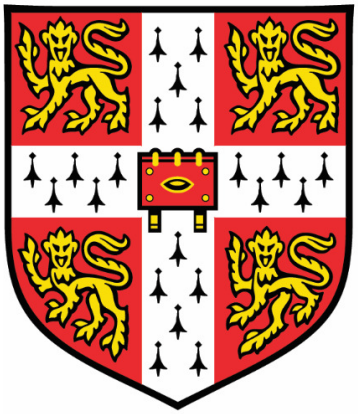
Exploring the Role of *BCL11A* in the Oesophagus and Mammary Gland Using *In Vivo* Lineage Tracing and 3D Organoid Models.



Rosemary Aylin Uğur

Department of Pharmacology

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Clare College

December 2020

Preface

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. The dissertation does not exceed the prescribed word limit for the Biological Degree Committee: it contains fewer than 60,000 words, excluding appendices, bibliography, footnotes, tables and equations, and has fewer than 150 figures.

Rosemary Uğur

December 2020

For my Auntie Cathy, not
a day goes by where I
don't wish you were here
with us.

Abstract

In the pursuit of effective cancer therapy, extensive disease heterogeneity remains one of the largest problems facing researchers. NGS sequencing has helped to identify specific markers of disease to deconvolute some of this heterogeneity, indeed it has also enabled the similarities of some epithelial tumours to be identified[1]. *BCL11A* has been identified as a triple negative breast cancer gene[2]. However, little is known about its function in the physiological mammary gland. In addition to its role in breast cancer, previous work in the lab has identified *BCL11A* as a selectively upregulated gene lung squamous carcinoma (LUSC)[3] and some preliminary data suggest that it is also upregulated in oesophageal squamous carcinoma (OSCC). With proven roles in malignancy and little knowledge about its physiological function, this work reports the combined use of *in vivo* lineage tracing techniques and *in vitro* 3D organoids to explore the expression and gene function of *Bcl11a* in the oesophageal epithelium and mammary gland. Gene function studies were undertaken using mouse models of *BCL11A* overexpression: *Rosa*^{CreERT2}; *Rosa-CAGG-LSL-BCL11A* (*BCL11A*^{OE}) and *Bcl11a* knock out: *Rosa*^{CreERT2}; *Bcl11a* flox/flox (*Bcl11a*^{KO}). *Bcl11a* lineage tracing was achieved using a *Bcl11a*^{CreERT2}; *Rosa-CAG-LSL-tdTomato* mouse model (*Bcl11a*^{Cre}). Within the murine mammary epithelium previous work in the lab revealed the localisation of *Bcl11a* expressing cells to the luminal progenitor compartment of the ductal network. Expansion of labelled cells in response to pregnancy was observed. To explore this physiological role further, mammary organoids were derived from single cells of the luminal progenitor compartment from both *BCL11A*^{OE} and *Bcl11a*^{KO} models. Organoids exhibiting these gene perturbations did not demonstrate phenotypic alterations. Further studies will be required to fully characterise the role of *Bcl11a* in the mammary gland. Oesophageal organoids cultivated from the *BCL11A*^{OE} model demonstrated a lack of differentiation and a superior colony forming ability after passaging. *Bcl11a*^{KO} organoids demonstrated an inhibition in growth and differentiation, inducing a loss of integrity in the organoid structures. Cells displayed a lower colony forming capacity after passage, suggesting *Bcl11a* expression was critical to organoid growth and differentiation. These findings were validated *in vivo* using lineage tracing. *Bcl11a* lineage tracing revealed the persistence and expansion of labelled cells in each layer of the squamous epithelium, for a period of up to 12 weeks post tamoxifen exposure. Labelling at 24 hours revealed cells present in both the basal layer (which houses progenitors) and in cells of the differentiated suprabasal layers. FACS sorted tdTomato-positive cells 24 hours post tamoxifen exposure demonstrated superior colony forming capacity compared to the tdTomato negative

population. Taken together these results suggest that *Bcl11a* has a role in progenitor cells of the oesophageal epithelium with a potential function in maintaining normal growth and differentiation. Taken together these results suggest that *Bcl11a* is a valuable candidate for further study in the murine oesophageal epithelium.

Exploring the Role of BCL11A in the Oesophagus and Mammary Gland using In Vivo Lineage Tracing and 3D Organoid Models

Rosemary Aylin Uğur

Acknowledgements

I would firstly like to thank my supervisor Dr Walid Khaled, thank you for believing in me and for giving me the space to pursue this project on my own terms. I am so grateful for this opportunity and I have been very proud to be a member of this lab. I must also thank Dr Maria Alcolea for her help and expertise in the early stages of the oesophagus project. Thank you to Dr Jon Holdich for teaching me about nearly every technical aspect of the Leica confocal and helping when things went wrong! I really owe you a lot. Without your expert knowledge many of my figures would not be what they are, so thank you.

My thanks and gratitude are also extended to my other scientific mentors: Dr Kyren Lazarus and Dr Sara Pensa. I am very grateful to Kyren of course for undertaking the experiments that have led to this work, without you it would not exist! But more generally I would like to say that your discipline and dedication to science was infectious during our time working together and contributed so much to my attitude and the way I approached this research. I also owe a huge amount of gratitude and thanks to Sara, one of the greatest scientists I know! Your help with key techniques in this project were vital to its success. It is so powerful to see a strong woman like you in science doing things on her own terms and doing them right! Thank you so much also for the level of support you have given me during these years and especially in these last months. I always appreciate your perspective and I hope to still have it long into the future! By the same token, thank you to Dr Alecia Jane Twigger for also being a strong woman in science and an excellent shoulder to lean on during this last year. You have been another great role model, your energy and passion is totally infectious! In addition, I also have to thank the other members of our group, in particular Karsten Bach for great chats and even better fruit breaks (chicken parties), as well as previous members Dr Fazal Hadi and Giacomo Borsari, for making the lab such an intellectually fertile environment in which to begin my research life, it has truly fuelled my passion over these years!

Thank you also to my wonderful Herts friends, particularly Jayna Patel, Rida Aslam, Dean French and Zakkia Sultana. Thanks for listening to me go on and on about this degree, I am sure it wasn't enjoyable, but I love you all. Thank you for all of your support and love over these years I am so grateful. I have to also give a huge thanks to Becca Swinburne. Thank you

for your love and support it has meant so much, our friendship is so precious to me. You are like a sister and I don't know where I would be without you.

To my parents Anne and Orhan, thank you for always being there to listen, support, love and tell me how it was in Jordan. I am wealthy beyond words to have you both as parents. To my sisters Sophie and Aimee, you have both helped me so much with your support and advice. Thank you so much for making home such a fun place to be whenever you come back to Cambridge, it was the best medicine during these tough years. I love you all so much!

Finally, last but not least thank you to Brendan Swain. You have been such an important presence in my life during these years. On a practical level, thank you for making me food and on a spiritual level, thank you for existing. I am so grateful for the support you have provided me, but I must formally apologise for knowing nothing about membrane proteins, and thus offering a minimal amount of help to your PhD!

Table of Contents

Preface i

Abstract iii

Acknowledgements v

List of Figures x

Abbreviations xvi

Chapter 1: Introduction

1.1 Targeted Therapy and Cancer 3

1.2. *BCL11A* is a Transcription Factor Activated in Malignancy 5

1.2.1 Breast cancer subtypes 10

1.2.2 Oesophageal Cancer 12

1.3 Mammary Gland Development 15

1.3.1 Pubertal Development and Ductal Morphogenesis 16

1.3.2 Cellular Hierarchy in the Mammary Gland 22

1.4. Oesophageal Development 30

1.4.1 Oesophageal Tissue Maintenance and Hierarchy 32

1.5. Models of Gene Expression/function 36

1.5.1 Primary 3D culture of the murine mammary gland 37

1.5.2 Primary 3D culture of the murine Oesophagus 39

1.6 Thesis Outline 40

Chapter 2: Experimental Procedures

2.1 Mouse Models 42

2.1.1 Mouse Genotyping PCR 43

2.2 Tissue Isolation and processing: Mammary Gland 48

2.3 Fluorescence Activated Cell Sorting in the Mammary Gland 48

2.4 Organoid Cell Culture for the Mammary Gland Using Matrigel ECM. 49

2.5 Organoid Cell Culture for the Mammary Gland Using a Matrigel and Collagen I mix 49

2.6 Tissue Isolation and Processing: Oesophagus. 50

2.7 Cell Culture for Oesophageal Organoids 50

2.8 *In vitro* 4-hydroxytamoxifen induction.....51

2.9 Preparation of Organoids for Whole Mount Immunofluorescence Imaging 54

2.10 RNA Extraction RT-qPCR 54

2.11 Quantitation of microscope images, cell counting.....56

2.12 Quantification of organoid number and size.....56

2.13 *In vivo* Tamoxifen Induction.....56

2.14 Lineage tracing Esophageal Tissue Preparation for Whole Mount Immunohistochemical Analysis.....57

2.15 Optical Clearing of the Esophageal Epithelium using CUBIC for lineage tracing. 57

2.16 Fluorescence Activated Cell Sorting for tdTomato Positive Cells 58

Chapter 3: Exploring *BCL11A* Function in Oesophageal Organoids

3.1 Introduction.....60

3.1.1 Exploring *BCL11A* gene function in the Oesophageal Epithelium Using 3D organoids 60

3.2 Results.....64

3.2.1Comparing Models of Oesophageal Organoids.....64

3.2.2 *BCL11A* is Required for Correct Differentiation in Oesophageal Organoids.....70

3.2.3 *Bcl11a* is Critical for Oesophageal Organoid Survival.....90

3.3 Discussion 106

Chapter 4: Exploring *Bcl11a* Function using *In Vivo* Lineage Tracing

4.1 Introduction.....108

4.1.1Exploring *Bcl11a* Gene Expression *In Vivo* 109

4.2 Results.....109

4.2.1 *Bcl11a* Lineage Tracing Marks Luminal Progenitor Cells in the Mammary Gland.109

4.2.2 Lineage tracing in the Oesophageal Epithelium 112
4.2.3 *Bcl11a* Lineage Tracing Labels Cells in the Oesophageal Epithelium 116
4.2.4 *Bcl11a* Labelled Cells Persist and Expand in the Oesophageal Epithelium 129
4.2.5 *Bcl11a* Marks a Long-Lived Population in the Oesophageal Epithelium 141
4.3. Discussion 153

Chapter 5: Exploring *BCL11A* Function in Mammary Organoids

5.1 Introduction 155
5.1.1 *BCL11A* Gene Function in the Mammary Gland 156
5.2 Results 159
5.2.1 Comparing Organoid Models of the Mammary Gland 159
5.2.2 Modelling *BCL11A* Overexpression in Mammary Organoids 175
5.2.3 Modelling *Bcl11a* deletion in Mammary Organoids 199
5.3 Discussion 209

Chapter 6: Discussion

6.0 Discussion 212
6.1 *Bcl11a* Has a Homeostatic Role in the Oesophagus. 212
6.2 Future Perspectives for Exploring Gene Function in the Oesophagus 214
6.3 *BCL11A* in the Mammary Gland 216
6.4 Future Perspectives for Exploring Gene Function in the Mammary Gland 218
6.5 Conclusion 219
7 Bibliography 220

List of Figures

Figure 1.1: Human *BCL11A* locus and predicted isoforms. 6

Figure 1.2: Cellular organisation of the mouse mammary gland.....21

Figure 1.3: A proposed model for cellular hierarchy in the mouse mamary gland. 29

Figure 1.4: Cellular organisation and proposed hierarchy of the mouse oesophageal epithelium. 35

Figure 2.1: *BCL11A*^{OE} and *Bcl11a*^{KO} mouse models used in this project.....45

Figure 2.2: Lineage tracing mouse models use in the project *Bcl11a*^{CreERT2} and *Bcl11a*^{WT} 46

Figure 3.1: Xenograft data from *BCL11A* Knock down. 62

Figure 3.2: Data obtained from the TCGA dataset identifies *BCL11A* as a selectively upregulated gene in OSCC Vs OSAC. 63

Figure 3.3: Oesophageal organoids cultivated using medium formulation A. 65

Figure 3.4: Oesophageal organoids cultivated using medium formulation B. 68

Figure 3.5: Fluorescence microscopy of Whole mount WT oesophageal organoids cultivated using medium formulation B. 69

Figure 3.6: Quantification of WT oesophageal organoid number and size. 72

Figure 3.7: Fluorescence microscopy of Whole mount WT oesophageal organoids exposed to 4-OHT vehicle control and left untreated. 74

Figure 3.8: Quantification of WT oesophageal organoid microscope images.....75

Figure 3.9: Multiple planes of WT oesophageal organoid displays correct differentiation 76

Figure 3.10: Brightfield view of *BCL11A*^{OE} oesophageal organoids after 14 days in culture. 79

Figure 3.11: Brightfield view of individual *BCL11A*^{OE} oesophageal organoids and quantification of organoid number and size.....80

Figure 3.12: Analysis of *BCL11A* expression in oesophageal organoids cultivated from *BCL11A*^{OE} mice. 81

Figure 3.13: Brightfield and GFP view of individual *BCL11A*^{OE} oesophageal organoids. 82

Figure 3.14: Fluorescence microscopy of whole mount *BCL11A*^{OE} oesophageal organoids. 83

Figure 3.15: Quantification of *BCL11A*^{OE} oesophageal organoid microscope images. ..84

Figure 3.16: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with 4-OHT. 85

Figure 3.17: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with vehicle. 86

Figure 3.18: Fluorescence microscope images of whole mount untreated *BCL11A*^{OE} organoids. 87

Figure 3.19: Multiple planes of *BCL11A*^{OE} organoids treated with 4-OHT or vehicle control. 88

Figure 3.20: Colony forming efficiency of passaged *BCL11A*^{OE} organoids. 89

Figure 3.21: Representative brightfield view of *Bcl11a*^{KO} oesophageal organoids and quantification of organoid number and size after 14 days in culture. 92

Figure 3.22: Individual *Bcl11a*^{KO} oesophageal organoids after 14 days in culture. 93

Figure 3.23: Analysis of *Bcl11a* expression in oesophageal organoids cultivated from *Bcl11a*^{KO} mice. 94

Figure 3.24: Fluorescence microscopy of whole mount *Bcl11a*^{KO} oesophageal organoids. 97

Figure 3.25: Quantification of *Bcl11a*^{KO} oesophageal organoid microscope images. 99

Figure 3.26: Fluorescence microscope images of whole mount *Bcl11a*^{KO} oesophageal organoids treated with 4-OHT. 100

Figure 3.27: Fluorescence microscope images of whole mount *Bcl1a*^{KO} oesophageal organoids treated with vehicle. 101

Figure 3.28: Fluorescence microscope images of whole mount untreated *Bcl11a*^{KO} oesophageal organoids. 102

Figure 3.29: Multiple planes of *Bcl11a*^{KO} organoids treated with 4-OHT or vehicle control. 104

Figure 3.30: Colony forming efficiency of passaged *Bcl11a*^{KO} oesophageal organoids. 105

Figure 4.1: *Bcl11a* lineage tracing in the mammary gland. 111

Figure 4.2: Schematic highlighting chase lengths for lineage tracing in the oesophageal epithelium. 113

Figure 4.3: Schematic of tissue orientation from different viewpoints. 114

Figure 4.4: Fluorescence microscope images of cleared *Bcl11a*^{WT} oesophagi. 118

Figure 4.5: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure, global view. 119

Figure 4.6: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure view 1. 120

Figure 4.7: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure view 1. 121

Figure 4.8: Quantification of microscope images from 2 *Bcl11a*^{CreERT2} mice 24 hours post tamoxifen exposure. 122

Figure 4.9: 3D maximum projection of Z stack view 2 of *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure. 123

Figure 4.10: 3D maximum projection of Z stack view 3 of *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure. 124

Figure 4.11: Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 24 hours post tamoxifen exposure. 125

Figure 4.12: tdTomato positive and negative cells were plated as organoids. 127

Figure 4.13: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure. 131

Figure 4.14: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure, view 1. 133

Figure 4.15: Quantification of microscope images from 2 *Bcl11a*^{CreERT2} mice 8 weeks post tamoxifen exposure. 135

Figure 4.16: 3D maximum projection of Z stack view 2 of *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure. 136

Figure 4.17: A 3D maximum projection of Z stack view 3 of *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure. 137

Figure 4.18: Fluorescence microscope step through of *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure. 138

Figure 4.19: Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 8 weeks post tamoxifen exposure. 139

Figure 4.20: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure, global view. 143

Figure 4.21: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure, global view. 145

Figure 4.22: Quantification of microscope images from 2 *Bcl11a*^{CreERT2} mice 12 weeks post tamoxifen exposure. 147

Figure 4.23: A 3D maximum projection of z stack view 2 of *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure. 148

Figure 4.24: A 3D maximum projection of z stack view 3 of *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure. 149

Figure 4.25: Fluorescence microscope step through of the oesophageal epithelium of a *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure. 150

Figure 4.26: Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 12 weeks post tamoxifen exposure. 151

Figure 5.1: Representative FACS plots for the selection of luminal progenitor cells. ..158

Figure 5.2: Luminal organoids cultivated using medium formulation 1. 161

Figure 5.3: Individual luminal organoids cultivated using medium formulation 1. 163

Figure 5.4: Fluorescence microscopy of whole mount WT mammary luminal organoids cultivated in medium formulation 1.....164

Figure 5.5: Luminal organoids cultivated using medium formulation 2. 167

Figure 5.6: Luminal organoids cultivated using medium formulation 2 seeded at a lower density. 168

Figure 5.7: Individual luminal organoids cultivated using medium formulation 2. 169

Figure 5.8: Fluorescence microscopy of whole mount WT mammary luminal organoids cultivated using medium formulation 2. 170

Figure 5.9: Luminal mammary organoids cultivated in a collagen and matrigel ECM mix using medium formulation 2. 173

Figure 5.10: Morphological comparison of organoids cultivated using different ECM compositions. 174

Figure 5.11: Quantification of WT luminal mammary organoid number and size. 177

Figure 5.12: Fluorescence microscopy of whole mount WT mammary luminal organoids exposed to 4-OHT vehicle control and left untreated. 179

Figure 5.13: Quantification of WT mammary organoid microscope images. 180

Figure 5.14: Fluorescence microscope images of whole mount WT organoids treated with 4-OHT. 181

Figure 5.15: Fluorescence microscope images of whole mount WT organoids treated with vehicle. 182

Figure 5.16: Brightfield view of *BCL11A*^{OE} mammary luminal organoids and quantification of organoid numnber an size after 14 days in culture. 186

Figure 5.17: Brightfield view of individual *BCL11A*^{OE} luminal mammary organoids cultivated for 14 days. 187

Figure 5.18: Analysis of *BCL11A* expression in luminal mammary organoids cultivated from *BCL11A*^{OE} mice. 188

Figure 5.19: Fluorescence microscopy of whole mount *BCL11A*^{OE} luminal mammary organoids. 190

Figure 5.20: Quantification of *BCL11A*^{OE} mammary organoid microscope images. 192

Figure 5.21: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with 4-OHT. 193

Figure 5.22: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with vehicle. 195

Figure 5.23: Fluorescence microscope images of untreated whole mount *BCL11A*^{OE} organoids. 197

Figure 5.24:Representative brightfield view of *Bcl11a*^{KO} mammary organoids and quantification of organoid number and size after 14 days in culture.....201

Figure 5.25: individual *Bcl11a*^{KO} mammary luminal organoids after 14 days in culture. 202

Figure 5.26: Fluorescence microscopy of whole mount *Bcl11a*^{KO} organoids. 203

Figure 5.27: Quantification of *Bcl11a*^{KO} mammary organoid microscope images. 205

Figure 5.28: Fluorescence microscope images of whole mount *Bcl11a*^{KO} organoids treated with 4-OHT.....206

Figure 5.29: Fluorescence microscope images of whole mount *Bcl11a*^{KO} organoids treated with vehicle.....207

Figure 5.30: Fluorescence microscope images of untreated whole mount *Bcl11a*^{KO} organoids.....208

Abbreviations

Bcl11a^{CreERT2}	*Bcl11a* lineage tracing mouse model
Bcl11a^{KO}	*Bcl11a* deletion mouse model
BCL11A^{OE}	*BCL11A* overexpression mouse model
BCL11A^{WT}	*Bcl11a* lineage tracing control genotype
BL1	Basal like 1
BL2	Basal like 2
BLBC	Basal like breast cancer
BPE	Bovine pituitary extract
cDNA	Complementary DNA
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle Medium
DMEM-F12	Dulbecco's Modified Eagle Medium with ham F12
DNA	Deoxyribose nucleic acid
ECM	Extra cellular matrix
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EpCAM	Epithelial cell adhesion molecule
ER	Estrogen receptor
ERα	Estrogen receptor alpha
ES	Embryonic stem cells
FACS	Fluorescence activated cell sorting
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
HBSS	Hank's balanced salt solution
HER2	Human epidermal growth factor receptor 2
HF	Hanks balanced and foetal bovine serum
IM	Immunomodulatory
KSFM	Keratinocyte serum free media
LAR	Luminal androgen receptor
LUSC	Lung squamous cell carcinoma
MaSC	Mammary stem cell
mRNA	Messenger RNA
MSL	Mesenchymal stem like
NC3Rs	National centre for the replacement, refinement and reduction of animals in research
NS	Not significant
NSG mouse	NOD scid gamma mouse
NurD	nucleosome remodelling and deacetylase
OAC	Oesophageal adenocarcinoma
OFU	Organoid forming unit
OSCC	Oesophageal squamous carcinoma
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
RIME	Rapid immunoprecipitation mass spectrometry of endogenous protiens

RNA	Ribonucleic acid
RTK	Receptor tyrosine kinase
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
ScRNASeq	Single cell RNA sequencing
shRNA	Short hairpin RNA
SWI/SNF	Switch sucrose non fermentable
TCGA	The cancer genome atlas
TEB	Terminal end bud
TF	Transcription factor
TGF B	Transforming growth factor beta
TNBC	Triple negative breast cancer
WAT	White adipose tissue
WNT	Wingless and Int-13
WT	Black 6 mouse with (C57BL/6J)
2D	2 Dimensional
3D	3 Dimensional
4-OHT	(Z)-4-Hydroxytamoxifen

Chapter 1: Introduction

1.1 Targeted Therapy and Cancer

One of the major hurdles faced by scientists and clinicians is the lack of targeted therapeutics in the treatment of cancer. Commonly, targeted therapy in cancer treatment refers to the use of drugs to target specific molecules to block the growth and spread of cancer cells[4]. Defining molecular targets for therapy can help to categorise and subtype tumours based on these molecular parameters, thereby ensuring a level of stratification or personalisation in the treatment received. This enables specificity for patients and therefore a greater chance of efficacy during treatment.

Despite the potential advantages of targeted cancer therapy, there remains a need to identify the ideal targets. Understanding the molecular mechanisms of disease is an important way to identify such proteins suitable for targeting. Cancer development is associated with changes in gene expression or mutations which ultimately result in cell survival and proliferation[4]. These alterations can be used to differentiate between cancer and normal cells and could therefore be ideal as cancer cell specific targets. Understanding tumour biology at this level can be reached by sequencing patient tumour samples with matched normal tissue to identify changes in gene expression patterns. Indeed, over the last decade large scale sequencing experiments have generated a wealth of genomic data on many types of tumour.

Deconvoluting cellular hierarchy is a major hurdle to our understanding of how cancer subtypes arise from healthy tissue. The understanding of developmental biology as well as the homeostasis of a tissue can greatly inform our understanding of how tumour initiating events take hold and develop into cancer. Epithelial malignancies account for up to 90% of all cancer cases [5]. Epithelial cells are found in either internal or external linings of the body and epithelial tumours include carcinomas of the oesophagus, breast, lung, and skin. Epithelial cells undergo relatively frequent divisions and as such can represent compartments with high cellular turnover. The mammary gland is one such example where continual remodelling (around 450 cycles of growth and involution in a woman's lifetime) leaves this tissue particularly vulnerable to neoplastic transformation [5]. Indeed, frequent cell division represents a strong risk factor for developing carcinomas. In the mouse intestine, an estimated 1.1 million crypts produce 25 million cells per hour, balanced by the loss of cells from the tissue [6]. Therefore, it is unsurprising that cancer incidence in these tissues increases with age [5]. It has been the goal of large scale multi-platform genomic data consortia to understand the translation of genomic alterations to phenotypic outcomes, with the overall aim of aiding precision medicine in cancer [7]. Such large-scale sequencing efforts have also identified genetic similarities between some

epithelial tumours [1]. This is of note as it could mean that epithelial tumours of different body sites share common markers of disease. Such phenomena allow the larger scale uncovering of shared mechanistic properties for these tumours which could ultimately result in more effective targeted therapeutics.

1.2. *BCL11A* is a Transcription Factor Activated in Malignancy

The transcription factor B-cell lymphoma/leukaemia 11A (BCL11A) is a zinc finger protein that is mainly expressed in the brain and hematopoietic cell lineages [8,9]. In general, zinc finger proteins represent the largest family of sequence-specific DNA binding proteins within the human genome [8]. As such there are a range of combinations and functions of zinc finger motifs, making these proteins extremely versatile, with roles in processes such as development, differentiation and lipid metabolism [10]. BCL11A is a C2H2 type zinc finger protein, the largest class of motifs, and has five different splice variants, all with different expression patterns [10]. The nomenclature of BCL11A isoforms pertains to the size of the variants (see figure 1.1). The most common isoform, BCL11A-XL, is expressed in normal B cells, whereas the BCL11A-S isoform is expressed in malignant B-cells (among others) and BCL11A-L is largely confined to the cerebral cortex [10]. Whilst the C2H2 zinc finger is the most dominant type, the isoforms also contain a C2HC zinc finger, thought to enable protein-protein interactions. The transcription factor is thought to localise with the mammalian SWI/SNF or BAF complex and is therefore classed as an associated protein of this adenosine triphosphate (ATP) dependent chromatin remodelling complex [11]. This complex is specialised to enable nucleosome restructuring, ensuring DNA is accessible for transcription replication and DNA repair [12].

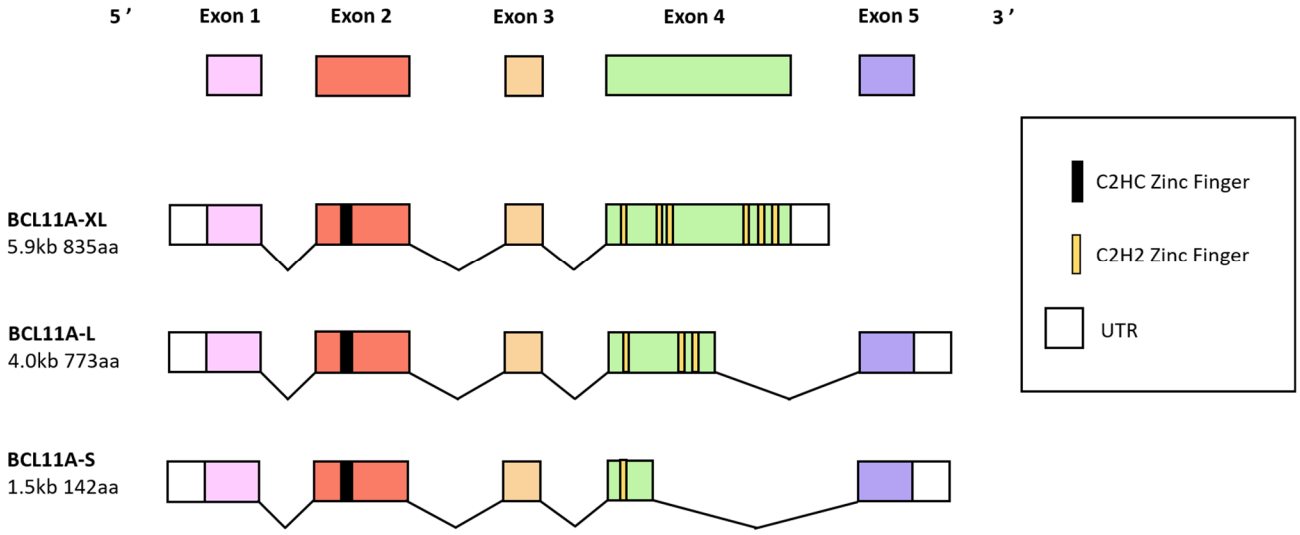


Figure 1.1: Human *BCL11A* locus and predicted isoforms.

Alternative splicing in the *BCL11A* gene locus leads to the generation of three dominant isoforms: X-L, L and S. All protein products contain exons 1 and 2, including a C2HC zinc finger likely enabling in protein- protein interactions. All isoforms also display a portion of exon 4 which encodes up to six C2H2 zinc fingers and are the most common DNA binding motifs found in transcription factors.

As well as localisation to the brain, *BCL11A* is also vital for B-cell development and as such is expressed in most hematopoietic cells such as hematopoietic stem cells, common lymphoid progenitors and early T-cell progenitors [13]. *BCL11A* is essential for the development of multiple cell lineages in this context. Examples include B-cell development, plasmacytoid dendritic cell maturation and the maintenance of cell stemness in hematopoietic stem cells [14–17]. The knock-out of *BCL11A* in mice was shown to reduce the self-renewal capacity of hematopoietic stem cells, delaying the cell cycle through the CDK6 pathway [14]. Another group [18] demonstrated that the absence of the transcription factor may lead to activation causing apoptosis of early B cells and lymphoid progenitors as well as eradicating the ability of the hematopoietic stem cells to give rise to B and T and NK cells [18].

BCL11A also plays a crucial role in the fetal to adult haemoglobin switch [19]. In human erythroid cells this transcription factor is critical for the silencing of the gamma globin gene and reduces fetal hemoglobin expression to promote adult hemoglobin. In order to perform this role, *BCL11A* is thought to interact with other transcription factors such as DNA methyl transferase 1 (*DNMT1*) and SRY-box 6 (*SOX6*) as well as complexes such as LSD1/CoREST and the NuRD complex. These are thought to bind to promoters of the gamma globin gene causing silencing [20].

BCL11A also interacts with CASK, a calcium/calmodulin-dependent serine kinase to regulate axon branching in the brain [21], as well as targeting deleted in colorectal carcinoma (DCC) and microtubule associated protein 1 (MAP1b), which are responsible for axon branching [22]. *BCL11A* has also been shown as the regulator of the sensory area development of layer four neurons [23]. Indeed, *BCL11A* acts in a dose-dependent manner to decide the subcerebral and corticothalamic fates in layer 5 and 6 neurons [24]. The mechanisms by which these events occur remain largely unclear.

There is also evidence of a developmental role for *BCL11A* in the skin [25]. Authors report the regulatory role in epidermal differentiation and lipid metabolism during embryonic skin development [25]. The epidermal permeability barrier functions to prevent dehydration and infection in organisms; *BCL11A* was reported as highly expressed in the developing murine

epidermis [25]. The authors report that the germline deletion of this transcription factor causes defects in the epidermal permeability barrier as well as poor epidermal differentiation. Deletion also causes a large reduction in profilaggrin processing and altered lipid composition. Upon analysing transcriptional data in *BCL11A* deficient embryonic skin cells, the authors report altered gene expression in lipid metabolism enzymes, skin barrier associated transcription factors and junctional proteins [25]. *BCL11A* was shown to interact with genomic elements within regulatory region of the genes Fos related antigen2 (Fosl2) and fatty acid elongase 4 (ELvol4). The expression levels of these genes were then reduced in the absence of *BCL11A*, leading to an increased risk of skin infection due to a defect in the epidermal permeability barrier [25].

