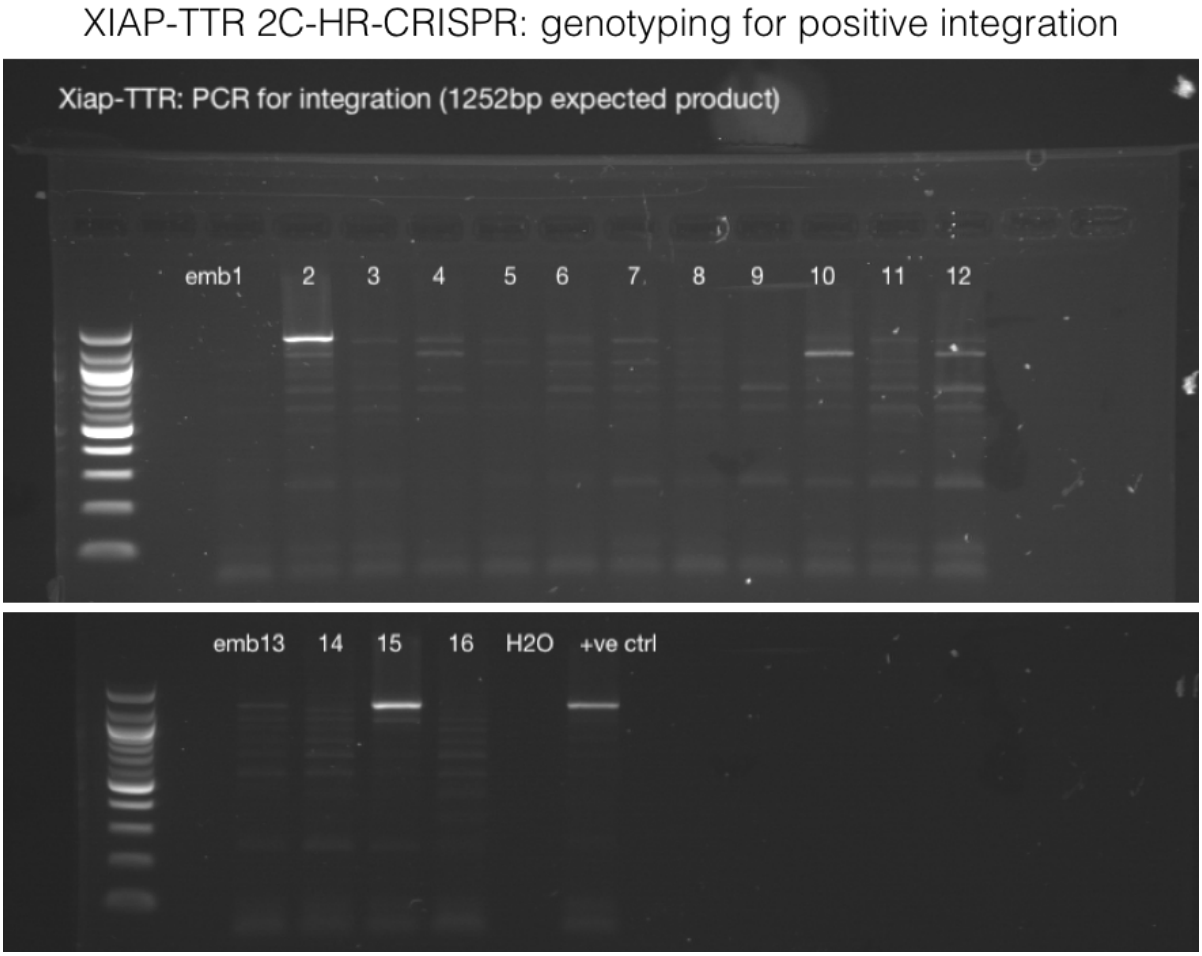


Figure 107. E6.5 embryo genotyping for positive integration of *Xiap-Ttr* from 2C-HR-CRISPR microinjections
Presence of clear positive integration for embryo 2 and 15 (1252 bp expected band, as seen in the positive control).

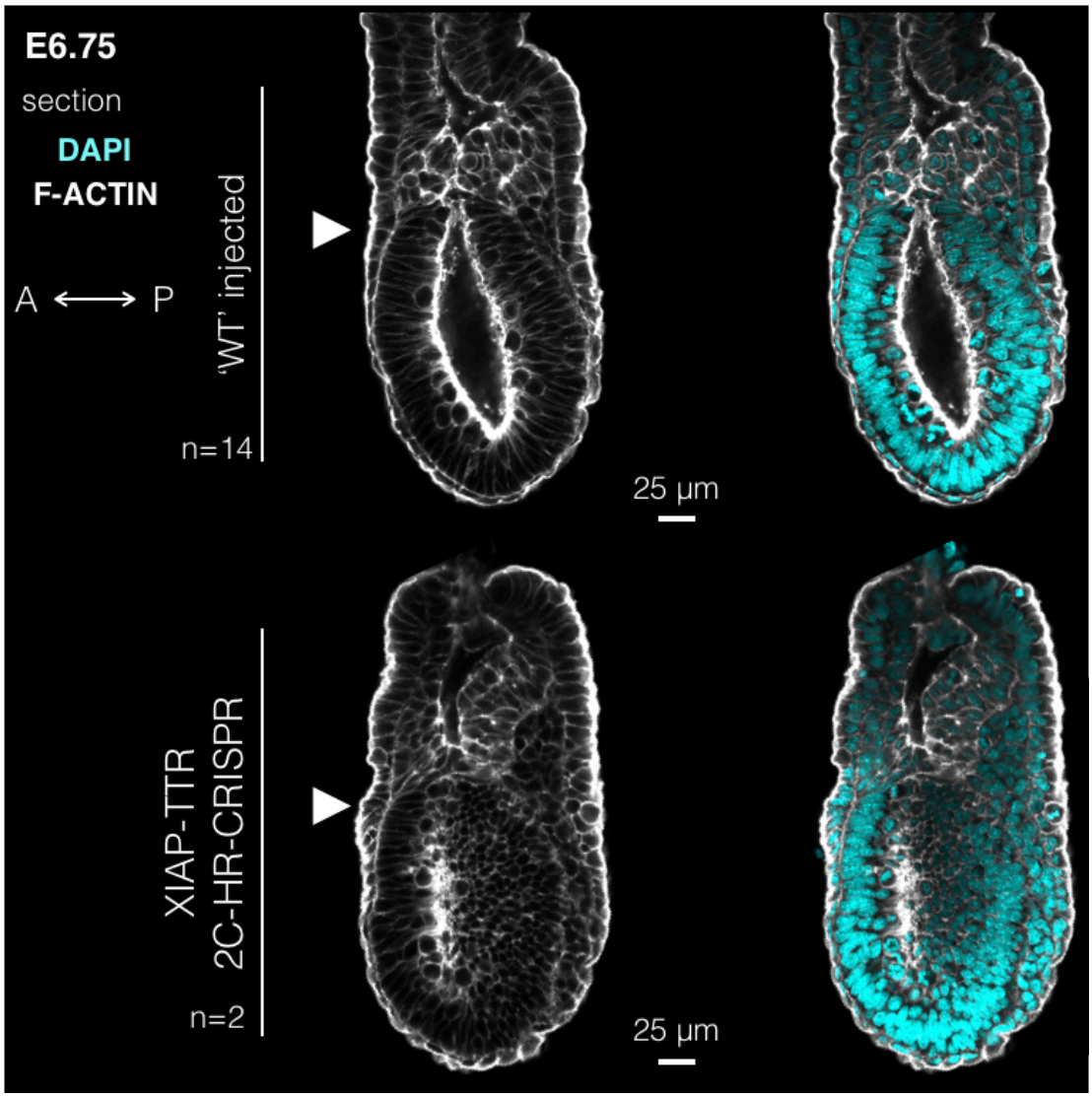


Figure 108. 2C-HR-CRISPR-mediated knock-in of *Xiap* into the *Ttr* endogenous locus
Representative mid-sagittal sections of whole mount staining of non-edited ('WT' injected) compared to *Ttr-Xiap* 2C-HR-CRISPR-edited embryos. The AVE here (arrowhead) was determined by cell morphology, as only F-ACTIN (gray) and DAPI (cyan) had been assessed. Again, no Td-Tomato signal was detected in all embryos. Scale bar: 25 µm.

Here, I have attempted 2C-HR-CRISPR, with technical assistance from Dr. Satish A. Jayaram and Dr. Jonathan Godwin, to misexpress top candidate WNT agonists in the VE of WT embryos to see if I could recapitulate the mutant AVE phenotype. However, results from both *Msn*- and *Xiap-Ttr* injections were inconclusive and require increased numbers to verify presence of mutant AVE phenotypic recapitulation concomitant with genotyped positive integration. New genotyping primers may be designed to amplify a region that spans the 3' end

of the gene of interest and a 3' region of *Ttr* genomic DNA, as opposed to the currently used 5' genotyping strategy.

5.3 Discussion

This Chapter probed the downstream effectors of ETS2 transcriptional activity by scRNA-seq. Single-cell technology enabled tissue-specific analysis of *Ets2*-mutant defects, which was essential since ETS2 appears to have different roles in the ExE and AVE. Furthermore, early embryonic origins of mutant defect origins meant that there would likely not be sufficient material for bulk-seq.

ScRNA-seq of mutant epiblast and ExE

In my first attempt to collect all cells from E5.5 embryos of the pure 129 background, I used enzymatic treatment without microdissection. This led to the loss of a large proportion of AVE and VE cells, potentially due to different cell populations having varied cell-cell adhesion strengths. Therefore, disaggregation of the entire embryo with a single approach was deemed to be suboptimal for recovering all cell types in equal proportions. With 6 minute TrypLE treatment at 37°C and mouth pipetting, the single cell recovery from individual E5.5 embryos was mostly epiblast. Another shortcoming of E5.5 pure 129 single cell collections was that control versus mutant embryos were phenotyped based on the degree of ExE present. Although the reduced mutant ExE phenotype is evident in the 129 background by E5.5, phenotyping is susceptible to human error. Mutant genotype could not be detected by scRNA-seq since the mutant allele retains the long 3' UTR of *Ets2* and Smart-seq2 is conducted from the 3' end (Picelli et al., 2014). Therefore, genotyping was added in the pipeline for future control versus mutant single-cell collections. Analysis of control versus mutant E5.5 epiblast revealed broadly normal epiblast transcriptome in mutants, shown by: 1. Absence of cluster

separation upon t-SNE and heatmap analysis, and 2. No significant differences from cluster stability analysis. Only a few genes were differentially expressed in the E5.5 mutant epiblast (6 upregulated and 6 downregulated), including downregulation of mesoderm-associated marker *Msx1*. Pluripotency factor *Nanog* and several members of the ID family of transcriptional repressors (*Id1/2/*3) were also downregulated. These findings could all relate to the mesoderm induction defect seen in the mutant epiblast at E6.5. However, the candidates have not been validated *in vivo* by antibody staining or ISH. Although the mutant epiblast is broadly unchanged in its transcriptome, a few DEGs present at E5.5 may play a significant role in later embryo patterning. The next step is to validate these DEGs.

A few ExE cells in both control and mutants were recovered from scRNA-seq, which enabled comparative analysis. We could confirm that mutant ExE cells at E5.5 retained signature marker gene expression upon evaluation of 17 independent ExE markers in mutant versus control. DE analysis had not been possible with the low numbers in our E5.5 ExE scRNA-seq dataset. However, we were able to adopt the Sinaplot approach (Sidiropoulos et al., 2018) to assess differentially regulated GO terms. We used this method to verify the absence of proliferation and apoptosis anomalies in the post-implantation mutant ExE (previously characterised by quantitative immunofluorescence in 3.2.6/3.2.7). We found no significant differences when evaluating GO terms for 'negative regulation of proliferation' and 'positive regulation of apoptosis' in the E5.5 mutant ExE scRNA-seq data, further suggesting that the post-implantation mutant ExE has normal growth and death rates. This lends additional support to the hypothesis that ETS2 is required during a narrow time window in ExE development (i.e. peri-implantation) to regulate its tissue size.

ScRNA-seq of mutant AVE

To purify for AVE cells in control versus mutant, the *Ets2-Hhex* mouse line was utilised to fluorescently label the AVE. The disadvantage here was that *Ets2-Hhex* had been maintained on a mixed genetic background, which differs from the pure 129 background used for collecting E5.5 control versus mutant epiblast and ExE. However, the AVE thickening phenotype appeared to be consistent amongst pure and mixed genetic backgrounds, as visualised using HHEX-GFP or AMOT, and quantified by assessing cell density, where both genetic backgrounds displayed excessively crowded E6.5 mutant AVE (4.2.9). A notable disadvantage is that the context-dependent transcriptional differences of ETS2 in the ExE versus AVE cannot be assessed by direct comparison of scRNA-seq data between these 2 different tissues due to the different genetic backgrounds used.

Single-cell collections for *Ets2-Hhex* mice were conducted at two time points spanning mutant AVE phenotypic initiation and progression: 1. At E5.5 prior to the onset of an overt AVE thickening phenotype, but when the mutant AVE proliferation defect is present; and 2. At E6.25 as the mutant AVE thickening is present and progressing to become more severe. Different dissociation conditions were used for these stages to achieve the mildest condition possible to fully disaggregate the VE without causing excessive cell death by enzymatic treatment (Methods 2.8). Fast genotyping of the Reichert's membrane was added to unequivocally discern embryo genotype.

Heatmap and t-SNE-based clustering of control versus mutant AVE at E5.5 and E6.25 revealed no separation amongst the genotypes. This suggests that the mutant AVE is broadly normal in its transcriptome, and has not transdifferentiated into another cell type. Similar to the E5.5 mutant epiblast, the mutant AVE at both stages was additionally confirmed to be transcriptionally comparable to control counterparts by cluster stability analysis.

I next analysed the DEGs in the mutant AVE using several GO tools: PANTHER, Princeton GO::TermFinder, Reactome, and g:Profiler. My input gene list for GO was candidates with FC > 2 and p-adj < 0.1, as per standard methodology (Ji et al., 2011). For all GO analyses reported in this thesis, I used the default background list, which is commonly employed for studies with genome-wide scope (Huang et al., 2007). A tissue-specific background list (i.e. AVE-specific) is recommended only for studies markedly below genome-wide scope, e.g. 500 genes. Recurrent from all analyses was the significant enrichment of the WNT pathway. Upon assessment of the top 124 most significant candidates defined by highest fold change and lowest p-adj value (cutoff = 0.001, which is very stringent) approximately 53.2% (43 WNT upstream + 23 WNT downstream out of 124 total top candidates) were found to agonise the canonical WNT/β-CATENIN pathway. Collectively, bioinformatics and GO analysis repeatedly implicated WNT as the top misregulated pathway, therefore making WNT signalling a promising candidate for *in vivo* validation. However, other top most significantly enriched GO pathways could be interesting for future investigation, i.e. genes and mechanisms associated with metabolism, cell cycle, chromatin organisation, and cytoskeletal processes.

Validation of mutant AVE scRNA-seq results by pharmacological intervention

To validate the contribution of upregulated WNT/β-CATENIN signalling in giving rise to the mutant AVE phenotype, I perturbed the pathway *ex vivo* via pharmacological cultures. Like with the MEK inhibitor studies, I treated WT embryos, from E5.5 to E6.5, with WNT agonising drug CHIR, which works by inhibiting the β-CATENIN destruction complex component GSK3β, thereby enabling β-CATENIN stabilisation in the absence of upstream signals (Clevers, 2006). Culturing with CHIR at the lowest recommended concentration (3 μM) led to recapitulation of the *Ets2*-mutant thickened AVE phenotype. Coincidentally, CHIR is the other ‘i' of the ‘2i’ inhibitor cocktail used for ESC maintenance, demonstrating that

culturing WT embryos at the same concentrations of 'ESC 2i' – 1 μM PD03, or 3 μM CHIR – led to mutant AVE defect recapitulation. 2i in ESC culture is used to promote self-renewal and prevent differentiation. Interestingly, it appears that 2i is also promoting the 'self-renewal' (increased proliferation) of AVE cells. Assessment of ETS2, GATA6, and β-CATENIN in CHIR-treated WT embryos further demonstrated recapitulation of molecular aspects of the mutant AVE phenotype by WNT agonist treatment. Together, these data lend support to the presence of an ETS2-WNT axis of signalling in regulating AVE cell numbers. This axis can be further tested by attempting to rescue the mutant AVE phenotype by treatment with a canonical WNT signalling inhibitor (e.g. XAV939; Huang et al., 2009). ETS2 appears to be normally acting to decrease WNT/β-CATENIN activity in the AVE. In the absence of ETS2, WNT could become overactive and result in increased AVE proliferation. Further analysis of normal anterior-posterior VE differences with respect to the WNT pathway, as well as how this differs in mutants can lend additional insight into the role of WNT signalling in regulating AVE cell numbers, as well as the levels of WNT required for normal AVE function and maintenance.

In vivo validation of mutant AVE scRNA-seq results by 2C-HR-CRISPR

Having recapitulated the mutant AVE thickening defect in WT embryos with a WNT agonist, I next tested whether I could demonstrate the relevance of this signalling cascade by *in vivo* genetics. Pharmacological treatment involves submerging the entire embryo in the drug and therefore lacks tissue specificity. The concentration of the drug is also difficult to titrate to physiological levels. Misexpressing candidate WNT/β-CATENIN agonists in WT using a VE-specific promoter was a preferred method to ectopically activate this signalling cascade at physiological levels. The *Ttr* promoter for the VE was selected, as the gene is first expressed at E5.0 (demonstrated using a *Ttr*-fluorescent reporter by Dr. Matthew Stower in the Srinivas

Group; unpublished), which is prior to the onset of the increased proliferation defect in mutants. Furthermore, since ETS2 appears to be expressed throughout the VE, but only the AVE appears to be sensitive to the loss of ETS2, activating WNT throughout the VE should in theory only affect the AVE.

To efficiently misexpress top WNT/β-CATENIN agonists in the VE, I attempted the 2C-HR-CRISPR approach for efficient knock-in gene targeting of a large donor DNA fragment via two advances: 1. CRISPR microinjection at the mouse embryonic 2-cell; and 2. Biotin-streptavidin tagging of donor PCR template and Cas9, respectively, to increase the probability of the CRIPSR components to interact and mediate editing (Gu et al., 2018). *Moesin* was selected as the top candidate to test by 2C-HR-CRIPSR, as its upregulation was validated *in vivo* by antibody stain, and it is the top upregulated candidate at E5.5, which is the initiation point of the mutant AVE defect. Moesin has been shown to activate the WNT/β-catenin pathway and result in increased proliferation upon entering the nucleus (Batchelor et al.,2004; Zhu et al., 2012). I cloned *Moesin*, as well as other candidate WNT agonists into the *Ttr* targeting backbone provided by Dr. Jayaram. This *Ttr* construct contained 5' and 3' HA for targeted knock-in to the endogenous *Ttr* locus, as well as *Td-Tomato* reporter and 2A peptide upstream of the insertion site of the gene of interest. I cloned my gene of interest into the *Ttr* backbone and conducted the biotinylated PCR (Gu et al., 2018) to generate the final linearised injection construct with biotin tags on either end. The donor template PCR construct was verified to lack mutations by Sanger sequencing prior to microinjection. I dissected embryos from 2C-HR-CRISPR injections at E6.5, which is a stage when the mutant AVE phenotype should be evident, and noticed no developmental delays in injected embryos compared to WT litters. I stained *Ttr-Msn* 2C-HR-CRISPR embryos with F-ACTIN, similar to what I have done to characterise the AVE phenotype to enable assessment of epithelial architecture; and AVE marker AMOT.

Ttr-Msn 2C-HR-CRISPR was inconclusive in this thesis and requires further testing. Although 13/34 embryos appeared to have an AVE thickening appearance, only 2/34 were genotyped for positive integration of *Ttr-Msn* by Dr. Jayaram. New primers may be designed to verify this result. Furthermore, Td-Tomato was not detected in any of the embryos, even in the ones that were genotyped for positive integration. Because, the *Td-Tomato* sequence was verified to be correct prior to injection, absence of fluorescent signal could be due to errors in protein folding, in which case another fluorescent protein tag may be considered. The *Td-Tomato* sequence also has multiple repeated elements, making it susceptible to genotyping errors.

Xiap, a top upregulated candidate at E6.25, was additionally evaluated by 2C-HR-CRISPR using the identical construct design, but swapping out *Msn* for *Xiap*. Misexpression of *Xiap* did not result in a visibly evident thickened AVE phenotype. However, I had much lower microinjected embryo numbers for *Xiap* as compared to *Msn*. When I genotyped for positive integration using the same primers as for *Msn*, I found that 2/16 embryos showed presence of gene targeting. These embryos may have slightly more crowded AVE, but this requires quantification (blinded to genotype). The next step is to increase numbers, conduct blinded genotyping (with new primers), as well as carry out blinded quantification of cell crowding in the AVE.

Other top upregulated candidates that agonise the WNT pathway at E5.5 and E6.25 can be tested by 2C-HR-CRISPR. Alternatively, candidates that do not classify with the WNT pathway, but regulate proliferation and cytoskeletal dynamics could also be interesting to misexpress. These candidates would first be ideally validated in control and mutant embryos by antibody stain, or by ISH. An *in vitro* system can additionally be utilised where control and mutant ESCs, in particular those that express HHEX-GFP to mark the AVE can be differentiated into embyroid bodies that overexpress MSN, or other WNT-agonising

candidates. A WNT reporter such as pTopflash (Korinek et al., 1997) may be employed to confirm overactivation of the WNT pathway *in vitro* upon upregulation of Msn in AVE-like cells.

In summary, 2C-HR-CRISPR had been attempted to misexpress top candidate WNT agonists in the VE of WT embryos to recapitulate the mutant AVE phenotype. These experiments require troubleshooting to attain presence of fluorescent reporter expression, and robust genotyping for positive integration. Though the technique may not be as efficient as reported in Gu et al., 2018, it does allow for targeted insertion of a large donor construct, which other CRISPR technologies do not permit (Quadros et al., 2017). In this case, a large donor construct (~ 5 kb) is required for whole gene misexpression under the regulation of an endogenous promoter. Once improved, the same CRISPR approach may be used to ectopically express other top upregulated candidates from scRNA-seq analysis *in vivo.* CRISPR-validated candidates can be further examined *in vitro* to probe, in greater depth, the mechanism by which ETS2 regulates the WNT pathway.

6. Technology development for investigating mechanistic underpinnings of early mouse embryogenesis

6.1 Introduction

The first 3 results Chapters have focused on the role of transcription factor ETS2 in early mouse embryogenesis. ETS2 has confounding functions in human cancer, as it may either promote, or suppress tumour growth depending on its transcriptional cofactors (Plotnik and Hollenhorst, 2017). The *Ets2* loss-of-function mouse mutant *Ets2* ^{db1} is embryonic lethal by E9.0, but interestingly has phenotypes related to both oncogenic and tumour suppressive roles of ETS2 at early post-implantation stages. Namely, the ExE is reduced and the AVE is overgrown in *Ets2*-mutants, suggesting that ETS2 may normally function like an oncogene in the ExE (e.g. prevent apoptosis) and like a tumour suppressor in the AVE (e.g. limit proliferation). Having characterised the phenotypic and molecular features of the mutant ExE and AVE defects in Chapter 3 and 4, I turned to scRNA-seq technology in Chapter 5 to investigate the downstream transcriptional effectors of ETS2 in a tissue-specific fashion. To do so, I adjusted published single-cell dissociation protocols (Mohammed et al., 2017; Wen et al., 2017) to enable recovery of sufficient numbers of AVE cells (without microdissection) at E5.5 and E6.25 to compare *Ets2*-mutant versus littermate control AVE, and conduct differential gene expression analysis.

6.1.1 ScRNA-seq of the VE lineage

While I was investigating ETS2 function in the AVE by scRNA-seq, two papers had been published revisiting the trajectory of VE development at single cell resolution (Nowotschin et al., 2019; Pijuan-Sala et al., 2019). While these scRNA-seq studies confirmed the presence of stereotypic tissue-specific markers, they also propose new hypotheses about VE contributions in development that require *in vivo* validation. In particular, their findings

challenge the conventional paradigm that the gut endoderm derives from the definitive endoderm (DE), which comes from the anterior primitive streak of the epiblast (Lawson and Pederson, 1987). ScRNA-seq analysis of E6.5 to E8.5 embryos by Pijuan-Sala et al., 2019 showed that the VE can contribute to gut endoderm formation at E8.5. They conducted diffusion pseudotime trajectory analysis (DPT; Haghverdi et al., 2016), and generated transport maps (Schiebinger et al., 2017) to predict the separate contributions of E7.0 VE and DE to the gut endoderm by E8.5. They found the convergence of VE (marked by *Ttr*) and definitive endoderm (marked by *Mixl1*) in giving rise to distinct regions of the gut endoderm. This study demonstrates the novel insights that scRNA-seq combined with new computational approaches can provide in understanding embryonic development in greater detail.

Another recent study tracing the ontogeny of gut endoderm formation involved large scale scRNA-seq (10x) to track endoderm development from E3.5 to E8.75 (Nowotschin et al., 2019). They used a different computation approach – Harmony (developed by Nowotschin et al., 2019) combined with Palantir (Setty et al., 2019) – to generate a map of endoderm developmental progression. They found different subclusters in the endodermal lineage over time eventually giving rise to distinct regions of the gut, including contribution from the embryonic VE.

The next step for both of these studies is to validate their findings by *in vivo* lineage tracing. While Nowotschin et al., 2019 confirmed spatial patterning, predicted from scRNA-seq, of different domains of gene expression in the E8.75 gut tube by ISH, quantitative hybridisation-chain-reaction (HCR) *in situ* analysis (Choi et al., 2018) could be additionally beneficial to examine transcript expression at higher resolution. HCR-ISH enables probing of multiple genes from a single embryo, and also allows for co-stain with DAPI and Phalloidin (F-ACTIN marker), so that cell and tissue boundaries can be clearly visualised. Namely, this would enable quantification of the precise number of cells with a particular expression pattern,

representing a specific subpopulation from scRNA-seq. Finally, the geographical positioning of cells from the embryo were predicted for scRNA-seq analysis in Pijuan-Sala et al., 2019 and Nowotschin et al., 2019. Knowledge of the precise positioning of distinct cell populations and how they change and relocate overtime could be a helpful next step in accurately tracking the developmental trajectory of the VE. Currently, this may be accomplished by Geo-seq, which combines modified Smart-seq2 and laser capture microdissection to preserve the embryo location which the cells came from (Peng et al., 2016; Chen et al., 2017). However, the technology is not yet at single-cell resolution.

In summary, scRNA-seq of the VE lineage overtime has provided new insights into the formation of the gut tube from mixed embryonic and extra-embryonic populations. However, these studies did not examine early post-implantation VE development in detail, i.e. from E5.5 to E6.5 – spanning different stages of AVE migration. As the anterior shift of the AVE establishes the anterior-posterior body axis, it would be interesting to investigate the molecular differences between the AVE versus VE. This can help elucidate how the AVE can distinctly exert its developmental signalling function in patterning the epiblast and localising the primitive streak to the posterior. It can also help address why the AVE undergoes directed migration while the rest of the VE does not. Furthermore, scRNA-seq can uncover novel markers of the AVE and VE, which can complement other single cell reports mentioned earlier in this section to interrogate the developmental function and consequences of VE contribution to the gut endoderm.

6.1.2 ESC derivation and technology

While investigating developmental mechanisms in the native context of the embryo is ideal, certain techniques are not yet amenable to the small amount of starting material the early embryo provides. These include chromatin immunoprecipitation sequencing (ChIP-seq;

O'Geen et al., 2011) to uncover regions of chromatin a particular transcription factor binds; and mass spectrometry (mass spec) to identify expression of proteins in a given context. While single-cell mass-spec is being developed, the technology is not yet sufficiently robust for studying embryos (Doerr, 2019). Both ChIP-seq and mass spec would be informative in relation to transcription factor ETS2 to elucidate direct downstream targets, and functional consequences of ETS2 transcriptional activity at the protein level. Therefore, a relevant *in vitro* system would be useful for providing sufficient material to address these questions immediately.

Embryonic stem cell (ESC) culture is a powerful *in vitro* approach to complement *in vivo* studies. It serves as a high throughput platform for conducting biochemical, genetic and molecular studies to investigate mechanism in depth. However, this is provided that ESCs, or the product of ESC directed differentiation is a suitable model of the *in vivo* cell type of interest. For the *Ets2*-mutant, AVE-like and ExE-like cells are required *in vitro* to investigate the molecular underpinnings, and in particular, tissue-specific transcriptional outputs of ETS2 that underlie the AVE and ExE mutant phenotypes.

Mouse ESC derivation

The first embryonic stem cell isolations from mouse blastocysts were reported by Evans and Kaufman, 1981 and Martin, 1981. Subsequently, Bryja et al., 2006 described a protocol for derivation of mouse ESCs from E3.5 blastocysts that utilises mouse embryonic fibroblast feeders (MEF) at early passage, as well as switching between fetal bovine serum and serum replacement-containing media. This method yielded a recovery of approximately 50%, i.e. ability to obtain 50 mESC lines from 96 total blastocysts. Here, the entire process from ESC extraction to being able to freeze down the ESC line takes 15 to 20 days. Mouse ESCs (mESCs) derived from ICM of the pre-implantation blastocyst can give rise to all cells of the organism,

including germ cells (Bradley et al., 1984). They can self-renew when provided with leukemia inhibitory factor (LIF) (Smith et al., 1988; Williams et al., 1988), and may be grown in culture indefinitely without reaching senescence (Miura et al., 2004).

Differentiation of ESCs may be blocked by inhibiting MAPK signalling (Kunath et al., 2007) using PD0325901 (PD03), and activating canonical WNT/β-catenin signalling through blocking GSK3 using CHIR99021 (CHIR; Ying et al., 2008). LIF may be added to improve efficiency, but is not necessary (Wray et al., 2011). Though ESCs are usually derived from the ICM of the E3.5-E4.5 blastocyst, they can technically be obtained, as well as propagated from any stage of the pre-implantation blastocyst (Tesar, 2005). ESC derivation and propagation is fully efficient from the E4.5 epiblast that features ERK-independent self-renewal *in vitro* (Boroviak et al., 2014).

The ESC derivation protocol utilised in this thesis was an updated version provided by Professor Jennifer Nichols. The protocol involves flushing out mouse embryos at the 8-cell stage and growing them in 2i media to prevent formation of the primitive endoderm. The hatched blastocyst is then subject to immunosurgery to rid of the TE, and to be left with the pre-implantation epiblast. Each isolated epiblast is then plated in a well of a 96-well plate, coated with gelatin and containing 2i and LIF. The ESCs are then passaged and expanded in the same conditions.

Trophoblast stem cell (TSC) derivation had not been possible with *Ets2*-mutants, as the cells could not propagate in culture (Wen et al., 2007), presumably due to the defective and reduced mutant ExE. The ICM and E5.5 epiblast, on the other hand, appear normal in size, and the latter in transcriptional profile, suggesting that ESC derivation should be possible with *Ets2*-mutants.

This Chapter is focused on generating resources, and utilising scRNA-seq sample preparation protocols refined in this thesis to address new mechanistic questions in early mouse

embryogenesis. It examines WT VE by scRNA-seq to investigate the differences between the VE and AVE from E5.5 to E6.5; generates *Ets2*-mutant and control ESCs to serve as a high throughput *in vitro* system for studying ETS2 function; and investigates mechanisms of epiblast cell competition via pharmacological culture and scRNA-seq.

6.2 Results

6.2.1 Enriching for VE cells by vital fluorescent labelling for scRNA-seq

In 5.2.5, I used the HHEX-GFP fluorescent reporter to purify for AVE cells by FACS-mediated single cell collection. As this approach does not enable recovery of the entire VE, I devised a method to enrich for both AVE and surrounding VE cells. Since we did not have a VE-reporter line in a suitable genetic background, i.e. compatible with the mutant AVE single-cell dataset, or to serve as the standard reference dataset (C57BL/6; e.g. used in Mohammed et al., 2017), I attempted a non-genetic approach to fluorescently mark the VE. As the VE is a highly trophic tissue (Bielinska et al., 1999), I reasoned that treating post-implantation embryos with a vital dye for a brief period of time at low concentration may be sufficient to label the surface VE without having the dye penetrate into the epiblast and ExE. I tested vital dyes PKH26 (555 nm; Sigma PKH26GL) and Claret (633 nm; Sigma Minclaret-1KT), which have long aliphatic tails that stably incorporate into lipid domains of the cell membrane. I assessed red (PKH26) and far red (Claret) vital dye labelling in post-implantation embryos (Figure 109, 110 – E5.5, E5.75, E6.5) and found both to have similar labelling efficiencies. The useful advantage of these dyes is their long *in vivo* half-life (more than 100 days). This persistence of fluorescence meant that the dyes would be suitable for single cell experiments where embryos are labelled, dissociated into single cells and sorted by FACS, which takes 2-3 hours. Through testing various concentrations, I found that half the recommended concentration, with the exact suggested time (5 minutes at room temperature) was suitable for labelling the VE at E5.5 and

E6.5 (Figure 110). To verify the retention of fluorescent signal for a long period of time, I cultured labelled embryos overnight and imaged them by confocal microscopy. I found retention of VE fluorescent signal following overnight culture of labelled embryos, applicable for both dyes (example with PKH26 only: Figure 109). I proceeded to carry out WT VE single-cell collections with far red vital dye labelling to keep the red channel free. This is because there are more relevant fluorescent reporter mouse lines with red fluorescence than far red in the Srinivas laboratory. I proceeded to collect VE cells from WT embryos at E5.5, E5.75, and E6.25. For E5.5 and E5.75 collections, I used only the C57BL/6 background. For E6.25, I used both C57BL/6 and the identical background as *Ets2-Hhex*, so that the WT VE data can be directly compared with the mutant AVE data.

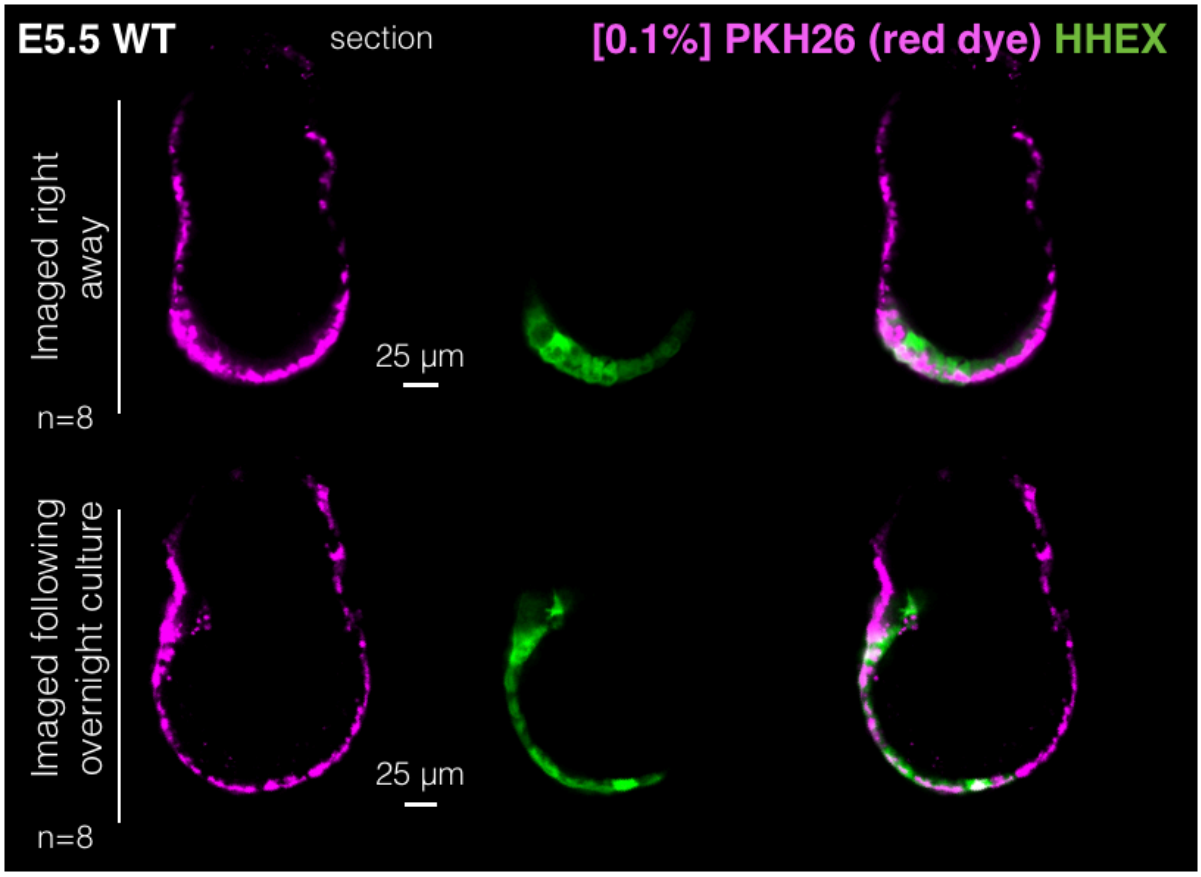


Figure 109. PKH26 red fluorescent vital membrane labelling (0.1% concentration) of E5.5 embryos, imaged right away and following overnight culture
Representative mid-sagittal sections of E5.5 embryos labelled with [0.1%] PKH26 (magenta); HHEX-GFP (green – AVE). PKH26 signal is retained in the VE directly after labelling, as well as after overnight culture. Scale bars: 25 µm.

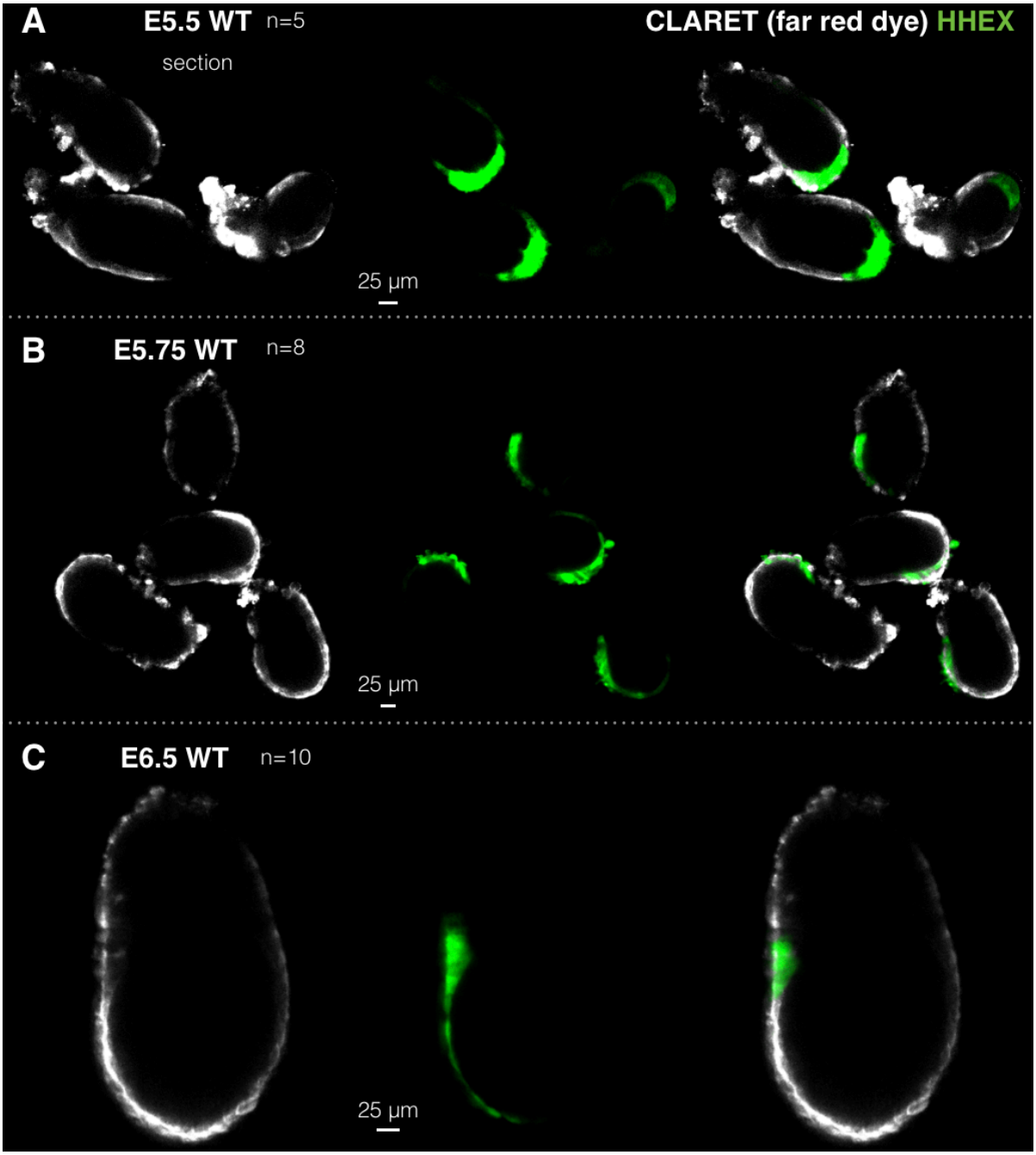


Figure 110. Claret far red fluorescent membrane labelling (0.1% concentration) of E5.5, E5.75 and E6.5 embryos

Representative mid-sagittal sections of (A) E5.5, (B) E5.75, and (C) E6.5 embryos labelled with [0.1%] Claret (gray); HHEX-GFP (green – AVE). Claret appears to be VE-specific at all stages. Scale bars: 25 µm.

6.2.2 Analysis of WT VE scRNA-seq data

WT VE scRNA-seq analysis was conducted by Dr. Maria Solovey (E6.25), Dr. Jonathan Fiorentino (E5.5), and Dr. Antonio Scialdone (Group Leader).

ScRNA-seq of E6.25 WT VE

Analysis of E6.25 WT VE presented in this thesis only encompasses cells from the *Ets2-Hhex* mixed background. Standard quality control steps were first conducted to rid of poor quality reads. Of the 379 WT VE cells collected, only 240 passed all quality control steps (63.3% recovery). Known layer markers were utilised to distinguish VE and AVE cells, and filter out other tissues that may have been accidentally labelled by the far red dye. HHEX-GFP fluorescent reporter was additionally used to discern AVE (HHEX-GFP-positive) versus surrounding VE. Clustering based on layer marker genes employed tissue-specific markers used in Mohammed et al., 2017: ExE – *Cdx2, Bmp4, Furin, Pcsk6, Eomes, Elf5, Esrrb, Ets2*; epiblast – *Pou5f1, Fgf4, Nodal*; AVE – *Hhex, Cer1, Lefty1, Dkk1, Lhx1*; and VE – *Gata6, Ttr Afp*. From clustering based on signature marker genes for each layer, the VE and AVE cells appeared to segregate into separate clusters. To test whether far red positive cells were indeed VE, correlation analysis was conducted to assess the proportion of these cells expressing VE markers *Gata6, Afp,* and *Ttr.* This analysis demonstrated that most far red-positive cells do indeed express *Ttr* and *Afp*, as well as *Gata6*, but to a lesser extent (data not shown). Clustering analysis based on highly variable genes (HVGs) revealed 4 distinct clusters (Figure 111), distinguishing VE, epiblast, AVE, and ExE.