More recently, this transcription factor has been recognised as a marker in disease [2,3,26], with seemingly wide-ranging functions in diseases from haematological malignancies and triple negative breast cancer, to lung and laryngeal squamous carcinomas. It is arguably the propensity of this transcription factor to act as a marker for disease which most draws attention to its mechanistic properties. This transcription factor, perhaps expectedly, is highly expressed in glioblastoma and neuroblastomas. Indeed, the importance of this gene as a biomarker for glioblastoma was identified using the analysis of single cell gene expression profiles. The data were also used to aid the development of a more comprehensive method of classification and subtyping of the disease [27]. The authors stress the importance of such biomarker identification as key candidates for drug targets that could perhaps switch clinical features or reverse the progression of the tumours. Continuing the theme of tumour regression, the targeting of *BCL11A* has been shown to induce apoptosis in neuroblastoma cells [28]. *BCl11A* was recognised as a target of the potentially tumour suppressive micro RNA miR-146a, where targeting *BCL11A* resulted in the downregulation of mRNA and protein expression, and promoted the apoptosis of neuroblastoma SK-N-SH cells [28]. This represents elucidation of the mechanisms by which BCL11A may be acting as an antiapoptotic protein in tumours, to better aid in the identification of a more suitable drug target if *BCL11A* itself is harder to therapeutically target. In this case the micro RNA miR-146a was deemed a more suitable candidate target for the treatment of human neuroblastoma [28].

In laryngeal squamous carcinomas (a carcinoma of the head and neck), *BCL11A* is shown to play a similar role, with high levels of expression in patients and its overexpression promoting

the progression of disease in cell lines [26,29]. In analysing patient data, researchers also reported an association with significantly higher co-local recurrence or mortality. *BCL11A* has been shown to be involved in the regulation of cell proliferation, transformation and apoptosis [30]. *In vitro* studies confirmed that the overexpression of *BCL11A* promoted cell proliferation with a higher level of colony formation. Overexpression was also associated with a significantly increased migration and invasion capacity of laryngeal squamous carcinoma cells [26]. Taken together, this evidence suggests a possible biomarker role for this protein in squamous carcinomas of the larynx, offering a subtype specific marker of the disease which could be exploited as a therapeutic target [26].

BCL11A is also a triple negative breast cancer gene [2]. The authors report that its genomic locus is amplified in up to 38% of BLBC tumours. Gene overexpression was found to promote tumour formation whereas its knockdown in cell lines was found to supress tumour formation in xenograft models [2]. The deletion of this gene was found to severely decrease tumour formation in a mouse model of TNBC [2]. The overexpression of *BCL11A* in a non-tumorigenic mammary epithelial cell line was found to both promote colony formation *in vitro* and tumour development when injected into the cleared fat pad of the mammary gland *in vivo* [2]. This evidence suggests that such elevated levels of gene expression promote tumour development. Knock down experiments in TNBC cell lines revealed no significant impact on cell viability or cell death, however it did significantly reduce the clonogenic capacity of cell lines *in vitro* [2]. The authors also report the requirement of this gene for progenitor cells of the mammary gland. Using a *Bcl11a-lacZ* knock in mouse coupled with X-gal staining the authors identify higher levels of Bcl11a in luminal progenitor cells of the mammary gland ($CD49b^{+}$/ $CD24^{high}$) as well as in a mammary stem cell enriched population ($CD49b^{+}$/ $CD24^{med}$) [2]. Despite this characterisation, still little is known about the exact function and mechanisms by which *BCL11A* acts in physiological and malignant contexts.

BCL11A has also been identified as a lung squamous carcinoma oncogene (LUSC) selectively regulated in this subtype as compared to lung adenocarcinoma [3]. The authors report the transcriptional regulation of *BCL11A* by *SOX2*, (a known LSCC oncogene), required for its oncogenic functions. *BCL11A* is identified as regulating key epigenetic factors such as SETD8. The authors report a significantly reduced colony forming capacity in LSCC cell lines with

BCL11A knocked down. The authors also report the use of a 3D organoid model to demonstrate functional effects of *BCL11A* gene perturbations. Organoids cultivated from cells over expressing *BCL11A* demonstrated a lack of hollow lumen, consistent with hyper proliferation and a loss of organisation [3]. Cells cultivated from organoids with *Bcl11a* knockout did not demonstrate any colony forming capacity, suggesting that this gene was required for the formation of organoids [3].

1.2.1 Breast cancer subtypes

Breast cancer is an extensively heterogenous disease, comprised of multiple tumour entities, all associated with distinct biology and histological features [31]. Traditional attempts to understand and classify tumours relied on clinical parameters such as tumour size, histological grade and lymph node involvement [32]. Such analysis aimed to integrate the level of cellular differentiation with the replicative potential of cells, relying on physical differences between normal and cancerous cells [33].

Current understanding compiles gene expression profiling of many tumours in order to uncover distinct molecular subtypes. Thus breast cancer may be divided into clinically relevant subgroups based on respective molecular portraits. These are Luminal A and B, HER2-positive, and basal like and triple negative [34]. Luminal A is associated with ER-positivity and PR-positivity. It represents 50-60% of all cases [32] and is therefore the most common of all breast cancer subtypes. The Luminal B subtype is also associated with ER-positivity. With a higher proliferation index, it is a more aggressive subtype [35]. These tumours are associated with either PR-positivity or negativity and represent 10-20% of all breast cancers [35]. Typically ER positive tumours are treated with Tamoxifen, a prodrug whose active metabolites compete with estrogen for ER binding [36]. The anticancer effect of this drug is therefore predicted through the level of expression of the ER receptors [36].

The HER2-positive subtype is characterized by the high expression of the genes associated with the HER2 pathway [34]. It accounts for 15-20% of cases with high levels of expression in genes associated with proliferation. The subtype also displays high frequency of p53 mutations, adding to its proliferative capacity and its aggressive nature [34]. Herceptin, a monoclonal antibody that binds to the extracellular domain of the HER2 protein, is used for the treatment of HER2 positive tumours [37]. Although initially effective, patients often develop resistance

within a year. This could potentially be due to mutations in regions coding for the extracellular domain. Additionally problems may occur in cell signaling, where Herceptin does not interact with other HER receptors which are thus able to continue downstream signaling [37]. In this case it may be preferential to target multiple HER receptors rather than only HER2 [37].

Basal like and triple negative cancers (BLBC and TNBC) are so called due to the high expression of genes characteristic of normal myoepithelial cells such as high molecular weight cytokeratins (Keratin 5 and Keratin 14) [38]. These cancers are 'triple negative' due to the lack of the hormone and HER2 receptors and characteristic of the above subtypes. This therefore renders conventional targeted therapy ineffective [38]. These tumours are extremely aggressive due to the higher incidence of p53 mutations. Basal like breast cancers are also associated with BRCA1 mutations [38,39].

This categorisation remains an issue of contention, as further analysis reveals yet more heterogeneity within subtypes. The development and optimisation of high throughput technologies has shed more light on this issue. An example includes the further subgrouping of triple negative tumours by Lehmann et al.. Through the analysis of gene expression profiles of 3,247 primary breast tumours, the group reports the presence of 6 subgroups with unique gene expression [40]. These are two basal like groups: BL1 and BL2, an immunomodulatory group (IM), a mesenchymal group (M), a mesenchymal stem cell like (MSL) and finally a luminal androgen receptor (LAR) group [40].

The BL1 and BL2 subgroups have high expression of both DNA damage response genes and those required for cell cycle progression and cell division [40]. The M and MSL subtypes are enriched for epithelial to mesenchymal gene expression as well as the genes involved in cell motility and differentiation pathways [40]. The LAR subtype is characterized by androgen receptor signaling as well as luminal cytokeratin expression whilst the IM subgroup is characterized by immune cell signaling genes [40]. This work highlights the importance of understanding the key initiating mutations that cause such extensive heterogeneity.

Interestingly, recent sequencing efforts and genomic analyses have identified similarities between BLBC and squamous carcinomas. In a recent study from Xia et al. hierarchical clustering of tumour types based on the predictability of gene signatures reveal informative tumour subgroupings. BLBC was found to cluster with lung squamous carcinomas as well as carcinomas of the head and neck [1]. This is particularly interesting owing to the shared molecular markers of these diseases. Indeed, authors Prat et al. highlight the importance of clinical trials focusing on common profiles and biomarker expression, especially in the case of BLBC and lung squamous carcinoma [41]. The authors report striking similarities between BLBCs and squamous cell lung cancers. Highlighting that some somatic mutations are also shared between squamous carcinomas of multiple body sites [42,43].

1.2.2 Oesophageal Cancer

Interestingly, lung squamous carcinomas share common markers with oesophageal squamous carcinomas, a subtype of oesophageal cancer [3,43]. This further highlights the apparent mechanistic similarities between squamous carcinomas of multiple body sites. Cancer of the oesophagus can largely be divided into two subtypes, oesophageal squamous cell carcinoma (OSCC) and oesophageal adenocarcinoma (OAC). These diseases are known to be epidemiologically and pathologically distinct, as biologically, OSCC is more related to squamous carcinomas of the head and neck than OAC. OAC is found to more closely resemble gastrointestinal tumours [44]. Globally, 70% of oesophageal cancer cases are OSCC [45]. The risk factors for OSCC include chemical or physical insult to the mucosa, for example in the form of hot drinks, tobacco or alcohol [46]. Conversely, OAC is thought to arise primarily from long term gastro-oesophageal reflux or Barrett's oesophagus [47].

Barrett's oesophagus is thought to arise as a result of the recurrent exposure of the oesophageal mucosa to the reflux of acid and bile from the stomach. This exposure injures the tissue because of the reactive oxygen species and nitric oxide production [48]. This is thought to cause DNA damage and particularly of note, the A>C transversion which has become a characteristic mutational signature of OAC attributed to reflux damage. Such a genetic profile is common to OAC and Barrett's oesophagus patients, signifying that such DNA damage could represent early causal events in malignancy [48]. Barrett's oesophagus occurs when the squamous oesophageal epithelium is replaced by columnar mucosa, similar to the intestine, and is thought

to be an adaptive response to the continual damage caused by reflux [49]. This mucosa is thought to contain somatic mutations which may predicate carcinogenesis in patients. The dysregulation of p53 signalling is thought to play a key role in the progression of Barrett's oesophagus to OAC [50,51].

Common genetic abnormalities in Barrett's include a loss of heterozygosity of multiple loci, specifically 17p which contains the tumour suppressor p53 [50,51]. Unsurprisingly this loss of heterozygosity has been associated with an increased potential for malignant progression. A paired sequencing study of between Barrett's and OAC samples revealed that a genome doubling event in a TP53 mutant cell commonly precedes OAC development [52]. Tumours that have undergone this doubling event have an increase in focal genomic amplifications, among other distinct characteristics. There are two proposed mechanisms by which Barrett's transitions to OAC: one involves the stepwise loss of tumour suppressor genes such as CDKN2A and TP53 without a genome doubling event; the second involves large scale genome instability associated with aneuploidy following loss of p53 regulation [53].

Conversely, OSCC risk is increased by recurrent chemical or physical insult to the mucosa. OSCC is particularly associated with tobacco and alcohol use [46]. High incidence rates of OSCC are found in China and areas of East Africa. A low intake of fruit and vegetables is also associated with increased risk along with specific regional marginal micronutrient deficiencies particularly in endemic areas [54]. Risk factors are associated with lower socioeconomic status and therefore this malignancy is more common in economically deprived areas [55].

OSCC is thought to develop from basal cell hyperplasia and dysplasia through to carcinoma *in situ* which can be identified using an endoscopy [56]. Although the molecular progression has not been well studied it is thought that dysregulation of TP53 and cell cycle regulators are key characteristics of the disease which can be identified as precursor lesions [56]. The accumulation of abnormal p53 protein has identified in oesophagitis adjacent to dysplasia and carcinoma. An increased expression of CDKN2A/RB1 has been associated with an incremental progression from inflammation to cancer in oesophageal lesions [57].

The evaluation of differentially expressed genes between OSCC and normal mucosa identified two candidate biomarkers TNFAIP3 and CHN which are thought to aid in future diagnosis of either dysplasia or invasive OSCC [58]. Indeed, large scale sequencing efforts have evaluated the mutational transcriptomic and epigenetic profile of invasive OSCC. Commonly mutated genes with point mutations or small indels from the TCGA data set include *TP53, MLL2, NFE2L2*. Identified amplifications include Sry box 2 (*SOX2) TP63* and *FGFR1* [44,59]. There are several dysregulated pathways that are of therapeutic interest in OSCC, including cell cycle regulators, tyrosine kinase receptors, chromatin remodelling and embryonic pathways [59].

Indeed, *SOX2* is a transcription factor that is thought to promote proliferation in OSCC cells [60].The authors demonstrated a role for this gene in proliferation through the use of knock out studies in OSCC cells [60]. Using immunohistochemical analysis, the authors identified elevation of SOX2 protein levels in 70% of OSCC tumours, which is involved in the development and progression of the disease. This amplification is shared with squamous carcinoma of the lung [43] It has been identified that *BCL11A* interacts with *SOX2* to drive the LUSC phenotype [3]. The *SOX2* gene is an amplification target in LUSC [61], and has been shown *in vitro* to promote the proliferation of human epithelial lung lines [61].

1.3 Mammary Gland Development

The classification of mammals relies partly on the presence of mammary glands, the primary role of which is to provide offspring with nourishment and immunity [62]. The mammary gland is particularly remarkable as, unlike any other organ, it undergoes most of its development in the post pubertal female, after which it will continually cycle in structural changes for the duration of the reproductively active life [62]. Therefore, the 'development' of the mammary gland uniquely refers not only to embryonic growth but also to the cellular orchestration of differentiation, expansion and regression which occurs as a result of local and systemic queues during the female reproductive cycle [63]. This tightly orchestrated series of events in the postnatal and adult mammary gland provides the perfect model in which to study a range of biological phenomena, but perhaps the field that stands to gain the most from this research is that of breast cancer and cancer development [64].

Postnatal mammary gland development can be categorised into distinct stages of pre-pubertal development, pubertal development and then the pregnancy stages of early gestation, mid gestation, lactation and involution [63–65]. The mouse has five pairs of mammary fat pads which are located along the mammary line. These fat pads are composed of specialised stroma including adipose tissue, and the rudimentary ductal tree; each fat pad is also host to a lymph node [65]. The major cellular compartments in the gland are therefore the epithelium and the surrounding stroma, between which much of the crosstalk to promote this functional oscillation occurs [64–66]. The primary epithelial duct is attached to an exterior nipple which enables the release of milk produced in the glands during lactation.

1.3.1 Pubertal Development and Ductal Morphogenesis

From the time of birth to the onset of puberty, the mouse mammary gland exhibits equivalent growth to the rest of the animal [63]. As mentioned earlier, a primary ductal tree fills only a small portion of the fat pad. These ducts are composed of two differentiated cell layers, with a layer of epithelial cells surrounding a lumen (luminal cells) and myoepithelial cells in contact with the stroma (basal cells). Together these form the mammary epithelium (See figure 1.2) [63]. Rising hormone levels during puberty induces the formation of club shaped terminal end buds (TEBs) rounding off the end of the ducts. TEBs are composed of two cell types: an inner layer of central body cells and an outer layer of cap cells. These cells give rise to luminal epithelium and myoepithelium respectively [66,67]. Puberty begins at 3 weeks of age during which the rudimentary epithelium begins to exhibit branching morphogenesis, displaying a rapid expansion of the pre-existing epithelium to invade the fat pad (see figure 1.2) [68,69]. This happens in response to the rise in the ovarian hormone estrogen, which stimulates such rapid proliferation and expansion [70]. The driver of the proliferative invasion into the fat pad is the proliferation within the TEB. This results in ductal elongation and bifurcation of the TEB and generates a network of primary ducts which then in turn generate secondary branching structures [71]. Apoptosis has been previously reported as the source of lumen clearing during this phase, citing that the central body cells undergo apoptosis at this stage to clear the lumen during elongation [72]. More recently however, a model of radial intercalation has been proposed as the source of bud evasion and the transition from a multi-layered to a bilayered epithelium in the TEB[73]. Radial intercalation refers to the phenomenon where cells exchange places through out a multi-layered tissue in order to drive tissue spreading. Neumann et al, observe radial intercalation of luminal cells through to the myoepithelial layer of the TEB, helping to elongate the tissue and to resolve the TEB into a bilayer with a single layer of luminal cells [73]. The authors attempted to identify apoptotic markers but did not detect an enrichment of apoptotic cells inside the TEB, concluding that radial intercalation was responsible for this phenomenon instead [73]. Once the extremity of the fat pad has been reached, the TEBs become motilically inactive which signals the end of pubertal growth. It is this branching morphogenesis which characterises the pubertal gland in mice [74]

As outlined, this ductal morphogenesis is characterised by a need for estrogen. The critical requirement of estrogen in ductal morphogenesis was first evidenced in studies which involved removing the ovaries of mice to ablate the release of steroid hormones [74]. These mice then

failed to develop this characteristic ductal network at the time of puberty. The phenotype was, however, rescued through the administration of slow release estrogen pellets, after which the glands displayed full invasion of the fat pad by the ductal tree [74]. Similarly, mice with epithelial cells deficient in estrogen receptor alpha (ERα), also fail to grow a ductal network. When ERα $^{-/-}$ epithelial cells were transplanted into ERα $^{+/+}$ mice, no outgrowth of the ductal network was visible, indicating the requirement of this signalling for proliferation and invasion to occur in epithelial cells. Similarly, when ERα $^{+/+}$ cells were transplanted into ERα $^{-/-}$ mice the ductal morphogenesis was observed, thus highlighting the need for ERα expression in the epithelial cells and not in the stromal cells.

In addition to these studies, others have been undertaken to ascertain the importance of the stroma in this setting. They have primarily involved mice lacking white adipose tissue (WAT). Although the rudimentary ductal tree is present in mice that lack WAT, they fail to display ductal morphogenesis, instead forming very few short and severely distended ducts [75]. What is perhaps particularly important to note about these mice is the apparent lack of this phenotype when mice undergo parturition. The authors note that, despite the absence of WAT, during pregnancy the glands do display the high levels of epithelial cell division and alveolar cell formation characteristic of the pregnant gland. Transplantation assays of cells from the WAT-lacking transgenic mice show that these cells do undergo ductal morphogenesis when subjected to a normal environment. The adipose tissue is therefore vital to the physiological development of the mammary gland at puberty but the gland does contain cells that are capable of alveolar differentiation [75]. Taken together we can infer that there is a great deal of stromal and epithelial crosstalk responsible for orchestrating the invasion of the TEB into the fat pad. The adipose tissue is required for pubertal development but, perhaps more importantly, the cells of the mammary gland are not reliant on this mechanism for differentiation and seemingly possess this capability intrinsically [75].

There are also many other growth factors and cytokines active at this stage of mammary gland development. Epidermal growth factor (EGF) and epidermal growth factor receptor (EGFR) are both expressed in the mammary gland during puberty [76]. They are also important for alveolar differentiation and lactation, both stages of pregnancy. EGFR is a receptor tyrosine kinase (RTK) [77]. The binding of ligand to an RTK results in dimerization of the receptor and

activation of the intracellular domain. This then results in trans auto-phosphorylation and continuation of downstream signalling, the effects of which are wide ranging. For example, the activation of RAS/MEK leads to mitogenic signals and the activation of PI3K-AKT leads to cell survival. EGFR is highly expressed in both the TEB and the surrounding stroma [77]. In a similar experiment to that used to assess estrogen dependency of the TEB, mice with their ovaries removed were used to understand the importance of EGF in ductal morphogenesis [78]. The implantation of an EGF releasing pellet in ovariectomised mice resulted in a re-initiation of ductal morphogenesis in otherwise growth static glands. The same study looked at competitive binding assays to characterise EGFR activity in the mammary glands. Labelling in the TEB was confined to the cap cells and the myoepithelial cells in the ducts [78]. Labelling in the stroma was heaviest adjacent to the epithelium which, the authors postulate, suggests an induction of EGF receptors by the epithelium. This suggests a model whereby stromal cells are able to respond to queues to promote ductal elongation [79], again reinforcing the idea of a dynamic exchange between epithelial cells and the surrounding stroma during puberty.

In addition to this, the fibroblast growth factor (FGF) ligand family members are important for the normal growth and development of the mammary gland [80]. The fibroblast growth factor family is comprised of 18 secreted proteins that interact with four FGF receptors (FGFRs) [81]. In the case of ductal elongation, FGF1, 2, 4, 7 and 10 are all highly expressed ligands in the mammary gland [82]. Ligand binding causes receptor dimerization in the FGFR, kinase activation and subsequent autophosphorylation of tyrosine residues in the cytoplasmic domain of the receptor [80]. The most characterised ligand-receptor pair in the FGF family is FGF10/ FGFR2b. Acting in a paracrine manner on the mammary epithelium, FGF10 is produced by the adipose tissue [83]. It is localised in the mesenchyme whilst the receptor is localised to the TEB. A conditional doxycycline inducible knockout of this receptor was used to characterise its role in this phase of mammary gland development. The knock out resulted in failure to form the TEBs or to grow past the lymph node [80]. This phenotype was then rescued with withdrawal of the doxycycline, resulting in the formation of the TEB and then successful outgrowth [84]. Taken together the authors conclude that FGFR2b signalling is crucial for the proper development of the gland in puberty, particularly the survival and proliferation of the mammary luminal epithelial cells but it does not affect the regenerative potential of the mammary epithelial progenitor cells, as the phenotype was recovered.

Transforming growth factor beta (TGF-β) ligands are also involved in normal mammary gland development and have 3 different isoforms named TGF-β1, 2 and 3. In the mammary gland, these isoforms are expressed in the epithelium at all phases of development [85]. TGF- β can inhibit DNA synthesis, induce apoptosis and modulate extra cellular matrix remodelling, therefore the effects of ligand binding depend also on the cellular context. The cell type, differentiation and microenvironment all have a bearing on the endpoint of the signalling; therefore, tissue specificity is important to consider. TGF- β1 and TGF- β3 have overlapping expression patterns in the epithelium during outgrowth, while TGF- β3 alone is expressed in the TEBs, with TGF- β2 expression upregulated during pregnancy. It has been shown that there is a dose dependency in the effects of pathway stimulation, where lower levels of TGF- β stimulate branching morphogenesis [86]. TGF- β also contributes to TEB regression via signalling in stromal cells to induce proliferation and production of ECM [84]. Interestingly, the high levels of estrogen during puberty actually inhibit TGF- β ligand expression [86]. This results in low levels of expression during ductal elongation. To assess the function of TGF- β, an overexpression model was developed using an activated form of TGF-β1. This resulted in hypoplastic growth.[87]

The Wnt pathway is another that is important for signalling in the mammary gland. The pathway components are temporally and spatially differentially expressed in the glands [88]. For example ligands Wnt4 and Wnt5a are localised to the luminal compartment, whereas Wnt6 is localised to the basal layer [89]. Canonical Wnt signalling, referring to the stabilisation of Beta catenin, is required in the mammary gland for stem cell maintenance, branching and alveologenesis [90]. It has been shown to be important in TEBs as there is a requirement for Lgr5 and Lgr6 in the canonical signalling. These are localised in the basal layer and their targeted deletion in mammary glands showed fewer TEBs developing and a reduction in side branching [89]. Wnt4 has been shown to act specifically in the myoepithelial cells, with canonical signalling also seen in the stroma directly surrounding the TEB [91]. The non-canonical Wnt signalling pathways also play a role in the developing mammary gland. Wnt2, Wnt5a and Wnt7b are all expressed in the TEB during puberty in mice [92]. Wnt5a is responsible for regulating branching and proliferation in the TEB by inhibiting the Wnt/β catenin axis. The non-canonical receptor Ror2 has been characterised as important for normal TEB morphology and for the control of proliferation. Ror2 expression correlates inversely with Wnt activity in the TEB where Ror2 expression remains in the ductal region, up to the neck of the bud. β catenin activity occurs in

the cap cells of the TEB. It has since been concluded that Wnt/ β catenin activity and Ror2 expression actually oppose each other in the TEB. This is in order to promote myoepithelial cell differentiation throughout the TEB, as well as to inhibit the proliferation of committed cells in the neck of the TEB and maintain stemness in the progenitor populations in the cap cell region [92].

Taken together, it is clear that the process of ductal elongation during puberty is an intricate and tightly regulated series of events. The pathways and networks involved in mammary gland development and maintenance are important to understand - not only from the perspective of biological knowledge, but also for its application to cell culture. The culture of mammary and indeed many other types of cells have been made possible by such developmental research as *in vitro* cultivation, requires the recapitulation of these niche environments. Through better understanding these processes and the requirements of cells in these stages, cell culture conditions that are more physiologically relevant may be developed.

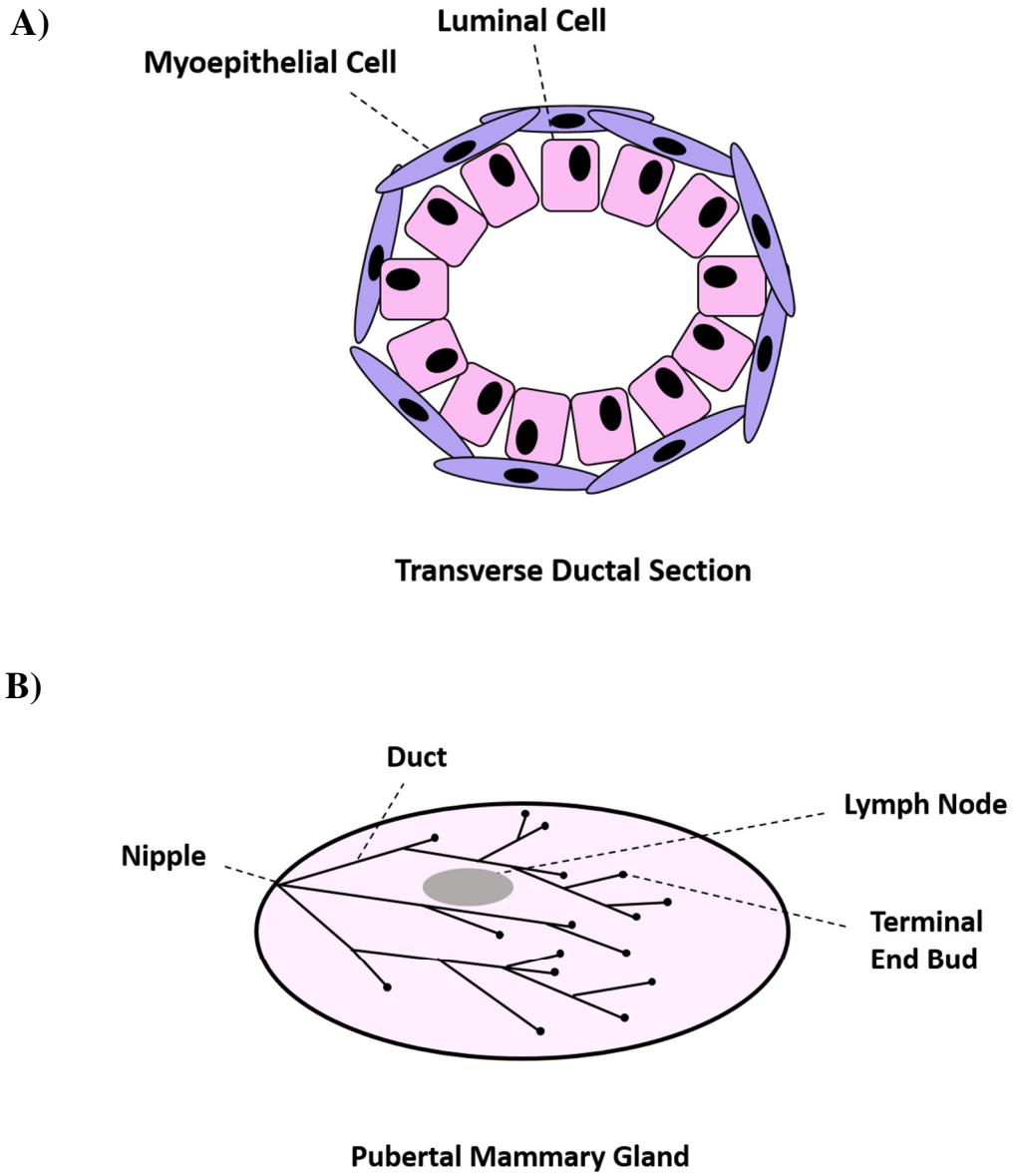


Figure 1.2: Cellular organisation of the mouse mammary gland.

Two alternative views of cellular organisation in the mouse mammary gland. **A)** Transverse section of the ducts found in the mammary gland with luminal cells and outer myoepithelial cells. **B)** Depicts the pubertal gland with the growing epithelium and terminal end bud invading the fat pad.

1.3.2 Cellular Hierarchy in the Mammary Gland

A thorough understanding of the developmental biology of an organ is important, not only for greater insight into the working of the organ itself, but particularly for the understanding it provides into the diseased state. Delineating the cellular hierarchy of the mammary gland remains an important hurdle in the understanding of the tissue itself and also how breast tumour subtypes arise. Although this remains a contentious topic, particularly on the point of mammary stem cells (MaSCs), recent targeting and sequencing advances have helped researchers to better conceptualise the differentiation trajectory of cells in the mammary gland. Proper understanding of this requires study at a single-cell resolution, ideally using techniques such as lineage tracing or single-cell RNA sequencing. When used together, these can be powerful tools, as exemplified in recent reviews [93,94].

In all tissues of the body, cell types are adapted to both repair injury and to maintain normal homeostasis [95,96]. Often these processes require the presence of a stem cell or potent progenitor. In understanding the dynamic and cyclical nature of mammary gland remodelling, it is easy to hypothesise that much of this could be accounted for by the presence of some type of MaSC. Indeed, many of the traditional experiments aimed at understanding cellular lineage in the mammary gland sought to define populations as 'stem' or 'progenitor'. These experiments investigated whether a population residing in the mammary gland could be defined using the traditional parameters of a stem cell - an unspecialised cell with an indefinite capacity for renewal [97] - as opposed to a progenitor with more restricted potential. Historically, the gold standard *in vivo* assay used was a transplantation assay, where the endogenous epithelium is removed from the mouse fat pad resulting in 'clearing' of the residing mammary epithelial cells so that tissue fragments may be transplanted. The extent of ductal outgrowth by the transplanted tissue was then used as an indicator of cellular potency [98]. These ductal outgrowths were identified as being able to repopulate glands of other mice in serial transplantation assays, implying the presence of long-lived stem cells throughout the ductal tree [98].

The further subdivision of mammary tissue into distinct populations based on cell surface markers enabled a more detailed classification of lineages in the mammary gland [99–101]. This allowed the fractionation of the mammary gland using flow cytometry and has come to be the foundation of much subsequent work on lineage. These experiments focussed on the isolation

of cell types distinguished by their surface markers and their transplantation into recipient mice, as explained previously. Researchers could then assess the extent of outgrowth, analysing the particular cell types present. Cells with the capacity to repopulate the mammary gland were found to be lineage negative ($CD24^{+}$, $CD29^{high}$, $CD49f^{high}$, $Sca1^{low}$). These cells were found to spatially reside in the basal compartment [99,100]. Conversely lineage negative cells found in the luminal compartment characterised $CD24^{+}$, $CD29^{low}$, $CD49f^{low}$, were unable to repopulate the gland. More recently, EpCam and CD49f have been used to separate basal and luminal cells and this provides a comparable enrichment of cells able to repopulate the mammary gland [102].

Although these early experiments do claim to provide evidence of a mammary stem cell compartment, it is important to consider that these experiments are particularly contrived. This is to say that none of these properties have been demonstrated sufficiently *in vivo* in the mammary gland. Authors of a recent review propose a new definition of MaSCs [103]. The authors suggest that these stem cells should not be defined by their capacity to repopulate a cleared fat pad or their cell surface markers, but rather the potential to demonstrate a bipotent nature with unlimited self-renewal (and replicative potential) *in situ*. Essentially, they state that, taken out of context, cells of the mammary gland may be induced or reprogrammed by their microenvironment which, while indicating an intrinsic plasticity, does not necessarily prove the capacity of these cells to behave as such in the normal gland. Therefore, to define these populations as stem cells would be inaccurate. As well as new terminology for MaSCs, the authors stress the need for accurate distinctions when defining stem and progenitor populations. They suggest that the progenitors must be long-lived with high levels of self-renewal and replicative potential, whilst being lineage restricted [103]. It is clear that work in defining these populations is still ongoing but the rate of progression within the field has been rapid.