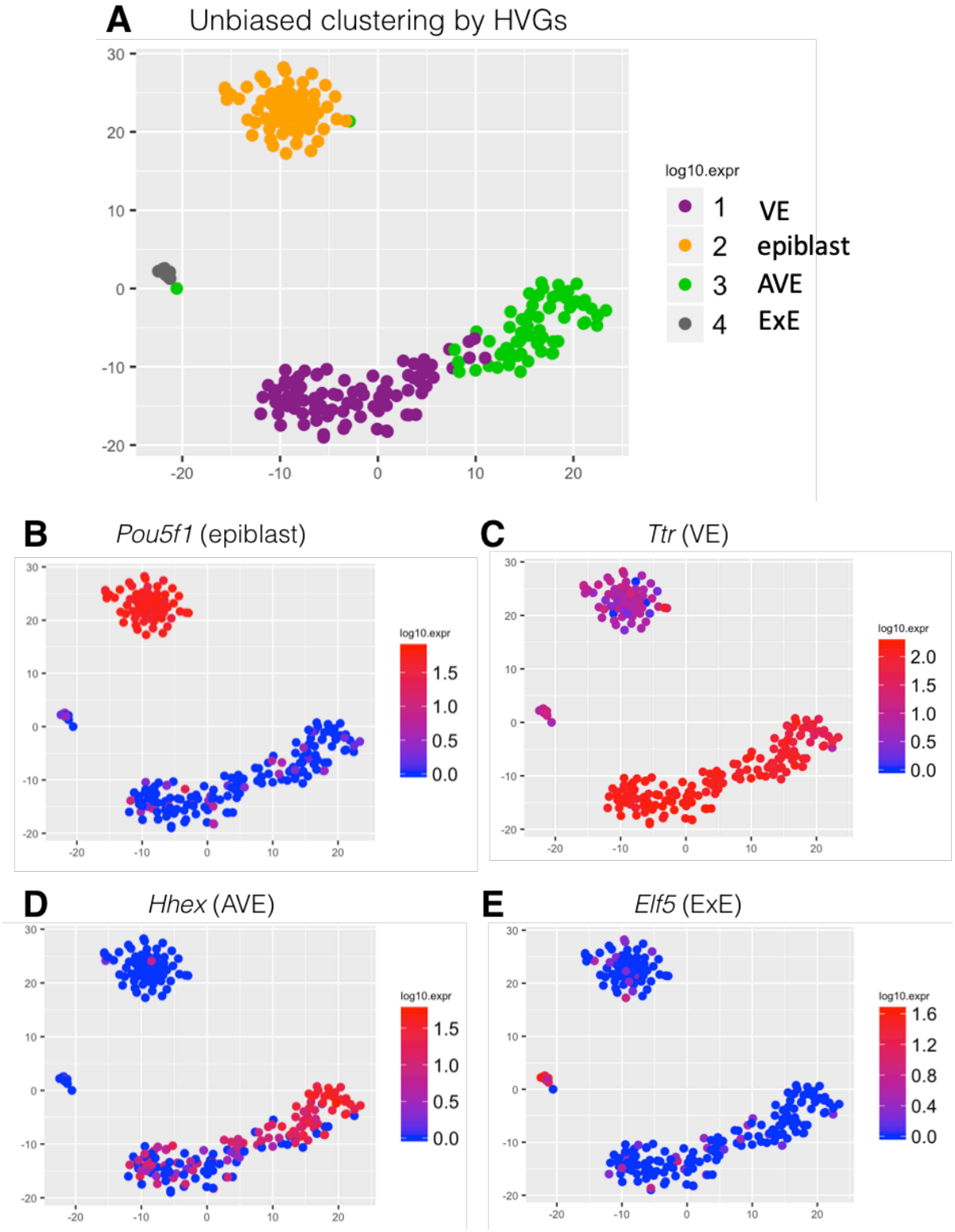


Figure 111. Unbiased clustering based on highly variable genes (HVGs) with t-SNE representation
(A) 4 distinct clusters appear by t-SNE representation upon clustering based on HVGs. Each cluster was then validated based on known tissue-specific markers shown as a red to blue heatmap where red is 'highly expressed': (B) epiblast – *Pouf51*, (C) VE – *Ttr*, (D) AVE – *Hhex*, (E) ExE – *Elf5*.

Following assignment and identification of clusters, the individual clusters could then be analysed to compare DEGs in different cell populations. Specifically, the focus here was to compare AVE versus surrounding VE. This was particularly interesting since other scRNA-seq datasets either had low cell numbers for the VE at E6.25, or did not intentionally use a specific marker for the AVE (Mohammed et al., 2017; Wen et al. 2017; Pijuan-Sala et al., 2019; Nowotschin et al., 2019). HHEX-GFP to mark the AVE served as an additional method to distinguish anterior versus posterior VE at E6.25.

Previous ISH, immunofluorescence, and genetic fluorescent reporter analysis has demonstrated that the AVE is a heterogeneous population, expressing different combinations and levels of its signature markers at different regions (Stower and Srinivas, 2014; Hoshino et al., 2015). This was supported by scRNA-seq analysis, which revealed non-overlapping populations of the AVE, demonstrated by plotting log10 gene expression of different AVE markers against each other to visualise and statistically analyse correlation (Figure 112).

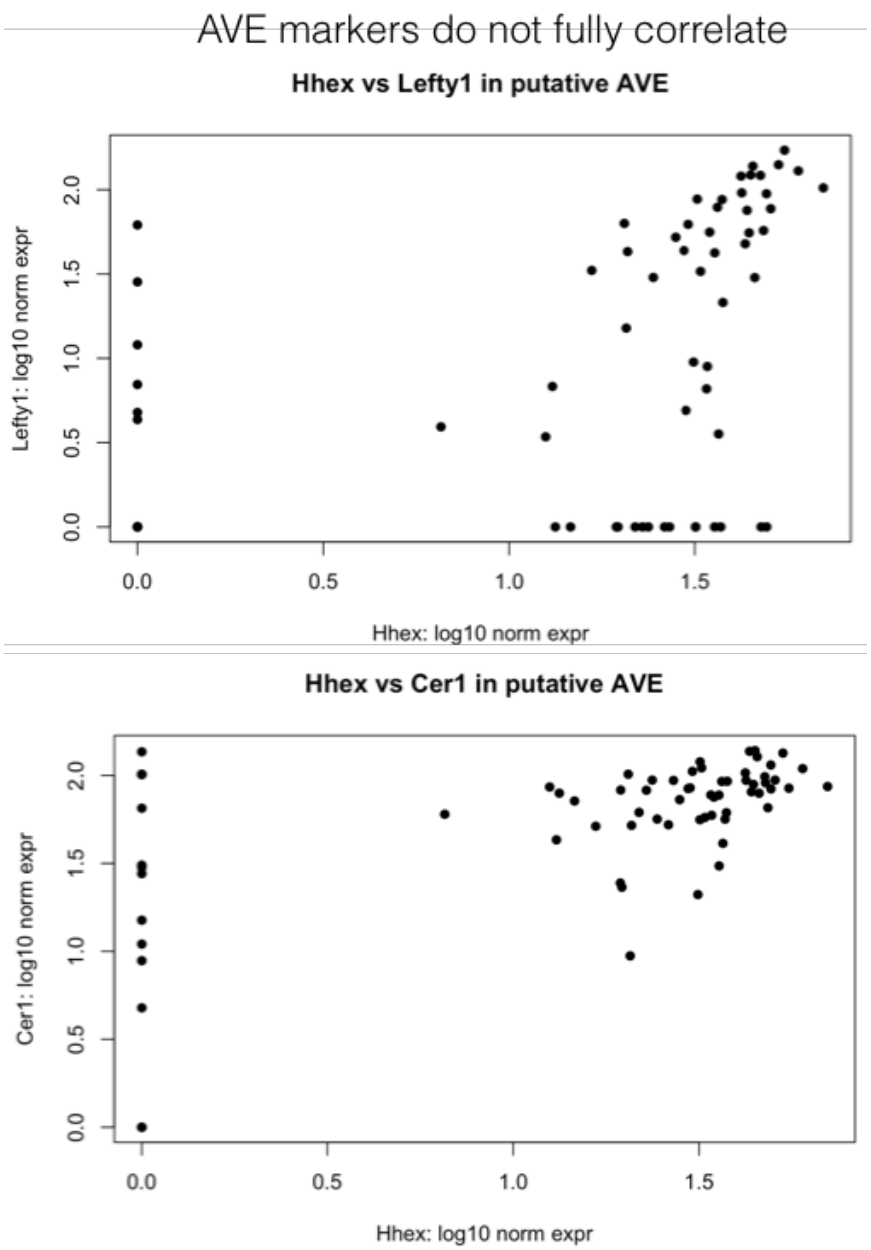


Figure 112. Plotting normalised expression levels of *Hhex*, *Lefty1*, and *Cer1* AVE markers against each other demonstrate that the markers do not fully overlap, supporting heterogeneity in the AVE
Hhex versus *Lefty1* and *Hhex* versus *Cer1* correlation plots show some cells that are not correlated in simultaneous expression of these AVE markers, e.g. there are *Lefty1*-high/*Hhex*-low, *Hhex*-high/*Lefty1*-low, and *Cer1*-high/*Hhex*-low cells.

DEG analysis between VE and AVE was subsequently conducted to interrogate significant differences between these two cell populations at E6.25, and what may account for their distinct behaviour and function. The top 50 genes upregulated in the AVE as compared to the VE upon DE analysis were as follows (Figure 113):

Top 50 DEGs upregulated in AVE

```
 [1] "Lefty1"  "Fgfbp1"  "Otx2"  "Cd8a"  "Shisa2"  "Slc39a2"
 [7] "Rgs8"  "Hesx1"  "Dkk1"  "Gsc"  "Nptx2"  "Tmem63a"
[13] "Fgf8"  "Cdc42ep5"  "Tph1"  "Rgs16"  "Cer1"  "Gata1"
[19] "Lhfp"  "Tuba8"  "Tox3" "Fzd8"  "Cald1"  "Krt19"
[25] "Mest"  "Sema5b"  "Adgrg1"  "Usp18"  "Upp1"  "Afap1l1"
[31] "Cdhr1"  "Stra6"  "Prg2"  "Spon2"  "Flt1"  "Gprc5a"
[37] "Hs3st6"  "ENSMUSG00000063646" "Capn6"  "Cd24a"  "Ovol2" "Gm10359"
[43] "Gm12671"  "Mobp"  "Krt7"  "Pcdh7"  "Fzd5"  "Piezo2"  "Coq8a"  "Akr1c19"
```

Figure 113. Top 50 DEGs upregulated in the AVE compared to surrounding VE
Lefty1, Otx2, Hesx1, Dkk1, Gsc, and *Cer1* have been previously reported as AVE markers. Other markers require *in vivo* validation.

GO analysis of upregulated DEGs with FC > 2 and p-adj < 0.1 in the AVE compared to VE using GOrilla (Eden et al., 2009), conducted by Dr. Solovey, revealed the following pathways enriched in the AVE (Table 21).

GO analysis of upregulated DEGs in the AVE

GO: Process

	GO Term	Description	P-value	FDR q-value
2	GO:0007275	multicellular organism development	2.75E-5	1.45E-1
3	GO:0043506	regulation of JUN kinase activity	1.07E-4	2.82E-1
4	GO:0009798	axis specification	2.24E-4	3.93E-1
5	GO:0009948	anterior/posterior axis specification	3.33E-4	4.38E-1
6	GO:0009653	anatomical structure morphogenesis	3.56E-4	3.75E-1
7	GO:0071900	regulation of protein serine/threonine kinase activity	4.82E-4	4.23E-1
8	GO:0019439	aromatic compound catabolic process	4.99E-4	3.75E-1
9	GO:0048731	system development	5.12E-4	3.37E-1
10	GO:0042074	cell migration involved in gastrulation	7.79E-4	4.55E-1

Table 21. GO analysis of upregulated DEGs in the AVE using GOrilla
Showing top 10 most enriched GO terms for biological process.

Amongst the pathways upregulated in the AVE, an interesting hit was 'regulation of JUN kinase activity', as this pathway had not been previously implicated in AVE development directly. However, Jun kinase is known to become activated downstream of TGFβ signalling, which has been characterised to be involved in A-P axis development via Nodal and Activin (Perea-Gomez et al., 2001; Clements et al., 2011) that can trigger Jun kinase activation (Figure 114; Zhang, 2009).

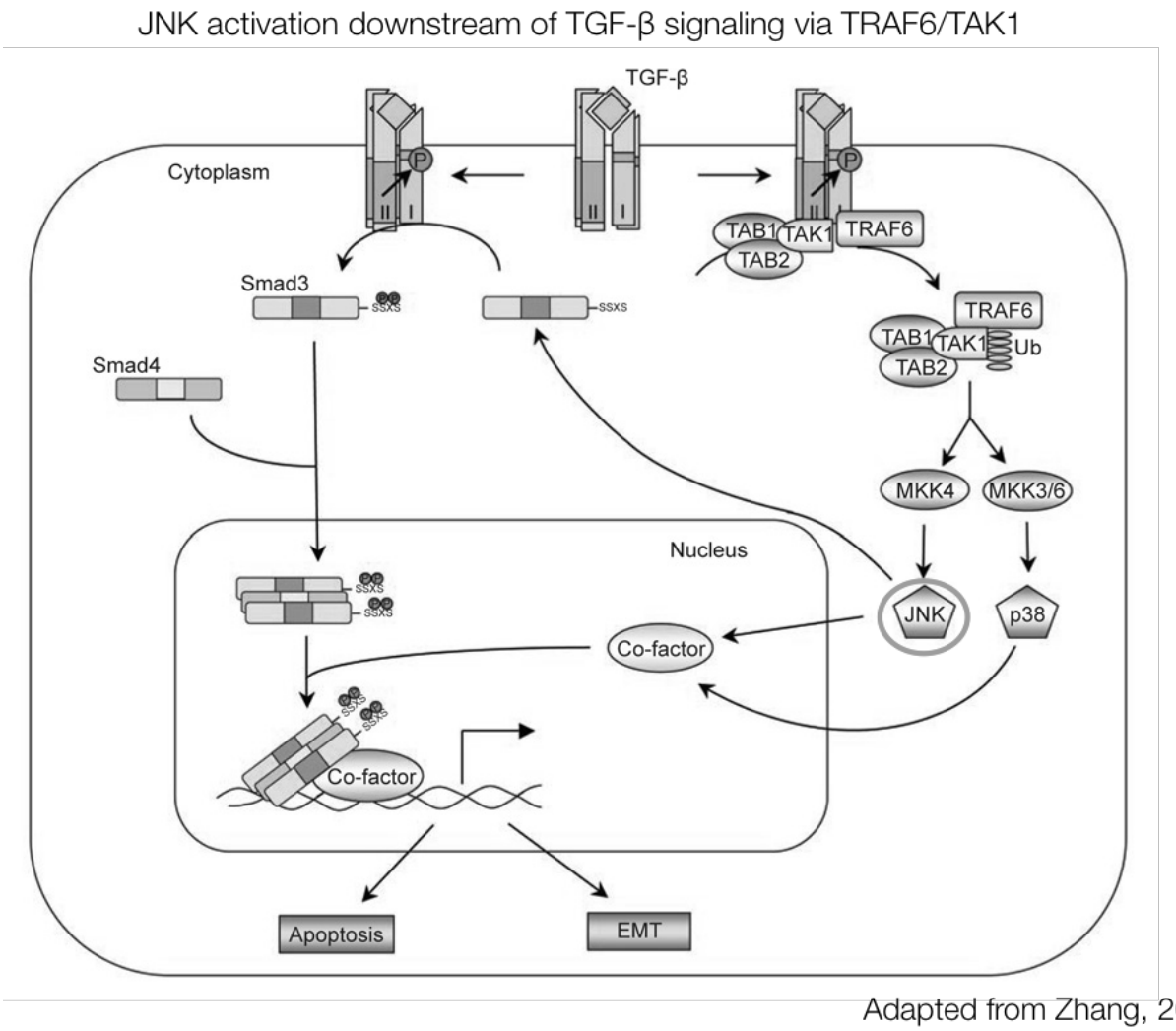


Figure 114. JUN kinase (JNK) is known to be activated downstream of TGFβ signalling
Both JNK and p38 have been shown to become activated downstream of TGFβ receptor activation via TRAF6 and TAK1 recruitment, differing from the canonical TGFβ-Smad signalling cascade. Activated JNK/p38 can subsequently interact with Smads in the nucleus along with transcriptional cofactors to regulate apoptosis and epithelial-to-mesenchymal transition (EMT).

Trajectory analysis via the DPT approach (Haghverdi et al., 2016) was next conducted to further probe differences between the AVE and VE at E6.25 (Figure 115). The technique allowed elucidation of novel markers of AVE and VE that may be more highly expressed and specifically enriched in the respective populations than markers conventionally used.

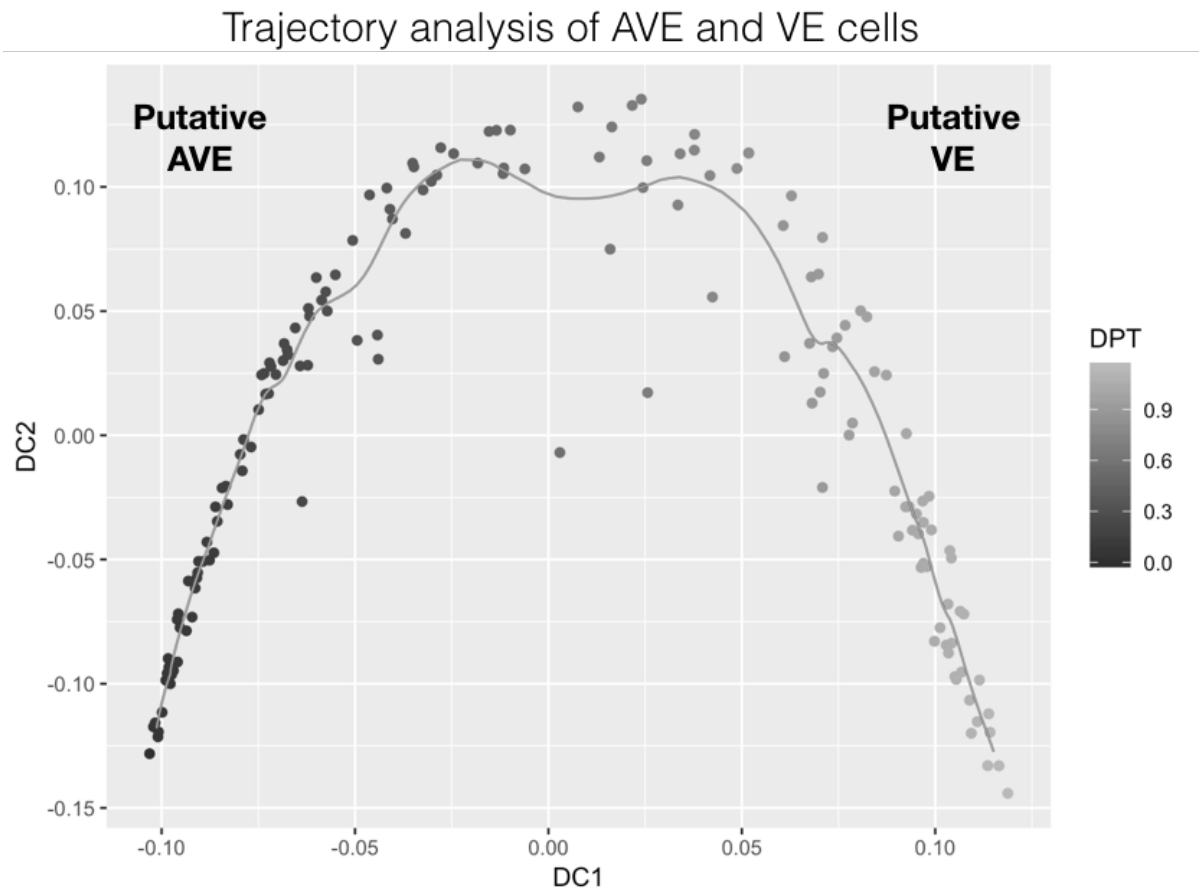


Figure 115. Trajectory formation of AVE and VE cells to spatially position them along the anterior-posterior embryo axis at E6.25
Pseudospatial plotting of clusters was conducted using diffusion pseudotime (DPT; Haghverdi et al., 2016) to simulate the anterior-posterior distribution of VE/AVE single cell transcriptomes.

Known marker genes were then assessed to see where they fall within the pseudospatial trajectory, and to validate whether the expected pattern of the gene's known spatial distribution is observed. Marker genes that overlap with trajectory DEGs verified the correct assignment of anterior versus posterior in the E6.25 trajectory, i.e. AVE markers were expressed in the anterior and *Wnt3* was present in the posterior (Figure 116).

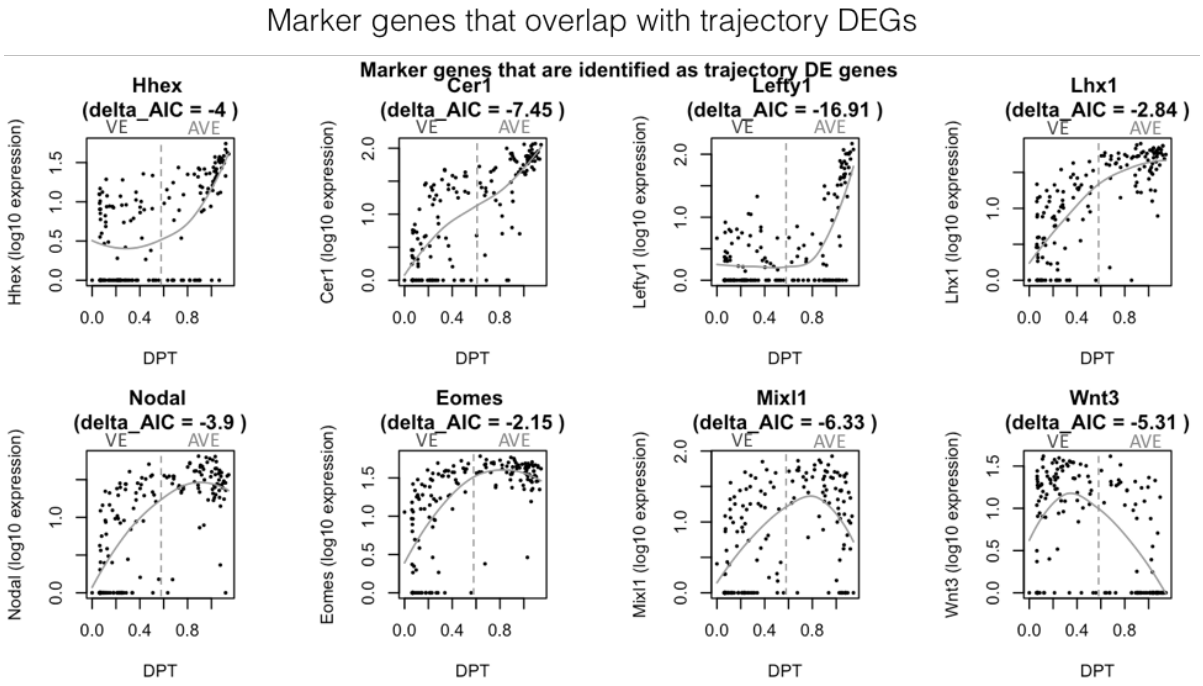


Figure 116. Marker genes that overlap with trajectory DEGs are mostly in accordance with previous literature apart from *Mixl1*

Hhex, Cer1, Lefty1, Lhx1, Nodal, and *Eomes* are upregulated in the AVE compared to VE (Stower and Srinivas, 2015; Hoshino et al., 2015) in the trajectory. *Wnt3* is known to be upregulated in the posterior VE (Yoon et al., 2015), also confirmed in the trajectory map. However, *Mixl1* is expected to be expressed in the posterior shown by ISH (Chu et al., 2004), but appears in this analysis to be upregulated in the AVE.