The functional studies above represent the classical understanding of cellular hierarchies in the mammary gland. Current and future understanding is more reliant on key technological advances, particularly high throughput sequencing capabilities and lineage tracing. However, despite the relative ease with which researchers may now undertake such projects, it seems that there are still key questions unanswered.

Lineage tracing represents another technique commonly used to assess the contribution a cell type might make to tissue development and homeostasis *in situ*. Such an approach allows the mapping of cell fate in its native environment. One specific approach represents a 'population-based' method where a predetermined cell type is traced through the selection of a highly specific promoter [104]. Lineage tracing in this manner is an inducible mechanism whereby upon the administration of tamoxifen, a lineage-specific promoter drives Cre-recombinase expression. Active Cre-recombinase then enables the recombination of *loxP* sites to remove a stop codon and activate gene expression. This causes activation in a reporter gene linked to a ubiquitous promoter and the cell, as well as all of its progeny, are labelled. This labelling lasts for the whole lifespan of the cell, allowing the reliable mapping of cellular populations depending on chase length[104,105]. Stochastic models of lineage tracing have also been used to provide a more unbiased approach to mapping cellular fate. Unlike population-based approaches, these studies involve labelling in a manner independent of a lineage specific promoter, thereby allowing a more native approach and more general conclusions about cell fate [106,107].

Work on lineage tracing in the mammary gland has aimed to follow the progeny of MaSCs and other progenitor cells to uncover more about their dynamics, particularly whether any cells have the capacity to replenish both the basal and luminal compartments. Lineage tracing has been applied to understand whether the repressor Blimp1 is activated in luminal progenitor populations. The repressor protein is thought to govern cell fate decisions in both embryonic and adult tissues and is highly induced in alveolar cells during pregnancy [108]. The results showed that Blimp1^{+} cells, first noticed in the embryonic mammary gland, are maintained as a constant pool of long-lived alveolar progenitors which aid the accompanying remodelling and expansion of alveolar cells in pregnancy [109]. There were rare Blimp1^{+} lineage restricted luminal progenitor cells that maintained their identity and demonstrated roles in duct formation, homeostasis and alveogenesis during pregnancy [109]. The Blimp1^{+} cells gave rise to a highly proliferative subset of luminal progenitors identified as Elf5^{+} ERalpha^{-} PR^{-}, which are those responsible for alveogenesis-induced expansion. These cells are also long-lived, surviving the involution stages post-weaning.

Another key question for lineage tracing in the mammary gland has been to identify whether there is a common progenitor from which ER^+ and ER^- lineages are derived, or whether the ER^+ lineage is maintained by a multipotent MaSC. Van Keymeulen et al. addressed this question using doxycicline-inducible ER-rtTA mice to trace the lineage of ER^+ luminal cells to study fate and long-term maintenance in the mammary gland [110]. Authors reported that ER^+ cells are maintained by lineage restricted cells and that these cells exclusively contribute to the expansion of the ER^+ lineage during puberty and maintenance in the adult gland. In another study by Wang et al., the authors report the determination of ER^+ and ER^- cells as independent lineages maintained by separate stem cells [111]. Using lineage tracing, the authors found that SOX9-expressing cells contributed to the development and maintenance of the ER^- luminal cells. PROM1 expressing cells gave rise to ER^+ luminal cells, demonstrating that these cell types my indeed be maintained by their own pools of progenitor cells [111]. The authors report that the cells also sustain their lineages after pregnancy, indicating the long-lived nature of these cells. This therefore further adds to the evidence of lineage restricted progenitors, which maintain requisite lineages in the mammary gland as postulated by some of the scRNA-seq data on cells of the same origin.

An understanding of the cellular dynamics of contributing cells in the lobuloalveolar structures is also important for cellular hierarchy. Using the 'slippage' $R26^{[CA]30EYFP}$ model, Davis et al. analysed lineage tracing in lactating mice. The authors report that clonal progeny contributes exclusively to either the luminal or basal lineages [106]. The distribution of these clones is sporadic at either branching ducts or alveoli. Upon analysis, the authors note that some regions were highly labelled whilst others were labelled at a lower density. This, the authors suggest, is indicative of the presence of multiple lineage restricted stem cells and the admixing of their progeny [106]. The lactation data revealed the presence of different subpopulations of alveolar cells. The authors postulate that the presence of a pool of alveolar stem or committed cells is likely due to the importance of alveolar expansion during pregnancy and the advent of successive pregnancies over a life time [106]

More recently, in aiming to functionally characterise basal and luminal cell lineages, authors Centonze et al. identified a degree of placticity within the basal cells of the mammary gland upon ablation of the luminal cell population[112]. The authors used a mouse model that enabled

the lineage tracing of basal cells and the ablation of luminal cells using doxycycline inducible diphtheria toxin: K5CreER/Rosa-tdTomato/K8rtTA/TetO-DTA. Upon the removal of the luminal cells, lineage tracing enabled the visualisation of tdTomato in cells in of both basal and luminal origin [112]. In a similar experiment the authors generated a mouse model to trace the lineage of luminal cells in the absence of basal cells: K14rtTA/TetO-DTA/K8CreER/Rosa-YFP. After the administration of doxycycline the authors did not observe any such replacement of the basal compartment from the remaining luminal cells [112]. Taken together this suggests that under regenerative conditions basal cells are more plastic than luminal cells. The authors further characterised this phenomenon using organoids to understand whether this was an intrinsic property of the cells in the absence of stroma and inflammatory cells. The authors performed lineage tracing and lineage ablation in organoids composed of both basal and luminal cells and 72 hours after administration of doxycycline and the subsequent apoptosis of luminal cells, 20% of luminal cells had been replaced by tdTomato positive basal cells [112]. This suggests that in the absence of stroma and inflammatory cells, luminal cell ablation is enough to promote basal cell multipotency. This further suggests that the presence of luminal cells actively restricts basal cell multipotency which is an intrisnsic mechanism and not dependent on stomal or immune cells as has been demonstrated in other organs [112]. The authors couple these functional studies with genomic analysis, performing bulk and single cell RNA sequencing on cells from these mouse models. The authors conclude that the regulation of multipotent vs unipotent basal cell fate are regulated by cell-cell communications that control the fate of stem cells. Instructing them the need to generate the cell lineages required in the tissue. The authors identify TNF as a key factor secreted by luminal cells, demonstrating under physiological conditions that this cytokine restricts the multipotency of basal stem cells in the mammary gland [112]. The authors also demonstrate the importance of the Notch, Wnt and ErbB pathways for the activation of basal stem cell multipotency in the absence of luminal cells and highlight the importance of further characterisation of these mechanisms for the parallels seen in tumorigenesis [112].

Single cell RNA sequencing (scRNA-seq) experiments have been used to compliment functional experiments and to predict lineage hierarchies. The dynamic nature of the gland has made it harder ensure that all stages are sufficiently captured, including all the cell types responsible for remodelling. It is advantageous therefore to capture several states of the gland by collecting cells at different time points for sequencing. In addition, there still remains a lack

of uniformity in sample processing and computational analysis which can lead to conflicting findings [104]. Nonetheless, the insights gained from these analyses have been extremely useful in delineating the differentiation trajectory in cells of the mammary gland. Indeed, this method was employed in the first scRNA-seq experiments of the mammary epithelium [113,114]. Both authors characterised cells from four different developmental stages. The stages were different between the studies, with Pal et al. focussing on pre-puberty and mid-puberty, and Bach et al. placing greater value in understanding the post pregnancy state with lactating and post-involution time points. The shared stages analysed were adult virgin or nulliparous and mid-gestation. The scRNA-seq analysis by Bach et al. revealed 15 clusters that were based on gene expression with no cluster identifiable by the expression of a single gene. The clusters were broadly divided into 11 luminal and 4 basal. The data indicated mammary epithelial cells display more of a differentiation trajectory as opposed to defined clusters. They suggest that common luminal progenitors give rise to restricted intermediates, restricted alveolar and hormone-sensitive progenitors. This was later supported by findings of another group [115] indicating that after birth there is a level of lineage restriction in the cells. The basal clusters were also further subdivided into differentiated, myoepithelial, and stem cell-like basal cells. However, the authors note that there were too few basal cells captured to produce a definitive hierarchy [114].

In their study Pal et al. had somewhat different conclusions, reporting the presence of basal gene expression throughout all developmental stages, whereas luminal expression was only detected at puberty through to adult hood. The authors also report a mixed lineage cluster [113] within the basal cluster which was not reported in either of the two previous studies [114,115]. Taken together these studies provide greater insights into the differentiation spectrum in the mammary gland. There are hints of potential plasticity in these cells, however more work is undoubtedly required to definitively prove whether such cells exist. This remains one of the biggest unanswered questions in the field owing to the recent work done to uncover the cellular origin of basal-like breast cancer in luminal cells [116,117].

In the human breast, scRNA-seq in cells isolated from reduction mammoplasties identified the existence of three distinct epithelial cell populations through unbiased clustering analysis [118]. These were defined as basal, luminal secretory and luminal hormone responsive. The authors

define the two luminal clusters as luminal progenitors and mature luminal cells, respectively, and postulate the presence of MaSCs within the basal cluster [118]. Interestingly, the authors report the detection of replicating cells in all three of the clusters, implying that the clusters may be self-maintained [118]. By extension, this could imply that the clusters have their own stem or progenitor cell populations sustaining these compartments in the human mammary gland.

Taken together, these data highlight the importance of lineage tracing in deconvoluting hierarchical relationships in the mammary gland. Combined with scRNA-seq data, these are powerful tools which have enabled greater insight into the cellular dynamics of the cyclical remodelling undergone by the mammary gland. Important conclusions from these publications include the presence of lineage restricted progenitors, including dedicated progenitors for ER^+ and ER^- lineages (figure 1.3). In addition, basal cells may also be maintained by a lineage restricted progenitor. Whilst the most contentious area of the mammary hierarchy remains the question of multipotent progenitors and MaSCs, these may participate in tissue remodelling during involution [119] (figure 1.3). However, some discordant findings indicate that more work must be done to draw stronger conclusions. Indeed, the rate of change within the field has meant that models of cellular relationships are constantly reviewed and refined.

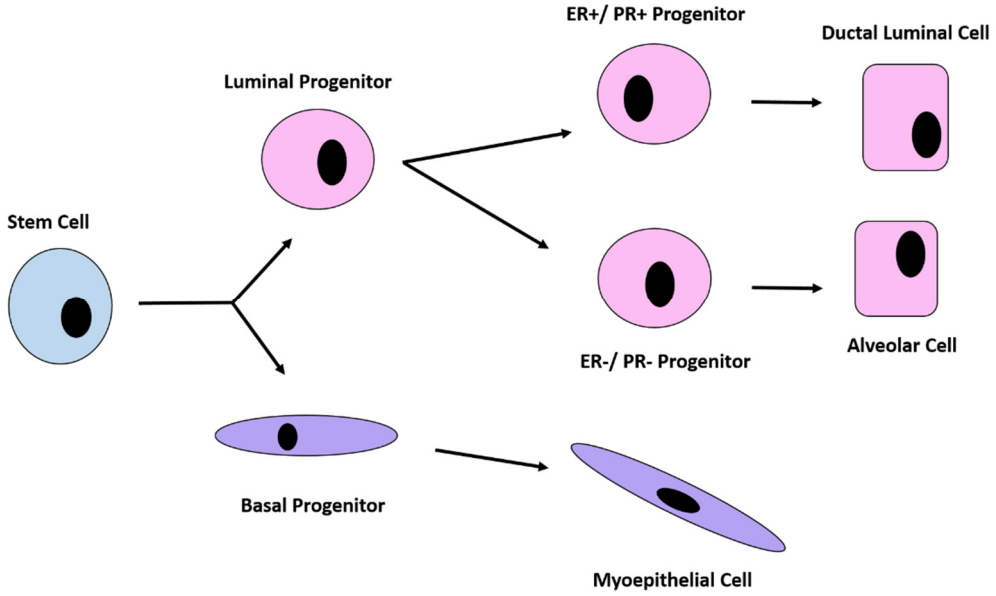


Figure 1.3: A proposed model for cellular hierarchy in the mouse mamary gland.

One model of cellular hierarchy in the mouse mammary gland. Multipotent MaSCs may give rise to progenitors responsible for postnatal population of the glands. Luminal ER $^{+}$ and ER $^{-}$ lineages as well as basal cells are thought to be sustained by their own pools of dedicated unipotent progenitor cells.

1.4. Oesophageal Development

The primary role of the oesophagus is to transport food from the mouth to the stomach. As such it is a muscular tube and the only part of the gastrointestinal tract that lacks any absorptive, metabolic or digestive capacity [120]. The oesophagus is derived from the anterior portion of the foregut [121] which also gives rise to the trachea, lung and stomach. During embryonic development the oesophagus and the trachea share a lumen [122]. Gene expression patterns differ between the dorsal and ventral regions which represent the oesophageal and tracheal regions respectively. The transcription factor *SOX2* is highly expressed in the dorsal foregut (prospective oesophagus), whereas NK homebox 1 (*NKX2.1*) is expressed in the ventral foregut (forming the trachea) [123,124]. Canonical Wnt/β-catenin signalling is a critical mediator of embryogenesis and also plays a role in foregut morphogenesis. Signalling activity from this pathway is active in the ventral side of the unseparated foregut tube. Dorsal Foregut division, the separation of the dorsal and ventral foregut components, forms the basis of gastro respiratory development [125].

After the process of separation and elongation, the endodermal and mesenchymal development is different for both the oesophagus and trachea. The trachea will go on to develop pseudostratified columnar epithelium, whereas the oesophagus will become a stratified squamous epithelium [120]. The oesophageal mesenchyme will develop the smooth muscle layer which encases the oesophagus to ensure the peristaltic movement of food through to the stomach. In order for the oesophageal epithelium to become stratified however, the already present columnar cells must undergo a conversion [120,122].

After foregut separation is completed, Sox2 remains highly expressed in the epithelial cells of the oesophagus [126]. This transcription factor is required for the stratification and differentiation of progenitor cells during the development of the oesophagus [126,127].The reduction of Sox2 protein levels has been demonstrated to block the formation of the stratified epithelium, instead mutants displayed a layer of columnar cells similar to those seen in mucus metaplasia, a pathology related to Barret's oesophagus [126]. This work highlights the developmental importance of this transcription factor in the oesophageal epithelium.

The transcription factor Trp63 (a member of the p53 family) specifically the ΔNTrp63 isoform, is also critical for the generation of the stratified squamous epithelium [128]. The loss of this transcription factor leads to the failed conversion of columnar cells to stratified squamous epithelium [128]. Bmp signalling also plays an interesting dual role in oesophageal development. During the embryonic stages E11.0 to E14.5 at which the epithelium is proliferating and stratifying, the tissue remains negative for Bmp activity [129]. However, on E15.0 Bmp signalling is detected in the oesophageal suprabasal cells [129]. These two stages correlate with Bmp function in the regulation of epithelial development, the first stage where Bmp is blocked through the expression of noggin (an inhibitor of Bmp) to allow stratification and the second where Bmp signalling promotes differentiation [129]. The epithelium gradually becomes stratified so that there are around 4 layers [130,131] of epithelial cells. The basal layer – what was the layer of columnar cells - then initiates expression of K5 and its binding partner Keratin 14, whilst simultaneously losing expression of Keratin 8 and 18. The upper layers do retain low levels of these cytokeratins but expression is lost postnatally [131]. The suprabasal cells then begin to express the pairs Keratin 1 and 10, and Keratin 4 and 13, whilst simultaneously losing their capacity to proliferate. The cells differentiate upwards with the topmost layers expressing the squamous cell differentiation marker involucrin. In a phenomenon specific only to the murine oesophagus and similar to skin morphogenesis, cells undergo a stage of enucleation, losing nuclei and becoming keratinised [122].

1.4.1 Oesophageal Tissue Maintenance and Hierarchy

The exact dynamics of tissue maintenance as well as contributing lineages of the oesophageal epithelium remain a contentious topic within the field. Differing paradigms have been proposed in recent years [95,127]. The murine oesophagus is a squamous epithelium that consists of layers of keratinocytes. Proliferation in the epithelium is confined to the basal layer, which contains cells that are differentiating and self-maintaining throughout the lifespan of the tissue [132]. Differentiated cells move to the suprabasal layers and become keratinised before being shed off into the lumen (See figure 1.4). In terms of tissue dynamics, it is critical that when dividing, cells generate proliferating and differentiating daughters to maintain the equilibrium of the tissue at homeostasis. An imbalance of this equilibrium could result in a loss of tissue integrity or an increase in proliferation which could then lead to malignancy [133].

Understanding these complex dynamics has been the topic of much work, with researchers broadly trying to identify whether all basal cells of the oesophagus are functionally equivalent or whether a dedicated subset of stem cells contribute to tissue maintenance. Work began on cells of the oesophageal epithelium in the late sixties and identified the localisation of proliferating cells to the basal layer using thymidine incorporation [134]. These studies argued against asymmetrical division, claiming that dividing cells were equipotent and that cellular commitment, and therefore exit of the basal layer, was stochastic. In 2002, a transit amplifying model of stems cells in the oesophagus was proposed [135]. This marked the beginning of an effort to identify discrete stem cell populations within the oesophageal epithelium trying to address whether there is heterogeneity in these basal cells to which proliferation is combined.

In a study by Croagh et al., the authors pose a model whereby the epithelium is maintained by a slower cycling, self-renewing stem cell population that generates a short-lived transit amplifying cell that terminally differentiates after only a few rounds of division [136]. The authors tried to isolate cells using markers previously used for the identification of stem cells in the epidermis [137]. However, efforts to assess the enhanced replicative potential of these cells *in vitro* failed. More recent work on this model by DeWard et al. claimed to have identified cell surface markers specific to distinct populations in the basal epithelium [127]. Their apparent stem cell potential was displayed using organoid forming efficiency *in vivo.* The authors also describe *SOX2* as an important marker for organoid formation and self-renewal in oesophageal

basal cells, identifying cells expressing the highest levels of the markers integrin alpha 6 (or CD49f) and p75 as a putative stem cell population based on organoid forming efficiency [127]. This work is some of the strongest evidence for heterogeneity amongst basal cells of the oesophageal epithelium. However, the application of lineage tracing techniques to address the question of basal cell heterogeneity within the oesophageal epithelium has identified a single progenitor model of tissue maintenance [95,96].

The authors Doupé et al. applied quantitative lineage tracing to label individual basal cells over the course of a year, concluding that a single progenitor population maintains the tissue, dividing stochastically to generate proliferating and differentiating daughters with equal probability [95]. This seems to support early work based on label incorporation, calculating the dilution of doxycycline induced Histone-2B-GFP fusion protein, which concluded that no slow cycling cells were present [138]. Using the same Histone-2B-GFP model, Doupé et al. confirmed the original observations of all basal cells dividing at the same rate. In addition, in injury studies using the same label retaining assays, Doupé et al. revealed that widespread activation of progenitor cells around the wound produced an excess of proliferating daughter cells [95].

A more recent lineage tracing study confirms these findings, reporting that basal cells divide on average 2.9 times every 2 weeks, describing the likelihood of basal cell division producing proliferating and differentiating daughters as on average, equal [96]. The authors conclude that the single progenitor model is the paradigm of epithelial maintenance for the oesophagus by applying quantitative lineage tracing mouse models [96]. Progenitor cells are described as sharing the basal layer with post mitotic cells that are in the early stages of differentiation (figure 1.4). These post mitotic cells are transiently maintained in the basal layer before they become stratified Suprabasal cells. The authors conclude that individual basal cells have the potential to divide or differentiate and exit the basal layer [96]. This decision is stochastic but it is skewed through spatial coordination between neighbouring and sibling cells [96]. This means the probability of each outcome is balanced but this can be tilted at any point for individual cells depending on their local niche.

Overall, this work suggests that the oesophageal epithelium is likely to be maintained by a single progenitor population. Despite some evidence of heterogeneity in cells of the basal layer, quantitative lineage tracing appears to demonstrate the equal probability of any given division to generate proliferating or differentiating daughters. Whilst it appears that post mitotic cells are transiently retained in the basal layer, this could account for a degree of heterogeneity observed in some of the data. A comprehensive scRNA-Seq analysis of the murine oesophageal epithelium remains to be published and these experiments would undoubtedly shed more light on the cellular hierarchy of the tissue.

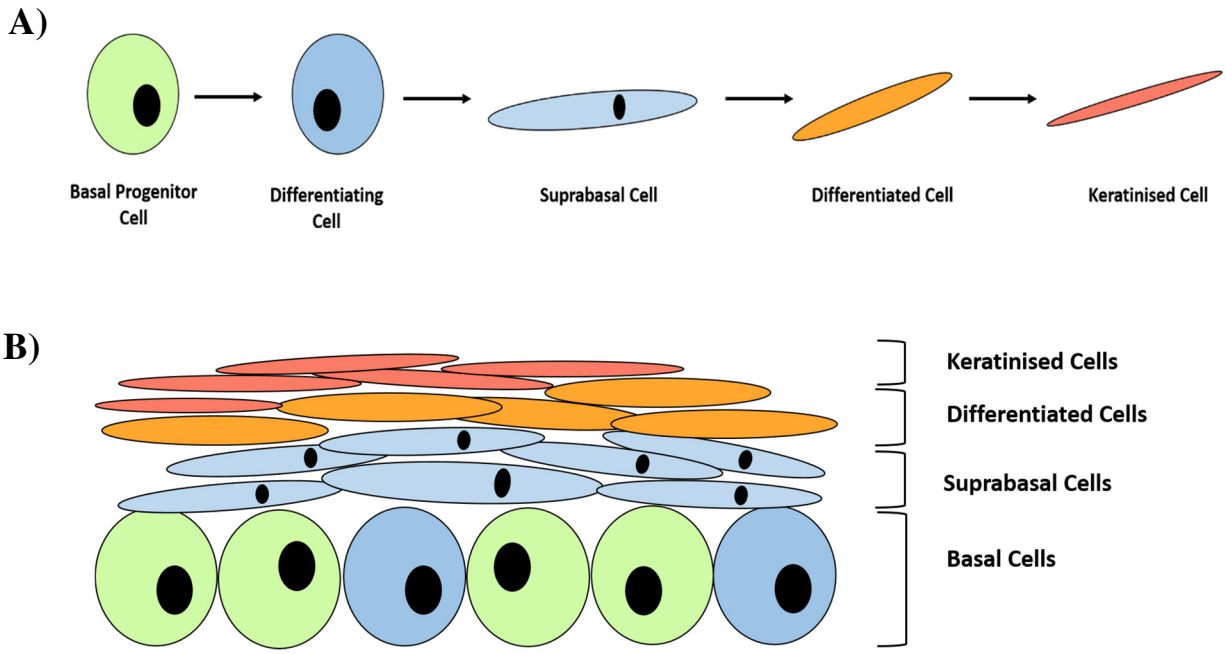


Figure 1.4: Cellular organisation and proposed hierarchy of the mouse oesophageal epithelium.

The cellular organisation of the mouse oesophageal epithelium and proposed cellular hierarchy as suggested by Piedrafita et al.[96]. **A)** Basal cells give rise to proliferating or differentiating daughters with equal probability on average. Differentiating cells stratify and become suprabasal before finally becoming keratinised and shed into the lumen. **B)** The cellular organisation of the oesophageal epithelium with basal progenitor cells located along the basement membrane. Differentiating or post mitotic cells are transiently maintained in the basal layer before becoming stratified suprabasal cells. After which a process of keratinisation occurs and cells are shed into the lumen.

1.5. Models of Gene Expression/function.

To fully understand gene function requires the study of genes in interaction with the complex physiological mechanisms of an organism. For this, *in vivo* research is the gold standard, and offers a vital resource for studying the causal link between genotypes and phenotypes in both healthy and diseased tissue. In the study of cancer and indeed wider disease biology, varying forms of mouse models are commonly used to understand these interactions. Mice have specific molecular, cellular and anatomical characteristics in common with humans which make them useful for cancer research [139]. In addition to small size, short gestation periods and relatively easy colony generation, mouse models also offer logistical advantages in research. Models have ranged from traditional xenograft models with immune-deficient mice to allow the subcutaneous transplantation of tumours, to the engineering of the blastocyst of mouse embryos to generate transgenic mouse lines [139].

The production of transgenic mouse lines can be time-consuming, often taking several years to establish the desired mutation in embryonic stem (ES) cells, generate colonies and validate the model. However, this method is still widely used in preclinical studies. This method requires very efficient targeting of the ES cells, at least two generations of crosses to obtain the desired genotypes which may take some years to achieve in addition to being expensive [139]. The national centre for replacement refinement and reduction of animals in research (NC3R) aims to reduce animal use in biological research, funding research into the development of technology to replace the use of animals. This particular method of mouse model generation does not align with the 3Rs principles, as it is a wasteful process, requiring the production of many generations before the correct genotype is reached. In light of this, much effort has gone into the generation of *ex vivo* culture and development of tissue culture methods, striving to reduce the number of animals used in research without compromising on scientific standard or impact.

Cell culture aims to comprehend cellular behaviours and mechanisms in many different settings. The culture of cells is not a new model of gene function, with the first experiments taking place in 1907 [140]. Traditional cell culture centres around the cultivation of cells in a 2D monolayer. Cultivation of cells in this manner forged a path for new assays as well as drug trials and screening, representing major scientific advances. The specific needs of different cell

types gave rise to the development of different culture techniques and media for supplementation [140]. The formulation of the media used for cell culture has an immediate effect on the development, nutrition and differentiation of cells, which all correspond to changes in gene expression pattern. Tissue culture practices are also important for the maintenance of cells. For example, passaging techniques have repercussions for gene expression and may therefore influence experimental outcomes. Similarly, media formulations can be supplemented with growth factors to better nourish cells. However, the native 3-dimensional (3D) structures in which cells are normally found are often continuously perfused by the blood stream, something which cannot be achieved with monolayer culture alone [140]. Similarly, tissue-specific architecture and cellular interactions, as well as biochemical and mechanical signals, are lost in such rudimentary conditions. Therefore, many of the efforts to develop cell culture techniques have aimed to mimic native environments more closely, taking into account the shortfalls of 2-dimensional (2D) culture.

The 3D culture of cells *in vitro* has represented an important advance in studies of this nature, enabling the growth of cells within a 3D matrix or scaffold to more closely mimic *in vivo* conditions. Frequently used in these cultures are hydrogel scaffolds to ensure the suspension of cells in the culture [141]. Matrigel matrix is an example of a scaffold which has been used for the 3D culture of many cell types [142]. This scaffold is particularly advantageous over 2D cell culture as it allows the maintenance of cell-cell adhesions as well as tissue-specific architecture. The 3D system is therefore able to exhibit biochemical and morphological features more closely related to the *in vivo* state.

1.5.1 Primary 3D culture of the murine mammary gland

The culture of 3D mammary cells were amongst the first cultures of this nature to have been attempted [143]. Cultured on floating collagen gels, mammary epithelial cells were found to maintain expression of differentiation markers as well as self-assembling into structures reminiscent of acini. Such culture conditions suggested that the extra cellular matrix (ECM) has an important effect on gene expression, and was the starting point for further research into cellular microenvironment. The composition of the ECM has therefore been the subject of intense research to identify the most physiologically relevant approach for 3D culture of the mammary gland. This work sought to understand the influence of microenvironment on

cellular morphology, representing some of the most important studies of ECM regulation. More recently, the adaptation of a laminin-rich ECM derived from Engelbreth-Holm-Swarm sarcoma (Matrigel) for 3D culture was found to allow epithelial cell polarisation and morphogenesis native to the mammary gland [142,144]. In addition to important decisions on ECM, mammary organoids can be cultivated in different ways. Broadly speaking, organoids from primary material can be cultured within a matrix from tissue fragments or from single cells [145]. There are also important examples of primary mammary co-cultures and even more technologically advanced examples, such as bioprinting and microfluidics chambers [146].

In addition to ECM composition, as in any cell culture method the media supplementation is also an important consideration. There have been many formulations of 3D mammary organoid culture media [145,147,148]. Although the presence of some factors are consistent in the culture, such as EGF and FGF, researchers have tried to optimise these further to produce robust culture systems that promote long term cultivation whist sustaining physiological phenotypes [148].

For murine mammary organoid culture, Jard**é** et al. adapted a medium formulation with the purpose of extending the culture period for up to a period of 2.5 months [148]. Authors describe the use of factors such as Neuregulin 1, a member of the EGF family of ligands, as well as R-spondin1 to stimulate the Wnt pathway. This medium formulation represents one that more closely resembles the recent epithelial 3D culture mediums of other tissues [148]. The authors describe the extended growth and maintenance with distinct basal and luminal compartments, of which the luminal compartment contained ER^+ and ER^- subtypes. Organoids were found to have full regenerative potential when implanted back into the cleared fat pad of the mouse [148]. The authors also report the cultivation of both tissue fragments and single cells as organoids, claiming to have developed a long-term organoid culture system for the murine mammary gland.

Soon after, Jamieson et al. continued to develop the 3D organoid model for the mammary gland [147]. The authors supplement extensively, in comparison to Jard**é** et al., including factors such as R-spondin2, Wnt3a to ensure a more robust culture condition was developed. The authors describe organoids cultivated from single cells of either basal or luminal origin [147]. The

importance of other factors such as FGFs have also been recently defined for the in vitro culture of mammary organoids. Indeed, 3D organoids are a powerful tool to study the developmental biology of the organ. A model of lactation and involution has recently been described [149].

1.5.2 Primary 3D culture of the murine Oesophagus.

The cultivation of oesophageal epithelial cells in 3D represents an important tool to study the biology of the tissue *ex vivo*. These cultures have been used to aid the study of processes such as cell proliferation motility, differentiation and stress response. In response to the original mammary experiments undertaken by Li et al.[150] to culture cells within a matrix of collagen I or Matrigel, researchers [151] began to culture oesophageal epithelial cell aggregates in these matrices. Authors defined the spheroid culture for the study of tumour microenvironment and cell adhesion [151]. These cultures formed the first rudimentary conditions for cells grown in 3D, with serum free media supplemented with factors such as EGF [152]. Such experiments were particularly useful for the study of tumour spheres, looking at cancer stem cells, tumour initiating cells and therapeutic response.

The robust culture of the oesophageal epithelium as 3D organoids with complete differentiation gradients has only recently become possible. The first 3D organoids were established from single cells by DeWard et al. As mentioned, the organoid system was developed by the authors to address questions about cellular hierarchy in the murine oesophageal epithelium [127]. Primary basal epithelial cells were seeded in Matrigel and cultivated using a media similar to that used for the culture of intestinal organoids. The resulting structures appeared to contain an outer layer of basal cells as well as inner suprabasal cells, complete with a keratinised core. Important factors to facilitate the establishment of organoids were used for cultivation, such as TGF-β inhibitor, p38 mitogen activated protein kinase inhibitor and rho-associated kinase inhibitor [127]. Media were further supplemented with noggin, EGF, Wnt3A, R-spondin and gastrin, factors used to recreate the cellular microenvironment *in vivo*. The authors use the serial passage of organoids as an indicator of stem cell potential in *SOX2* expressing cells of the oesophagus [127].

Contrary to DeWard et al., authors Kasagi et al. demonstrated the redundancy of such factors after defining a simplified media formulation for the same culture. The authors use only EGF and bovine pituitary extract (BPE) as well as Rho-kinase inhibitor to generate sufficiently differentiated organoids from single cells [153]. Authors claim many factors from DeWard et al.

are dispensable, demonstrating organoid cultivation from oesophageal cells of both murine and human epithelia.