The pseudospace trajectory was further validated using marker genes that are expected to not overlap with trajectory DEGs, i.e. known to be expressed throughout the VE. These include: *Gata6, Afp,* and *Ttr* (Figure 117).

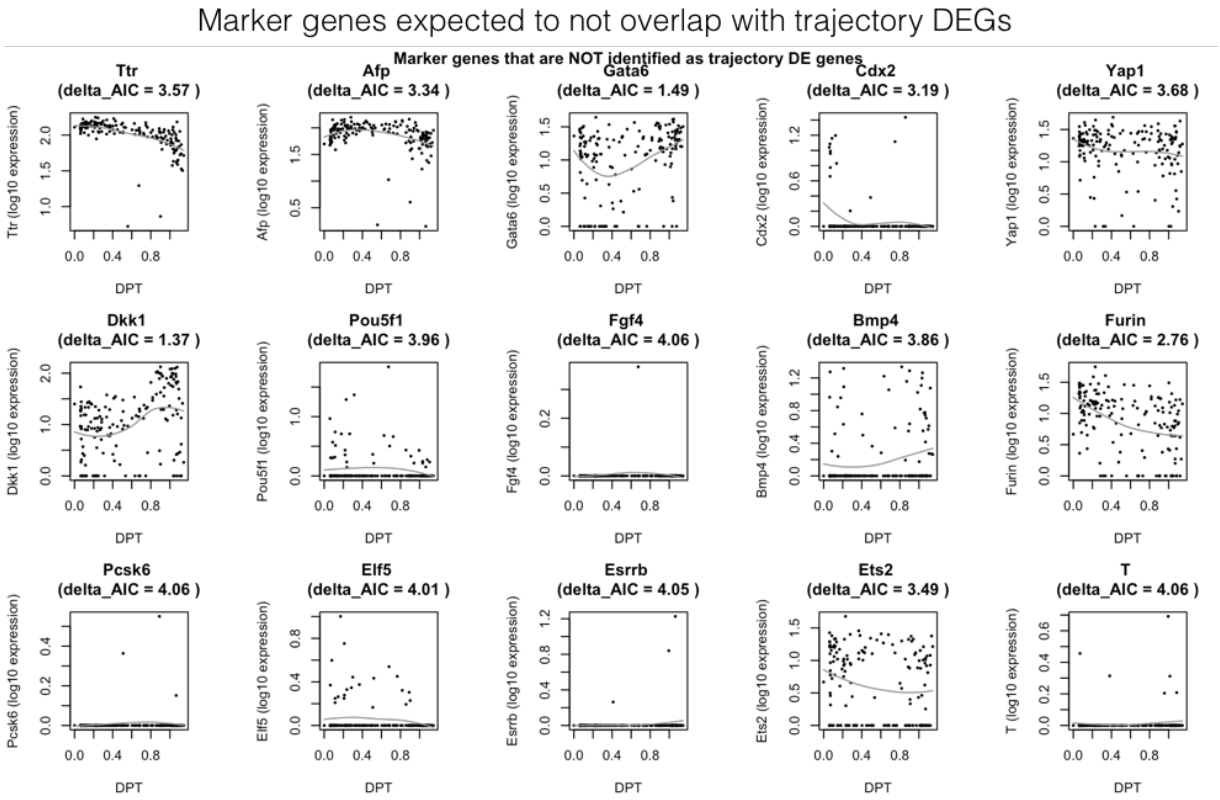


Figure 117. Expected marker genes that do not overlap with trajectory DEGs

Once the trajectory was validated using known marker gene expression patterns, the top 10 trajectory DEGs were identified, and novel VE-specific markers were elucidated (Figure 118).

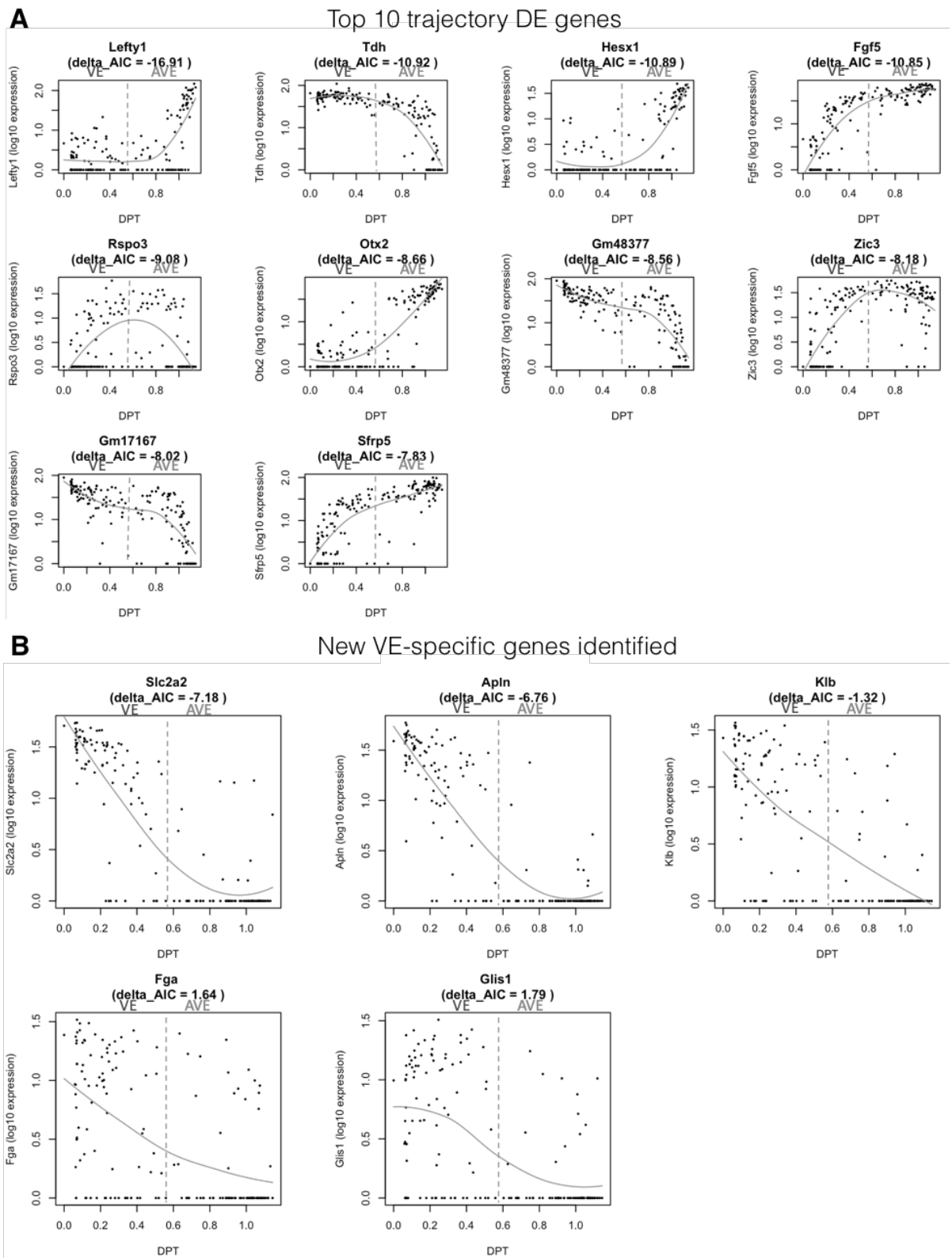


Figure 118. Utilising the validated trajectory plot to identify the top 10 trajectory DEGs
(A) Top trajectory DEGs include known AVE marker *Lefty1*.
(B) Novel VE-specific genes are identified using validated diffusion plot. These genes include: *Slc2a2, Apln* – receptor/Nodal signalling agonist (Deshwar et al., 2016), *Klb, Fga,* and *Glis1*.

This analysis uncovered several novel marker genes for the anterior versus posterior VE. The following plot shows the top 20 best markers for AVE and VE, with the most specific

AVE markers being at the furthest bottom right, while the most specific VE markers are at the top left (Figure 119).

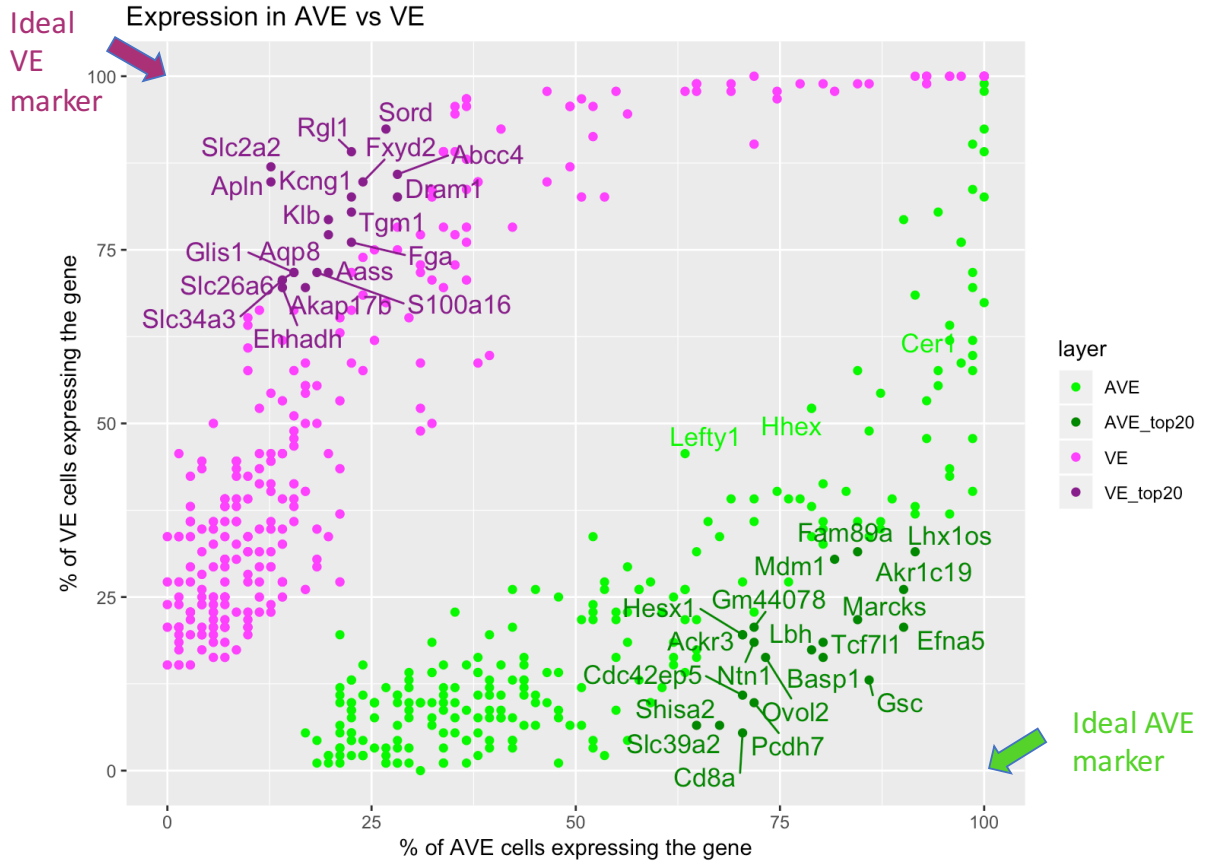


Figure 119. Top 20 markers for the AVE and VE based on trajectory pseudospatial analysis

Interestingly, conventional AVE markers were not present in the list of top 20 markers for the AVE from trajectory analysis. This analysis represents an unbiased, high resolution single-cell approach to assess differences between WT AVE versus VE at E6.25. The next step is to validate these novel markers by antibody staining, or HCR-ISH if a suitable antibody is unavailable. I have validated one of these markers TCF7L1 – an inhibitor of ESC self-renewal and a key regulator of the pluripotency GRN in ESCs (Tam et al., 2008; Hoffman et al., 2013) – by immunostaining from early to late post-implantation (Figure 120, 121). Indeed, TCF7L1 displays higher nuclear expression in the anterior (marked by HHEX-GFP) compared to the posterior by E6.5. Additional analysis that can be conducted with this dataset involves investigating the receptor-ligand interactions between the anterior VE and anterior epiblast, as

well as that for the posterior. A similar type of approach to uncover direct cellular communication via receptor-ligand interactions between different cell types has been previously conducted with scRNA-seq data in the context of the human fetal-maternal interface (Vento-Tormo et al., 2018). As a starting point, receptors and ligands have been identified amongst the top 50 DEGs from the AVE-VE diffusion pseudotime trajectory. These DEGs are presented in a heatmap where the top 29 genes (*Gpc4* to *Per2*) most likely demarcate the AVE, while the bottom 21 genes (*Tdh* to *Myc*) most likely demarcate the VE (Figure 122). Detailed receptor-ligand analysis, involving computational and mathematical approaches to investigate cell-cell communication has been conducted by Dr. Maria Solovey and will be validated by future members of the Srinivas laboratory.

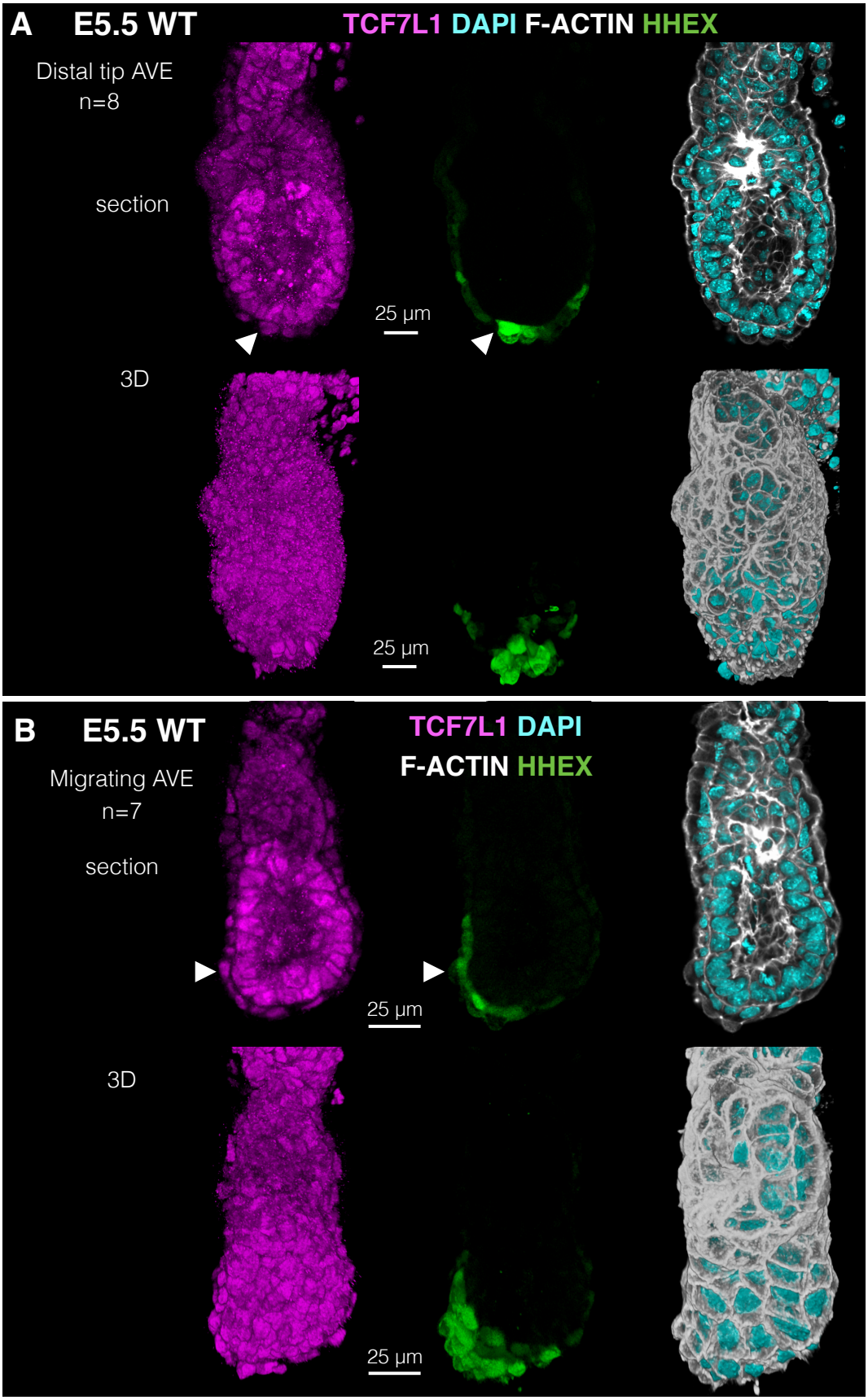


Figure 120. TCF7L1 is slightly upregulated in the AVE compared to the surrounding VE at 5.5
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescence of WT embryos at E5.5 distal tip AVE stage (A) and migrating AVE stage (B) stained for TCF7L1 (magenta); F-ACTIN (gray); DAPI (cyan). HHEX-GFP (green – AVE; arrowhead). Scale bars: 25 µm.

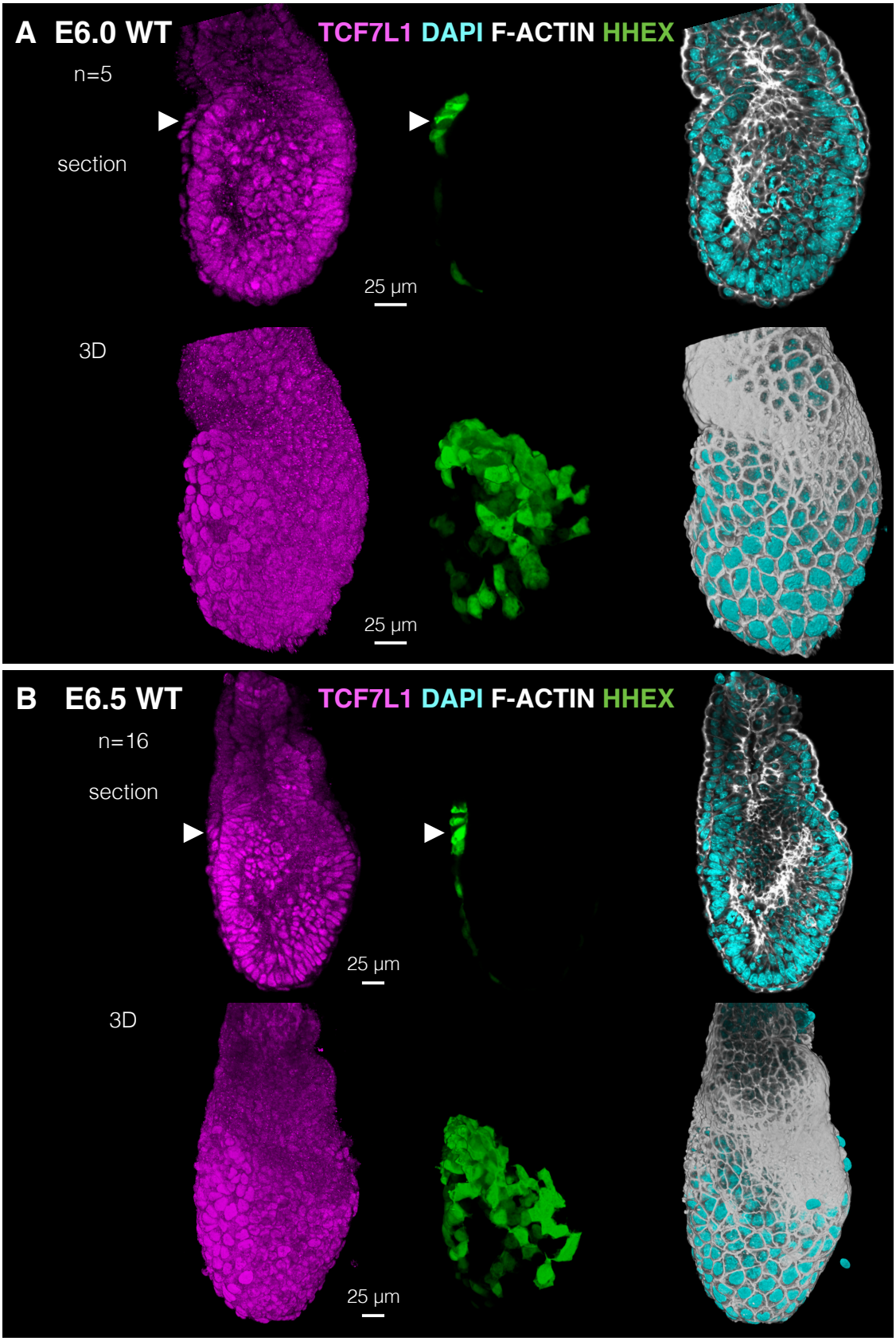


Figure 121. TCF7L1 is upregulated in the AVE compared to the surrounding VE by E6.5
Representative mid-sagittal sections and 3D opacity renderings of whole mount immunofluorescence of WT embryos at E6.0 (A) and E6.5 (B) stained for TCF7L1 (magenta); F-ACTIN (gray); DAPI (cyan). HHEX-GFP (green – AVE; arrowhead). Scale bars: 25 µm.