1.6 Thesis Outline

In this thesis I will explore the use of 3D organoids and lineage tracing to further understand the role of *Bcl11a* in tissues of the mammary gland and the oesophagus. In chapter three, I will outline the setup of oesophageal organoids and demonstrate the use of these cultures to identify a potential progenitor cell function for *Bcl11a* in in the oesophageal epithelium. In chapter four, I will show the use of a lineage tracing mouse model ($Bcl11a^{CreERT2}$) to validate *in vitro* data. I will demonstrate that *Bcl11a* marks long lived cells in the oesophagus, with labelled cells identified in each layer of the squamous epithelium. I will demonstrate how these cells are seen to expand over time, ultimately suggesting a role for *Bcl11a* in progenitor cells of the murine oesophageal epithelium. In chapter five, I will describe the setup and optimisation of the 3D culture system of the mammary gland from single cells of the luminal progenitor compartment. I will apply this system to gene function studies, using organoids cultivated from both *BCL11A* overexpression ($BCL11A^{OE}$) and *Bcl11a* knock out ($Bcl11a^{KO}$) mouse models.

Chapter 2: Experimental Procedures

2.1 Mouse Models

All mice used in this project were maintained at the Sanger Institute. Housing and breeding of mice and experimental procedures were performed according to the UK 1986 Animals Scientific Procedure Act and local institute ethics committee regulations. Mice were culled in accordance with the Home Office protocols and in all cases method of culling was administration of carbon dioxide followed by cervical dislocation.

In this study wild type (WT) refers to non-transgenic C57BL/6 (Black 6) mice. Out of the transgenic mice used in this study one mouse model is a *BCL11A* overexpression mouse model. The *BCL11A-XL* coding sequence was knocked into the *Rosa26* locus and downstream of a Lox-Stop-Lox (LSL) cassette and driven by the *CAGG* promoter. Downstream of the *BCL11A* cDNA is *eGFP* under the control of an internal ribosomal entry site *(IRES)* (figure 2.1). In the presence of an active Cre recombinase the LSL cassette is removed and *BCL11A* overexpression is driven by the CAGG promoter. This model will be referred to as $BCL11A^{OE}$ herein. The other mouse model used in this study enables the conditional deletion of *Bcl11a* in a Cre dependent manner. $Bcl11a^{KO}$ mice were generated using *lox*P site upstream of exon 4 of *Bcl11a* (figure 2.1). This model will be referred to as $Bcl11a^{KO}$ herein. For both these lines I used the inducible *CreERT2* which is knocked-in at the *Rosa26* locus which enables temporal induction of overexpression or deletion of *Bcl11a* upon 4-OHT treatment. Genotyping was undertaken using the primers specified in table 1.

Lineage tracing mice contain *CreERT2* in the 3' UTR of the *Bcl11a* gene. The tdTomato gene is preceded by a *LoxP* flanked stop codon driven by a CAGG promoter in the *Rosa26* locus, thus, allowing tdTomato expression upon tamoxifen administration in only the cells also expressing *Bcl11a* (figure 2.2). This model will be referred to as $Bcl11a^{CreERT2}$ herein. Control mice for the lineage tracing experiments were *Bcl11a* wildtype without *CreERT2* and contained the tdTomato construct in the Rosa26 locus (figure 2.2). This model will be called $Bcl11a^{WT}$ herein. Genotyping was undertaken using the primers specified in table 1.

2.1.1 Mouse Genotyping PCR

Genotyping was undertaken using ear notches taken by research technicians at the Sanger Institute. DNA extraction was performed by incubating the ear notches in 50 µl lysis buffer and heated to 95 °C for 1 hour using a tetrad thermal cycler. The Lysis buffer contained 100 mM Tris-HCL (pH 8.0) with 50 mM EDTA and 1% (w/v) SDS (Sodium dodecyl sulfate). This solution was then mixed with 50 µl of a neutralisation buffer composed of 1 M Tris (pH 7.4) and 1.5 M NaCl. This solution was then diluted 1:10 for amplification using the polymerase chain reaction (PCR). DNA primers were ordered bespoke from Sigma and reconstituted to 100 µM (table 1). A 10 µM stock solution was then prepared for the forward and reverse primers using nuclease free H_2O (Qiagen). The PCR was undertaken using 10 µl ReddyMix PCR master mix (Thermo Fisher Scientific), 2 µl primer stock solution, 7 µl H_2O and 2 µl genomic DNA per reaction. In the case of the reaction to amplify Cre^{ERT2} site in the lineage tracing colony, Phusion PCR master mix was used (Thermo Fisher Scientific). The reaction was then run on a thermal cycler using the following program (specific to the primers used).

PCR program for: Rosa26, Cre ERT2 (*BCL11A*^{OE} and *Bcl11a*^{KO}), tdTomato site amplification.

Incubate at 95 °C for 2.0 minutes

Incubate at 95 °C for 0.25 minutes
Incubation at 58 °C for 0.35 minutes
Incubate at 72 °C for 1.0 minutes
} Cycle 35 times

Incubate at 72 °C for 5.0 minutes

Hold at 4 °C

PCR program for: BCL11A (*BCL11A*^{OE} colony) site amplification.

Incubate at 95 °C for 2.0 minutes

Incubate at 95 °C for 0.25 minutes
Incubation at 49 °C for 0.35 minutes
Incubate at 72 °C for 1.05 minutes
} Cycle 35 times

Incubate at 72 °C for 5.0 minutes

Hold at 4 °C

PCR program for: Cre ERT2 (*Bcl11a* ^{CreERT2} colony) site amplification.

Incubate at 98 °C for 0.30 minutes

Incubate at 98 °C for 0.10 minutes
Incubation at 58 °C for 0.15 minutes
Incubate at 72 °C for 0.30 minutes
} Cycle 35 times

Incubate at 72 °C for 5.0 minutes

Hold at 4 °C

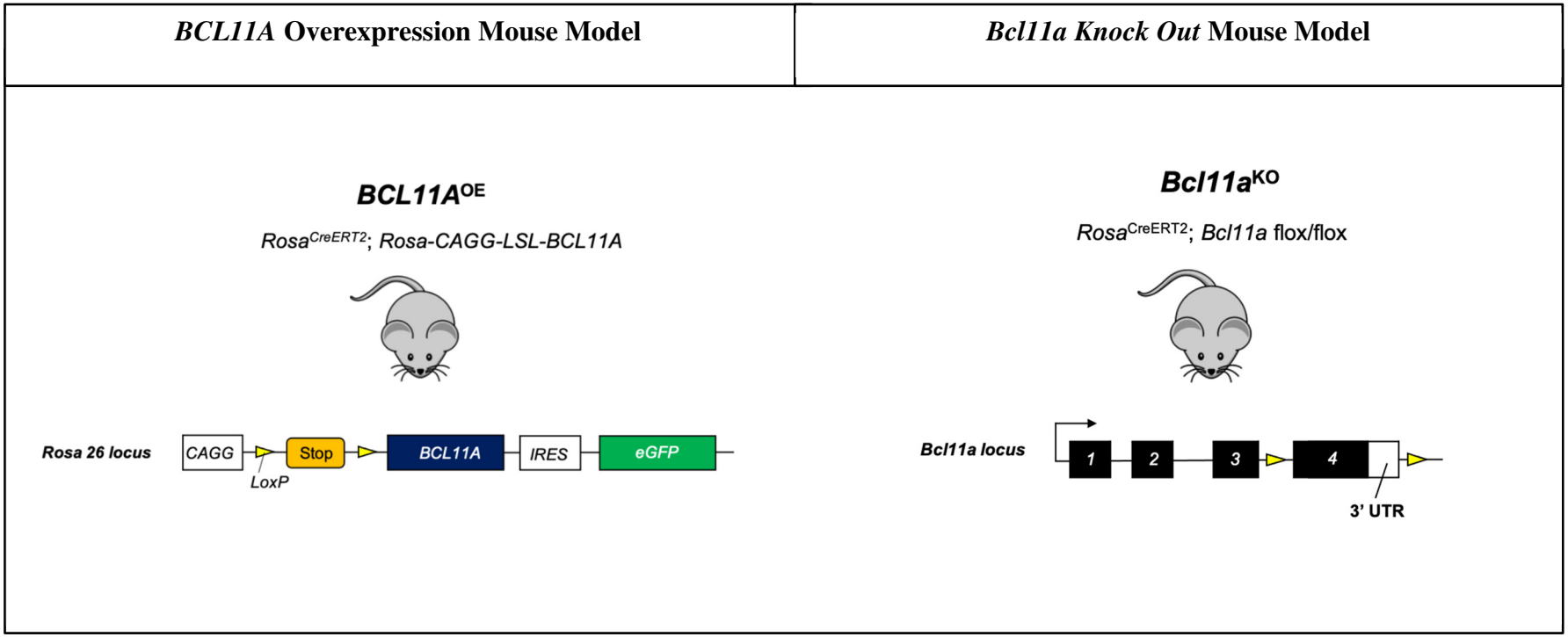


Figure 2.1: *BCL11A*^{OE} and *Bcl11a*^{KO} mouse models used in this project.

BCL11A^{OE} is an overexpression model generated by knocking in a *BCL11A-XL* coding sequence to the *Rosa26* locus. It is downstream of a Lox-Stop-Lox (LSL) cassette and driven by the *CAGG* promoter. Downstream of the *BCL11A* cDNA is *eGFP* under the control of an internal ribosomal entry site *(IRES)*. *Bcl11a*^{KO} mice were generated using *lox*P site upstream of exon 4 of *Bcl11a.* Both mice used inducible *CreERT2* which is knocked-in at the *Rosa26* locus.

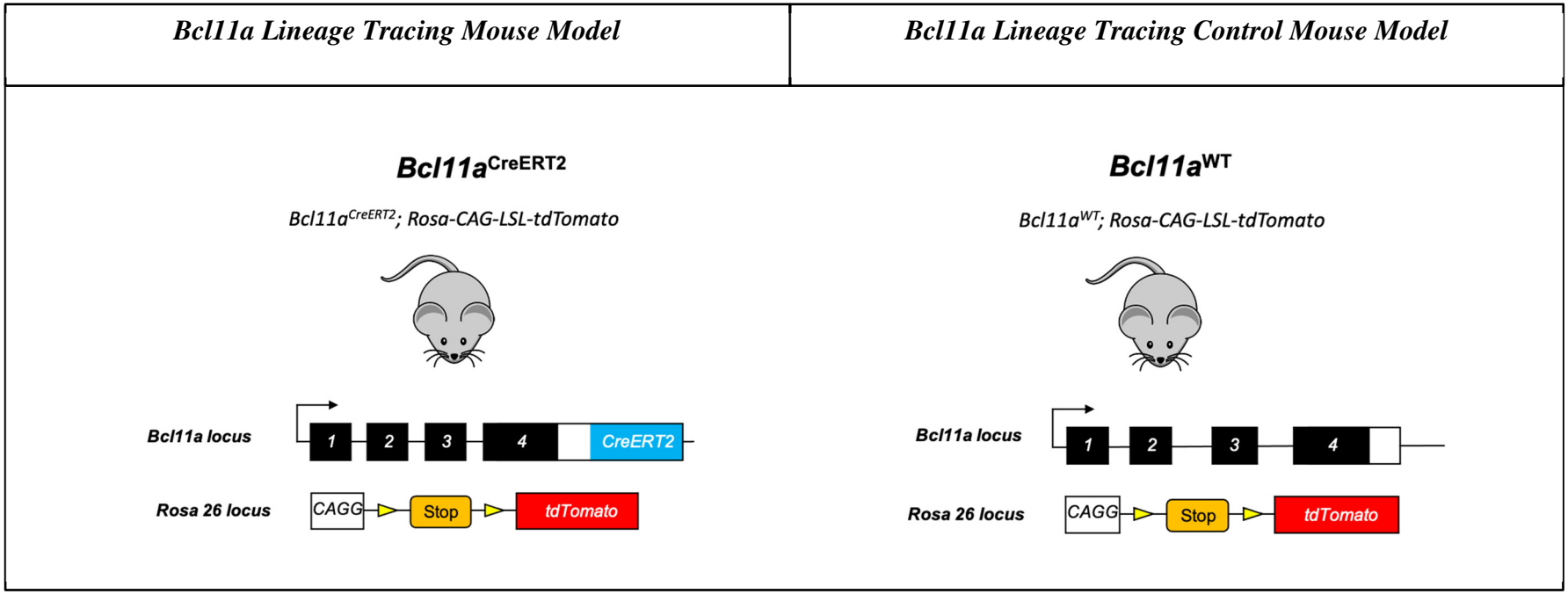


Figure 2.2: Lineage tracing mouse models use in the project *Bcl11a*^{CreERT2} and *Bcl11a*^{WT}

Bcl11a^{CreERT} mice have the *CreERT2* transgene in the 3' UTR of the *Bcl11a* gene. The tdTomato gene is preceded by a *LoxP* flanked stop codon driven by a CAGG promoter in the *Rosa26* locus. *Bcl11a*^{WT} mice had no *CreERT2* expression in the *Bcl11a* locus and instead expressed only the tdTomato reporter construct as described.

Table 1, Primer list for Genotyping the mouse models used in this study.

Colony	Amplified region	Forward	Reverse	Expected Size
BCL11A^{OE}	*BCL11A*	5' CAGTACCCTGGAGAAACACATG 3'	5' GCAATATGGTGGAAAATAAC 3'	230bp
	Rosa26	5' GGAGTGTTGCAATACCTTTCTGGGAGTTC 3'	5'TGTCCCTCCAATTTTACACCTGTTCAATTC 3'	210bp
	CreERT2	5' TGTGGACAGAGGAGCCATAAC 3'	5' AAGACCCAACCAACAGCAG 3'	500bp WT
		5' CATCACTCGTTGCATCGACC 3'		300bp MUT
Bcl11a^{KO}	*Bcl11a*	5'GCCGGGTACCCCACTGGCCAAAGAGTTGAC 3'	5'GCCCGGATCCTTGGTCTGTCTTGGGTTCCT 3'	492bp MUT 405bp WT
Bcl11a^{CreERT2}	*CreERT2*	5' ATTTGTAAGATGCCTTTTAGCGTG 3'	5' GGGGGTTCCGCTGCCTG 3'	400bp MUT
		5' ATTTGTAAGATGCCTTTTAGCGTG 3'	5' GCGTTGTTTTCTACCCTCCC 3'	585bp WT
	tdTomato	5' AAGGGAGCTGCAGTGGAGTA 3'	5' CCGAAAATCTGTGGGAAGTC 3'	196 MUT
		5' GGCATTAAAGCAGCGTATCC 3'	5' CTGTTCCTGTACGGCATGG 3'	297bp WT

2.2 Tissue Isolation and processing: Mammary Gland

Mammary glands were collected from 8-12 week old female virgin mice, the genotype of which differed depending on the experiment and is outlined in the results section. The lymph nodes were removed from the glands and placed in 5 ml Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM-F12, Thermo Fisher) with 1% HEPES buffer (Sigma Aldrich) and 1% Gentamicin (Sigma Aldrich). A digestion mixture of 1 mg/ml Collagenase (Roche) and 100 U/ml Hyaluronidase (Sigma Aldrich) was added for 16 hour digestion at 37°C. Following the digestion, the mix was vortexed to ensure complete dissociation of the tissue. Red blood cells were lysed using NH_4Cl. For dissociation to a single cell level after red blood cell lysis the pellet was resuspended in 0.05% trypsin (Thermo Fisher) and pipetted gently 20 times for two rounds on each sample. After washing with HBSS plus 1% foetal bovine serum (FBS, HF), 5 mg/ml Dispase II (Stem cell technologies) and 0.1 mg/ml DNase (Sigma) mix was added to ensure further digestion and DNA degradation for 2 minutes. The solution was mixed with 10 ml HF followed by filtration through a 40um mesh to obtain a single cell suspension. This solution was then washed with HF and re-suspended in 500 µl blocking solution (HF with 10% normal rat serum).

2.3 Fluorescence Activated Cell Sorting in the Mammary Gland

After blocking mammary gland cell suspensions, small volumes of cells were removed for single colour controls. The cell samples were stained with zombie Aqua (Thermo Fisher Scientific) live dead stain in PBS for 20 minutes in the dark at room temperature. Cells were then taken for the zombie single colour control and the remainder of samples were then washed and incubated with the following antibodies: EpCam-APC-Cy7 1:500, CD49f- BV421 1:100, CD45-PE-Cy7, Ter119-PE-Cy7, CD31-PE-Cy7 all 1:100 SCA1-AF647 1:100 and CD49b-PE 1:100 (all obtained from Biolegend). These were incubated in the dark at 4 °C for 10 minutes. After this time cells were washed and suspended in 500 µl HF. Finally, before sorting, cells were filtered through a 40µm filter to ensure a single cell suspension. Samples were sorted using the BD Fusion machine at the BRC phenotyping hub at Addenbrookes. Manual compensation was applied before every sort to ensure accuracy. Cells were sorted on the basis of these surface markers. Luminal progenitors were visualised via gating the bulk luminal population: $EpCam^{high}$ $CD49f^{low}$, with $CD49b^{positive}$ and $Sca1^{negative}$ sorted as progenitors within this population. Each sorted sample represented a single mouse to ensure plots were

comparable between each animal, however luminal progenitors were ultimately pooled between mice to ensure sufficient numbers of cells for plating in culture.

2.4 Organoid Cell Culture for the Mammary Gland Using Matrigel ECM.

Cells were sorted to enrich for the luminal progenitor lineage from multiple mice and seeded as single cells in growth factor reduced Matrigel (Corning). Luminal progenitor cells were plated at a density of 2000 cells or 3000 per 50 µl of Matrigel in a 24 well plate. After the deposition of matrigel within the wells the plate was allowed to solidify by being placed at 37° C for 30 minutes. The culture medium particular to the experiment (see table 2) was overlaid at a volume of 500 µl. All media was prepared by first creating a base medium using either Dulbecco's Modified Eagle's Medium with F-12 Ham supplement (DMEM-F12) or advanced DMEM-F12, supplemented with antibiotics (see table 2 for details) and HEPES buffer. Fresh growth factors were then supplemented to the base medium, fresh for each change of medium. Cells cultured were maintained at 37 ° C with 5% Oxygen and 5% Carbon dioxide, with medium changed every 3-4 days.

2.5 Organoid Cell Culture for the Mammary Gland Using a Matrigel and Collagen I mix

This protocol was adapted for use from Nguyen-Ngoc etal. Enriched luminal progenitor cells were seeded in a gel mixture of Matrigel and Collagen I (Corning). Before plating, the collagen was neutralised using a mixture of sodium hydroxide and Dulbecco's Modified Eagle's Medium (DMEM, used for its phenol red content). The sodium hydroxide and DMEM were mixed together, until a basic pH was reached. The Collagen I was added until the solution reached a neutral pH, indicated by a pale pink colour. This neutralised collagen solution was then used to coat the wells of each plate to aid the attachment of the top gel. The plate was incubated at 37 ° C to allow the collagen solution to solidity. The Collagen I solution was pre-incubated at 4 ° C for 60-120 minutes, in order to allow preassembly of the collagen fibrils. Preassembled collagen was then mixed with the Matrigel matrix at a ratio of 3:7 (v/v) Matrigel to collagen I [145]. The gel mixture was added to the appropriate cell pellet in volumes to ensure 3000 cells in each 50 µl of Matrigel deposited in a well. The plates were then incubated at 37 ° C to allow the gel to solidify for 45 minutes before the medium was overlaid in a volume of 500 µl. The medium used in this experiment was medium formulation 2, adapted from

Jamieson et al (table 2) and cells were maintained at 37 °C with 5% Oxygen and 5% Carbon dioxide.

2.6 Tissue Isolation and Processing: Oesophagus.

Tissue was isolated from both male and female mice aged between 8-13 weeks unless otherwise indicated. As before the genotype of these mice differed depending on the experiment undertaken but is outlined in the results section. The esophagus was collected via an incision made across the abdomen with care taken to avoid disruption or contact with the thoracic region of the tissue. All collected tissue was placed in Hanks Balanced Salt Solution (HBSS, Thermo Fisher Scientific) containing 1% Penicillin and streptomycin (Thermo Fisher Scientific) on ice until downstream processing.

The esophagus was visualized under a dissecting microscope where it was opened longitudinally with the epithelial layer facing upwards. The muscle layer was then gently removed from the mucosal layers with care taken to avoid any contact or disruption of the length of the tissue. The tissue comprising the mucosal layers was then cut using a scalpel into 4 pieces and incubated with 0.5 mg/ml Dispase I (Sigma) shaking at 37 °C for 10 minutes. The tissue fragments were then visualized again under a dissecting microscope and the stromal layers removed from the epithelial sheets. These were then incubated with 0.05% trypsin (Thermo Fisher Scientific) shaking at 37 °C for 16 minutes to facilitate dissociation to a single cell level. The trypsin was then inhibited using HF and the cell suspension was filtered through a 30 µm cell strainer. Epithelial sheets were then further incubated in fresh trypsin shaking at 37 °C for a further 10 minutes. The trypsin containing the epithelial fragments was then neutralised and passed through the same filters. Cells were then pelleted by spinning at 650 g for 5 minutes at 4 °C. Where cells were plated directly, the pellet was re-suspended and cells were counted and plated. If cells were to be used for FACS as in the case of the lineage tracing experiment, these were taken for sorting.

2.7 Cell Culture for Oesophageal Organoids

For plating a suspension of cells without sorting, cells were counted and the desired number then mixed with the appropriate volume of Matrigel to ensure that there were 1000 cells per 50

µl of Matrigel deposited in each well. For cells that were sorted these were counted and the desired number mixed with an appropriate volume of Matrigel. The Matrigel was then allowed to solidify for 20 minutes at 37 ° C. The medium appropriate to the experiment was then overlaid in a volume of 500 µl (see table 3). In the case of medium formulation B, the base medium composed of advanced DMEM-F12 with glutimax, HEPES and Penicillin, streptomycin supplementation this was kept at 4 ° C for up to one week and all other growth factors and cytokines were added fresh at the time of media change. Cells were maintained at 37 ° C in a humidified atmosphere with 5% Carbon dioxide. Cells were maintained for a maximum of 2 weeks with media changed every 3-4 days.

2.8 *In vitro* 4-hydroxytamoxifen induction

For experiments involving cells harvested from any of the *Bcl11a* knock out or *BCL11A* overexpression mouse models media was supplemented with (Z)-4-Hydroxytamoxifen (4-OHT) to induce cre mediated recombination in vitro. 4-OHT (Santa Cruz Biotechnology) was diluted in ethanol (Sigma Aldrich) to reach a concentration of 1 mM, aliquoted and then stored at -20 ° C, minimising the exposure to light. In the case of both mammary and oesophageal organoids, colonies were left to form for 7 days after which 1 µM of 4-OHT was supplemented to the media for 24 hours. In addition to this, organoids were also treated with a vehicle control of ethanol only. After 24 hours, the 4-OHT and vehicle control supplemented media was removed and fresh media was overlaid with normal maintenance continued for a further7 days. At the end of the 14-day period organoids were imaged and taken for analysis via whole mount immunohistochemical analysis or nucleic acid extraction.

Table 2: Mammary Organoid cell culture media recipes.

	Supplier	Adapted from Jardé et al. 2016 **Medium formulation 1**	Adapted from Jamieson et al. 2017 **Medium formulation 2**
Base Medium		DMEM-F12	Advanced DMEM-F12
Penicillin, Streptomycin	Thermo Fisher scientific	1x	1x
HEPES	Thermo Fisher scientific	1x	1x
Glutamate	Thermo Fisher scientific	1x	1x
B27 Supplement	Thermo Fisher scientific	1x	1x
N2 Supplement	Thermo Fisher scientific	1x	-
Insulin	Thermo Fisher scientific	-	5 µg/ml
EGF	R and D systems	50 ng/ml	50 ng/ml
FGF2	R and D systems	-	5 ng/ml
FGF10	Peprotech	-	10 ng/ml
Hydrocortisone	Arcos	-	100 ng/ml
N-acetyl-L-cysteine	Merck Millipore	-	1.25 µM
Y27632 (Rock Inhibitor)	Merck Millipore	10 µM (first 3 days)	5 µM (first 3 days)
Heparin	Arcos	-	4 µg/ml
Wnt3a	R and D systems	-	10 ng/ml
Rspondin-2	Peprotech	-	2.6 ng/ml
Noggin	Peprotech	100 ng/ml	-
Neuregulin-1	Peprotech	100 ng/ml	-
Rspondin-1	Peprotech	2.656 ng/ml	-

Table 3: Oesophageal Organoid media recipes

	Supplier	Adapted from Kasagi et al. 2018 **Medium formulation A**	Adapted from DeWard et al. 2014 **Medium formulation B**
Base Medium	Thermo Fisher Scientific	Keratinocyte Serum Free Medium	Advanced DMEM-F12
Penicillin, Streptomycin	Thermo Fisher Scientific	1x	1x
HEPES	Thermo Fisher Scientific	-	1x
Glutamate	Thermo Fisher Scientific	-	1x
N2 supplement	Thermo Fisher Scientific	-	1x
B27 supplement	Thermo Fisher Scientific	-	1x
N-acetyl-L-cysteine	Arcos Scientific	-	1 mM
Gastrin I (human)	R and D systems	-	100 µM
EGF	R and D systems	1 ng/ml	50 ng/ml
Nicotinamide	Arcos Scientific	-	10 mM
SB202190 (MapK inhibitor)	Millipore	-	10 µM
Noggin	Peprotech	-	100 ng/ml
Wnt3A	R and D systems	-	100 ng/ml
Rspondin2	Peprotech	-	100 ng/ml
A8301 (TGFβ inhibitor)	Merck Millipore	-	500 nM
Y27632 (Rock Inhibitor)	Merck Millipore	10 µM	-
Bovine Pituitary Extract	Thermo Fisher Scientific	5µg/ml	-
$CaCl_2$	Merck Millipore	0.6 mM	-

2.9 Preparation of Organoids for Whole Mount Immunofluorescence Imaging

After cultivation *in vitro,* both mammary and oesophageal organoids were collected for whole mount analysis. In order to avoid organoids adhering to the pipette, all P1000 tips were widened by cutting off the end of the tip and pre coated with 100 % FBS before use. Matrigel was dispersed using ice cold PBS and the organoids were collected into a 15 ml falcon. The organoids were then left to settle and taken out of the tissue culture hood. The supernatant was then removed, and the organoids were washed with cold PBS and incubated at 4 ^{o}C to ensure complete removal of the gel matrix. The excess was removed, and the organoids were washed a further 2 times. Organoids were then fixed in 4% Paraformaldehyde (PFA) for 30 minutes at room temperature. The PFA was then removed and the organoids were washed using PBS as above 3 times for 10 minutes on an orbital shaker. Organoids were then incubated in blocking solution, consisting of PBS with 5% FBS 2% BSA and 0.2% Triton X (Sigma Aldrich), for 2 hours at room temperature. Organoids were then incubated with primary antibody in blocking buffer overnight at 4 ^{o}C on an orbital shaker. Primary antibodies used were keratin 5 1:500 (biolegend) and Keratin 8 1:250 (Troma I, Merck Millipore) or Keratin 13 1:100 (ab92551, abcam). The following day organoids were washed in PBS containing 0.1% Triton X, 3 times for 10 minutes on an orbital shaker. The secondary antibodies were diluted in blocking buffer and incubated for 2 hours at room temperature on an orbital shaker. The secondary antibodies used were: goat anti rabbit IgG AF647 1:250 and Goat Anti Chicken- AF488 1:250 (both obtained from Thermo Fisher). After this time, organoids were again washed 3 times for 10 minutes with PBS triton solution. For nuclear visualisation organoids were stained with 2 µl of 5mg/ml 4′,6-diamidino-2-phenylindole (DAPI) stock solution, diluted in 1 ml PBS and incubated for 30 minutes at room temperature. Organoids were then washed and resuspended in vectashield mounting medium (vector laboratories) and stored at 4 ^{o}C ready for visualisation. All microscopy was performed using the Leica SP4 Confocal microscope. To visualise, organoids were mounted onto a circular 5 mm coverslip enclosed in a coverslip cell chamber. Once organoids were viewed optical z planes were acquired.

2.10 RNA Extraction RT-qPCR

In order to extract RNA from both mammary and oesophageal organoids grown in culture, ice cold PBS was used to disperse the Matrigel. Organoids were then collected in 15 ml falcon tubes and incubated on ice with fresh PBS to ensure complete removal of Matrigel. Organoids were then transferred to an Eppendorf tube and spun maximum speed for 2 minutes. All

supernatant was removed and organoids were resuspended in RLT lysis buffer from the RNeasy Micro kit (Quagen) according to the manufacturers instructions. In order to ensure complete lysis, organoids were mechanically dissociated using a piston before being transferred to RNA-easy spin column. After spinning, the column was washed using RW1 buffer and incubated with 20U Rnase-free DNaseI (Roche) for 15 minutes at room temperature. Finally, the column was washed with RW1 and the RNA was eluted into RNase and DNase free water. Total RNA was diluted to a final volume of 11µl. 2µl of random primers (Promega) were added after which the mixture was incubated at 65 ^{o} C for 5 minutes. A master mix containing Transcriptor Reverse Transcriptase (Roche), Reverse Transcriptase buffer, 2mM dNTP mix and RNasin Ribonuclease inhibitors (Promega). This was then added to the RNA mix and incubated at 25 ^{o} C for 10 minutes, then 42 ^{o} C for 40 minutes and finally 70 ^{o} C for 10 minutes. The resulting cDNA was then diluted 1:2.5 in H2O for subsequent use. qPCR was performed using a Step-One Plus Real Time PCR System (Thermo Fisher Scientific). Taqman (Thermo Fisher Scientific) probes (see table 4) with GoTaq Real Time qPCR Master Mix (Promega) were used. The enrichment was normalised with control mRNA levels of *Gapdh* and relative mRNA levels were calculated using the ΔΔCt method comparing to control group.

Table 4, Taqman Probes for qPCR

Gene Target	Assay
BCL11A	Hs00256254_m1
Bcl11a	Mm00479358_m1
Gapdh	Mm99999915_g1

2.11 Quantitation of microscope images, cell counting

All analysis was performed using Fiji 4 [154]. In the case of organoid quantitation, cells were counted using the cell counter plug-in and scored as either Keratin 5 positive, Keratin 8 or 13 positive or double positive. Cells with no marker were labelled as 'none'. Graphs were constructed using GraphPad Prism 9 to depict numerical data. For quantitating lineage tracing, two microscope images were used per time point, one from each biological specimen. Cells were counted as either tdTomato positive, Keratin 5 positive or double positive. Again, graphs were constructed using GraphPad Prism 9.

2.12 Quantification of organoid number and size

All quantification was performed using Fiji 4 [154].Merged brightfield images of whole wells were used to count organoid number and quantify size. To count the number of organoids merged images were manually quantified using the cell counter plug-in. To assess the size of the organoids, the brightfield images were converted to binary colour mode by applying a luminance threshold. Colour was then inverted to highlight organoids in white and the number of white pixels were measured. The total value per well was divided by the number of organoids in the well to calculate a mean organoid area in pixels. Graphs were constructed using GraphPad Prism 9 to depict numerical data.

2.13 *In vivo* Tamoxifen Induction

For lineage tracing experiments requiring *in vivo* tamoxifen mediated recombination, the tamoxifen (T5648, Sigma Aldrich), was reconstituted in corn oil (C8267, Sigma Aldrich) at a concentration of 20 mg/ml by shaking overnight at 37 ^{o} C. The reagent was subsequently aliquoted and stored at -80 ^{o} C. Upon administration, 50ul of tamoxifen was injected intraperitoneally in the lower quadrant using a 25G needle. Animals were then subject to the requisite post procedure checks to ensure no adverse events had occurred, in line with the scientific procedures at the Sanger Institute.