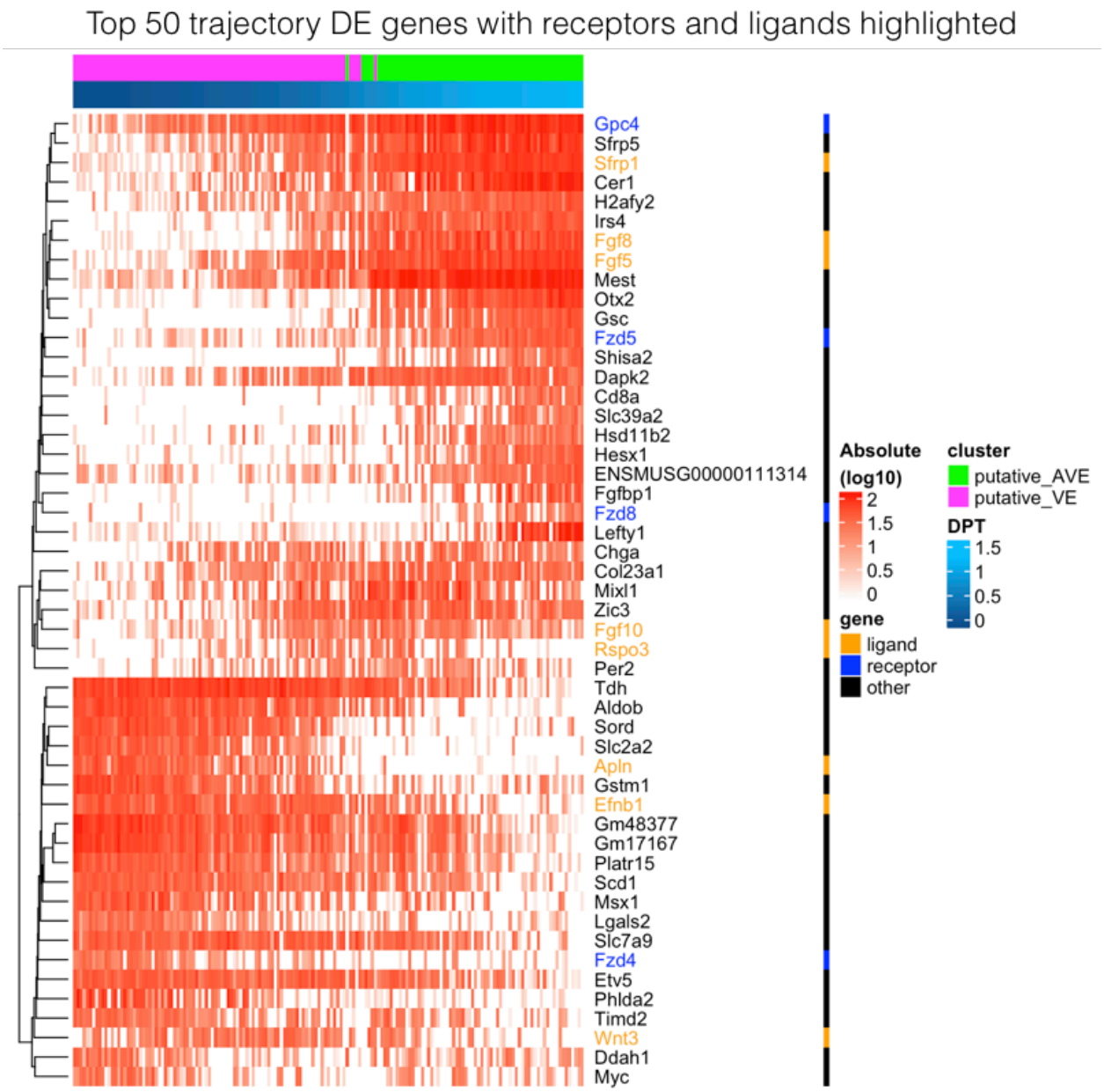


Figure 122. Top 50 trajectory DEGs highlighting receptors (blue) and ligands (orange)
Top 29 genes (*Gpc4* to *Per2*) most likely demarcate the AVE, while bottom 21 genes (*Tdh* to *Myc*) most likely demarcate the VE upon AVE versus VE DEG analysis of the diffusion pseudotime trajectory presented in Figure 115.

ScRNA-seq of E5.5 and E5.75 WT VE

I have also collected single cells for scRNA-seq of the VE at E5.5 and E5.75 using only the C57BL/6 strain (recovery: Table 22). I employed the same vital dye labelling approach to purify for VE cells as that used for the E6.25 WT VE dataset. Initial quality control analysis and unbiased t-SNE by Dr. Jonathan Fiorentino revealed recovery of 214/296 (72.3%) AVE/VE cells (73 of which were AVE) (Figure 123). Analysis of VE cells at E5.5 and E5.75

using trajectory analysis can help uncover: 1. Novel markers of the AVE that are consistent throughout the different stages of AVE migration; 2. DE genes and pathways between the AVE and VE over developmental time that may account for their differential behaviours and functions; and 3. How the epiblast changes via inspecting the direct receptor-ligand interactions between different regions of the VE and epiblast at various stages of AVE anterior shift.

Early post-implantation VE single cell collections (pure C57BL/6 background)

Stage	Number of far red-positive VE cells collected (DAPI-negative live single cells)
E5.5	144
E5.75	152

Table 22. Number of far red-positive live single VE cells collected at E5.5 and E5.75 from pure C57BL/6 background for scRNA-seq

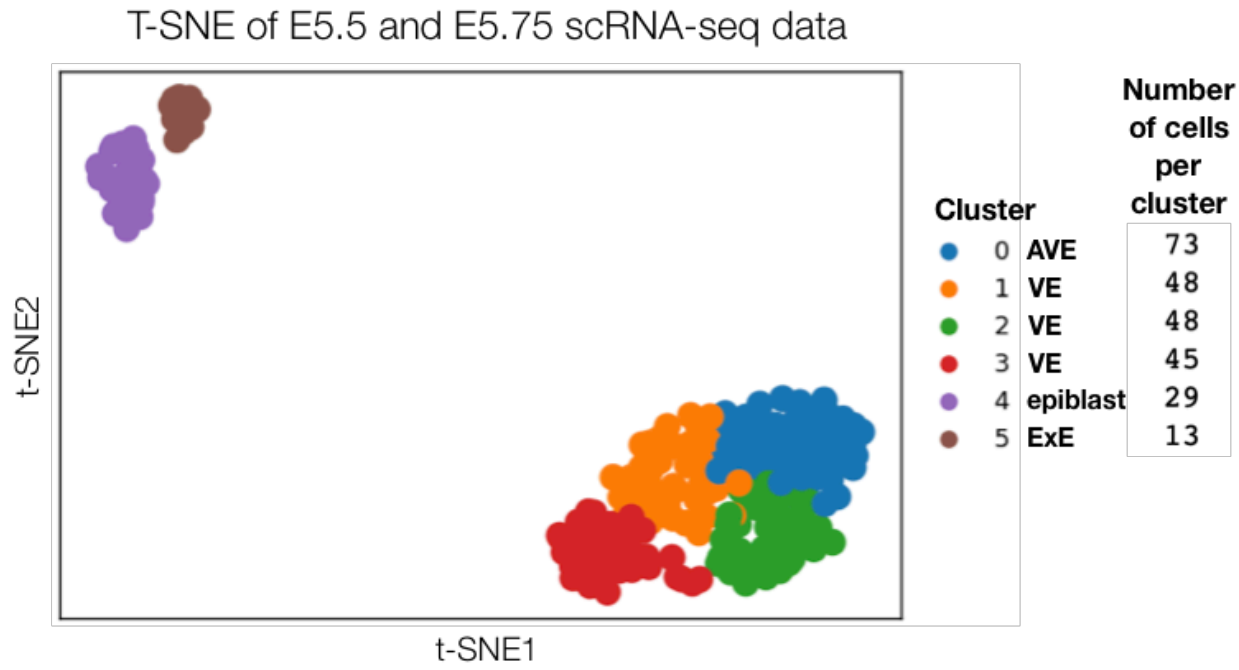


Figure 123. Unbiased t-SNE of E5.5 and E5.75 scRNA-seq data
Clusters were identified based on tissue-specific markers. 73 AVE and 141 VE cells were recovered.

Overall, this section employed a vital dye labelling approach to purify for VE cells, which enabled scRNA-seq of the WT VE at E5.5, E5.75, and E6.25 – different stages of AVE migration. ScRNA-seq DE and trajectory analysis of E6.25 WT VE data elucidated

upregulation of the Jun kinase signalling pathway, as well as identified novel AVE and VE markers that require *in vivo* validation.

6.2.3 *Ets2*-mutant ESC derivation

I derived ESCs from *Ets2*-mutants and control litter mates in order to establish an *in vitro* system for high throughput testing of hypotheses arising from scRNA-seq. I conducted ESC derivations from both pure 129 and mixed *Ets2-Hhex* genetic backgrounds (Table 23). The former should have a more severe defect in differentiating into ExE-like cells, and the latter has the advantage of HHEX-GFP for marking AVE-like cells generated *in vitro*. Mutant ESCs were not significantly different from control in terms of growth rates upon expansion. ESCs were validated by antibody stain to confirm presence of pluripotency factors SOX2 and OCT4 (Figure 124). Mutant ESCs were also shown to lack ETS2, while nuclear ETS2 was present in controls (Figure 124B). No further experiments were conducted with ESCs in this thesis.

Mutant and control ESC colonies derived

Ets2-Hhex (mixed background)		Pure 129 background	
Genotype	**ID and number of vials (2 ml cryogenic) frozen in liquid nitrogen**	**Genotype**	**ID and number of vials (2 ml cryogenic) frozen in liquid nitrogen**
Mutant	X2.2 - 2 vials X2.8 - 12 vials	**Mutant**	E2.3 - 4 vials E2.4 - 4 vials
Het	X2.1 - 3 vials	**Het**	E1.2 - 6 vials E1.4 - 1 vial

Table 23. Number of mutant and heterozygous (het) control ESCs derived from *Ets2*-mutants in both the mixed (HHEX-GFP-positive) and pure 129 genetic backgrounds
Number of vials frozen down in liquid nitrogen are shown. These frozen vials are of low passage number (4 or 5).

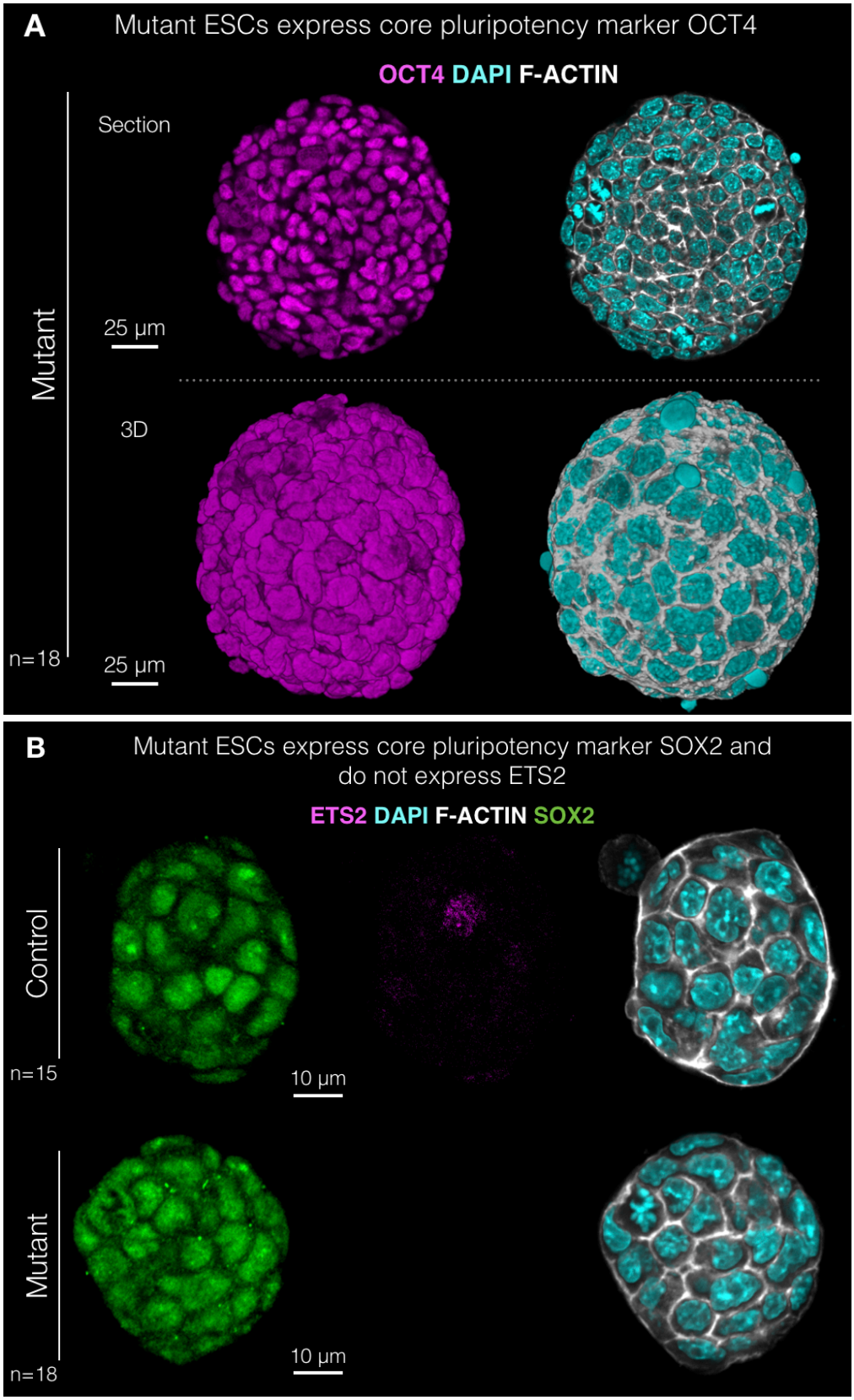


Figure 124. Antibody stain validation of mutant and control ESCs to confirm presence of pluripotency markers OCT4 and SOX2

Representative mid-sagittal sections and 3D opacity renderings of ESCs in (A) mutant only, stained for pluripotency factor OCT4 (magenta), and (B) of control and mutant stained for another pluripotency marker SOX2 (green), as well as ETS2 (magenta) present in controls and absent in mutants. DAPI (cyan) and F-ACTIN (gray). Scale bars: 25 µm and 10 µm as indicated.

6.2.4 Using Pan-Caspase inhibition and scRNA-seq to investigate mechanisms of epiblast cell competition

In collaboration with Professor Tristan Rodriguez at Imperial College London, we investigated the wave of cell death characterised to occur in the epiblast from E5.75 to E6.5 (Di Gregorio et al., 2016; Bowling et al., 2018) using scRNA-seq. This process is thought to involve Myc-dependent cell competition (Claveria et al., 2013). However, detailed comparison of transcriptional profiles between surviving (winner) and dying (loser) cells in the epiblast in native context has not been conducted. The *in vivo* mechanisms governing determination of winner versus loser cell fate in the epiblast from E5.75 to E6.5 remain unclear. To probe how epiblast cells that end up undergoing apoptosis differ from those that progress in development, I used the FMK inhibitor (tested in 3.2.10) to block the apoptotic pathway from E5.75 to E6.5. To assess this wave of apoptosis and validate the efficacy of the FMK inhibitor, I cultured embryos in the FMK inhibitor at different concentrations: 200 μM, 50 μM, and 25 μM, and found the lowest concentration for a robust effect was 25 μM with minimal cell death in the DMSO condition (0.125%). I next performed the Click-TUNEL assay as a readout for presence of apoptosis, and found FMK-treated embryos to have markedly reduced Click-TUNEL signal compared to DMSO-controls (Figure 125).

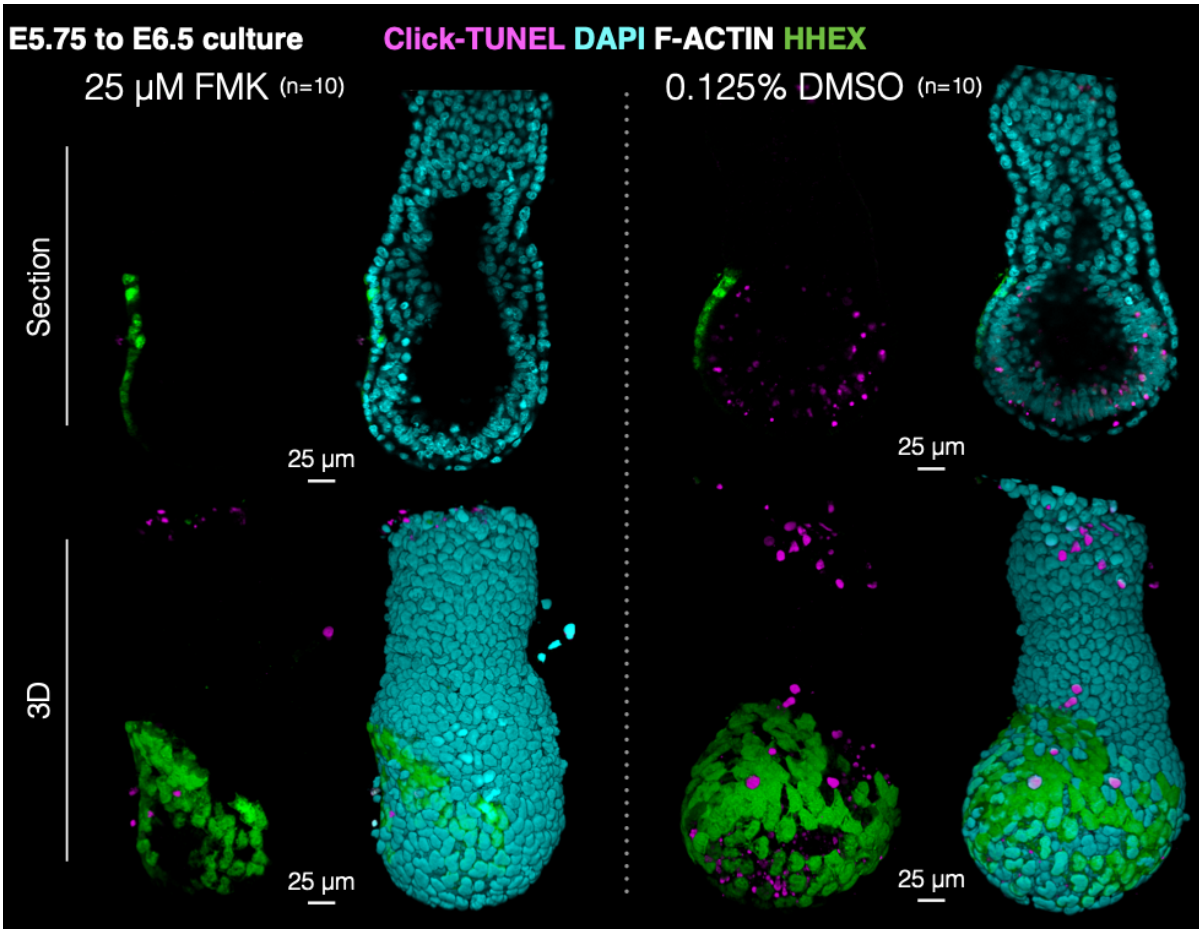


Figure 125. Click-TUNEL assay to show absence of apoptotic cell death upon treatment with FMK Pan-Caspase inhibitor
Representative mid-sagittal sections and 3D opacity renderings of whole mount staining of 25 µM FMK-treated (left) versus 0.125% DMSO control (right). Click-TUNEL (fragmented DNA – magenta); F-ACTIN (gray); HHEX (AVE – green); DAPI (nuclei – cyan). Scale bars: 25 µm.

To quantitatively characterise the change in live and dead cell number in FMK-treated versus DMSO control conditions, I counted epiblast cell number using a hemacytometer. The range of error in automated segmentation of nuclei at E6.5 was deemed to be too high for obtaining an accurate measure, and thus counting all the cells per 'VE-removed' embryonic portion of each embryo was a preferred approach. Prior to hemacytometer examination, cells were treated with Trypan Blue, which labelled dead cells blue. Quantification by hemacytometer revealed a 33.1% increase in live epiblast cell number in FMK-treated versus DMSO controls, suggesting absent cell elimination and therefore more cells in the treated condition (Figure 126A). The number of dead cells was decreased by 51.2% in the FMK-treated condition, supporting successful apoptosis inhibition (Figure 126B). The ratio of live

to dead cells was not significantly different in FMK-treated versus DMSO control (Figure 126C). This shows that total cell number including live and dead cells is unchanged in these culture conditions.