2.14 Lineage tracing Esophageal Tissue Preparation for Whole Mount Immunohistochemical Analysis

The mice used for whole mount immunohistochemical analysis were aged between 12-15 weeks. The esophagus was harvested from two *Bcl11a*^{CreERT2} and one *Bcl11a*^{WT} mouse per time point. The esophagus was visualized under a dissecting microscope where it was cut longitudinally with the epithelial layer facing upwards. In some instances, where it was preferrable to ensure the length of the tissue was intact, no incision was made. The muscle layer was then gently removed from the mucosal layers with care taken to avoid any contact or disruption of the length of the tissue. The tissue comprising the mucosal layers was then incubated with 5 mM EDTA (sigma) in phosphate buffered saline (PBS) for 2 hours to facilitate the removal of the mucosa from the submucosa through calcium binding (reducing integrin adhesion). The tissue was then visualized under a dissection microscope once more, and the epithelial sheet removed from the remaining tissue. The tissue was then fixed using 4% PFA (Cell Path, Fisher Scientific) for 30 minutes at room temperature.

2.15 Optical Clearing of the Esophageal Epithelium using CUBIC for lineage tracing.

The tissue fragments were then washed in PBS 3 times for 10 minutes. Each fragment was then submerged in CUBIC reagent 1 for 3 days at 37 ^{o} C. Cubic reagent 1 was prepared using urea (25% (w/w)), N,N,N',N'-tetrakis(2-hydroxypropyl)ethylenediamine (25% (w/w)), triton-X100 (15% (w/w)) in distilled water. Tissue fragments were washed in PBS for 1 hour, 3 times and subsequently blocked for 2 days in triton-X100 (0.5% (w/v)) and goat serum (10% (v/v)) at 4°C. Tissue fragments were then incubated with primary antibodies, depending on their use. The primary antibodies used were: RFP (600-401-379, Rockland), Keratin 5 (905904, Biolegend) diluted in blocking buffer and incubated for 4 days at 4°C with gentle rocking. Tissue fragments were then washed in PBS for 1 hour, 3 times and incubated with Alexa Flour conjugated secondary antibodies: goat anti rabbit IgG AF647 or goat anti Mouse AF647 and goat anti Chicken- AF488 (all obtained from Thermo Fisher) for 2 days at 4°C with gentle rocking. Samples were then washed with PBS for 1 hour, and incubated in DAPI for 1 hour at room temperature in the dark with gentle rocking. Samples were then washed for a further hour at room temperature in PBS. Finally, tissue fragments were submerged in CUBIC Reagent 2 prepared using sucrose (44% (w/w)), urea (22% (w/w)), 2,2',2"-nitrilotriethanol (9% (w/w)),

triton-X100 (0.1% (w/w)) in distilled water, for a minimum of 2 days before imaging. All imaging was undertaken using a leica Sp4 Confocal microscope

2.16 Fluorescence Activated Cell Sorting for tdTomato Positive Cells

Bcl11a^{CreERT2} mice were collected 24 hours post tamoxifen exposure and were digested as in section 2.4.1. Cells were sorted on the basis of endogenous tdTomato expression using the Sony SH800Z sorter. The instrument was set up using cells from Black 6 WT mice as an unstained control. Both tdTomato positive and negative populations were sorted for plating and plated at a density of 500 cells per well.

Chapter 3: Exploring *BCL11A* Function in Oesophageal Organoids

3.1 Introduction

3.1.1 Exploring *BCL11A* gene function in the Oesophageal Epithelium Using 3D organoids

The importance of *BCL11A* as a proto-oncogene in epithelial tumours [2,3,26] makes this transcription factor a valuable candidate for further study with the hope of exploiting its biology for therapeutic purposes. As part of a previous study in the lab which identified *BCL11A* as a gene selectively upregulated in lung squamous carcinoma (LUSC), multiple cell lines were analysed using shRNA against BCL11A. Serendipitously, one of the cell lines analysed was later identified to be an oesophageal squamous carcinoma (OSCC) cell line. In work performed by Dr Kyren Lazarus, the OSCC cell line KYSE 140 was transfected with two different shRNAs targeting the *BCL11A* gene. Transfected cell lines were then analysed using a tumour xenograft model. After 15 days, tumour size was measured, and a significant reduction or in some cases abolition of tumour growth in the knockdown cell lines was detected (Figure 3.1a and b). This suggested that *BCL11A* might be oncogenic in OSCC, which was supported by its upregulation in OSCC patients in the TCGA dataset [44] (Figure 3.2).

This observation prompted many questions, perhaps the most important of all being related to gene function. Little is known about this gene in the context of normal tissue maintenance in the oesophagus. Indeed, identifying whether *BCL11A* gene expression is important for progenitor and/or differentiation as it has been in other cell types[14] is an exciting prospect. Having used the organoid system for functional studies in the lung [3], it was decided to test this approach to study the role of *BCL11A* in the oesophageal epithelium. Cells were cultivated from two transgenic mouse models of *Bcl11a* that enable the overexpression or deletion of *Bcl11a* (figure2.1). One model is a knock-in of the *BCL11A-XL* coding sequence into the *Rosa26* locus and downstream of a Lox-Stop-Lox (LSL) cassette and the *CAGG* promoter. In the presence of an active Cre recombinase the LSL cassette is removed and *BCL11A* overexpression is driven by the CAGG promoter. This model is referred to as *BCL11A*^{OE}. The other mouse model used in this study enables the conditional deletion of *Bcl11a* in a Cre dependent manor. This model is referred to as *Bcl11a*^{KO}. For both these lines we used the inducible *CreERT2* which is knocked-in at *Rosa26* locus which enables temporal induction of overexpression or deletion of *Bcl11a* upon 4-OHT treatment *in vitro*.

Organoid culture of the oesophageal epithelium has been described using varying media formulations and methods [127,153]. Therefore, the first aim was to compare published models of cell culture and to determine the one most suitable for downstream applications. Using cells cultivated from the oesophageal epithelium of transgenic mouse models: *Bcl11a*^{KO} and *BCL11A*^{OE}, the second phase of the project sought to understand more about *BCL11A*'s function. Aiming to understand the effects if any, on cellular and molecular dynamics induced by these perturbations, would help to gain insight into the role this transcription factor plays in both physiological and malignant tissue.

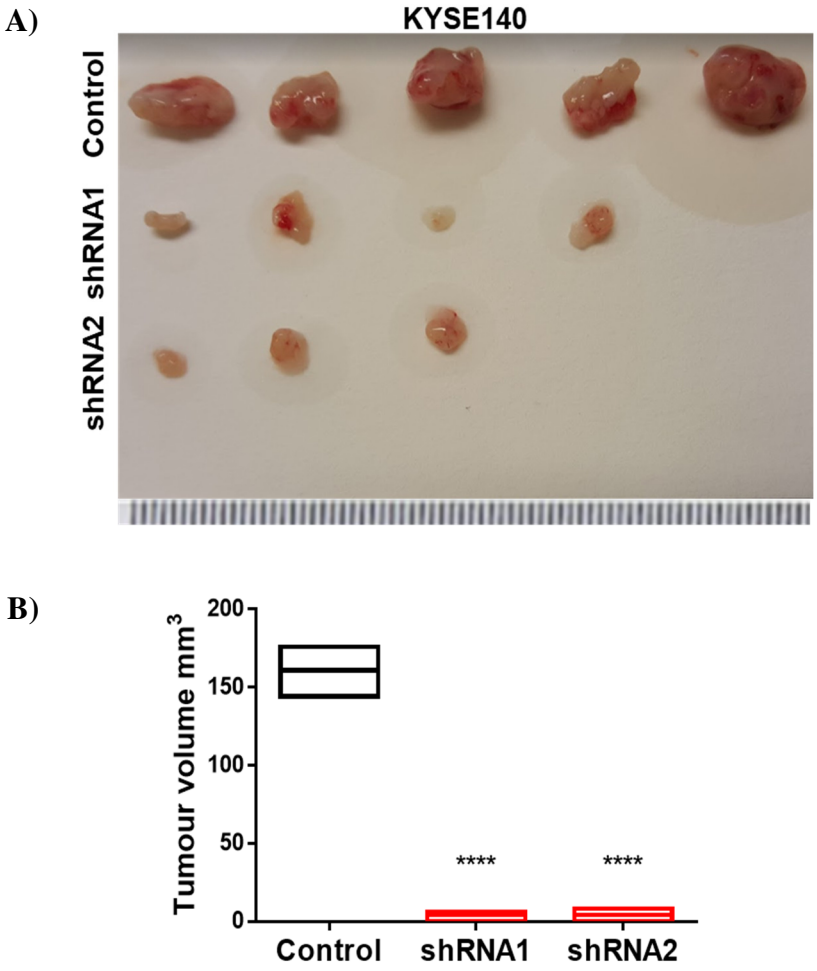


Figure 3.1: Xenograft data from *BCL11A* Knock down.

All data obtained from Dr Kyren Lazarus. Using 2 different shRNAs *BCL11A* was knocked down in an oesophageal squamous carcinoma cell line (KYSE 140). Control refers to cells treated with the empty vector. One million cells of each condition were suspended in matrigel and HBSS and subcutaneously injected in 6–12-week-old NSG mice. **A)** Images of actual tumours measured. Five mice per cell line were monitored for 15 days after which tumours were removed and measured. **B)** Data presented as mean ± S.D.. ($n = 5$). One-way ANOVA test performed.

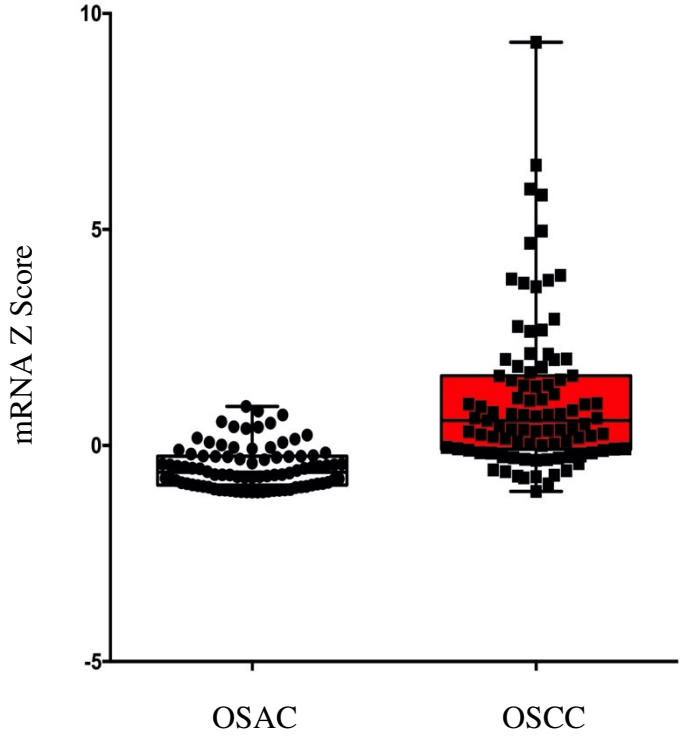


Figure 3.2: Data obtained from the TCGA dataset identifies *BCL11A* as a selectively upregulated gene in OSCC Vs OSAC.

Figure obtained from Dr Kyren Lazarus. Relative expression levels of *BCL11A* in oesophageal squamous carcinoma (OSCC) and oesophageal adenocarcinoma (OSAC) were downloaded from the TCGA database.

3.2 Results

3.2.1 Comparing Models of Oesophageal Organoids

The culture of oesophageal organoids has been described using different media formulations and indeed methods [127,153]. To identify one system that faithfully mimicked the murine oesophageal epithelium whilst also being technically robust was the first objective for oesophageal cell culture. Organoid medium was compared from Kasagi et al. termed 'medium formulation A' and from DeWard et al. termed 'medium formulation B'. Organoids were firstly cultivated using medium formulation A. As seen in table 3 this medium formulation was conservative in its supplementation. This was advantageous as it was both economical and simple to prepare.

As with the oesophageal epithelium *in vivo,* the hallmark of effective organoid differentiation is the stratification of the epithelium. *In vivo,* this is demonstrated by the differentiation of basal keratinocytes into suprabasal cells, becoming keratinised before being shed off into the lumen. Within the organoid however, there is no such action, resulting in the accumulation of keratinised cells in the centre of the organoid to form a keratinised core. The demonstration of this morphology was key to culture success. In addition, more technical aspects of the culture systems were also considered, such as colony forming unit efficiency.

Single cells dissociated from the oesophageal epithelium were seeded at a density of 1000 cells per well, embedded in matrigel and cultivated using medium formulation A. Authors Kasagi et al. describe the generation of organoids (structures greater than 50 μm) after a period of 7 days [153]. However, when cultured in this manner, cells did not appear to generate structures after 7 days. These cells were left for a further week to assess whether any growth was visible. After a period of 14 days organoids did appear to be visible, although at a severely reduced organoid forming unit (OFU) frequency of 0.3% (Figure 3.3a). Closer inspection of these structures revealed layers of differentiated cells resulting in a keratinised core in line with the *in vivo* stratified epithelium (figure 3.3b). This would suggest that Keratinocyte serum free medium (KSFM) was able to induce the differentiation and stratification observed in the oesophageal epithelium *in vivo.* These organoids closely mimicked those reported by Kasagi et al. in their morphology, however the OFU frequency was below that which the authors reported.

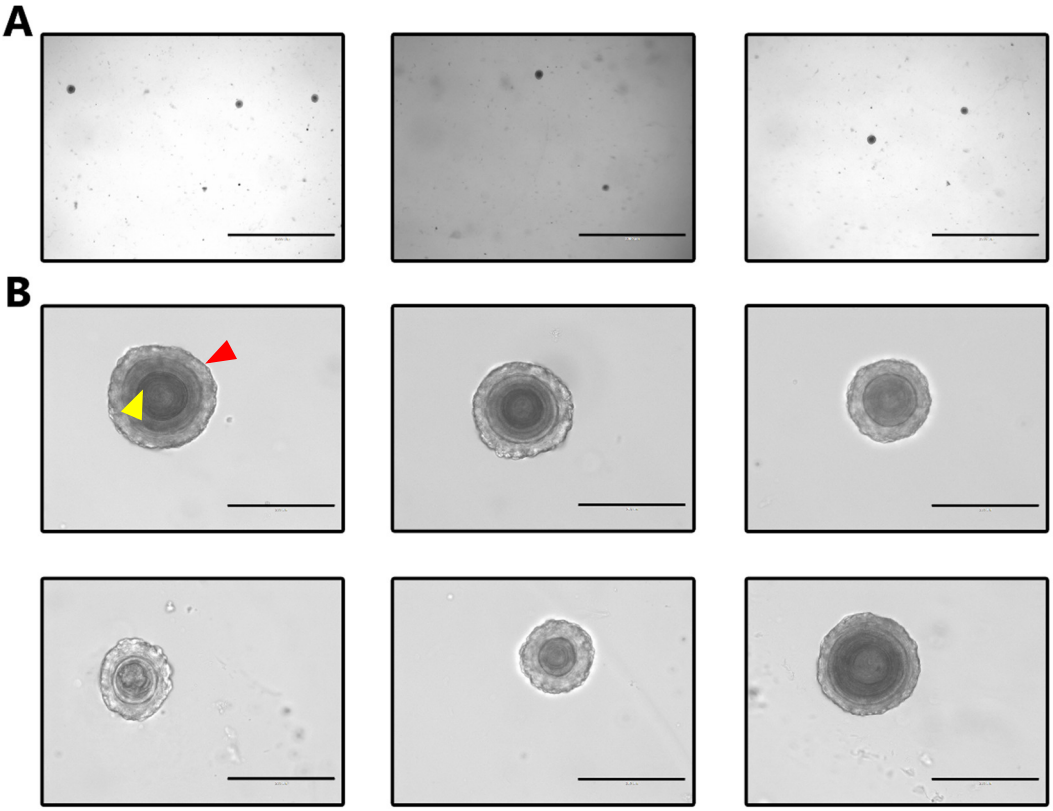


Figure 3.3: Oesophageal organoids cultivated using medium formulation A.

Organoids cultivated from single cells of the oesophageal epithelium of WT mice at 14 days post seeding. **A)** Representative wells, cells plated at 1000 cells per well. Scale bar represents 2000 µm. **B)** Individual organoids cultivated using medium formulation A. Scale bar represents 400 µm. Red arrow denotes outer basal layer, yellow arrow indicates keratin core.

The contribution of reagents to such a poor OFU frequency was ruled out as far as was possible. Some of the key factors and reagents were also used for the effective cultivation of mammary organoids, whilst others such as the keratinocyte base medium and BPE were prepared fresh. The cultivation of organoids at a higher density did not yield an increase in efficiency and ultimately this led to the use of many animals in each experiment which was not sustainable for long term use. The *Bcl11a*^{KO} and *BCL11A*^{OE} mouse colonies were not large enough to support the use of so many experimental animals due to the COVID-19 pandemic. Organoids were too few to successfully analyse using confocal staining. Therefore, this medium was not considered suitable based on the poor OFU frequency.

The second medium sought for the culture of oesophageal organoids was formulation B. As seen in table 3 this medium was extensively supplemented and represented a media formulation more closely related to organoid cultures of the intestine [155]. Factors such as Wnt3A and Rspondin-1 were added for stimulation of the Wnt pathway, whilst factors such as TGF-β inhibitor and Map Kinase inhibitors were added to aid the inhibition of apoptosis [156].

After a period of 7 days in culture, single cells of the oesophageal epithelium plated at a density of 1000 cells per well gave rise to organoids. When cultivated for a period of 14 days organoids were well established with a higher OFU frequency of roughly 20%. As seen in figure 3.4a after a period of 14 days cells cultivated in medium formulation 2 were well established as organoids. Closer inspection of organoid morphology demonstrated the apparent stratification of the epithelial cells into a central keratinised core (figure 3.4b). This morphology and OFU efficiency was consistent with that demonstrated by DeWard et al [127]. In addition to organoids demonstrating spherical morphology, some organoids displayed an apparent protrusion of the keratin (figure 3.4b). Although this morphology was somewhat unexpected it was seen to be encountered by other authors attempting the culture of oesophageal organoids in 3D [127].Although the authors do not describe in detail the cause of such a phenomenon, it could be the result of an accumulation of the keratin causing it to expand outwards or a particular area of weakness in the cells basal cell layer allowing the growth to expand in that direction.

Whole mount analysis revealed control organoids displayed cells positive for Keratin 5 staining on the outside basal layer as well as Keratin 13 staining closer to the keratin core, consistent with the differentiation gradient expected in the oesophageal epithelium (figure 3.5) [95,127].

Some organoids however, contained cells that were double positive for these markers, which was not an expected result (figure 3.5). There were also some cells that were not positive for any markers inside the organoids. Again, this was inconsistent with the organoids described by DeWard et al. [127]. This could have been as a result of poor antibody penetrance as the dedicated Keratin 5 and Keratin 13 positive cells did suggest that these cells were able to differentiate *in vitro.* As organoids cultivated using this medium displayed a higher OFU efficiency as well as the generation of organoids largely with the correct morphology, medium formulation B was taken forward for use with cells from both *Bcl11a*^{KO}and *BCL11A*^{OE} mouse models. For better observation of the differentiation phenotype in response to gene perturbations, the cells were counted, allowing any significant changes to be characterised.

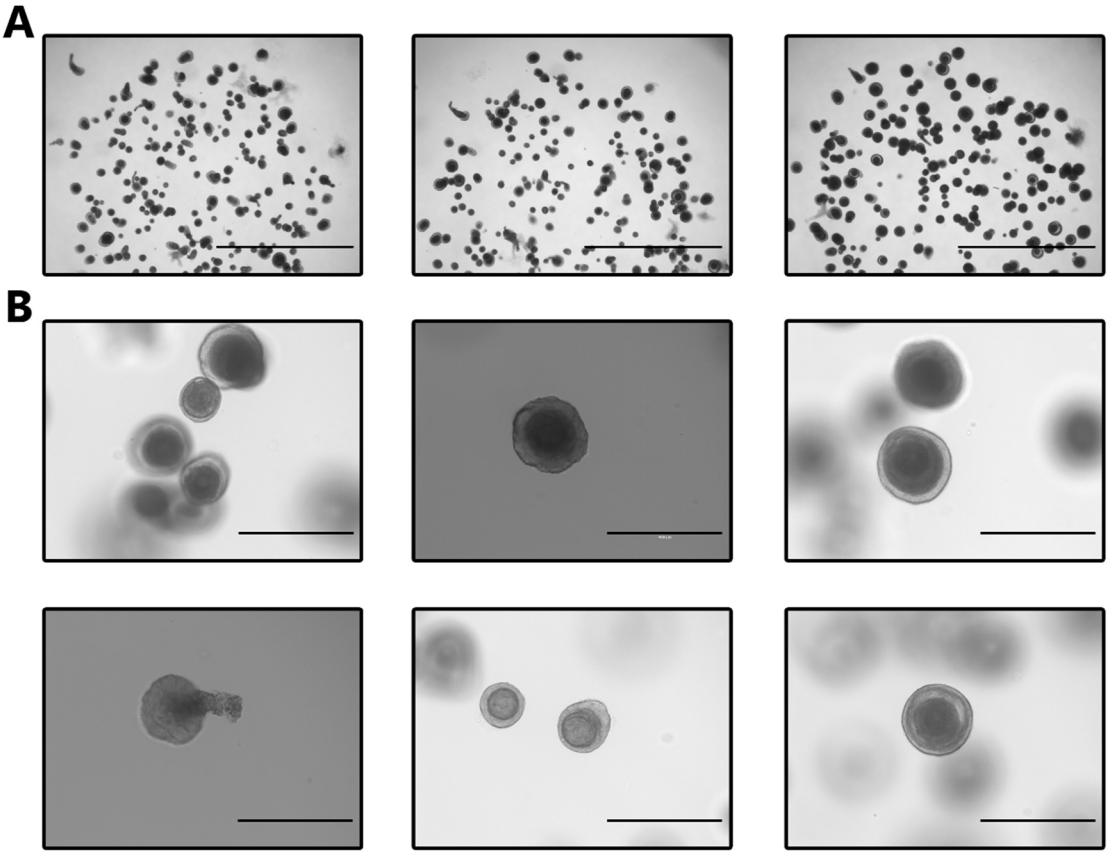


Figure 3.4: Oesophageal organoids cultivated using medium formulation B.

Figure shows oesophageal organoids cultivated from WT mice for 14 days and supplemented with medium B. **A)** Representative wells, cells plated at 1000 cells per well. Scale bar represents 2000 µm. **B)** Individual organoids cultivated using medium formulation A. Scale bar represents 400 µm.

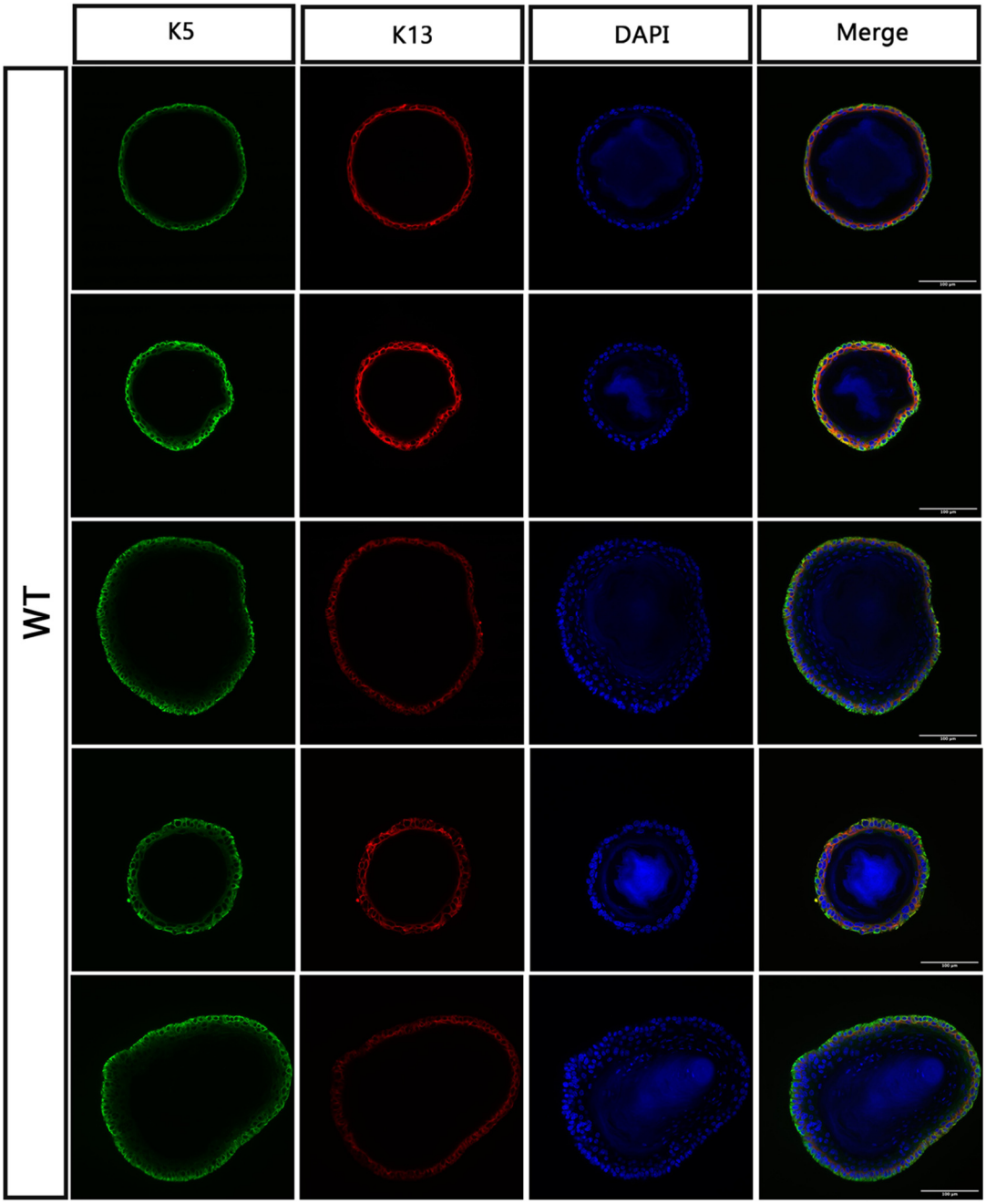


Figure 3.5: Fluorescence microscopy of Whole mount WT oesophageal organoids cultivated using medium formulation B.

Oesophageal organoids were cultivated for a period of 14 days and subjected to whole mount analysis. Organoids were stained with a marker of basal cells (Keratin 5 in green), suprabasal cells (Keratin 13 in red) and DAPI to visualise nuclei in blue. Scale bar indicates 100 µm.

3.2.2 *BCL11A* is Required for Correct Differentiation in Oesophageal Organoids

In order to assess the morphological effects of *BCL11A* overexpression on established organoids, cells were cultivated from the oesophageal epithelium of *BCL11A*^{OE} mice. Induction of the transgene was mediated through the *in vitro* administration of 4-OHT after 7 days in culture, whilst control groups were treated with vehicle control or left untreated. Organoids were maintained in culture for a further 7 days after induction. As mentioned, in section 1.4.1, proliferation of cells in the oesophagus is confined to progenitors in the basal layer, allowing the expansion of these cells over time *in vitro*. Cells plated at the time of seeding were therefore a heterogeneous dissociation mix of both basal and suprabasal, (progenitors and differentiated) cells. After 14 days, organoids were analysed to determine the average size and number of organoids per condition. In addition, RT-qPCR analysis was undertaken to confirm relative gene expression. Whole mount analysis was also undertaken to confirm cellular composition and the effect, if any, *BCL11A* expression changes caused to cellular differentiation in the organoids.

Firstly, to determine any effect of 4-OHT exposure alone, organoids from WT mice were cultivated for 7 days, after which they were treated with either 4-OHT, vehicle control or left untreated. After a further 7 days in culture, organoids were imaged using a bright field microscope. Oesophageal organoids displayed no apparent difference between the treatment conditions as demonstrated in figure 3.6 a. To ensure this was the case, merged bright field images were analysed for organoid size and number (Figure 3.6 b and c). No difference was found between the groups, indicating that 4-OHT exposure alone did not cause any phenotypic changes visible at this stage (figure 3.6b and c).

Organoids were collected for whole mount analysis to ensure that exposure to these reagents did not have an effect on cellular composition. As seen in figure 3.7, organoids broadly appeared to be similar in cellular composition across groups whilst maintaining a spherical morphology. All organoids displayed an outer layer of Keratin 5 positive cells as expected in the basal layer of the oesophageal epithelium. Organoids also displayed Keratin 13 positivity within the suprabasal layer, in line with the normal differentiation spectrum observed in the oesophageal epithelium [95,127]. Interestingly, Keratin 13 positive cells were only seen in a single layer within the organoids and some cells remained negative for both markers. Keratin 13 positive cells are expected to be present throughout the centre of the organoid. In order to

characterise this further and to quantitate the cell types observed for organoids grown in these conditions, cells were counted and scored as either Keratin 5, Keratin 13 or double positive. Cells with no markers were scored as 'none' (figure 3.8). The organoids treated with 4-OHT largely appeared to display a greater proportion of cells without any markers as well as fewer cells positive for Keratin 13 compared to the other groups (figure 3.8 a and b). This could suggest a mild effect of 4-OHT exposure on the differentiation gradient within the organoids. Another observation was the presence of cells positive for both markers. Such cells have not been described in the oesophageal epithelium and are therefore likely to be artefacts of the cell culture system. Interestingly, the untreated organoids appeared to contain the greatest proportion of these double positive cells (figure 3.8b). This suggested that the culture conditions alone could be responsible for this phenomenon.

All organoids displayed a keratin core (Blue, DAPI) suggesting that cells within the organoids were differentiating. Images were acquired at multiple planes of view to demonstrate organoid morphology (figure 3.9). Taken together this suggested that exposure to 4-OHT or vehicle control alone did not induce any overall phenotypic changes in the organoids however, fractionally, there were very small changes in cellular composition. Although this result was not anticipated, it did not preclude experiments aimed at understanding *Bcl11a* transcription factor function in the oesophageal epithelium, from taking place.

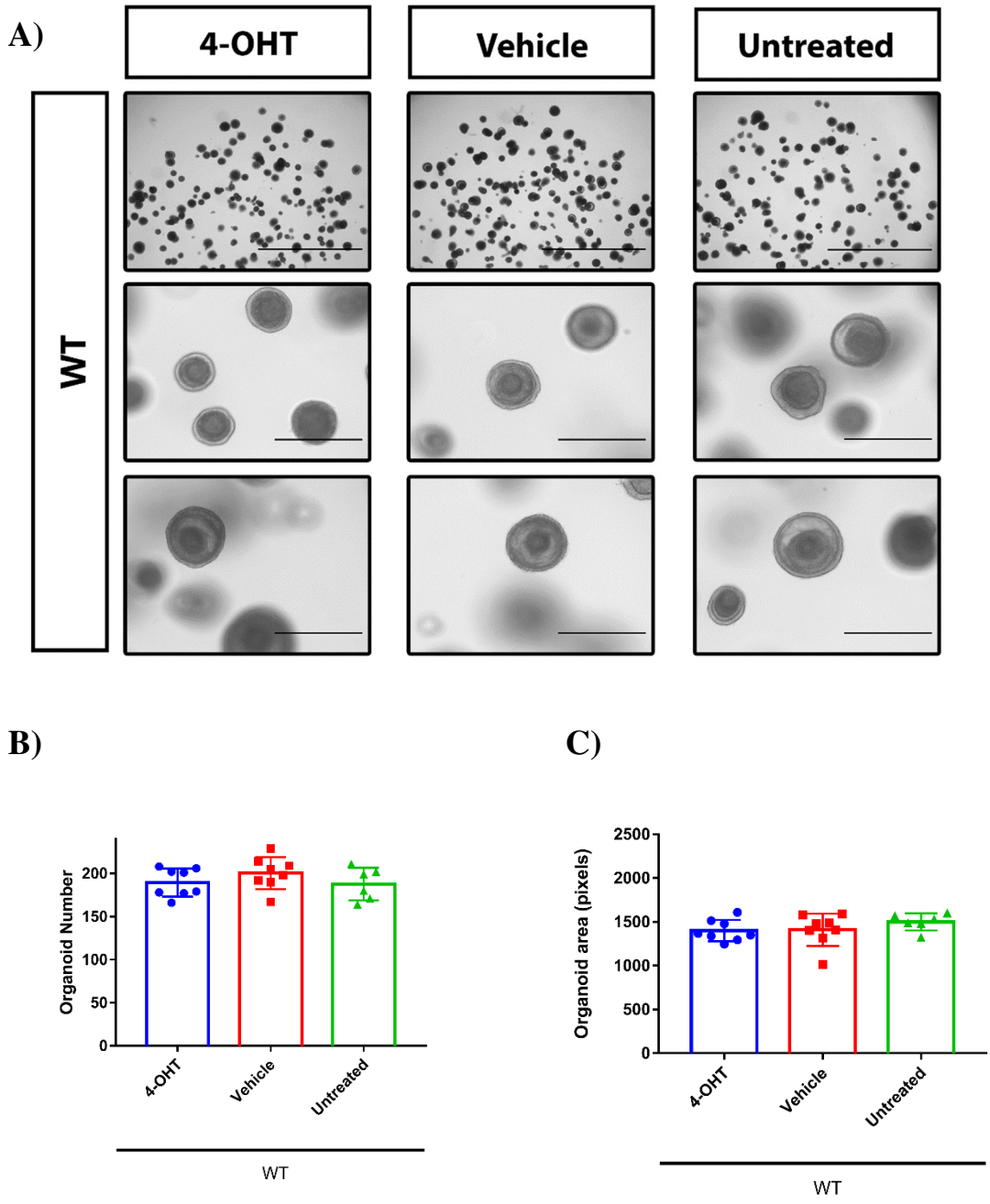


Figure 3.6: Quantification of WT oesophageal organoid number and size.

Figure 3.6 (cont.): Quantification of WT oesophageal organoid number and size.

Oesophageal organoids were cultivated from WT mice for 7 days after which they were exposed to either 4-OHT, vehicle control or left untreated and cultivated for a further 7 days. Data are representative of one biological replicate. **A)** Organoids were cultivated for a total of 14 days and subjected to different treatment. Scale bar represents either 2000 μm for the top row of images and 400 μm for the middle and bottom row. **B)** Shows organoid number cultivated from 1000 cells after 14 days in culture for the varying conditions. **C)** Shows he size of organoids cultivated in each condition. Data are represented as the mean size of organoids in pixels per well ± S.D. of n=6-8 technical replicates.

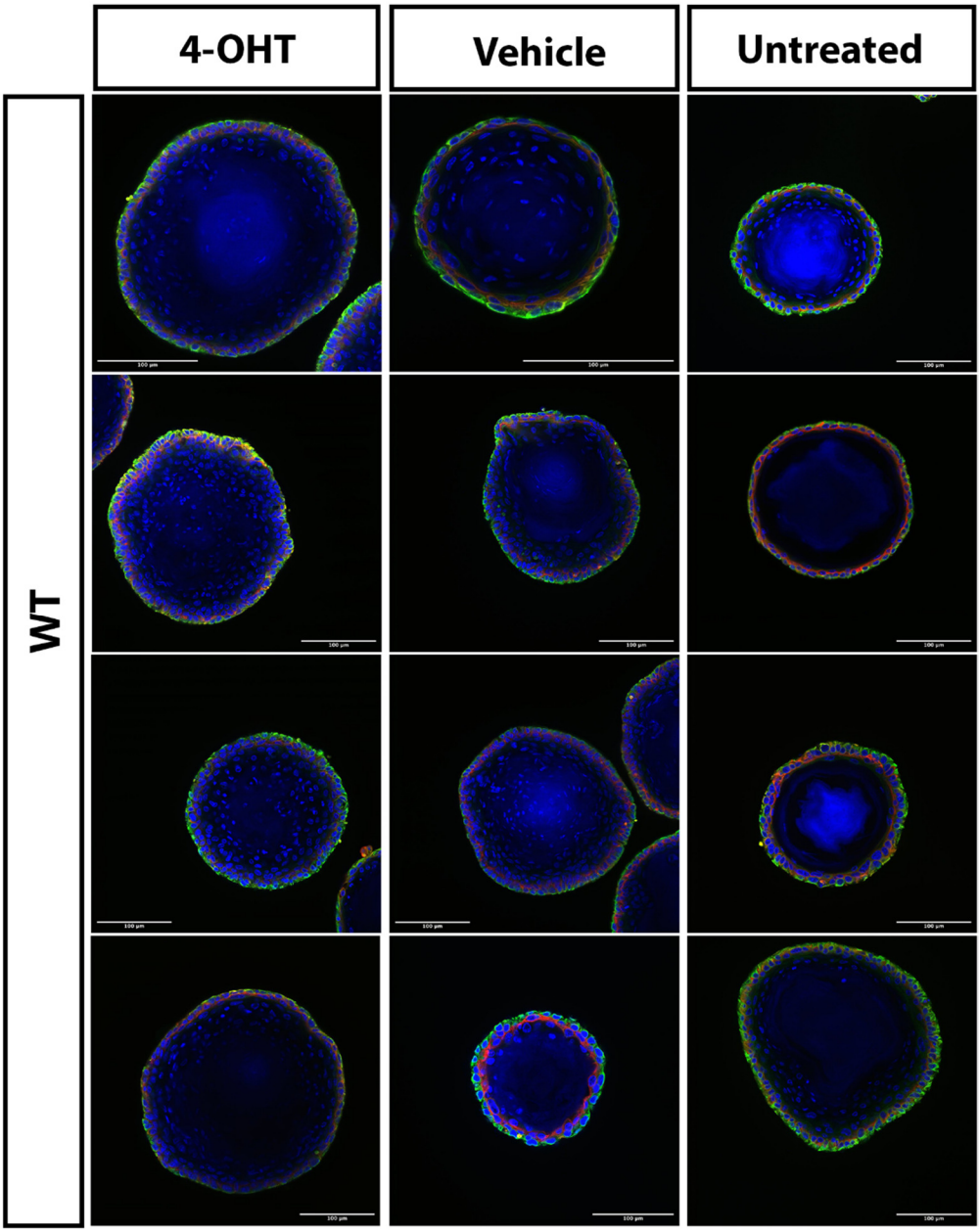


Figure 3.7: Fluorescence microscopy of Whole mount WT oesophageal organoids exposed to 4-OHT vehicle control and left untreated.

Organoids were cultivated from oesophageal epithelial cells of WT mice exposed to 4-OHT or vehicle control after 7 days in culture. Organoids were cultivated for a further 7 days and taken for whole mount analysis. Organoids were stained with a marker of basal cells (Keratin 5 in green), suprabasal cells (Keratin 13 in red) and DAPI in blue to visualise nuclei and keratin core. Scale bar indicates 100 μm.

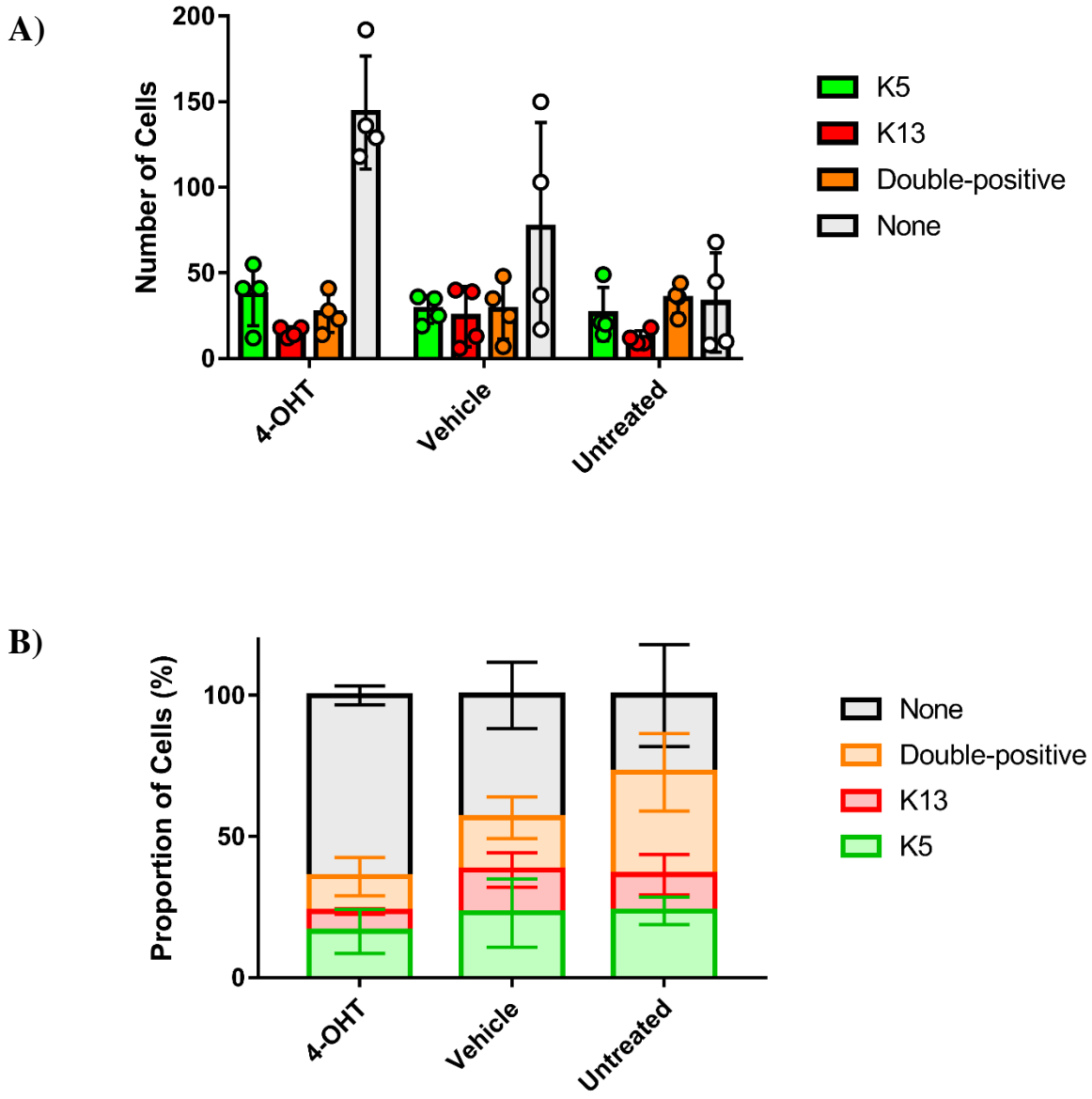


Figure 3.8: Quantification of WT oesophageal organoid microscope images.

Cells from the organoids displayed in figure 3.7 were counted and scored as either Keratin 5, Keratin 13 or double positive. Cells with no marker or indistinguishable in their identity were scored under 'none'. **A)** Depicts the mean number of cells displaying each marker ± S.D. of n=4 technical replicates. **B)** Data presented from (A) as the mean proportion of cells positive for each marker in the organoids ± S.D. of n=4 technical replicates.

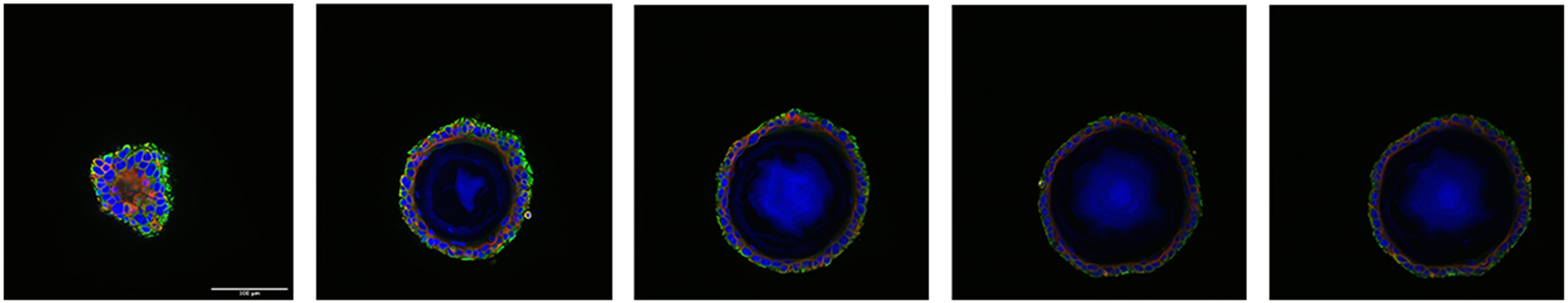


Figure 3.9: Multiple planes of WT oesophageal organoid displays correct differentiation

Using immunofluorescence WT organoids were stained for Keratin 5 (in green) and Keratin 13 (in red) markers of basal and suprabasal cells respectively. For nuclear visualisation DAPI was used, shown in blue. Multiple planes of view were acquired to demonstrate the correct morphology and differentiation gradient expected in organoids of the oesophageal epithelium. Scale bar is 100 µm

Next, organoids were cultivated from the oesophageal epithelium of *BCL11A*^{OE} mice and analysed using a bright field microscope to compare phenotypic differences between groups. Organoids treated with 4-OHT broadly appeared to be smaller in size (figure 3.10). Closer inspection of this group showed evidence of a smaller keratinised core, suggesting overexpression caused an inhibition in differentiation (figure 3.11). To assess whether *BCL11A*^{OE} organoids were smaller in area, average area of organoids was measured using brightfield images (figure 3.11b and c). This analysis demonstrated a decrease in size of *BCL11A*^{OE} organoids compared to the control groups. To confirm that *BCL11A* was indeed over expressed, organoids were subjected to RT-qPCR gene expression analysis (figure 3.12). Analysis revealed over 100-fold increase in *BCL11A* expression when normalised to *Gapdh,* indicating effective induction of the transgene in 4-OHT treated cells.

As mentioned previously, the construct in this transgenic line contained *eGFP* (see figure 2.1), which should be expressed once recombination has occurred. To assess whether cells were fluorescent, 4-OHT treated organoids were analysed for *eGFP* expression (figure 3.13). Organoids treated with 4-OHT did display strong fluorescence levels suggesting *eGFP* was expressed. The 4-OHT induction protocols were identical between both mammary gland and oesophageal cell culture, suggesting that tissue specific 4-OHT induction protocols may be a necessary consideration for future work as in the mammary organoids, no *eGFP* expression was observed.

To assess alterations, if any, *BCL11A* overexpression caused in the cellular composition of oesophageal organoids, they were taken for whole mount analysis. The analysis of *BCL11A* over expressing organoids displayed a phenotype distinct from that viewed in the control groups. Over expressing organoids appeared to have irregular nuclei and a less well-defined gradient of Keratin 5 and Keratin 13 expression (figure 3.14). Some organoids also displayed a lack of keratinisation or keratinisation inconsistent with that seen in the control groups. Together this was consistent with a lack of the differentiation normally observed in oesophageal organoids. Cells appeared disorganised and not in the uniform conformation observed in control groups. In many cases, cells around the outermost layer of the organoids displayed positivity for both the basal Keratin 5 and suprabasal Keratin 13 markers. This phenomenon was observed, albeit to a lesser extent, in some organoids of both control groups (figure 3.14). Closer analysis of the control groups revealed some evidence of this double positive expression more clearly, however this tended to be to a lower extent than in the over

expressing organoids (figure 3.14). To try and characterise the cellular composition of the organoids more comprehensively, the microscope images were quantitated (figure 3.15). Interestingly, as can be seen in figure 3.14, none of the organoids displayed a clear differentiation gradient, including a lack of cells positive for only Keratin 13. Although this was not an expected result, again, it did not preclude meaningful conclusions from the data. Organoids over expressing *BCL11A* largely contained cells that were not positive for any markers, as well as those that were double positive. This suggested that *BCL11A* over expression had an effect on differentiation within the organoids. This can be seen in closer detail in figures 3.16, 3.17 and 3.18. Despite organoids from all groups containing double positive cells, there were clear morphological differences between over expressing organoids and control groups. *BCL11A* over expressing organoids appeared not to have the smooth outer layer of cells found in the control groups.

To understand more about the keratinisation status of over expressing organoids, multiple planes of single organoids were analysed (Figure 3.19). The over expressing organoid appeared to show a lack of the keratinisation observed in the vehicle control group (figure 3.19). In addition, the nuclei of cells in the over expressing organoids appear to show many areas of high density. Nuclei were not uniform in size and appeared disorganised in comparison to the vehicle control group. Taken together with the lack of defined markers as well as the high proportion of non-labelled cells observed in the *BCL11A* over expressing organoids, this could represent the loss of internal organisation and differentiation seen in the control organoids. This could suggest that the over expression of this transcription factor affected cellular differentiation.

Finally, in order to assess whether *BCL11A* overexpression had any effect on the colony forming capacity of cells, over expressing and vehicle treated organoids were passaged. Cells were seeded at a density of 500 cells per well and cultivated for 14 days. Over expressing organoids tended to display a greater colony forming capacity as compared to vehicle control. This suggested that the overexpression of *BCL11A* aided colony formation, increasing cell viability and stemness when induced in oesophageal organoids (figure 3.20).

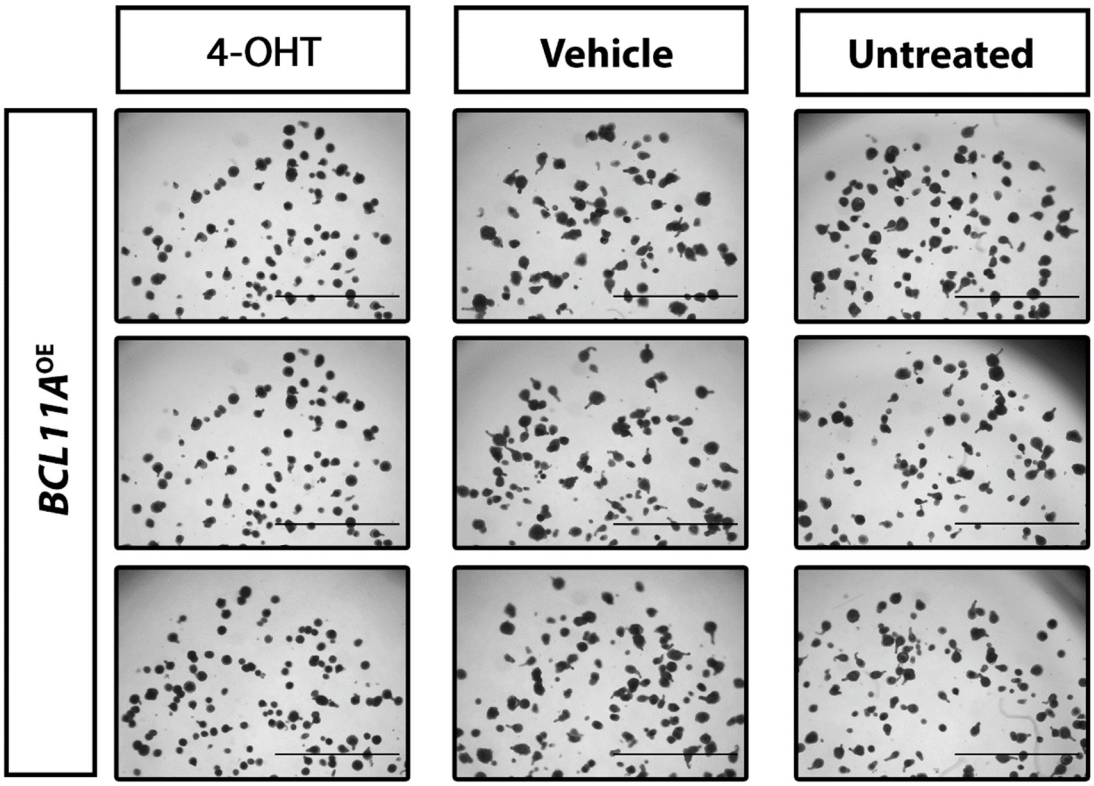


Figure 3.10: Brightfield view of *BCL11A*^{OE} oesophageal organoids after 14 days in culture.

Images show representative wells. Organoids were cultivated from oesophageal cells of *BCL11A*^{OE} mice for 7 days. Organoids were then exposed to 4-OHT, vehicle control and cultivated for a further 7 days or left untreated for the duration. Scale bars represent 2000 μm.

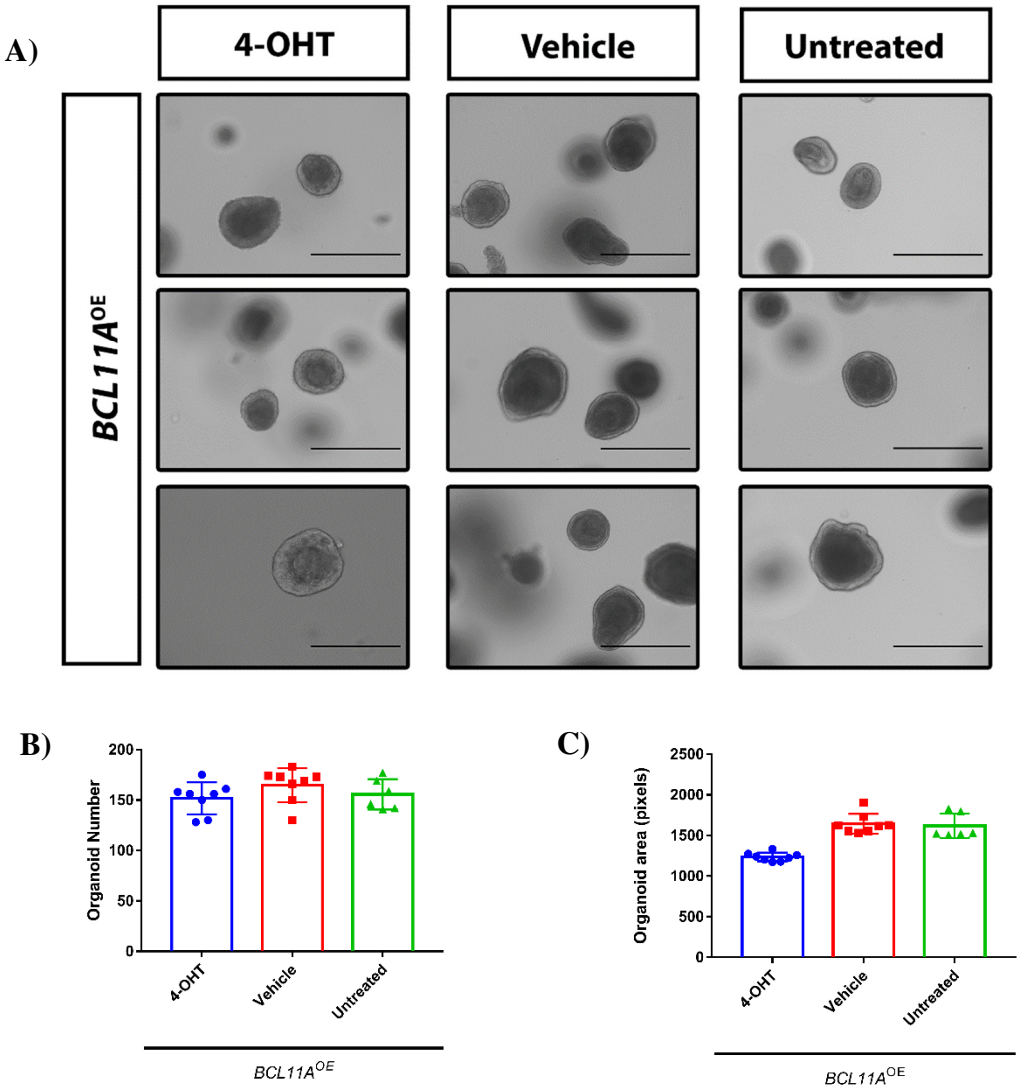


Figure 3.11: Brightfield view of individual *BCL11A*^{OE} oesophageal organoids and quantification of organoid number and size.

Organoids were cultivated from *BCL11A*^{OE} mice for 7 days and exposed to 4-OHT, ethanol vehicle control or left untreated. Organoids were then cultivated for a further 7 days. **A)** Figure shows individual organoids from each condition. Scale bar represents 400 µm. **B)** Shows organoid number cultivated from 1000 cells after 14 days in culture for varying conditions **C)** Shows the size of organoids in pixels. Data are represented as mean ± S.D. n=6-8 technical replicates.

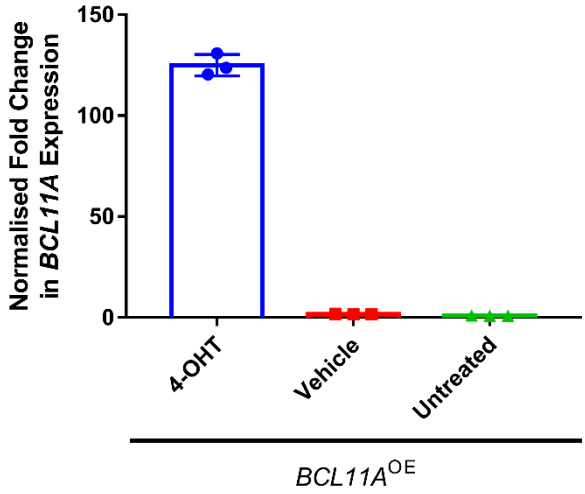


Figure 3.12: Analysis of *BCL11A* expression in oesophageal organoids cultivated from *BCL11A*^{OE} mice.

After 14 days, organoids were collected and RNA isolated. After reverse transcription, the relative abundance of cDNA for the *BCL11A* gene was quantified using qPCR. Shown is the expression of *BCL11A* relative to that of *Gapdh* for organoids treated with 4-OHT, vehicle control or left untreated (by the same protocol as in figure 3.9).

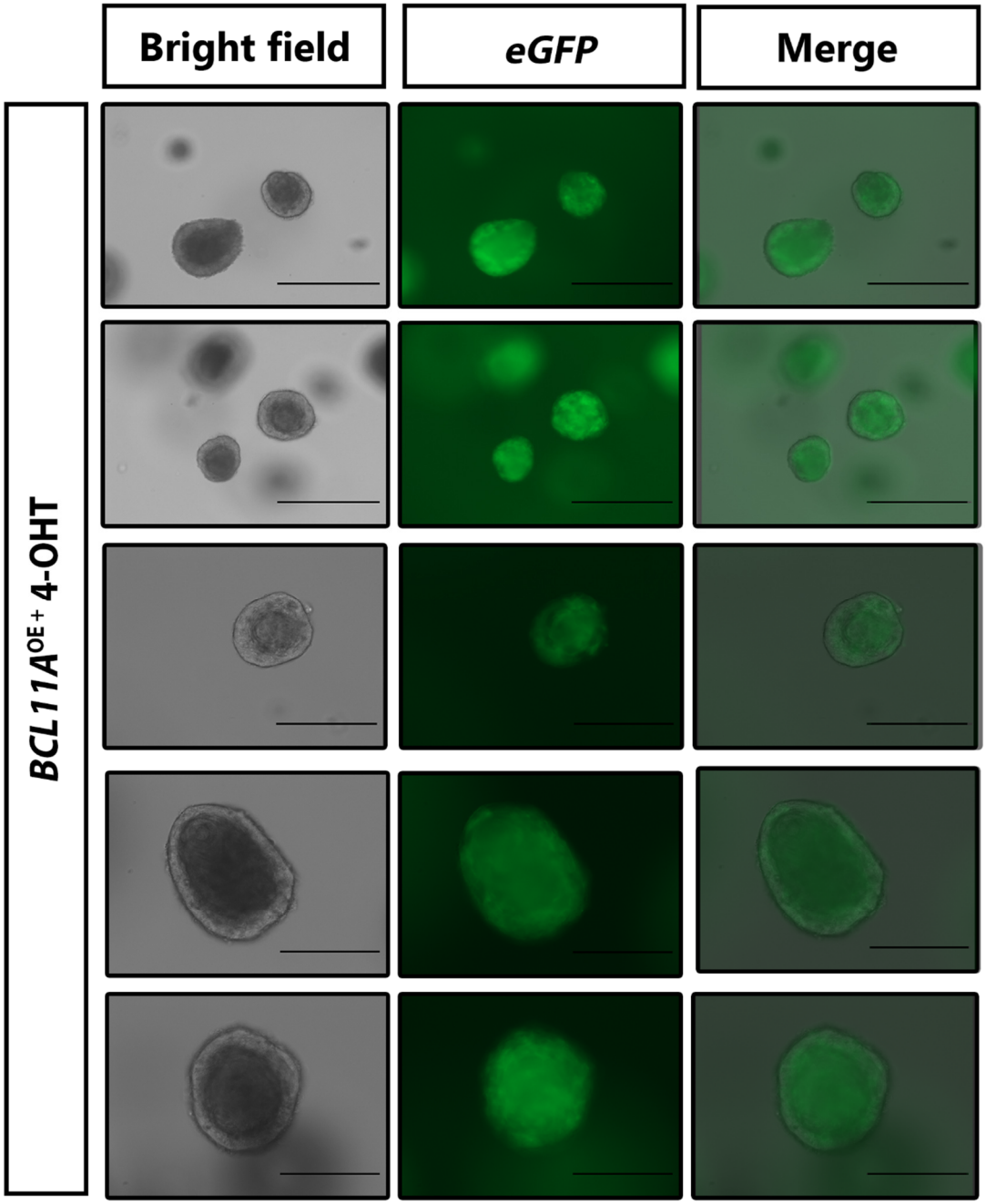


Figure 3.13: Brightfield and GFP view of individual *BCL11A*^{OE} oesophageal organoids.

Oesophageal organoids were cultivated from *BCL11A*^{OE} mice as described in figure 3.10. Organoids were visualised through a GFP channel. Scale bar represents 400 µm in the top 3 rows and 200 µm in the bottom 2 rows.

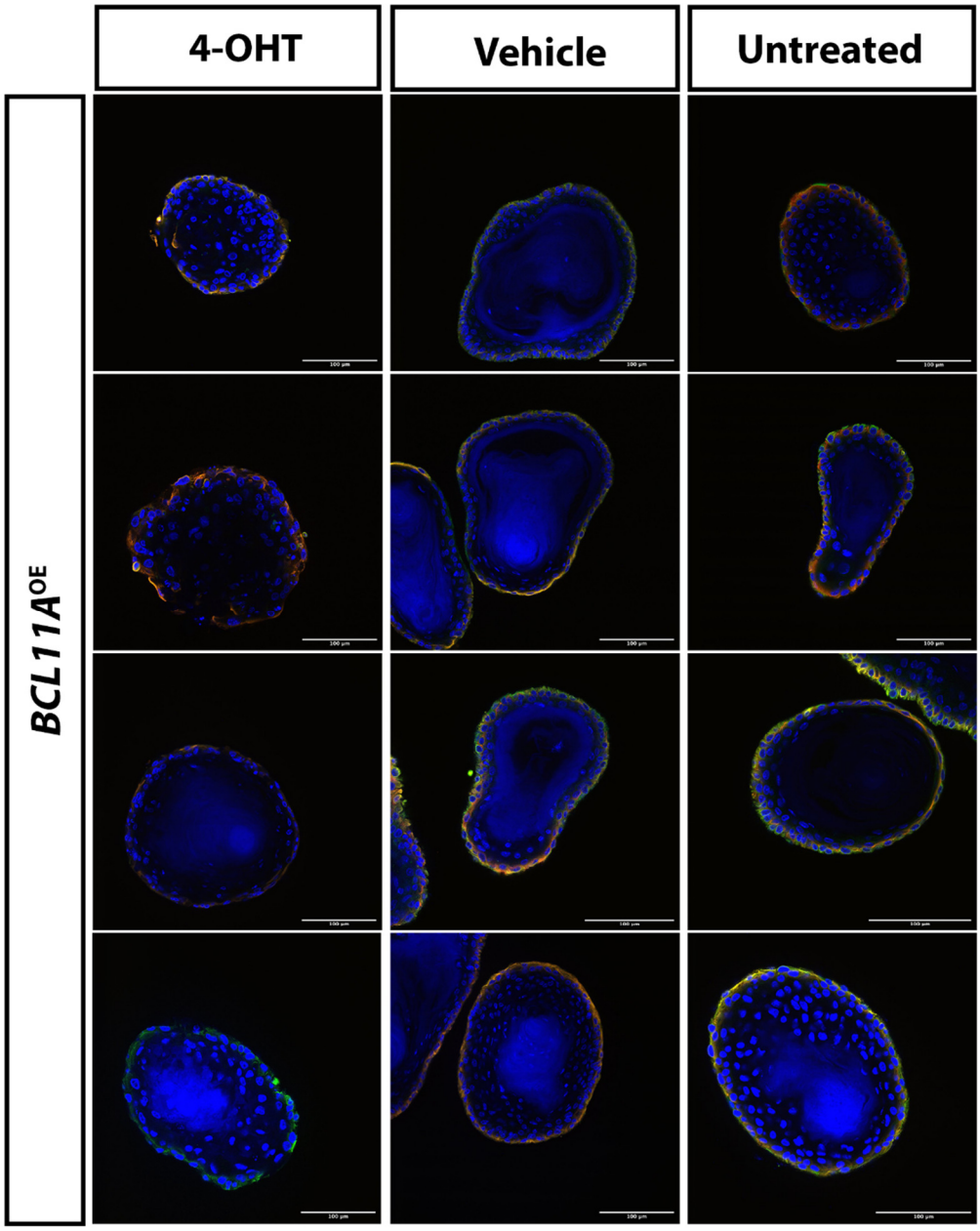


Figure 3.14: Fluorescence microscopy of whole mount *BCL11A*^{OE} oesophageal organoids.