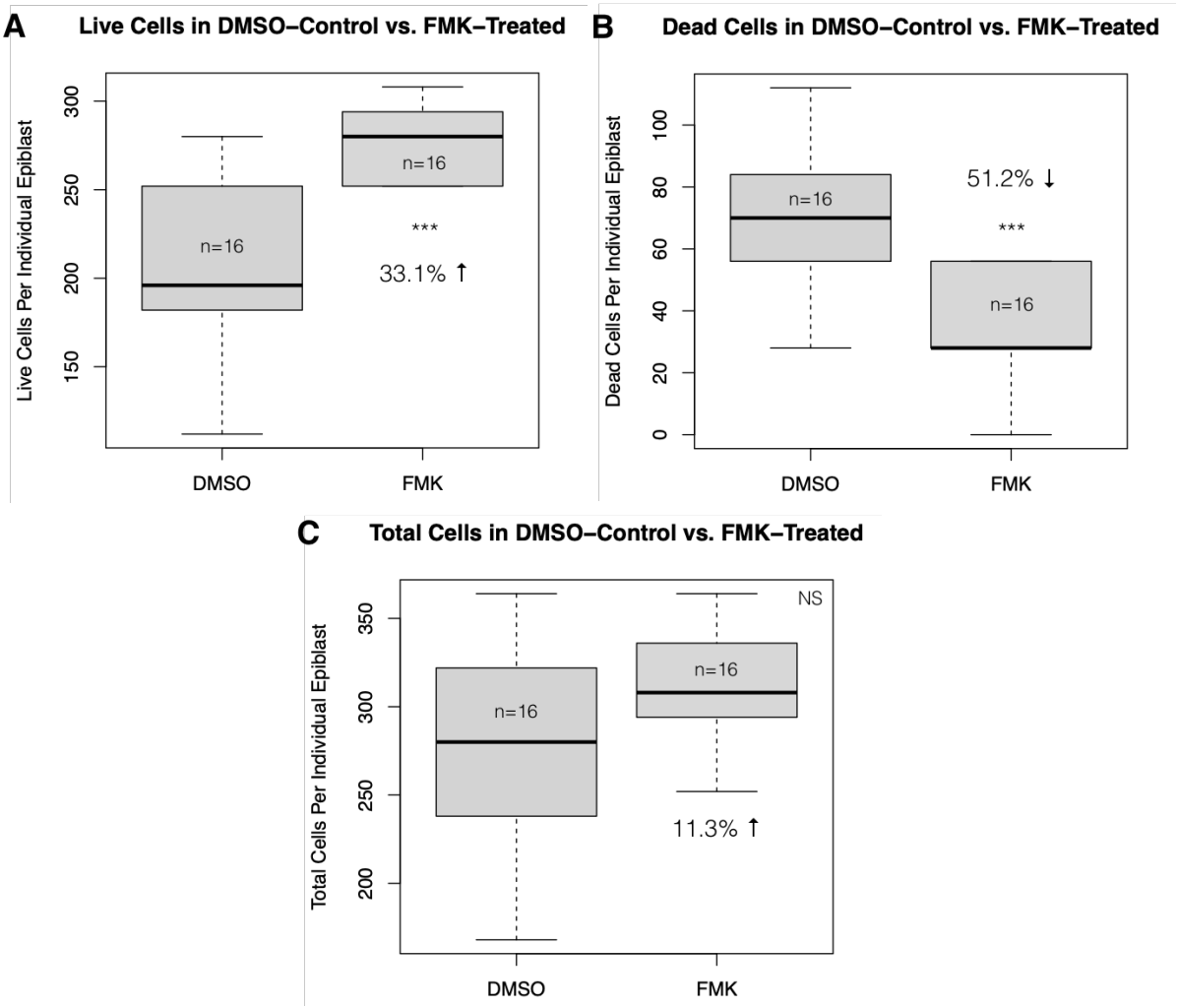
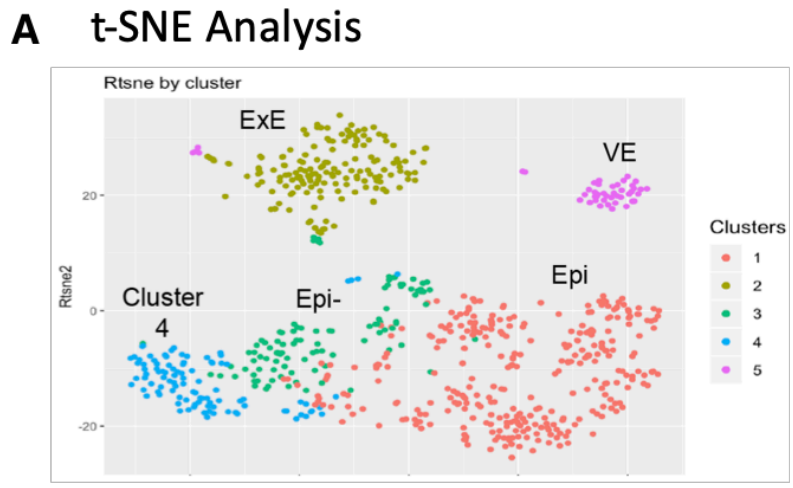


Figure 126. Comparison of epiblast cell number in DMSO control versus FMK-treated
(A) Live cells (B) dead cells (C) total cells. (Statistical test: hierarchical ANOVA. NS = not significant; *** $p < 0.001$).

FMK-treated embryos versus DMSO controls were subsequently analysed by scRNA-seq. Detailed protocol for single cell dissociation and sample preparation is described in Methods 2.8. In brief, embryos were dissected at E5.75 then placed in 25 μM FMK- or 0.125% DMSO-containing media. Embryos were cultured for 16 hours to E5.5 the next day. As embryos in our culture conditions have a tendency to be more proliferative in the EPC compared to regular development, I dissected away the EPC to enrich for epiblast cells. I

subsequently dissociated all embryos per condition separately, then collected cells by FACS. This was conducted at the exact same time with same conditions, but over the course of 4 days. All plates were filled with 50% FMK-treated and 50% DMSO-control cells.

Analysis was conducted by Gabriele Lubatti of Dr. Antonio Scialdone's laboratory. Details of scRNA-seq analysis for this study will not be elaborated on in this thesis. However, the unbiased clustering is of relevance, as it revealed a novel cluster in FMK-treated embryos compared to DMSO controls (cluster 4; Figure 127A). This further supported that FMK inhibition was effective in recovering a population of cells that are not normally recoverable by E6.5 in the epiblast. Cells of the ExE and VE were also attained, verifying that culture conditions and the dissociation protocol did not obliterate any particular population specifically. Analysis of the novel 'cluster 4', presumably recovered from apoptosis inhibition, revealed differential expression of genes and pathways that support 'decreased fitness' of these cells compared to normal epiblast (Figure 127B, C). These include enrichment in apoptosis, cell death, and necrosis GO terms, as well as activated TP53 and inhibited MYC. Remainder of bioinformatics analysis and functional experiments are being carried out by the Scialdone and Rodriguez Laboratories, respectively.



B **IPA analysis**
Functional terms for genes mis-regulated in cluster 4

Functional Category	Predicted Activation	Activation Z-Score	*p*-value
Cell death	Increased	2.334	2.95E-19
Cell cycle	-	1.751	1.75E-18
Necrosis	Increased	2.397	4.73E-18
Apoptosis	Increased	2.088	3.50E-17

C **IPA analysis**
Factors with targets enriched in cluster 4

Upstream regulator	Predicted Activation	Activation Z-Score	*p*-value
TP53	Activated	3.725	4.2E-33
MYC	Inhibited	-4.529	1.18E-28
KRAS	-	1.770	5.07E-19

Figure 127. t-SNE analysis to show presence of a novel cluster upon apoptosis inhibition, suggesting recovery of cells that would otherwise have died by cell competition
(A) Unbiased t-SNE-based clustering where distinct clusters were annotated by known layer marker genes. Identification of a statistically significant distinct and novel 'cluster 4' near the epiblast cluster.
(B, C) Ingenuity Pathway Analysis (IPA) of top misregulated pathways and genes in cluster 4.

6.3 Discussion

ScRNA-seq to investigate the difference between WT AVE versus VE

In this Chapter, I have adapted vital dye labelling to enrich for VE cells for scRNA-seq. As the dissociation approach for HHEX-GFP-positive AVE cell collections had been effective, similar dissociation conditions were used to purify for the VE. To achieve the desired genetic backgrounds for WT VE single cell collections, a vital dye was preferred to avoid complex additional breeding with genetic fluorescent reporters of the VE. The Claret far red dye, as well as the PKH26 red dye, both functioning with the same chemistries, were selected due to their long half-life (>100 days), high signal to noise ratio, and their vital property, i.e. cells labelled remain alive.

Only the WT VE was analysed by scRNA-seq in this thesis. Analysis of E6.25 WT VE (mixed background) revealed novel markers for the AVE and VE. One novel AVE-specific marker TCF7L1, which is a nuclear repressor of WNT/β-catenin signalling (Murphy et al., 2016), was validated by antibody stain. Other markers can be validated by antibody stain, or HCR-ISH. Availability of E6.25 VE and epiblast datasets enables advanced computational analysis of direct receptor-ligand interactions between different regions of the epiblast and VE. This may aid in regionalised analysis of the epiblast where different subdomains can be distinguished with scRNA-seq. Currently, different regions of the epiblast do not cluster separately at E6.5 (Scialdone et al., 2016). Yet, the embryo is known to have molecular asymmetries in the epiblast by this stage (Rossant and Tam, 2009).

Derivation of *Ets2*-mutant ESCs to serve as a high throughput *in vitro* system to probe mechanism

I derived ESCs from control and mutant embryos of both pure 129 and mixed (with HHEX-GFP) genetic backgrounds to provide an *in vitro* tool for future mechanistic investigations. The ability to derive and expand mutant ESCs in the first place lends further support to the finding that the epiblast transcriptome is broadly normal in mutants. Mutant ESCs may be converted into embryoid bodies (EBs; Evans, 2011) to see whether a subset of the outer layer of VE-like cells have HHEX-GFP (or AMOT by antibody stain), and then whether mutant EBs develop a thickened AVE. Subsequently, mutant versus control EBs may be purified for HHEX-GFP by FACS, then analysed by qPCR with top scRNA-seq candidates as a starting point. The next question could be to verify that the AVE thickening defect is independent of epiblast-like cells. This may be interrogated by lentiviral-mediated knockdown of *Ets2* in the outerlayer of VE-like cells in control EBs, or by attempting to differentiate ESCs into AVE-like cells directly. However, differentiation protocols for the latter are currently still

suboptimal for mouse. Other experiments in mutant versus control ESCs to further probe the mechanistic basis of ExE and AVE regulation by ETS2 include:

1. Attempt to differentiate control and mutant ESCs into TSCs. Mutants are likely to fail in this process, supported by the inability to derive TSCs from *Ets2*-mutants (Wen et al., 2007). Using this assay, one could attempt to rescue mutant TSC differentiation first by Pan-Caspase inhibition, then by a more focused candidate approach based on known ETS2 targets that regulate apoptosis, starting with anti-apoptotic factors. This system may provide a platform to elucidate the precise effectors that allow ETS2 to protect the peri-implantation ExE from apoptosis.
2. Use mutant ESCs as a negative control for co-immunoprecipitation and ChIP-seq experiments to assess the interacting partners with ETS2, and regions of DNA occupied by ETS2, respectively, in (rescued) ExE and AVE-like cells *in vitro*.
3. Conduct mass spec in control versus mutant (rescued) ExE and AVE-like cells *in vitro* to investigate differences in protein expression amongst the different genotypes and cell types.

ScRNA-seq to investigate mechanisms of cell competition in the epiblast

Finally, scRNA-seq and culture methods refined in this thesis were used to investigate mechanisms of cell competition in the epiblast. This study demonstrated: 1. The ability to block the wave of apoptosis in the epiblast from E5.75 to E6.5 with FMK Pan-Caspase inhibition; 2. The ability to recover scRNA-seq data following embryo culture for 16 hours in DMSO or FMK inhibitor; and 3. Presence of a unique t-SNE cluster upon 'loser fate inhibition', suggesting that winner and loser cells in the epiblast are indeed transcriptionally distinct. Investigating the differences between the unique cluster and normal epiblast can help uncover the transcriptional basis for cell competition in the E5.75 epiblast.

In summary, this Chapter contributed a WT VE dataset that focuses on the VE versus AVE at different stages of AVE migration; generated *Ets2*-mutant versus control ESCs for ChIP-seq, mass spec, and high throughput *in vitro* studies; and elucidated a transcriptional basis for epiblast cell competition. These resources can be used to uncover further mechanistic details of early mouse embryogenesis to guide the understanding of normal and abnormal development.

7. Closing remarks

The opposing tissue size phenotypes in the *Ets2*^{db1}-mutant reported by Georgiades and Rossant, 2006 motivated this thesis. ETS2 activation downstream of MAPK signalling has been shown to have oncogenic or tumour suppressive effects in different cancers. This means that commonly used MAPK pathway therapeutic agents for cancer treatment may be effective in some cases, but not others (Tetsu and McCormick, 2017). Understanding the mechanism by which ETS2 exerts these opposing effects on tumorigenesis can help direct more targeted cancer treatment strategies. Because tumours are complex involving many different cell types, varying physiological conditions and timings for tumour progression, it is difficult to study the role of ETS2 precisely in a particular cell population in cancer. Therefore, the *Ets2*-mutant in early mouse embryogenesis presents a useful model for investigating how ETS2 can differentially regulate tissue size in a more defined and simplistic *in vivo* system.

I have traced the origins of both the ExE and AVE phenotypes. Though descriptive, these characterisations elucidated the narrow developmental time windows during which the defects arise. This knowledge was then used to select the appropriate stages for interventional and single cell genomics studies. The finding that ETS2 displays nuclear expression in all cell types at peri- and post-implantation means that both ExE and AVE phenotypes may be a result of cell-autonomous roles of ETS2.

For the ExE, I found peri-implantation stages to be the time frame during which ETS2 is required to prevent apoptosis. Different readouts for the progression of the apoptotic pathway detected earlier signs of apoptosis at E4.75 (upregulated Cleaved Caspase-3) and the ultimate DNA fragmentation phase of apoptosis at E5.0 (upregulated TUNEL signal). Though I have rescued the mutant ExE phenotype by apoptosis inhibition at peri-implantation, I have not uncovered the precise anti-apoptotic factor(s) downstream of ETS2 transcriptional activity. This could be done via scRNA-seq of mutant versus control ExE at E5.0, or testing of candidate

anti-apoptotic factors known to be downstream of ETS2 *in vitro*. I have, however, conducted scRNA-seq of the post-implantation reduced mutant ExE. This ExE subpopulation appears to be insensitive to the absence of ETS2, as demonstrated by retention of signature marker gene expression, and normal proliferation and apoptosis. These few mutant ExE cells at the embryonic-extra-embryonic interface appear sufficient for a transcriptionally normal epiblast at E5.5 in a broad sense (i.e. no cluster segregation amongst genotypes by t-SNE). Overt epiblast developmental effects are only manifested by late E6.5 when mutants failed to induce *T*-positive mesoderm, while all littermate controls had *T*-induction. At an earlier stage, the reduced ExE at post-implantation does not affect the stereotypic complete anterior shift of the AVE.

For the mutant AVE, the opposite phenotype of tissue overgrowth is characterised by epithelial stratification and increased cell crowding, similar to an epithelial carcinoma *in situ*. The timing of defect initiation and progression, i.e. burst in proliferation at E5.5 and overt thickening by E6.5, helped direct the stages to capture for scRNA-seq. In this instance, scRNA-seq of the mutant AVE at E5.5 and E6.25 did uncover potential mechanisms for the mutant AVE thickening. Though abnormal in growth, the mutant at both stages did not cluster separately from control by t-SNE, suggesting a broadly normal transcriptome and retention of AVE character in the mutant. Upon assessment of differentially expressed genes (DEGs), we found only a handful at E5.5, and many more at E6.25. GO analysis of E6.25 mutant AVE DEGs revealed WNT to be the most significantly enriched signalling pathway. Assessment of the top DEGs at E5.5 and E6.25 showed upregulation of WNT pathway agonist *Moesin* as the most significant candidate at the earlier stage; followed by a number of WNT upstream agonists and downstream targets upregulated at the later stage. This then served as the basis for WNT agonist pharmacological culture in WT to recapitulate the mutant AVE phenotype. The next step is to rescue the mutant AVE defect with a canonical WNT inhibitor (e.g. XAV939), and

to validate the top candidates misregulated in the mutant AVE from scRNA-seq *in vivo* by antibody stain or ISH, and tissue-specific genetics. Following this, genetic rescue of the mutant AVE phenotype may be attempted by dampening the WNT pathway via misexpression of top downregulated candidates in the VE/AVE. If successful, the same candidates may be used in attempts to rescue intestinal stem cell compartment driven tumorigenesis resultant from tissue-specific *Ets2* deficiency (Munera et al., 2011).

An interesting aspect about the mutant AVE overgrowth phenotype is that ETS2 is expressed throughout the VE, but only the AVE cells are affected. In development, the AVE and VE have different functions and behaviours. I started to further probe these differences in a WT scenario by scRNA-seq of the entire VE at different stages of AVE migration (E5.5, E5.75, E6.25). Trajectory analysis of the E6.25 VE by Dr. Antonio Scialdone's group revealed previously unknown markers and signalling pathways that distinguish the AVE from the rest of the VE. These findings require *in vivo* validation. Combining E5.5, E5.75, and E6.25 WT VE scRNA-seq data can help track the development progression of AVE migration in relation to the surrounding VE and adjacent epiblast with high resolution molecular detail. Once differences between the WT AVE versus VE have been elucidated in greater depth via validation of scRNA-seq findings *in vivo*, we may then look into how the AVE versus VE differ in the *Ets2*-mutant. This will help us understand, in a developmental context, why the AVE is sensitive to the loss of ETS2, while the VE is not. In line with this, ETS2 is thought to exert context-dependent roles via tissue-specific transcriptional cofactors. ETS2 co-IP experiments using relevant *in vitro* systems for the AVE, VE, and ExE can help address this. These same systems can enable elucidation of direct tissue-specific downstream targets of ETS2 by ChIP-seq. Future new technologies that allow us to answer these questions *in vivo* will be highly beneficial for understanding more precisely how transcription factor ETS2 can promote or suppress early embryonic tissue growth.

In broader context, perturbations of the MAPK signalling cascade are frequent in cancer, and MEK inhibition therapeutic strategies have shown inconsistent efficacy (Tetsu and McCormick, 2017). ETS2 activation downstream of MAPK signalling could explain why MEK inhibition may be effective in certain cancers but not others. Mechanistic insights from the early embryo may be applied to examine oncogenic versus tumour suppressive roles of ETS2 in cancer. While transcriptional signatures in cancer are likely to differ from the embryo, in particular with the immune response in cancer, underlying transcriptional co-factors and signalling pathways linked to ETS2 may be conserved. The stereotypic early embryo system can provide a framework of ETS2 function, which may be difficult to uncover directly from cancer systems that involve multiple mutations, cell types, and stages of progression. Therefore, further probing the mechanistic bases for the opposing tissue size phenotypes of the *Ets2*-mutant in early mouse embryogenesis can be invaluable for understanding ETS2 in disease.

8. References

Ang, S. L. and Constam, D. B. (2004). A gene network establishing polarity in the early mouse embryo. *Semin Cell Dev Biol* **15**, 555-561.

Anjum, R. and Blenis, J. (2008). The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* **9**, 747-758.

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J. K. and Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. *Proc Natl Acad Sci U S A* **95**, 5082-5087.

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.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29.

Auman, H. J., Nottoli, T., Lakiza, O., Winger, Q., Donaldson, S. and Williams, T. (2002). Transcription factor AP-2gamma is essential in the extra-embryonic lineages for early postimplantation development. *Development* **129**, 2733-2747.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-140.

Batchelor, C. L., Woodward, A. M. and Crouch, D. H. (2004). Nuclear ERM (ezrin, radixin, moesin) proteins: regulation by cell density and nuclear import. *Exp Cell Res* **296**, 208-222.

Beck, S., Le Good, J. A., Guzman, M., Ben Haim, N., Roy, K., Beermann, F. and Constam, D. B. (2002). Extraembryonic proteases regulate Nodal signalling during gastrulation. *Nat Cell Biol* **4**, 981-985.

Bedzhov, I. and Zernicka-Goetz, M. (2014). Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032-1044.

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.

Bielinska, M., Narita, N. and Wilson, D. B. (1999). Distinct roles for visceral endoderm during embryonic mouse development. *Int J Dev Biol* **43**, 183-205.

Biggins, J. S., Royer, C., Watanabe, T. and Srinivas, S. (2015). Towards understanding the roles of position and geometry on cell fate decisions during preimplantation development. *Semin Cell Dev Biol* **47-48**, 74-79.

Boroviak, T., Loos, R., Bertone, P., Smith, A. and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat Cell Biol* **16**, 516-528.

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A. and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev Cell* **35**, 366-382.

Boulukos, K. E., Pognonec, P., Rabault, B., Begue, A. and Ghysdael, J. (1989). Definition of an Ets1 protein domain required for nuclear localization in cells and DNA-binding activity in vitro. *Mol Cell Biol* **9**, 5718-5721.

Bowling, S., Di Gregorio, A., Sancho, M., Pozzi, S., Aarts, M., Signore, M., Schneider, M. D., Martinez-Barbera, J. P., Gil, J. and Rodriguez, T. A. (2018). Author Correction: P53 and mTOR signalling determine fitness selection through cell competition during early mouse embryonic development. *Nat Commun* **9**, 3123.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956.

Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M. and Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**, 3710-3715.

Bradley, A., Evans, M., Kaufman, M. H. and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-256.

Brennecke, P., Anders, S., Kim, J. K., Kolodziejczyk, A. A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S. A., Marioni, J. C., et al. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* **10**, 1093-1095.

Brison, D. R. and Schultz, R. M. (1998). Increased incidence of apoptosis in transforming growth factor alpha-deficient mouse blastocysts. *Biol Reprod* **59**, 136-144.

Brown, N. H. (2011). Extracellular matrix in development: insights from mechanisms conserved between invertebrates and vertebrates. *Cold Spring Harb Perspect Biol* **3**.

Bryja, V., Bonilla, S. and Arenas, E. (2006). Derivation of mouse embryonic stem cells. *Nat Protoc* **1**, 2082-2087.

Buettner, F., Natarajan, K. N., Casale, F. P., Proserpio, V., Scialdone, A., Theis, F. J., Teichmann, S. A., Marioni, J. C. and Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* **33**, 155-160.

Carroll, D. (2014). Genome engineering with targetable nucleases. *Annu Rev Biochem* **83**, 409-439.

Cavey, M. and Lecuit, T. (2009). Molecular bases of cell-cell junctions stability and dynamics. *Cold Spring Harb Perspect Biol* **1**, a002998.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655.

Chazaud, C. and Rossant, J. (2006). Disruption of early proximodistal patterning and AVE formation in Apc mutants. *Development* **133**, 3379-3387.

Chazaud, C., Yamanaka, Y., Pawson, T. and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell* **10**, 615-624.

Chen, J., Suo, S., Tam, P. P., Han, J. J., Peng, G. and Jing, N. (2017). Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. *Nat Protoc* **12**, 566-580.