Oesophageal organoids cultivated from *BCL11A*^{OE} mice were cultured for 7 days and treated with 4-OHT or ethanol vehicle control or left untreated. Organoids were cultured for a further 7 days and taken for whole mount analysis. Organoids were stained with a marker of basal cells (Keratin 5 in green), suprabasal cells (Keratin 13 in red) and DAPI in blue to visualise nuclei and Keratin core. Scale bar indicates 100 μm.

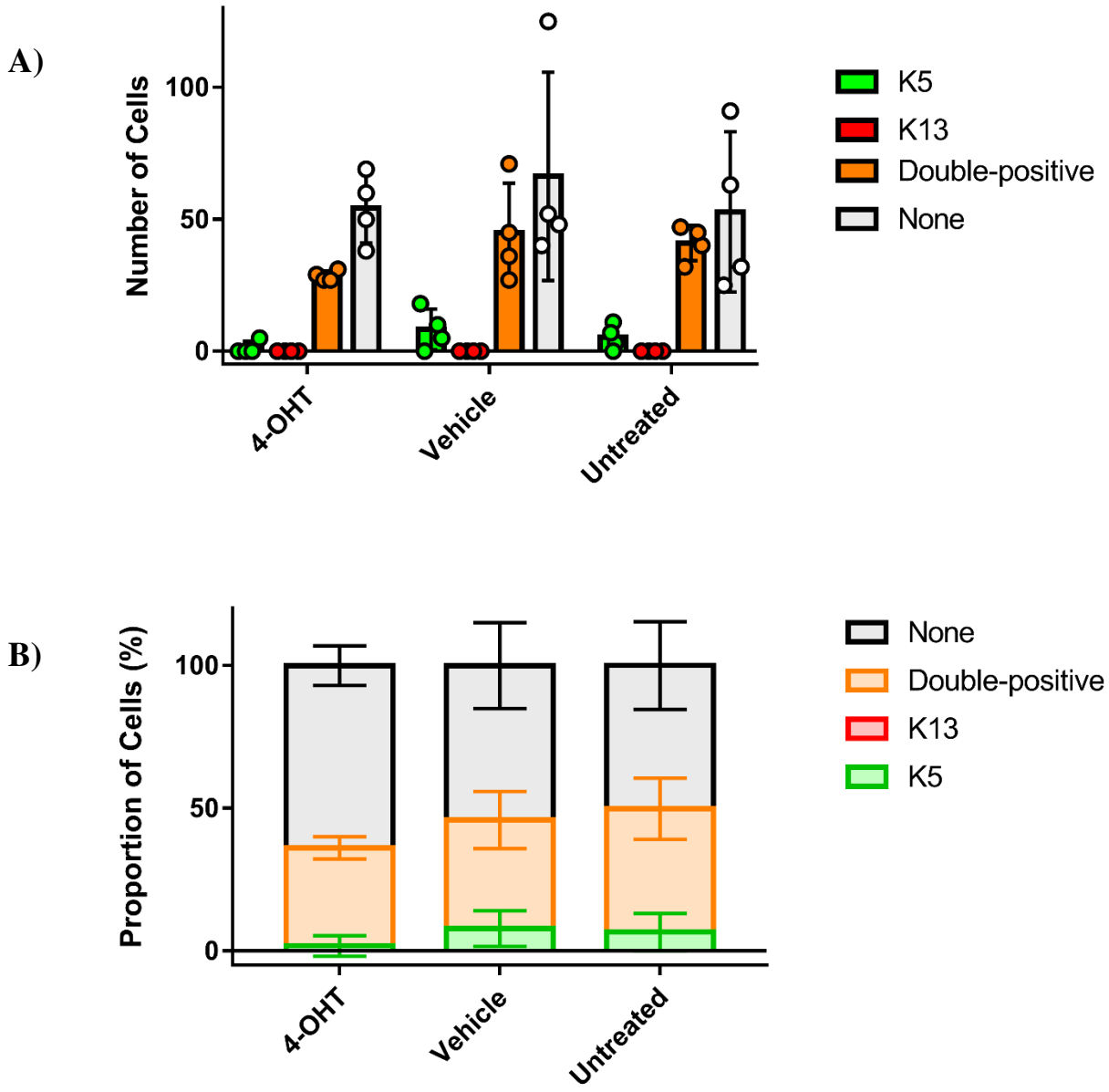


Figure 3.15: Quantification of *BCL11A*^{OE} oesophageal organoid microscope images.

Cells from the organoids displayed in figure 3.14 were counted and scored as either Keratin 5, Keratin 13 or double positive. Cells with no marker or indistinguishable in their identity were scored under 'none'. **A)** Depicts the mean number of cells displaying each marker ± S.D. of n=4 technical replicates. **B)** Data from (A) presented as the mean proportion of cells positive for each marker in the organoids ± S.D. of n=4 technical replicates.

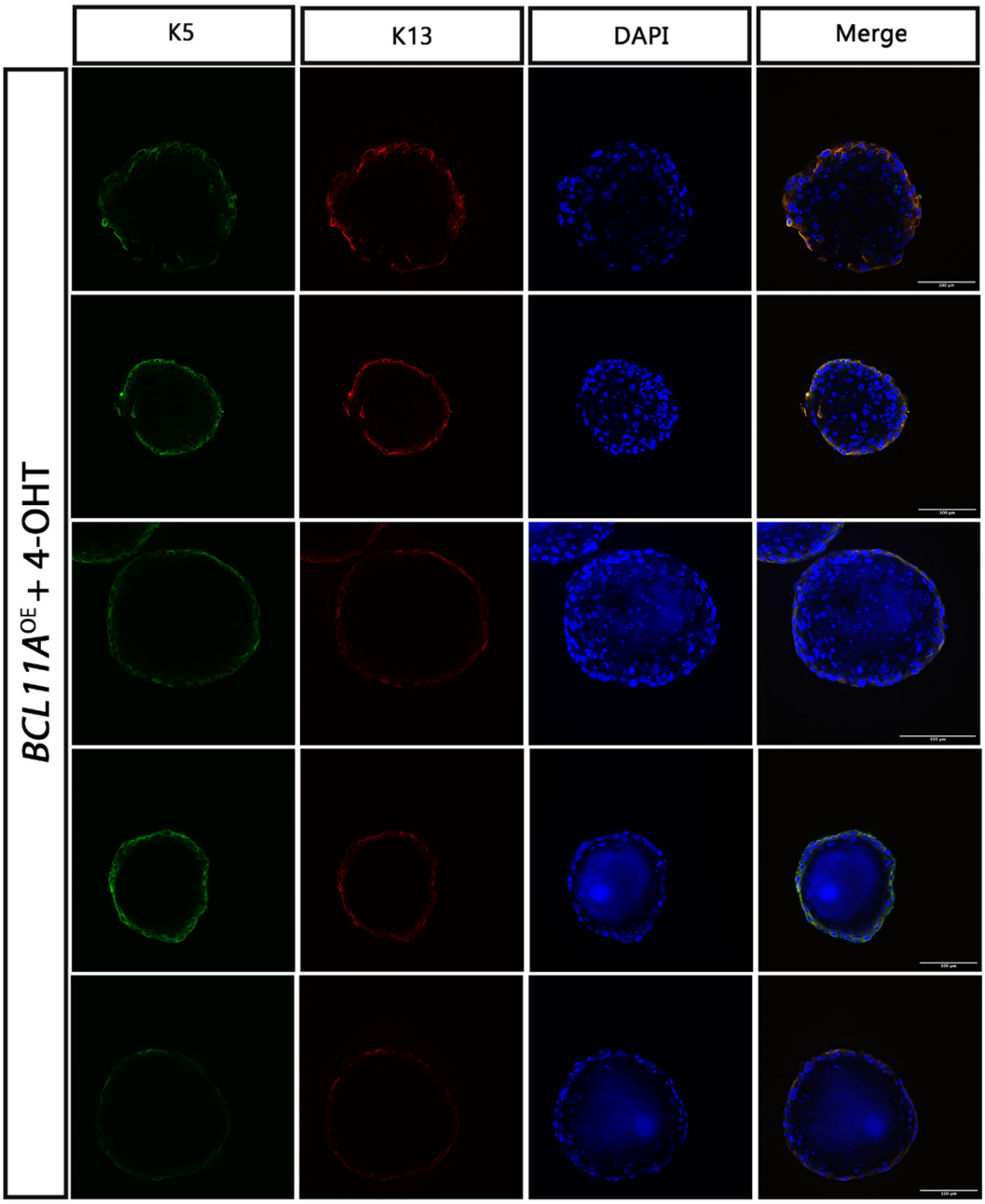


Figure 3.16: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with 4-OHT.

Organoids were cultivated from *BCL11A*^{OE} mice and treated with 4-OHT, (see figure 3.14 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of basal cells (Keratin 5 in green), a marker of suprabasal cells (Keratin 13 in red) and DAPI in blue for nuclear and keratin visualisation. Scale bar indicates 100 µm.

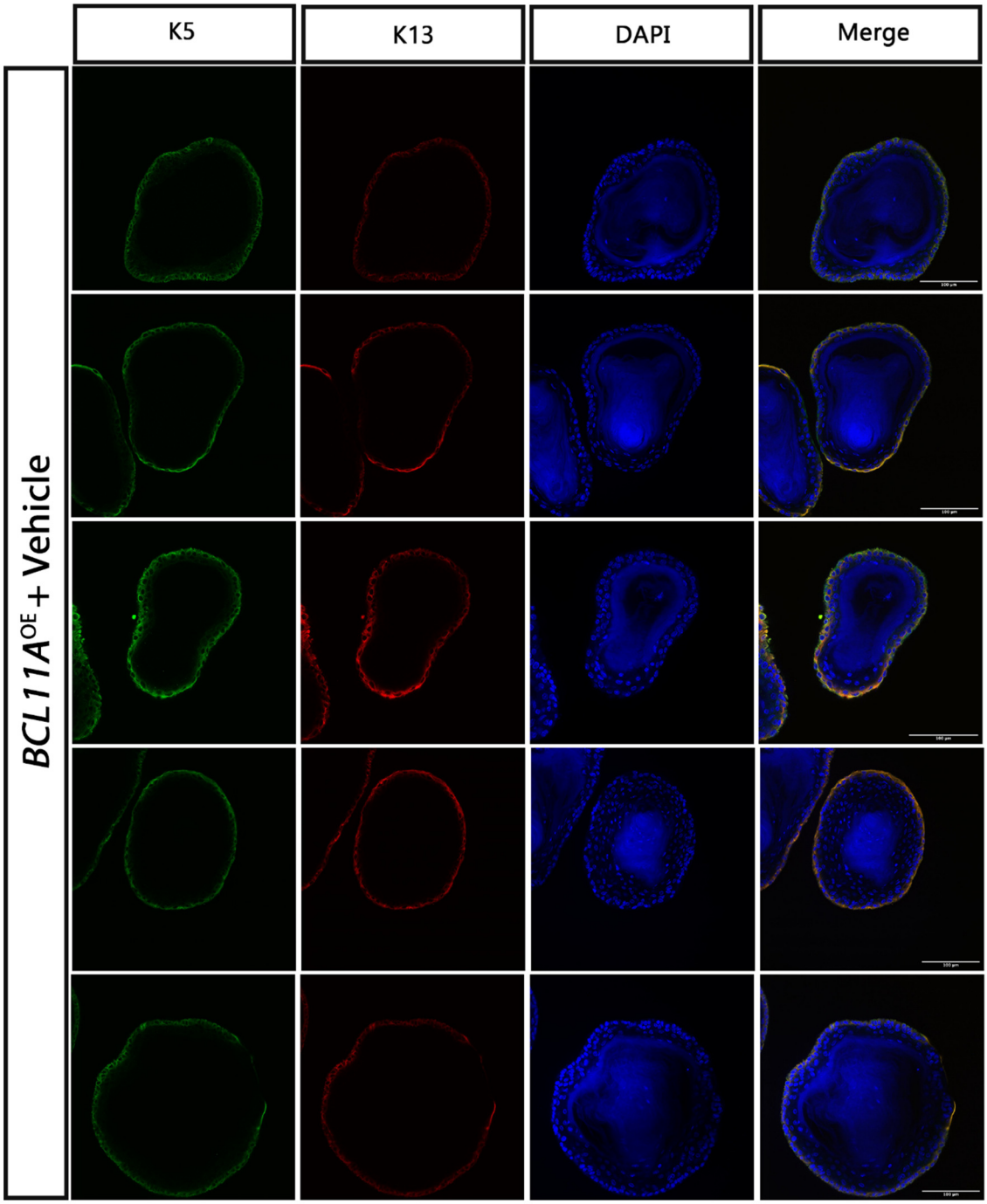


Figure 3.17: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with vehicle.

Organoids were cultivated from *BCL11A*^{OE} and treated with ethanol vehicle control, (see figure 12 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of basal cells (Keratin 5 in green), a marker of suprabasal cells (Keratin 13 in red) and DAPI in blue for nuclear visualisation. Scale bar indicates 100 µm

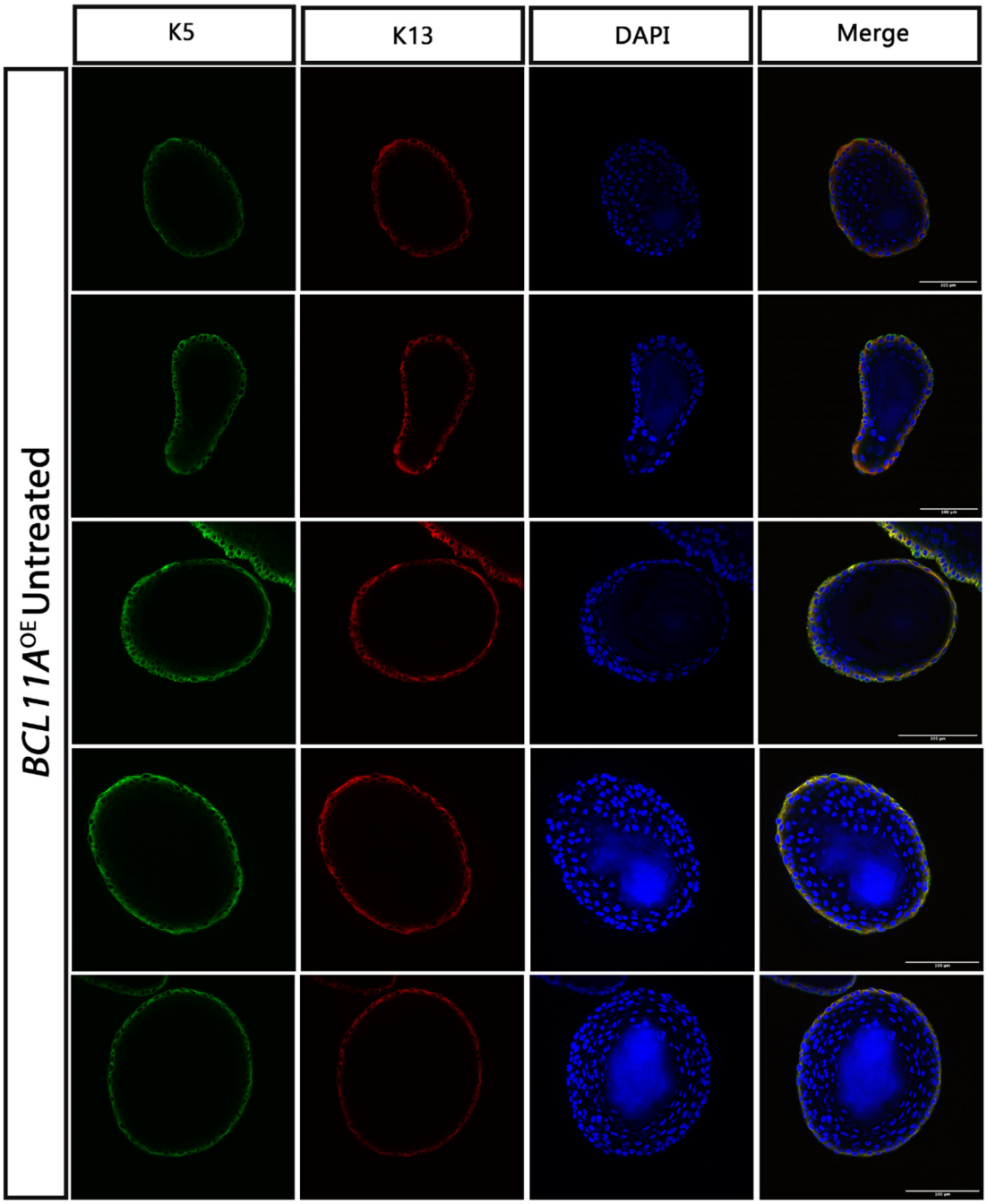


Figure 3.18: Fluorescence microscope images of whole mount untreated *BCL11A*^{OE} organoids.

Organoids were cultivated from *BCL11A*^{OE} mice and left untreated and cultivated for a period of 14 days. Organoids were stained with a marker of basal cells (Keratin 5 in green), a marker of suprabasal cells (Keratin 13 in red) and DAPI for nuclear and keratin visualisation. Scale bar indicates 100 µm

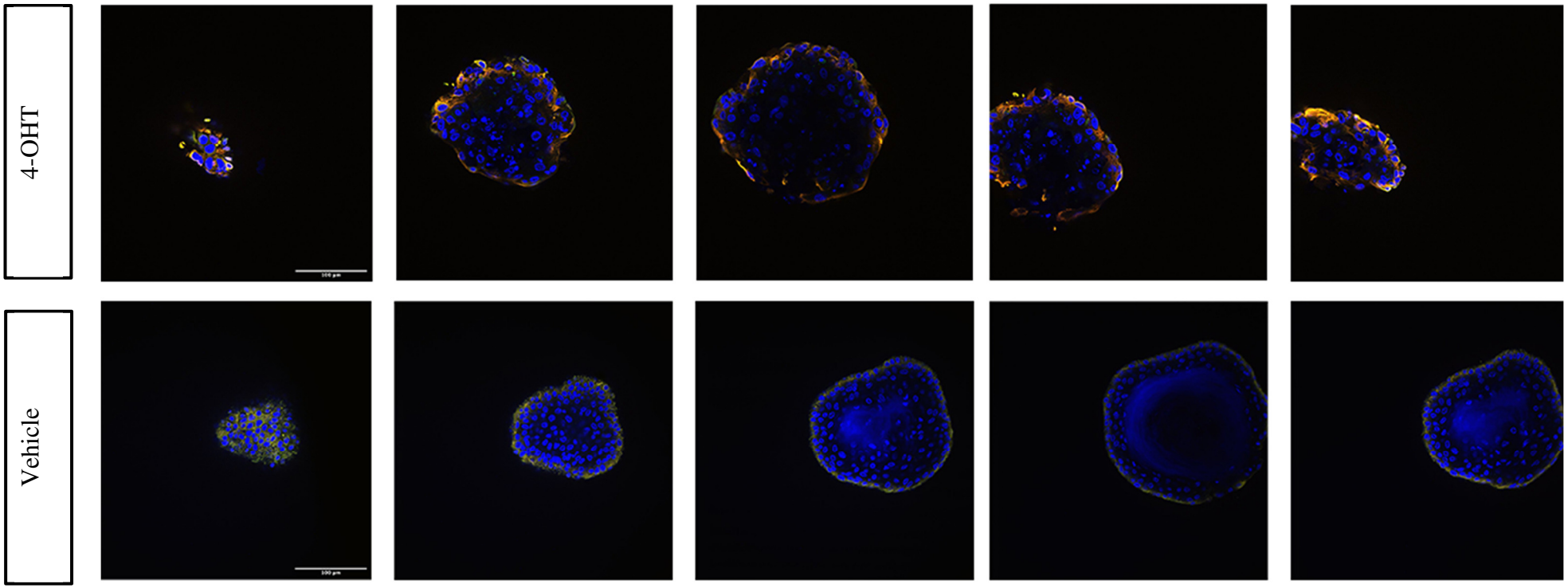
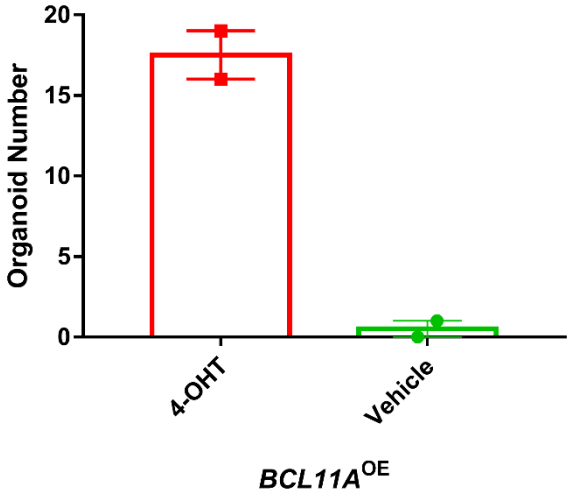


Figure 3.19: Multiple planes of *BCL11A*^{OE} organoids treated with 4-OHT or vehicle control.

Organoids were cultivated and stained as in figure 3.14 from 4-OHT and vehicle treated groups, multiple planes of view (Z stack) were acquired in a single field using a fluorescence microscope. Scale bar is 100 µm. Keratin 5 is shown in green, Keratin 13 in red and DAPI in blue.



`

Figure 3.20: Colony forming efficiency of passaged *BCL11A*^{OE} organoids.

BCL11A^{OE} organoids treated with 4-OHT or vehicle were passaged and 500 cells plated in 2 wells per condition. Shown are the number of organoids counted after 14 days in culture.

3.2.3 *Bcl11a* is Critical for Oesophageal Organoid Survival

In order to assess the effect of *Bcl11a* deletion on oesophageal organoids, cells were isolated from the *Bcl11a*^{KO} mice and seeded as described above. Similar to the cultivation of *BCL11A*^{OE} organoids, after 7 days in culture, organoids were treated with either 4-OHT, vehicle control or left untreated. Organoids were then cultivated for a further 7 days after which they were subjected to the same analyses of whole mount staining and RT-qPCR, as well as analysis of colony size and number.

Bcl11a^{KO} oesophageal organoids cultivated in this manner appeared to display phenotypic differences between 4-OHT treated and control wells (Figure 3.21). *Bcl11a*^{KO} organoids treated with 4-OHT appeared to be restricted in their size while organoids in the control groups appeared larger in diameter. Closer inspection of the 4-OHT treated organoids revealed divergent morphology between these and control groups (figure 3.22). Organoids from control groups appeared to display a uniform spherical morphology and had the appearance of a keratinised core; the 4-OHT treated organoids looked far smaller in size and seemed to exhibit a blebbing phenotype. To assess whether this was representative, other fields of vision were captured (figure 3.22). In all fields, organoids appeared to display evidence of blebbing to some extent. Although there was some variability, other organoids appeared to show extensive blebbing and a loss of structural integrity (figure 3.22b). This phenotype could be indicative of cells within the organoid dying, thus causing a loss in integrity of the overall structure. This suggests that *Bcl11a* could be acting in a pro-proliferative capacity in cells of the oesophageal epithelium. To measure the organoid size difference between conditions, the average pixel area of each organoid was calculated and compared (figure 3.21b and c). *Bcl11a* deleted organoids tended to be smaller than those in the control groups. As this experiment consisted of multiple technical replicates and only one biological replicate, further experimentation will be required to fully characterise this phenotype. However, the overall smaller size in the knock out organoids, coupled with the blebbing phenotype could suggest a critical role for *Bcl11a* in the growth and maintenance of organoids.

To confirm that *Bcl11a* was indeed deleted, organoids were collected for gene expression analysis via RT-qPCR. This analysis, after normalising to a *Gapdh* internal control, confirmed the knockout of *Bcl11a* in organoids treated with 4-OHT evident in comparison to organoids

cultivated in both control conditions (figure 3.23). This suggested that *Bcl11a* knockout was responsible for causing arrested growth and extensive blebbing in oesophageal organoids.

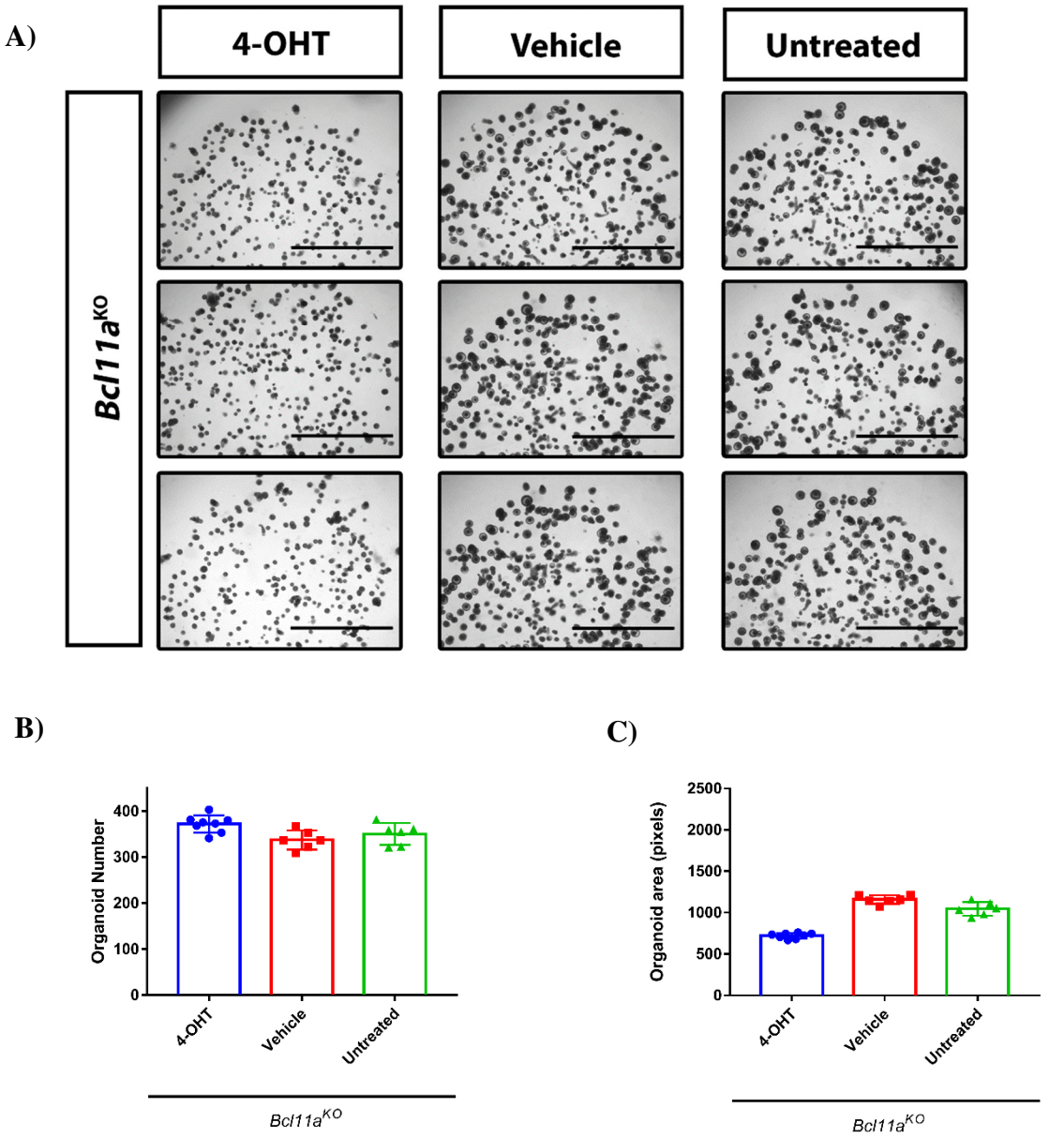


Figure 3.21: Representative brightfield view of *Bcl11a*^{KO} oesophageal organoids and quantification of organoid number and size after 14 days in culture.

Organoids were cultivated from oesophageal epithelial cells of *Bcl11a*^{KO} mice for 7 days and treated with either 4-OHT or ethanol vehicle control or left untreated and cultured for a further 7 days. **A)** Data represents an overview of organoids cultivated in respective conditions Scale bar indicates 2000 µm. **B)** shows the number of organoids derived from 1000 cells per well. **C)** shows the mean size of organoids in pixels per well. Data are represented as mean ± S.D. of n=6-8 technical replicates.

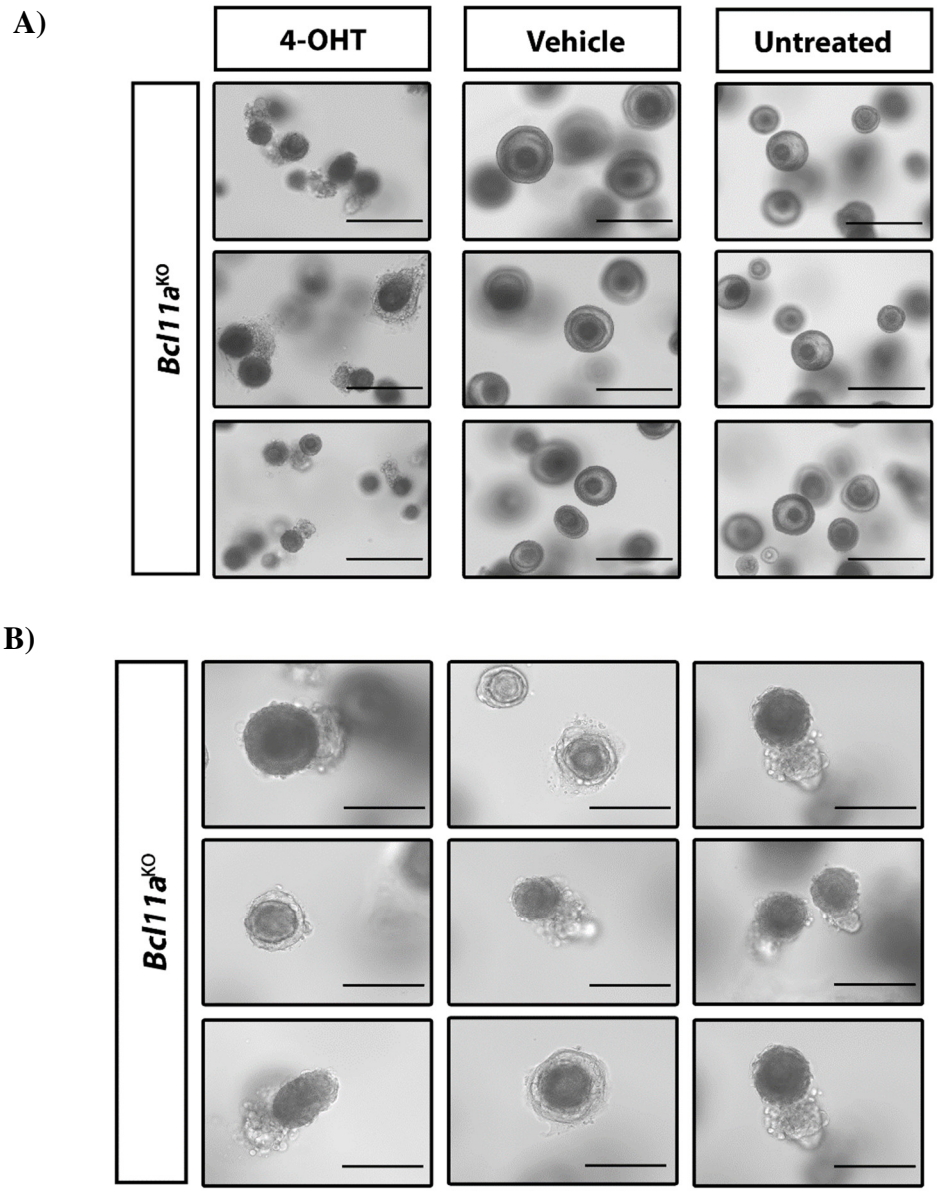


Figure 3.22: Individual *Bcl11a*^{KO} oesophageal organoids after 14 days in culture.