Cheng, S., Pei, Y., He, L., Peng, G., Reinius, B., Tam, P. P. L., Jing, N. and Deng, Q. (2019). Single-Cell RNA-Seq Reveals Cellular Heterogeneity of Pluripotency Transition and X Chromosome Dynamics during Early Mouse Development. *Cell Rep* **26**, 2593-2607.e2593.

Chew, J. L., Loh, Y. H., Zhang, W., Chen, X., Tam, W. L., Yeap, L. S., Li, P., Ang, Y. S., Lim, B., Robson, P., et al. (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* **25**, 6031-6046.

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A. and Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**.

Chu, G. C., Dunn, N. R., Anderson, D. C., Oxburgh, L. and Robertson, E. J. (2004). Differential requirements for Smad4 in TGFbeta-dependent patterning of the early mouse embryo. *Development* **131**, 3501-3512.

Claveria, C., Giovinazzo, G., Sierra, R. and Torres, M. (2013). Myc-driven endogenous cell competition in the early mammalian embryo. *Nature* **500**, 39-44.

Clements, M., Pernaute, B., Vella, F. and Rodriguez, T. A. (2011). Crosstalk between Nodal/activin and MAPK p38 signaling is essential for anterior-posterior axis specification. *Curr Biol* **21**, 1289-1295.

Clevers, H. (2006). Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469-480.

Corson, L. B., Yamanaka, Y., Lai, K. M. and Rossant, J. (2003). Spatial and temporal patterns of ERK signaling during mouse embryogenesis. *Development* **130**, 4527-4537.

Czabotar, P. E., Lessene, G., Strasser, A. and Adams, J. M. (2014). Control of apoptosis by the BCL-2 protein family: implications for physiology and therapy. *Nat Rev Mol Cell Biol* **15**, 49-63.

Davidson, A. J. and Insall, R. H. (2011). Actin-based motility: WAVE regulatory complex structure reopens old SCARs. *Curr Biol* **21**, R66-68.

Davidson, B., Shi, W., Beh, J., Christiaen, L. and Levine, M. (2006). FGF signaling delineates the cardiac progenitor field in the simple chordate, Ciona intestinalis. *Genes Dev* **20**, 2728-2738.

Degnan, B. M., Degnan, S. M., Naganuma, T. and Morse, D. E. (1993). The ets multigene family is conserved throughout the Metazoa. *Nucleic Acids Res* **21**, 3479-3484.

Deng, Q., Ramskold, D., Reinius, B. and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193-196.

Di Gregorio, A., Bowling, S. and Rodriguez, T. A. (2016). Cell Competition and Its Role in the Regulation of Cell Fitness from Development to Cancer. *Dev Cell* **38**, 621-634.

Dietrich, J. E. and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. *Development* **134**, 4219-4231.

Ding, J., Yang, L., Yan, Y. T., Chen, A., Desai, N., Wynshaw-Boris, A. and Shen, M. M. (1998). Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature* **395**, 702-707.

Dittmer, J. and Nordheim, A. (1998). Ets transcription factors and human disease. *Biochim Biophys Acta* **1377**, F1-11.
Doerr, A. (2019). Single-cell proteomics. *Nature Methods* **16**, 20-20.
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.
Donnison, M., Broadhurst, R. and Pfeffer, P. L. (2015). Elf5 and Ets2 maintain the mouse extraembryonic ectoderm in a dosage dependent synergistic manner. *Dev Biol* **397**, 77-88.
Drukker, M., Tang, C., Ardehali, R., Rinkevich, Y., Seita, J., Lee, A. S., Mosley, A. R., Weissman, I. L. and Soen, Y. (2012). Isolation of primitive endoderm, mesoderm, vascular endothelial and trophoblast progenitors from human pluripotent stem cells. *Nat Biotechnol* **30**, 531-542.
Dunn, N. R., Vincent, S. D., Oxburgh, L., Robertson, E. J. and Bikoff, E. K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development* **131**, 1717-1728.
Dupont, S. (2016). Role of YAP/TAZ in cell-matrix adhesion-mediated signalling and mechanotransduction. *Exp Cell Res* **343**, 42-53.
Eden, E., Navon, R., Steinfeld, I., Lipson, D. and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48.
Elmore, S. (2007). Apoptosis: a review of programmed cell death. *Toxicol Pathol* **35**, 495-516.
Episkopou, V., Maeda, S., Nishiguchi, S., Shimada, K., Gaitanaris, G. A., Gottesman, M. E. and Robertson, E. J. (1993). Disruption of the transthyretin gene results in mice with depressed levels of plasma retinol and thyroid hormone. *Proc Natl Acad Sci U S A* **90**, 2375-2379.
Escalante-Alcalde, D., Hernandez, L., Le Stunff, H., Maeda, R., Lee, H. S., Jr Gang, C., Sciorra, V. A., Daar, I., Spiegel, S., Morris, A. J., et al. (2003). The lipid phosphatase LPP3 regulates extra-embryonic vasculogenesis and axis patterning. *Development* **130**, 4623-4637.
Evans, M. (2011). Discovering pluripotency: 30 years of mouse embryonic stem cells. *Nat Rev Mol Cell Biol* **12**, 680-686.
Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B., et al. (2018). The Reactome Pathway Knowledgebase. *Nucleic Acids Res* **46**, D649-d655.
Fan, H. C., Fu, G. K. and Fodor, S. P. (2015). Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science* **347**, 1258367.
Fenrick, R., Amann, J. M., Lutterbach, B., Wang, L., Westendorf, J. J., Downing, J. R. and Hiebert, S. W. (1999). Both TEL and AML-1 contribute repression domains to the t(12;21) fusion protein. *Mol Cell Biol* **19**, 6566-6574.
Fietz, S. A., Kelava, I., Vogt, J., Wilsch-Brauninger, M., Stenzel, D., Fish, J. L., Corbeil, D., Riehn, A., Distler, W., Nitsch, R., et al. (2010). OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nat Neurosci* **13**, 690-699.
Fletcher, D. A. and Mullins, R. D. (2010). Cell mechanics and the cytoskeleton. *Nature* **463**, 485-492.
Foulds, C. E., Nelson, M. L., Blaszczak, A. G. and Graves, B. J. (2004). Ras/mitogen-activated protein kinase signaling activates Ets-1 and Ets-2 by CBP/p300 recruitment. *Mol Cell Biol* **24**, 10954-10964.
Fowles, L. F., Martin, M. L., Nelsen, L., Stacey, K. J., Redd, D., Clark, Y. M., Nagamine, Y., McMahon, M., Hume, D. A. and Ostrowski, M. C. (1998). Persistent activation of mitogen-activated protein kinases p42 and p44 and ets-2 phosphorylation in response to colony-stimulating factor 1/c-fms signaling. *Mol Cell Biol* **18**, 5148-5156.
Frisch, S. M. and Francis, H. (1994). Disruption of epithelial cell-matrix interactions induces apoptosis. *J Cell Biol* **124**, 619-626.
Frumkin, D., Wasserstrom, A., Itzkovitz, S., Harmelin, A., Rechavi, G. and Shapiro, E. (2008). Amplification of multiple genomic loci from single cells isolated by laser micro-dissection of tissues. *BMC Biotechnol* **8**, 17.
Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J. and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* **16**, 784-789.
Garcia, M. I., Ghiani, M., Lefort, A., Libert, F., Strollo, S. and Vassart, G. (2009). LGR5 deficiency deregulates Wnt signaling and leads to precocious Paneth cell differentiation in the fetal intestine. *Dev Biol* **331**, 58-67.
Georgescu, M. M. (2010). PTEN Tumor Suppressor Network in PI3K-Akt Pathway Control. *Genes Cancer* **1**, 1170-1177.

Georgiades, P. and Rossant, J. (2006). Ets2 is necessary in trophoblast for normal embryonic anteroposterior axis development. *Development* **133**, 1059-1068.
Gillies, T. E. and Cabernard, C. (2011). Cell division orientation in animals. *Curr Biol* **21**, R599-609.
Grasso, C. S., Wu, Y. M., Robinson, D. R., Cao, X., Dhanasekaran, S. M., Khan, A. P., Quist, M. J., Jing, X., Lonigro, R. J., Brenner, J. C., et al. (2012). The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239-243.
Gu, B., Posfai, E. and Rossant, J. (2018). Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol* **36**, 632-637.
Gu, Z., Nomura, M., Simpson, B. B., Lei, H., Feijen, A., van den Eijnden-van Raaij, J., Donahoe, P. K. and Li, E. (1998). The type I activin receptor ActRIB is required for egg cylinder organization and gastrulation in the mouse. *Genes Dev* **12**, 844-857.
Haghverdi, L., Buttner, M., Wolf, F. A., Buettner, F. and Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods* **13**, 845-848.
Hayashi, T., Shibata, N., Okumura, R., Kudome, T., Nishimura, O., Tarui, H. and Agata, K. (2010). Single-cell gene profiling of planarian stem cells using fluorescent activated cell sorting and its "index sorting" function for stem cell research. *Dev Growth Differ* **52**, 131-144.
Henkel, G. W., McKercher, S. R., Yamamoto, H., Anderson, K. L., Oshima, R. G. and Maki, R. A. (1996). PU.1 but not ets-2 is essential for macrophage development from embryonic stem cells. *Blood* **88**, 2917-2926.
Hermiston, M. L. and Gordon, J. I. (1995). In vivo analysis of cadherin function in the mouse intestinal epithelium: essential roles in adhesion, maintenance of differentiation, and regulation of programmed cell death. *J Cell Biol* **129**, 489-506.
Hiramatsu, R., Matsuoka, T., Kimura-Yoshida, C., Han, S. W., Mochida, K., Adachi, T., Takayama, S. and Matsuo, I. (2013). External mechanical cues trigger the establishment of the anterior-posterior axis in early mouse embryos. *Dev Cell* **27**, 131-144.
Hoffman, J. A., Wu, C. I. and Merrill, B. J. (2013). Tcf7l1 prepares epiblast cells in the gastrulating mouse embryo for lineage specification. *Development* **140**, 1665-1675.
Hollenhorst, P. C., Jones, D. A. and Graves, B. J. (2004). Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res* **32**, 5693-5702.
Hollenhorst, P. C., McIntosh, L. P. and Graves, B. J. (2011). Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem* **80**, 437-471.
Home, P., Ray, S., Dutta, D., Bronshteyn, I., Larson, M. and Paul, S. (2009). GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. *J Biol Chem* **284**, 28729-28737.
Hong, S. H., Lee, J. H., Lee, J. B., Ji, J. and Bhatia, M. (2011). ID1 and ID3 represent conserved negative regulators of human embryonic and induced pluripotent stem cell hematopoiesis. *J Cell Sci* **124**, 1445-1452.
Hoshino, H., Shioi, G. and Aizawa, S. (2015). AVE protein expression and visceral endoderm cell behavior during anterior-posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. *Dev Biol* **402**, 175-191.
Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., et al. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* **35**, W169-175.
Huang, S. M., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614-620.
Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C. and Birchmeier, W. (2000). Requirement for beta-catenin in anterior-posterior axis formation in mice. *J Cell Biol* **148**, 567-578.
Ichim, G. and Tait, S. W. (2016). A fate worse than death: apoptosis as an oncogenic process. *Nat Rev Cancer* **16**, 539-548.
Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lonnerberg, P. and Linnarsson, S. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* **11**, 163-166.
Islas, J. F., Liu, Y., Weng, K. C., Robertson, M. J., Zhang, S., Prejusa, A., Harger, J., Tikhomirova, D., Chopra, M., Iyer, D., et al. (2012). Transcription factors ETS2 and MESP1 transdifferentiate human dermal fibroblasts into cardiac progenitors. *Proc Natl Acad Sci U S A* **109**, 13016-13021.
Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., et al. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776-779.

Javali, A., Misra, A., Leonavicius, K., Acharyya, D., Vyas, B. and Sambasivan, R. (2017). Co-expression of Tbx6 and Sox2 identifies a novel transient neuromesoderm progenitor cell state. *Development* **144**, 4522-4529.

Ji, R. R., Ott, K. H., Yordanova, R. and Bruccoleri, R. E. (2011). FDR-FET: an optimizing gene set enrichment analysis method. *Adv Appl Bioinform Chem* **4**, 37-42.

Jiang, L., Schlesinger, F., Davis, C. A., Zhang, Y., Li, R., Salit, M., Gingeras, T. R. and Oliver, B. (2011). Synthetic spike-in standards for RNA-seq experiments. *Genome Res* **21**, 1543-1551.

Joswig, A., Gabriel, H. D., Kibschull, M. and Winterhager, E. (2003). Apoptosis in uterine epithelium and decidua in response to implantation: evidence for two different pathways. *Reprod Biol Endocrinol* **1**, 44.

Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P., et al. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* **129**, 2367-2379.

Karim, F. D., Urness, L. D., Thummel, C. S., Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C., Maki, R. A., Gunther, C. V., Nye, J. A., et al. (1990). The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev* **4**, 1451-1453.

Kattman, S. J., Witty, A. D., Gagliardi, M., Dubois, N. C., Niapour, M., Hotta, A., Ellis, J. and Keller, G. (2011). Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* **8**, 228-240.

Kawamura, M., Toiyama, Y., Tanaka, K., Saigusa, S., Okugawa, Y., Hiro, J., Uchida, K., Mohri, Y., Inoue, Y. and Kusunoki, M. (2012). CXCL5, a promoter of cell proliferation, migration and invasion, is a novel serum prognostic marker in patients with colorectal cancer. *Eur J Cancer* **48**, 2244-2251.

Kawano, Y. and Kypta, R. (2003). Secreted antagonists of the Wnt signalling pathway. *J Cell Sci* **116**, 2627-2634.

Keays, K. M., Owens, G. P., Ritchie, A. M., Gilden, D. H. and Burgoon, M. P. (2005). Laser capture microdissection and single-cell RT-PCR without RNA purification. *J Immunol Methods* **302**, 90-98.

Kemp, C. R., Willems, E., Wawrzak, D., Hendrickx, M., Agbor Agbor, T. and Leyns, L. (2007). Expression of Frizzled5, Frizzled7, and Frizzled10 during early mouse development and interactions with canonical Wnt signaling. *Dev Dyn* **236**, 2011-2019.

Kerr, J. F., Wyllie, A. H. and Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* **26**, 239-257.

Kim, J. Y., Jeong, H. S., Chung, T., Kim, M., Lee, J. H., Jung, W. H. and Koo, J. S. (2017). The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: A comparative study with Ki67. *Oncotarget* **8**, 65064-65076.

Kimura-Yoshida, C., Nakano, H., Okamura, D., Nakao, K., Yonemura, S., Belo, J. A., Aizawa, S., Matsui, Y. and Matsuo, I. (2005). Canonical Wnt signaling and its antagonist regulate anterior-posterior axis polarization by guiding cell migration in mouse visceral endoderm. *Dev Cell* **9**, 639-650.

Klambt, C. (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**, 163-176.

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. and Teichmann, S. A. (2015). The technology and biology of single-cell RNA sequencing. *Mol Cell* **58**, 610-620.

Korinek, V., Barker, N., Morin, P. J., van Wichen, D., de Weger, R., Kinzler, K. W., Vogelstein, B. and Clevers, H. (1997). Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**, 1784-1787.

Kubaczka, C., Senner, C. E., Cierlitza, M., Arauzo-Bravo, M. J., Kuckenberg, P., Peitz, M., Hemberger, M. and Schorle, H. (2015). Direct Induction of Trophoblast Stem Cells from Murine Fibroblasts. *Cell Stem Cell* **17**, 557-568.

Kunath, T., Arnaud, D., Uy, G. D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R. L., Avner, P. and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132**, 1649-1661.

Kunath, T., Saba-El-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S. and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**, 2895-2902.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K. D., Yamada, R. G., Ueda, H. R. and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res* **34**, e42.

Kwon, G. S. and Hadjantonakis, A. K. (2009). Transthyretin mouse transgenes direct RFP expression or Cre-mediated recombination throughout the visceral endoderm. *Genesis* **47**, 447-455.

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.

Lancaster, M. A. and Knoblich, J. A. (2012). Spindle orientation in mammalian cerebral cortical development. *Curr Opin Neurobiol* **22**, 737-746.
Latos, P. A. and Hemberger, M. (2016). From the stem of the placental tree: trophoblast stem cells and their progeny. *Development* **143**, 3650-3660.
Lawson, K. A. and Pedersen, R. A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development* **101**, 627-652.
Lawson, K. A. and Pedersen, R. A. (1992). Clonal analysis of cell fate during gastrulation and early neurulation in the mouse. *Ciba Found Symp* **165**, 3-21; discussion 21-26.
Le Gallic, L., Sgouras, D., Beal, G., Jr. and Mavrothalassitis, G. (1999). Transcriptional repressor ERF is a Ras/mitogen-activated protein kinase target that regulates cellular proliferation. *Mol Cell Biol* **19**, 4121-4133.
Lechler, T. and Fuchs, E. (2005). Asymmetric cell divisions promote stratification and differentiation of mammalian skin. *Nature* **437**, 275-280.
Lee, H. O. and Norden, C. (2013). Mechanisms controlling arrangements and movements of nuclei in pseudostratified epithelia. *Trends Cell Biol* **23**, 141-150.
Lee, J. E., Pintar, J. and Efstratiadis, A. (1990). Pattern of the insulin-like growth factor II gene expression during early mouse embryogenesis. *Development* **110**, 151-159.
Lelievre, E., Lionneton, F., Soncin, F. and Vandenbunder, B. (2001). The Ets family contains transcriptional activators and repressors involved in angiogenesis. *Int J Biochem Cell Biol* **33**, 391-407.
Li, H., Zhu, H., Xu, C. J. and Yuan, J. (1998). Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell* **94**, 491-501.
Lim, K. H., Huang, H., Pralle, A. and Park, S. (2013). Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. *Biotechnol Bioeng* **110**, 57-67.
Liu, Y. and Schwartz, R. J. (2013). Transient Mesp1 expression: a driver of cardiac cell fate determination. *Transcription* **4**, 92-96.
Liu, Z., Chen, O., Wall, J. B. J., Zheng, M., Zhou, Y., Wang, L., Ruth Vaseghi, H., Qian, L. and Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **7**, 2193.
Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.
Llambi, F., Moldoveanu, T., Tait, S. W., Bouchier-Hayes, L., Temirov, J., McCormick, L. L., Dillon, C. P. and Green, D. R. (2011). A unified model of mammalian BCL-2 protein family interactions at the mitochondria. *Mol Cell* **44**, 517-531.
Lu, C. C. and Robertson, E. J. (2004). Multiple roles for Nodal in the epiblast of the mouse embryo in the establishment of anterior-posterior patterning. *Dev Biol* **273**, 149-159.
Luo, X., Budihardjo, I., Zou, H., Slaughter, C. and Wang, X. (1998). Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors. *Cell* **94**, 481-490.
Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, 7634-7638.
Martinez, L. A. (2016). Mutant p53 and ETS2, a Tale of Reciprocity. *Front Oncol* **6**, 35.
Mattila, P. K. and Lappalainen, P. (2008). Filopodia: molecular architecture and cellular functions. *Nat Rev Mol Cell Biol* **9**, 446-454.
McCarthy, S. A., Chen, D., Yang, B. S., Garcia Ramirez, J. J., Cherwinski, H., Chen, X. R., Klagsbrun, M., Hauser, C. A., Ostrowski, M. C. and McMahon, M. (1997). Rapid phosphorylation of Ets-2 accompanies mitogen-activated protein kinase activation and the induction of heparin-binding epidermal growth factor gene expression by oncogenic Raf-1. *Mol Cell Biol* **17**, 2401-2412.
Mesnard, D., Filipe, M., Belo, J. A. and Zernicka-Goetz, M. (2004). The anterior-posterior axis emerges respecting the morphology of the mouse embryo that changes and aligns with the uterus before gastrulation. *Curr Biol* **14**, 184-196.
Mi, H., Muruganujan, A., Casagrande, J. T. and Thomas, P. D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* **8**, 1551-1566.
Mihajlovic, A. I. and Bruce, A. W. (2017). The first cell-fate decision of mouse preimplantation embryo development: integrating cell position and polarity. *Open Biol* **7**.
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642.
Miura, T., Mattson, M. P. and Rao, M. S. (2004). Cellular lifespan and senescence signaling in embryonic stem cells. *Aging Cell* **3**, 333-343.

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017). Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Rep* **20**, 1215-1228.

Morin, X. and Bellaiche, Y. (2011). Mitotic spindle orientation in asymmetric and symmetric cell divisions during animal development. *Dev Cell* **21**, 102-119.

Munera, J., Cecena, G., Jedlicka, P., Wankell, M. and Oshima, R. G. (2011). Ets2 regulates colonic stem cells and sensitivity to tumorigenesis. *Stem Cells* **29**, 430-439.

Muroyama, A. and Lechler, T. (2012). Polarity and stratification of the epidermis. *Semin Cell Dev Biol* **23**, 890-896.

Murphy, M., Chatterjee, S. S., Jain, S., Katari, M. and DasGupta, R. (2016). TCF7L1 Modulates Colorectal Cancer Growth by Inhibiting Expression of the Tumor-Suppressor Gene EPHB3. *Sci Rep* **6**, 28299.

Nancy, P., Tagliani, E., Tay, C. S., Asp, P., Levy, D. E. and Erlebacher, A. (2012). Chemokine gene silencing in decidual stromal cells limits T cell access to the maternal-fetal interface. *Science* **336**, 1317-1321.

Nichols, J. and Ying, Q. L. (2006). Derivation and propagation of embryonic stem cells in serum- and feeder-free culture. *Methods Mol Biol* **329**, 91-98.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H. and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K. and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech Dev* **125**, 270-283.

Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S. and Hedge, P. (1991). Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **253**, 665-669.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R. and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917-929.

Nowotschin, S., Setty, M., Kuo, Y. Y., Liu, V., Garg, V., Sharma, R., Simon, C. S., Saiz, N., Gardner, R., Boutet, S. C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature*.

Nunn, M. F., Seeburg, P. H., Moscovici, C. and Duesberg, P. H. (1983). Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* **306**, 391-395.

Oltersdorf, T., Elmore, S. W., Shoemaker, A. R., Armstrong, R. C., Augeri, D. J., Belli, B. A., Bruncko, M., Deckwerth, T. L., Dinges, J., Hajduk, P. J., et al. (2005). An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677-681.

Omelchenko, T., Rabadan, M. A., Hernandez-Martinez, R., Grego-Bessa, J., Anderson, K. V. and Hall, A. (2014). beta-Pix directs collective migration of anterior visceral endoderm cells in the early mouse embryo. *Genes Dev* **28**, 2764-2777.

Orlando, V. (2000). Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem Sci* **25**, 99-104.

Ovcharenko, I., Loots, G. G., Giardine, B. M., Hou, M., Ma, J., Hardison, R. C., Stubbs, L. and Miller, W. (2005). Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome Res* **15**, 184-194.

Parr, E. L., Tung, H. N. and Parr, M. B. (1987). Apoptosis as the mode of uterine epithelial cell death during embryo implantation in mice and rats. *Biol Reprod* **36**, 211-225.

Patel, A. P., Tirosh, I., Trombetta, J. J., Shalek, A. K., Gillespie, S. M., Wakimoto, H., Cahill, D. P., Nahed, B. V., Curry, W. T., Martuza, R. L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401.

Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G., et al. (2016). Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-gastrula Mouse Embryo. *Dev Cell* **36**, 681-697.

Perea-Gomez, A., Camus, A., Moreau, A., Grieve, K., Moneron, G., Dubois, A., Cibert, C. and Collignon, J. (2004). Initiation of gastrulation in the mouse embryo is preceded by an apparent shift in the orientation of the anterior-posterior axis. *Curr Biol* **14**, 197-207.

Perea-Gomez, A., Meilhac, S. M., Piotrowska-Nitsche, K., Gray, D., Collignon, J. and Zernicka-Goetz, M. (2007). Regionalization of the mouse visceral endoderm as the blastocyst transforms into the egg cylinder. *BMC Dev Biol* **7**, 96.

Perea-Gomez, A., Rhinn, M. and Ang, S. L. (2001). Role of the anterior visceral endoderm in restricting posterior signals in the mouse embryo. *Int J Dev Biol* **45**, 311-320.

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, P. N., Dobreva, M. P., Graham, L., Huylebroeck, D., Lawson, K. A. and Zwijsen, A. N. (2011). Amnion formation in the mouse embryo: the single amniochorionic fold model. *BMC Dev Biol* **11**, 48.

Peyre, E., Jaouen, F., Saadaoui, M., Haren, L., Merdes, A., Durbec, P. and Morin, X. (2011). A lateral belt of cortical LGN and NuMA guides mitotic spindle movements and planar division in neuroepithelial cells. *J Cell Biol* **193**, 141-154.

Pfister, S., Steiner, K. A. and Tam, P. P. (2007). Gene expression pattern and progression of embryogenesis in the immediate post-implantation period of mouse development. *Gene Expr Patterns* **7**, 558-573.

Picelli, S., Bjorklund, A. K., Faridani, O. R., Sagasser, S., Winberg, G. and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* **10**, 1096-1098.

Pijuan-Sala, B., Griffiths, J. A., Guibentif, C., Hiscock, T. W., Jawaid, W., Calero-Nieto, F. J., Mulas, C., Ibarra-Soria, X., Tyser, R. C. V., Ho, D. L. L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490-495.

Pijuan-Sala, B., Guibentif, C. and Gottgens, B. (2018). Single-cell transcriptional profiling: a window into embryonic cell-type specification. *Nat Rev Mol Cell Biol* **19**, 399-412.

Pitarresi, J. R., Liu, X., Sharma, S. M., Cuitino, M. C., Kladney, R. D., Mace, T. A., Donohue, S., Nayak, S. G., Qu, C., Lee, J., et al. (2016). Stromal ETS2 Regulates Chemokine Production and Immune Cell Recruitment during Acinar-to-Ductal Metaplasia. *Neoplasia* **18**, 541-552.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J. and Hadjantonakis, A. K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**, 3081-3091.

Polydorou, C. and Georgiades, P. (2013). Ets2-dependent trophoblast signalling is required for gastrulation progression after primitive streak initiation. *Nat Commun* **4**, 1658.

Price, M. A., Cruzalegui, F. H. and Treisman, R. (1996). The p38 and ERK MAP kinase pathways cooperate to activate Ternary Complex Factors and c-fos transcription in response to UV light. *Embo j* **15**, 6552-6563.

Quadros, R. M., Miura, H., Harms, D. W., Akatsuka, H., Sato, T., Aida, T., Redder, R., Richardson, G. P., Inagaki, Y., Sakai, D., et al. (2017). Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol* **18**, 92.

Raff, M. C., Barres, B. A., Burne, J. F., Coles, H. S., Ishizaki, Y. and Jacobson, M. D. (1993). Programmed cell death and the control of cell survival: lessons from the nervous system. *Science* **262**, 695-700.

Ragkousi, K. and Gibson, M. C. (2014). Cell division and the maintenance of epithelial order. *J Cell Biol* **207**, 181-188.

Ralston, A. and Rossant, J. (2008). Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Dev Biol* **313**, 614-629.

Rayon, T., Menchero, S., Rollan, I., Ors, I., Helness, A., Crespo, M., Nieto, A., Azuara, V., Rossant, J. and Manzanares, M. (2016). Distinct mechanisms regulate Cdx2 expression in the blastocyst and in trophoblast stem cells. *Sci Rep* **6**, 27139.

Razy-Krajka, F., Gravez, B., Kaplan, N., Racioppi, C., Wang, W. and Christiaen, L. (2018). An FGF-driven feed-forward circuit patterns the cardiopharyngeal mesoderm in space and time. *Elife* **7**.

Reddy, B. V. and Irvine, K. D. (2008). The Fat and Warts signaling pathways: new insights into their regulation, mechanism and conservation. *Development* **135**, 2827-2838.

Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H. and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res* **44**, W83-89.

Richardson, L., Torres-Padilla, M. E. and Zernicka-Goetz, M. (2006). Regionalised signalling within the extraembryonic ectoderm regulates anterior visceral endoderm positioning in the mouse embryo. *Mech Dev* **123**, 288-296.

Riedl, J., Flynn, K. C., Raducanu, A., Gartner, F., Beck, G., Bosl, M., Bradke, F., Massberg, S., Aszodi, A., Sixt, M., et al. (2010). Lifeact mice for studying F-actin dynamics. *Nat Methods* **7**, 168-169.

Rivera-Perez, J. A., Mager, J. and Magnuson, T. (2003). Dynamic morphogenetic events characterize the mouse visceral endoderm. *Dev Biol* **261**, 470-487.

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.

Robb, L., Hartley, L., Begley, C. G., Brodnicki, T. C., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Elefanty, A. G. (2000). Cloning, expression analysis, and chromosomal localization of murine and human homologues of a Xenopus mix gene. *Dev Dyn* **219**, 497-504.

Rodriguez, T. A., Casey, E. S., Harland, R. M., Smith, J. C. and Beddington, R. S. (2001). Distinct enhancer elements control Hex expression during gastrulation and early organogenesis. *Dev Biol* **234**, 304-316.
Rodriguez, T. A., Srinivas, S., Clements, M. P., Smith, J. C. and Beddington, R. S. (2005). Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm. *Development* **132**, 2513-2520.
Roehl, H. and Nusslein-Volhard, C. (2001). Zebrafish pea3 and erm are general targets of FGF8 signaling. *Curr Biol* **11**, 503-507.
Rossant, J., Chazaud, C. and Yamanaka, Y. (2003). Lineage allocation and asymmetries in the early mouse embryo. *Philos Trans R Soc Lond B Biol Sci* **358**, 1341-1348; discussion 1349.
Rossant, J. and Cross, J. C. (2001). Placental development: lessons from mouse mutants. *Nat Rev Genet* **2**, 538-548.
Rossant, J. and Tam, P. P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701-713.
Sah, V. P., Attardi, L. D., Mulligan, G. J., Williams, B. O., Bronson, R. T. and Jacks, T. (1995). A subset of p53-deficient embryos exhibit exencephaly. *Nat Genet* **10**, 175-180.
Saiz, N. and Plusa, B. (2013). Early cell fate decisions in the mouse embryo. *Reproduction* **145**, R65-80.
Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N. K., Macaulay, I. C., Marioni, J. C. and Gottgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature* **535**, 289-293.
Scott, E. W., Simon, M. C., Anastasi, J. and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573-1577.
Segditsas, S., Sieber, O., Deheragoda, M., East, P., Rowan, A., Jeffery, R., Nye, E., Clark, S., Spencer-Dene, B., Stamp, G., et al. (2008). Putative direct and indirect Wnt targets identified through consistent gene expression changes in APC-mutant intestinal adenomas from humans and mice. *Hum Mol Genet* **17**, 3864-3875.
Seidel, J. J. and Graves, B. J. (2002). An ERK2 docking site in the Pointed domain distinguishes a subset of ETS transcription factors. *Genes Dev* **16**, 127-137.
Sementchenko, V. I. and Watson, D. K. (2000). Ets target genes: past, present and future. *Oncogene* **19**, 6533-6548.
Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L. and Pe'er, D. (2019). Characterization of cell fate probabilities in single-cell data with Palantir. *Nat Biotechnol* **37**, 451-460.
Sevilla, L., Aperlo, C., Dulic, V., Chambard, J. C., Boutonnet, C., Pasquier, O., Pognonec, P. and Boulukos, K. E. (1999). The Ets2 transcription factor inhibits apoptosis induced by colony-stimulating factor 1 deprivation of macrophages through a Bcl-xL-dependent mechanism. *Mol Cell Biol* **19**, 2624-2634.
Shalek, A. K., Satija, R., Adiconis, X., Gertner, R. S., Gaublomme, J. T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498**, 236-240.
Sharrocks, A. D. (2001). The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol* **2**, 827-837.
Sharrocks, A. D., Brown, A. L., Ling, Y. and Yates, P. R. (1997). The ETS-domain transcription factor family. *Int J Biochem Cell Biol* **29**, 1371-1387.
Shimono, A. and Behringer, R. R. (2003). Angiomotin regulates visceral endoderm movements during mouse embryogenesis. *Curr Biol* **13**, 613-617.
Sidiropoulos, N., Sohi, S. H., Pedersen, T. L., Porse, B. T., Winther, O., Rapin, N. and Bagger, F. O. (2018). SinaPlot: An Enhanced Chart for Simple and Truthful Representation of Single Observations Over Multiple Classes. *Journal of Computational and Graphical Statistics* **27**, 673-676.
Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-690.
Snow, M. H. L. (1977). Gastrulation in the mouse: growth and regionalization of the epiblast. *J. Embryol. Exp. Morphol.* **42**, 293-303
Srinivas, S. (2010). Imaging cell movements in egg-cylinder stage mouse embryos. *Cold Spring Harb Protoc* **2010**, pdb.prot5539.
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.
Stegle, O., Teichmann, S. A. and Marioni, J. C. (2015). Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* **16**, 133-145.
Stephenson, R. O., Rossant, J. and Tam, P. P. (2012). Intercellular interactions, position, and polarity in establishing blastocyst cell lineages and embryonic axes. *Cold Spring Harb Perspect Biol* **4**.

Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F. and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**, 2093-2102.

Stuckey, D. W., Clements, M., Di-Gregorio, A., Senner, C. E., Le Tissier, P., Srinivas, S. and Rodriguez, T. A. (2011). Coordination of cell proliferation and anterior-posterior axis establishment in the mouse embryo. *Development* **138**, 1521-1530.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550.

Sussan, T. E., Yang, A., Li, F., Ostrowski, M. C. and Reeves, R. H. (2008). Trisomy represses Apc(Min)-mediated tumours in mouse models of Down's syndrome. *Nature* **451**, 73-75.

Szymczak, A. L. and Vignali, D. A. (2005). Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* **5**, 627-638.

Tait, S. W. and Green, D. R. (2010). Mitochondria and cell death: outer membrane permeabilization and beyond. *Nat Rev Mol Cell Biol* **11**, 621-632.

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, W. L., Lim, C. Y., Han, J., Zhang, J., Ang, Y. S., Ng, H. H., Yang, H. and Lim, B. (2008). T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. *Stem Cells* **26**, 2019-2031.

Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A. and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072-2075.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**, 377-382.

Taylor, R. C., Cullen, S. P. and Martin, S. J. (2008). Apoptosis: controlled demolition at the cellular level. *Nat Rev Mol Cell Biol* **9**, 231-241.

Tepass, U., Tanentzapf, G., Ward, R. and Fehon, R. (2001). Epithelial cell polarity and cell junctions in Drosophila. *Annu Rev Genet* **35**, 747-784.

Tesar, P. J. (2005). Derivation of germ-line-competent embryonic stem cell lines from preblastocyst mouse embryos. *Proc Natl Acad Sci U S A* **102**, 8239-8244.

Tetsu, O. and McCormick, F. (2017). ETS-targeted therapy: can it substitute for MEK inhibitors? *Clin Transl Med* **6**, 16.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S. and Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386.

Treutlein, B., Brownfield, D. G., Wu, A. R., Neff, N. F., Mantalas, G. L., Espinoza, F. H., Desai, T. J., Krasnow, M. A. and Quake, S. R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371-375.

Trichas, G., Begbie, J. and Srinivas, S. (2008). Use of the viral 2A peptide for bicistronic expression in transgenic mice. *BMC Biol* **6**, 40.

Trichas, G., Joyce, B., Crompton, L. A., Wilkins, V., Clements, M., Tada, M., Rodriguez, T. A. and Srinivas, S. (2011). Nodal dependent differential localisation of dishevelled-2 demarcates regions of differing cell behaviour in the visceral endoderm. *PLoS Biol* **9**, e1001019.

Trichas, G., Smith, A. M., White, N., Wilkins, V., Watanabe, T., Moore, A., Joyce, B., Sugnaseelan, J., Rodriguez, T. A., Kay, D., et al. (2012). Multi-cellular rosettes in the mouse visceral endoderm facilitate the ordered migration of anterior visceral endoderm cells. *PLoS Biol* **10**, e1001256.

Trimboli, A. J., Cantemir-Stone, C. Z., Li, F., Wallace, J. A., Merchant, A., Creasap, N., Thompson, J. C., Caserta, E., Wang, H., Chong, J. L., et al. (2009). Pten in stromal fibroblasts suppresses mammary epithelial tumours. *Nature* **461**, 1084-1091.

Tynan, J. A., Wen, F., Muller, W. J. and Oshima, R. G. (2005). Ets2-dependent microenvironmental support of mouse mammary tumors. *Oncogene* **24**, 6870-6876.

van de Wetering, M., Sancho, E., Verweij, C., de Lau, W., Oving, I., Hurlstone, A., van der Horn, K., Batlle, E., Coudreuse, D., Haramis, A. P., et al. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241-250.

Van der Flier, L. G., Sabates-Bellver, J., Oving, I., Haegebarth, A., De Palo, M., Anti, M., Van Gijn, M. E., Suijkerbuijk, S., Van de Wetering, M., Marra, G., et al. (2007). The Intestinal Wnt/TCF Signature. *Gastroenterology* **132**, 628-632.

van der Flier, L. G., van Gijn, M. E., Hatzis, P., Kujala, P., Haegebarth, A., Stange, D. E., Begthel, H., van den Born, M., Guryev, V., Oving, I., et al. (2009). Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* **136**, 903-912.

Vary, C. P., Li, V., Raouf, A., Kitching, R., Kola, I., Franceschi, C., Venanzoni, M. and Seth, A. (2000). Involvement of Ets transcription factors and targets in osteoblast differentiation and matrix mineralization. *Exp Cell Res* **257**, 213-222.

Vento-Tormo, R., Efremova, M., Botting, R. A., Turco, M. Y., Vento-Tormo, M., Meyer, K. B., Park, J. E., Stephenson, E., Polanski, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347-353.

Viotti, M., Nowotschin, S. and Hadjantonakis, A. K. (2014). SOX17 links gut endoderm morphogenesis and germ layer segregation. *Nat Cell Biol* **16**, 1146-1156.

Wallace, J. A., Li, F., Balakrishnan, S., Cantemir-Stone, C. Z., Pecot, T., Martin, C., Kladney, R. D., Sharma, S. M., Trimboli, A. J., Fernandez, S. A., et al. (2013). Ets2 in tumor fibroblasts promotes angiogenesis in breast cancer. *PLoS One* **8**, e71533.

Watanabe, M., Hitomi, M., van der Wee, K., Rothenberg, F., Fisher, S. A., Zucker, R., Svoboda, K. K., Goldsmith, E. C., Heiskanen, K. M. and Nieminen, A. L. (2002). The pros and cons of apoptosis assays for use in the study of cells, tissues, and organs. *Microsc Microanal* **8**, 375-391.

Watson, D. K., Kitching, R., Vary, C., Kola, I. and Seth, A. (2000). Isolation of target gene promoter/enhancer sequences by whole genome PCR method. *Methods Mol Biol* **130**, 1-11.

Watson, D. K., Mavrothalassitis, G. J., Jorcyk, C. L., Smyth, F. E. and Papas, T. S. (1990). Molecular organization and differential polyadenylation sites of the human ETS2 gene. *Oncogene* **5**, 1521-1527.

Wen, F., Tynan, J. A., Cecena, G., Williams, R., Munera, J., Mavrothalassitis, G. and Oshima, R. G. (2007). Ets2 is required for trophoblast stem cell self-renewal. *Dev Biol* **312**, 284-299.

Wen, J., Zeng, Y., Fang, Z., Gu, J., Ge, L., Tang, F., Qu, Z., Hu, J., Cui, Y., Zhang, K., et al. (2017). Single-cell analysis reveals lineage segregation in early post-implantation mouse embryos. *J Biol Chem* **292**, 9840-9854.

Whitmarsh, A. J. and Davis, R. J. (2000). Regulation of transcription factor function by phosphorylation. *Cell Mol Life Sci* **57**, 1172-1183.

Whitmarsh, A. J., Shore, P., Sharrocks, A. D. and Davis, R. J. (1995). Integration of MAP kinase signal transduction pathways at the serum response element. *Science* **269**, 403-407.

Williams, J. R., Little, J. B. and Shipley, W. U. (1974). Association of mammalian cell death with a specific endonucleolytic degradation of DNA. *Nature* **252**, 754-755.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A. and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**, 684-687.

Williams, S. E., Beronja, S., Pasolli, H. A. and Fuchs, E. (2011). Asymmetric cell divisions promote Notch-dependent epidermal differentiation. *Nature* **470**, 353-358.

Williams, S. E., Ratliff, L. A., Postiglione, M. P., Knoblich, J. A. and Fuchs, E. (2014). Par3-mInsc and Galphai3 cooperate to promote oriented epidermal cell divisions through LGN. *Nat Cell Biol* **16**, 758-769.

Winnier, G., Blessing, M., Labosky, P. A. and Hogan, B. L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev* **9**, 2105-2116.

Wolf, B. B., Schuler, M., Echeverri, F. and Green, D. R. (1999). Caspase-3 is the primary activator of apoptotic DNA fragmentation via DNA fragmentation factor-45/inhibitor of caspase-activated DNase inactivation. *J Biol Chem* **274**, 30651-30656.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R. and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol* **13**, 838-845.

Xu, F., Li, H. and Jin, T. (1999). Cell type-specific autoregulation of the Caudal-related homeobox gene Cdx-2/3. *J Biol Chem* **274**, 34310-34316.

Yamamoto, M., Beppu, H., Takaoka, K., Meno, C., Li, E., Miyazono, K. and Hamada, H. (2009). Antagonism between Smad1 and Smad2 signaling determines the site of distal visceral endoderm formation in the mouse embryo. *J Cell Biol* **184**, 323-334.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* **20**, 1131-1139.

Yang, H., Wang, H., Shivalila, C. S., Cheng, A. W., Shi, L. and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-1379.

Yang, Q. and Guan, K. L. (2007). Expanding mTOR signaling. *Cell Res* **17**, 666-681.

Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523.

Ying, Y., Liu, X. M., Marble, A., Lawson, K. A. and Zhao, G. Q. (2000). Requirement of Bmp8b for the generation of primordial germ cells in the mouse. *Mol Endocrinol* **14**, 1053-1063.

Yoon, Y., Huang, T., Tortelote, G. G., Wakamiya, M., Hadjantonakis, A. K., Behringer, R. R. and Rivera-Perez, J. A. (2015). Extra-embryonic Wnt3 regulates the establishment of the primitive streak in mice. *Dev Biol* **403**, 80-88.

Zakin, L., Reversade, B., Virlon, B., Rusniok, C., Glaser, P., Elalouf, J. M. and Brulet, P. (2000). Gene expression profiles in normal and Otx2-/- early gastrulating mouse embryos. *Proc Natl Acad Sci U S A* **97**, 14388-14393.

Zhang, J., Tam, W. L., Tong, G. Q., Wu, Q., Chan, H. Y., Soh, B. S., Lou, Y., Yang, J., Ma, Y., Chai, L., et al. (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* **8**, 1114-1123.

Zhang, Y. E. (2009). Non-Smad pathways in TGF-beta signaling. *Cell Res* **19**, 128-139.

Zhu, X., Morales, F. C., Agarwal, N. K., Dogruluk, T., Gagea, M. and Georgescu, M. M. (2013). Moesin is a glioma progression marker that induces proliferation and Wnt/beta-catenin pathway activation via interaction with CD44. *Cancer Res* **73**, 1142-1155.

Appendix A

ExE signature marker gene references

Marker	Reference
Acvr1b	Gu et al., 1998
Bmp4	Mohammed et al., 2017
Bmp8b	Ying et al., 2000
Cdx2	Mohammed et al., 2017
Elf5	Mohammed et al., 2017
Eomes	Mohammed et al., 2017
Esrrb	Mohammed et al., 2017
Ets2	Mohammed et al., 2017
Fgfr2	Arman et al., 1998
Furin	Mohammed et al., 2017
Gata3	Home et al., 2009
Igf2	Lee et al., 1990
Pcsk6	Mohammed et al., 2017
Smad3	Dunn et al., 2004
Sox2	Georgiades and Rossant, 2006
Spry4	Donnison et al., 2015
Tfap2c	Auman et al., 2002