Organoids were cultivated from the oesophageal epithelial cells of *Bcl11a*^{KO} mice and treated as described in figure 3.20. **A)** Data are representative images of individual organoids from each condition. Scale bar is 400 µm. **B)** Shown are representative wells of organoids treated with 4-OHT only. Scale bar is 200 µm.

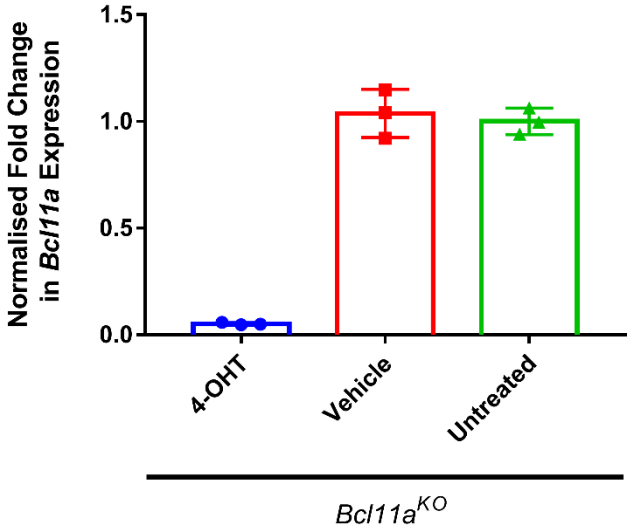


Figure 3.23: Analysis of *Bcl11a* expression in oesophageal organoids cultivated from *Bcl11a*^{KO} mice.

After 14 days, organoids were collected and RNA isolated. After reverse transcription, the relative abundance of cDNA for the *Bcl11a* gene was quantified using qPCR. Shown is the expression of *Bcl11a* relative to that of *Gapdh* for organoids treated with 4-OHT, vehicle control or left untreated (by the same protocol as in figure 3.21).

Next, to understand whether *Bcl11a* deletion had any effect on the cellular composition of the organoids, these were subjected to whole mount analysis. Similar to the previous conditions, organoids were stained with Keratin 5, Keratin 13 and DAPI to visualise basal and suprabasal cells as well as keratinisation and nuclei respectively. This analysis revealed divergent morphology across knock out organoids and control groups (figure 3.24). Knock out organoids appeared to demonstrate a lack of sufficiently differentiated epithelial cells. Some organoids displayed positivity for both cytokeratin markers in cells on the outermost basal layer, whilst others appeared to show few cells stained with Keratin 13 only (figure 3.24). In order to quantitate this observation, cells were counted and scored as either Keratin 5, Keratin 13 or double positive. Cells with no discernible marker were scored under 'none' (figure 3.25). The knock out organoids tended to be composed of fewer cells, consistent with the size based analysis in figure 3.21. The proportions of cell subpopulations appeared to be largely similar across all organoids, suggesting there was little difference in cellular composition caused by *Bcl11a* knock out. Closer inspection of control organoids demonstrated many cells in the suprabasal layers in comparison with the knockout organoids which contained fewer cells in this layer as was evident from the DAPI staining (figure 3.26, 3.27 and 3.28). Interestingly, some of the organoids from the control groups demonstrated oval and elongated morphology. This was consistent with the gross morphology observed and the presence of what appeared to be the protrusion of the keratin layer.

Closer inspection of 4-OHT treated organoids revealed the abnormal appearance of nuclei (Figure 3.26). The nuclei in cells of these organoids appeared distended with points of density in the DAPI staining. This could potentially indicate the condensation of chromatin, which is a hallmark of apoptotic cells, further reinforcing the notion that *Bcl11a* may play a pro-proliferative role in the oesophageal epithelium. To understand the extent of keratinisation in the 4-OHT treated organoids, multiple planes of view were acquired and compared to vehicle control (figure 3.29). Although knock out organoids do possess some evidence of a keratinised core, the nuclei continue to be visible inside the organoid. The visibility of the DAPI stain here may also suggest a defect in differentiation caused by *Bcl11a* knock out. In addition, 4-OHT treated organoids were composed of fewer cells than the vehicle control group. This was consistent with the notion that *Bcl11a* could play a pro-proliferative role in the oesophageal epithelium as its knock out is seen to severely impact organoid growth

To understand more about the effect of *Bcl11a* deletion on the seeding capacity of the epithelial cells, 4-OHT and vehicle control treated organoids were dissociated and reseeded as single cells. For each condition, 500 cells were plated and cultivated for 14 days (figure 3.30). *Bcl11a* knock out cells generated fewer organoids in comparison to the vehicle control cells (figure 3.30). Taken together this suggests that the ablation of *Bcl11a* expression hinders organoid formation, suggesting that *Bcl11a* may play a role in maintaining oesophageal progenitor function.

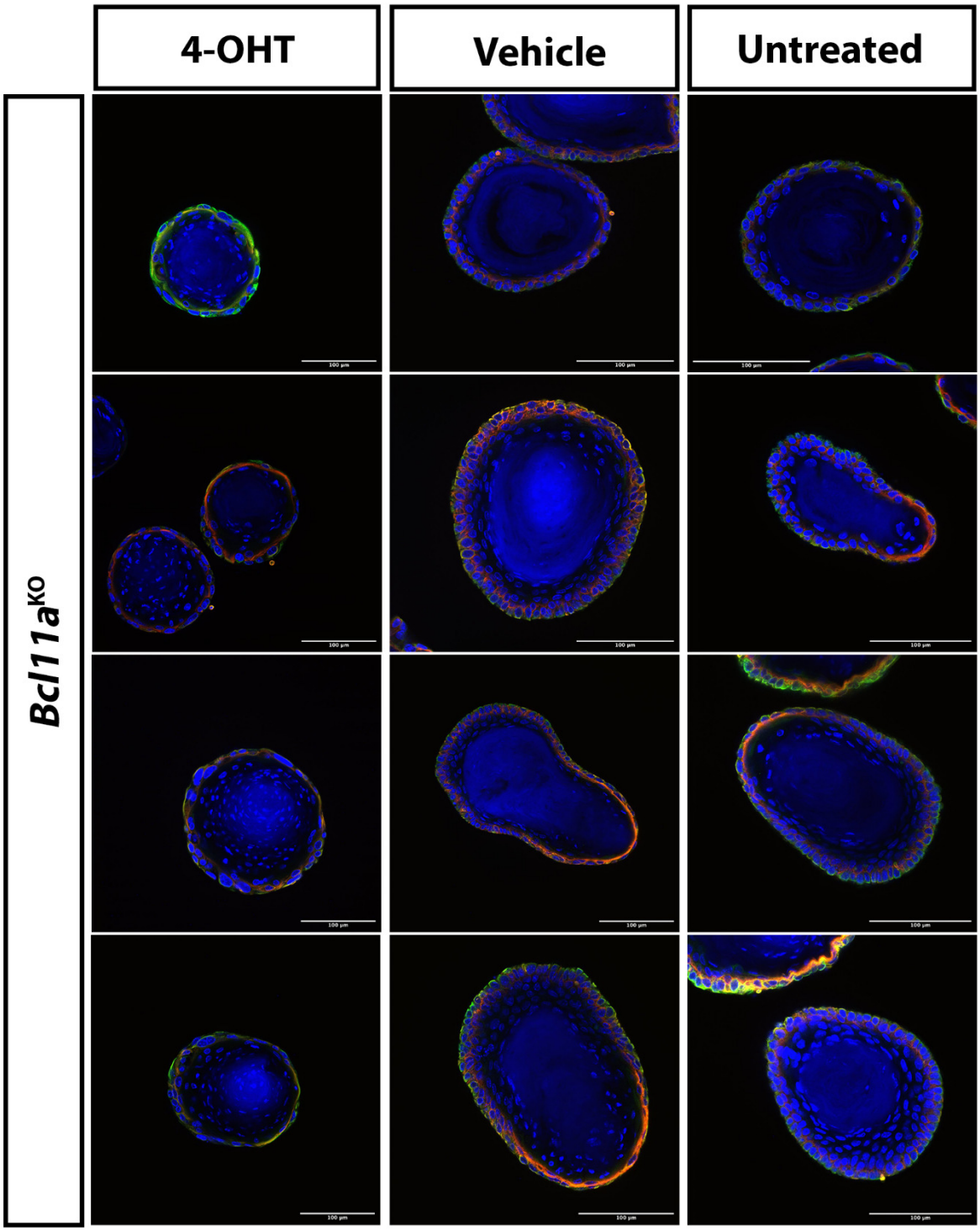


Figure 3.24: Fluorescence microscopy of whole mount *Bcl11a*^{KO} oesophageal organoids.

Figure 3.24 (cont.): Fluorescence microscopy of whole mount *Bcl11a*^{KO} oesophageal organoids.

Oesophageal organoids were cultivated for 14 days and treated with 4-OHT, vehicle control or left untreated. Organoids were cultured for a further 7 days and taken for whole mount analysis. Organoids were stained with a basal marker (Keratin 5 in green) a suprabasal marker, (Keratin 13 in red) and DAPI in blue to aid nuclear visualisation. Scale bar represents 100 μm.

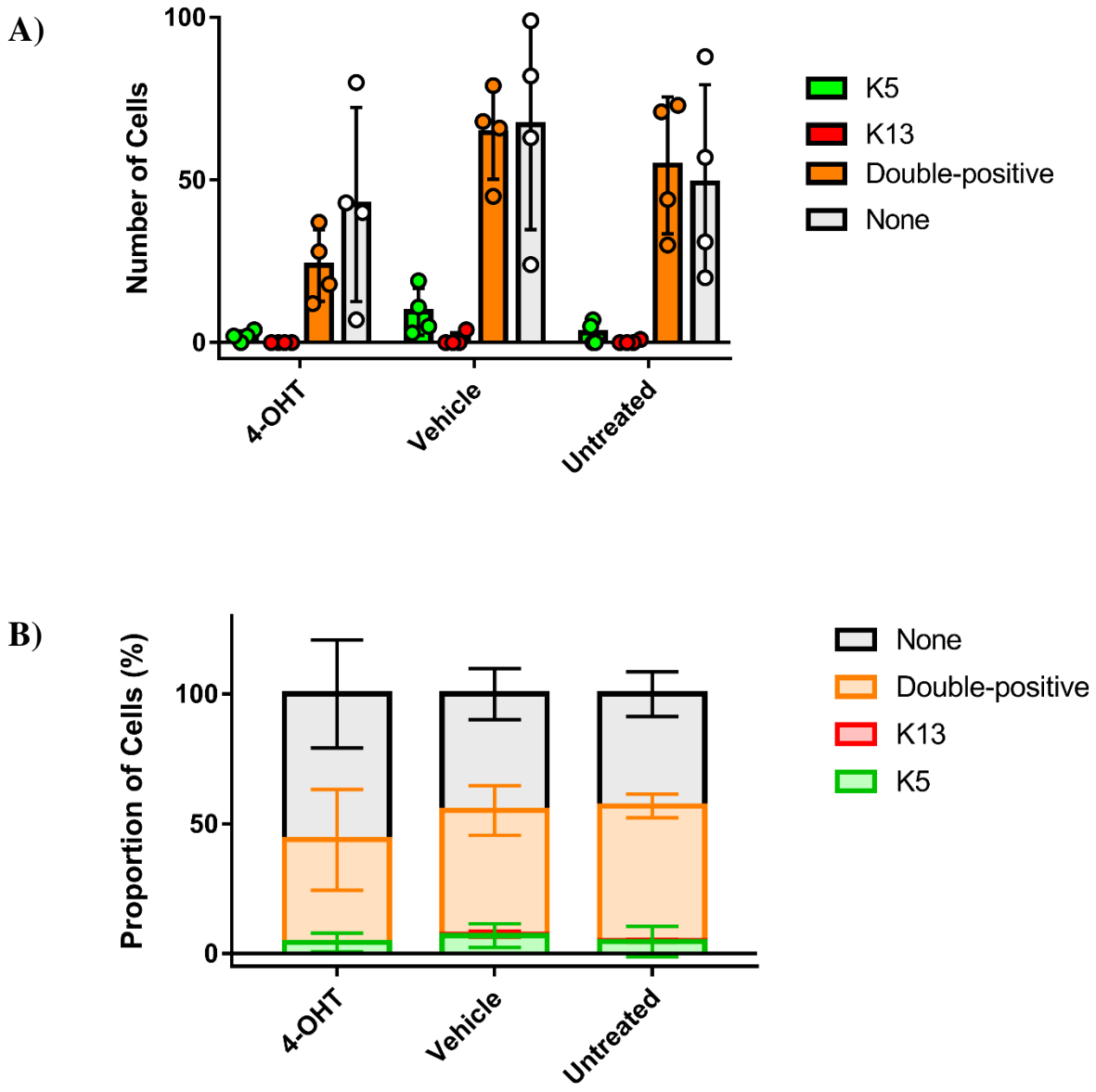


Figure 3.25: Quantification of *Bcl11a*^{KO} oesophageal organoid microscope images.

Cells from the organoids displayed in figure 3.24 were counted and scored as either Keratin 5, Keratin 13 or double positive. Cells with no marker or indistinguishable in their identity were scored under 'none'. **A)** Depicts the mean number of cells displaying each marker ± S.D. of n=4 technical replicates. **B)** Numerical data from (A) presented as the mean proportion of cells positive for each marker in the organoids ± S.D. of n=4 technical replicates.

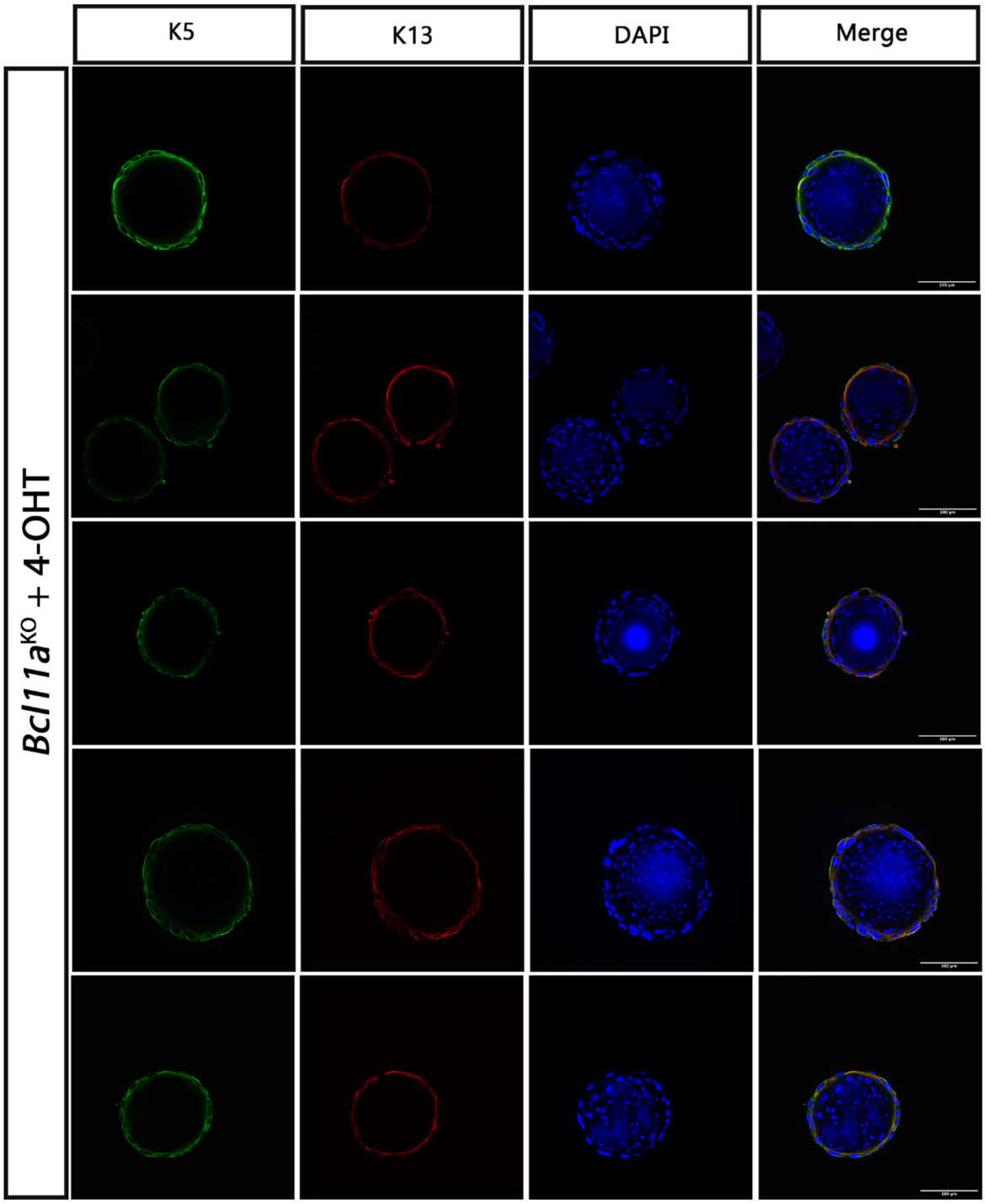


Figure 3.26: Fluorescence microscope images of whole mount *Bcl11a*^{KO} oesophageal organoids treated with 4-OHT.

Organoids were cultivated from *Bcl11a*^{KO} mice and treated with 4-OHT, (see figure 3.25 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of basal cells (Keratin 5 in green), a marker of suprabasal cells (Keratin 13 in red) and DAPI in blue for nuclear and keratin visualisation. Scale bar represents 100 µm.

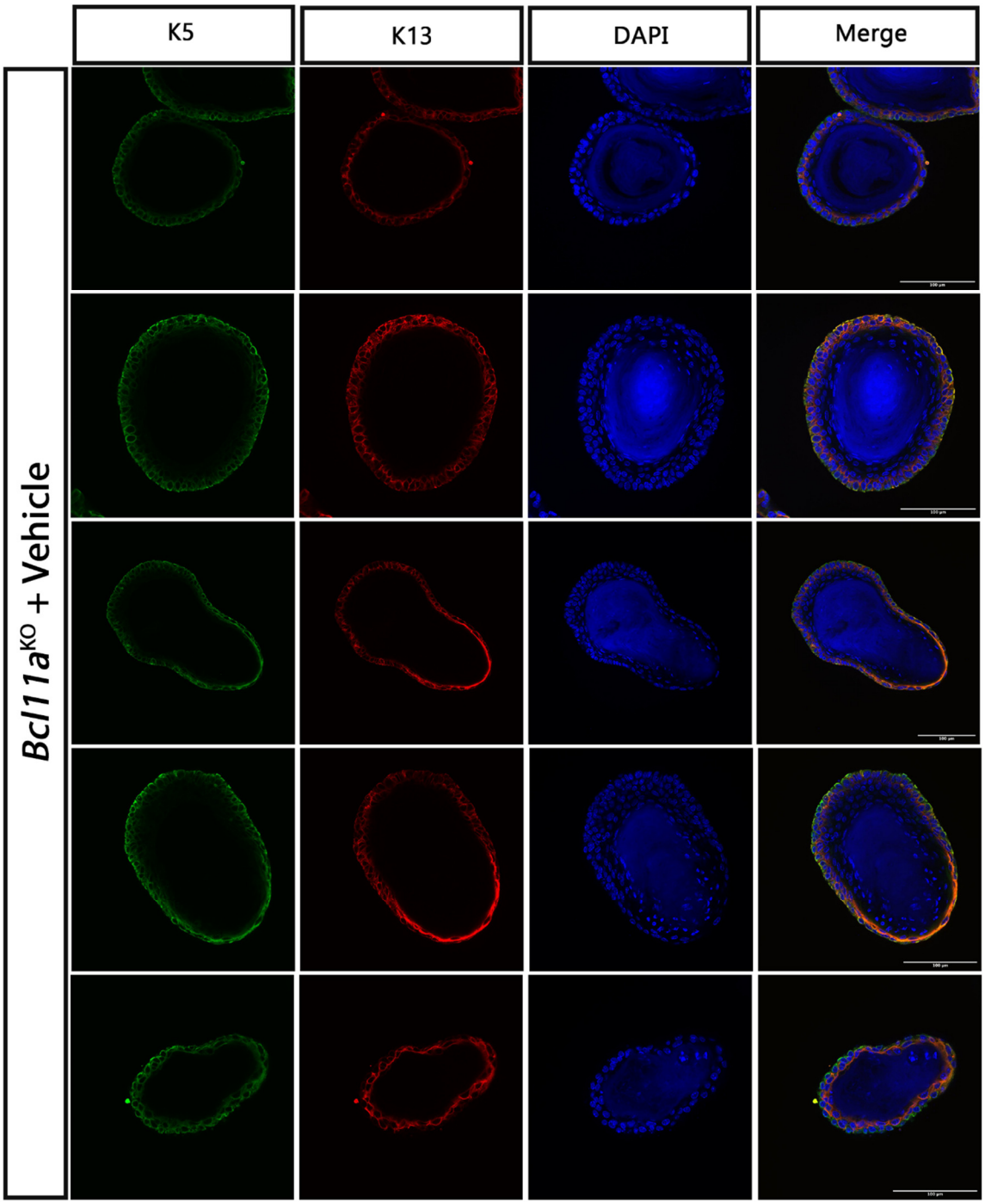


Figure 3.27: Fluorescence microscope images of whole mount *Bcl1a*^{KO} oesophageal organoids treated with vehicle.

Organoids were cultivated from *Bcl11a*^{KO} and treated with ethanol vehicle control, (see figure 3.25 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of basal cells (Keratin 5 in green), a marker of suprabasal cells (Keratin 13 in red) and DAPI in blue for nuclear and keratin visualisation. Scale bar represents 100 μm.

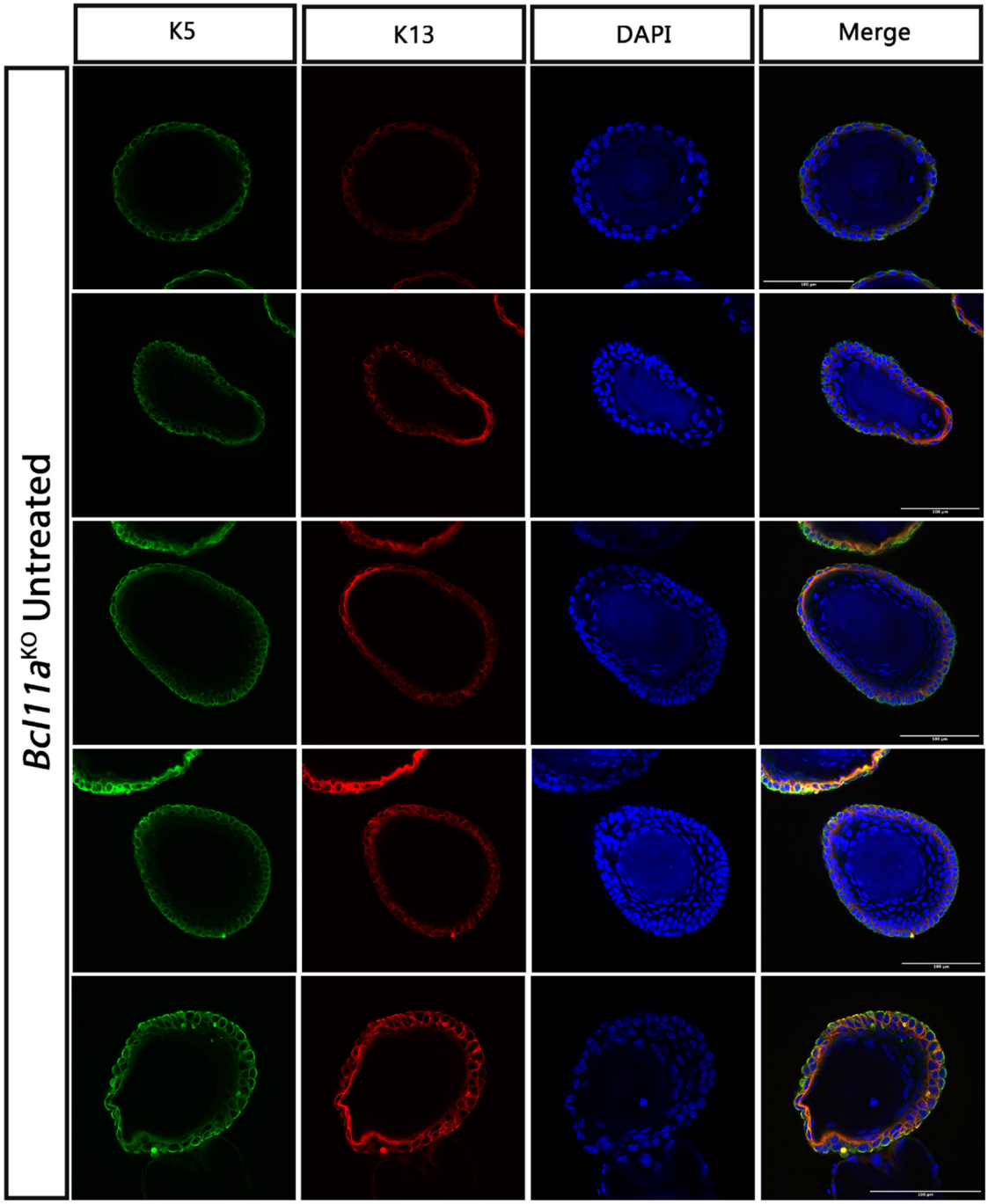


Figure 3.28: Fluorescence microscope images of whole mount untreated *Bcl11a*^{KO} oesophageal organoids.

Figure 3.28 (cont): Fluorescence microscope images of whole mount untreated *Bcl11a*^{KO} oesophageal organoids.

Organoids were cultivated from *Bcl11a*^{KO} mice and left untreated and cultivated for a period of 14 days. Organoids were stained with a marker of basal cells (Keratin 5 in green), a marker of suprabasal cells (Keratin13 in red) and DAPI in blue for nuclear and keratin visualisation. Scale bar represents 100 μm.

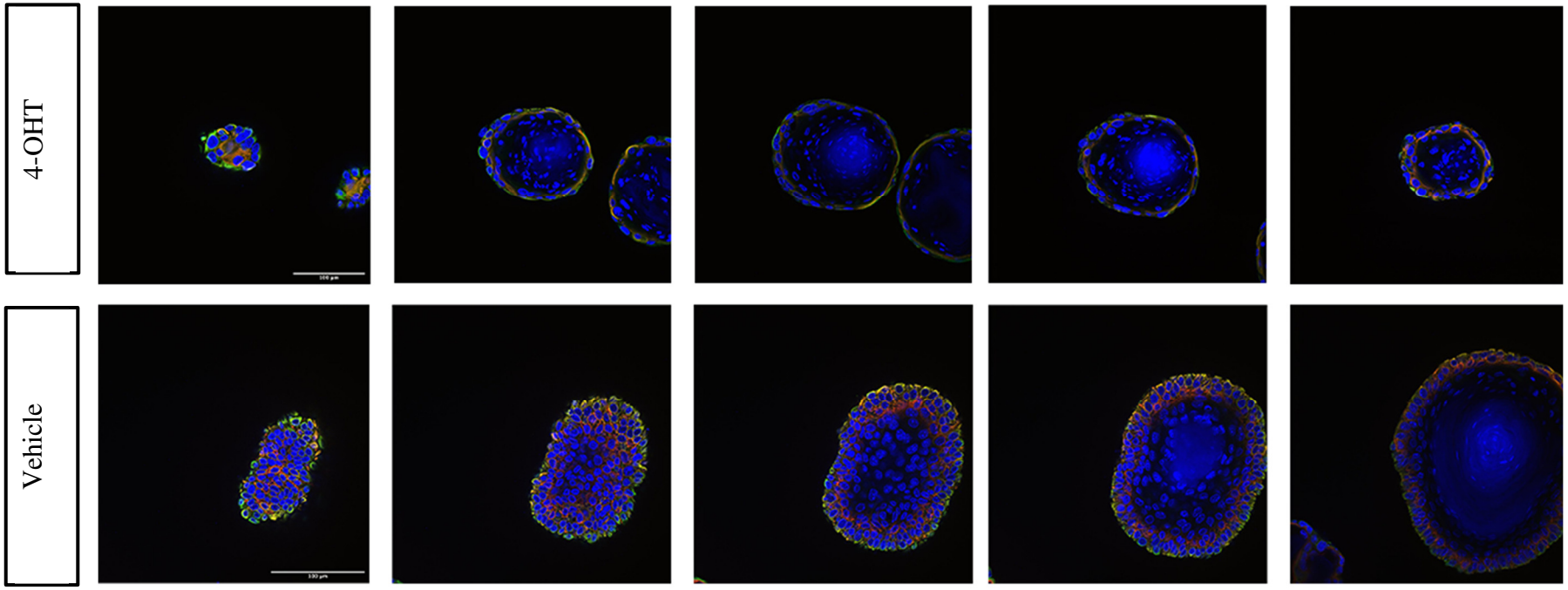


Figure 3.29: Multiple planes of *Bcl11a*^{KO} organoids treated with 4-OHT or vehicle control.

Organoids were cultivated and stained as in figure 3.25 from both 4-OHT and vehicle groups. Organoids were visualised using a confocal microscope and multiple planes of view were acquired (Z stack) to demonstrate the structural differences between organoids of the different groups. Keratin 5 is shown in green, Keratin 13 in red and DAPI nuclear stain in blue. Scale bar is 100µm

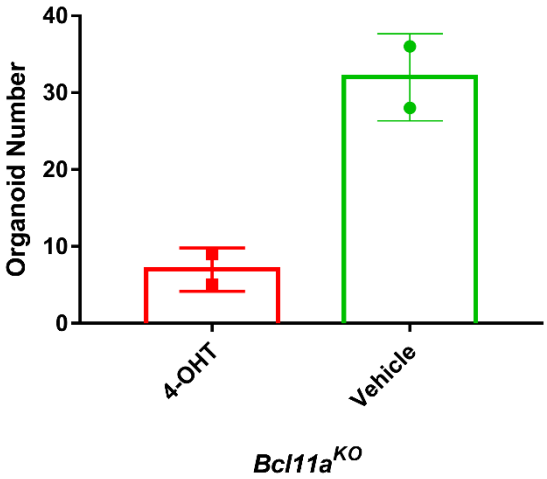


Figure 3.30: Colony forming efficiency of passaged *Bcl11a*^{KO} oesophageal organoids.

Bcl11a^{KO} organoids treated with 4-OHT or vehicle were passaged and 500 cells plated in 2 wells per conditions. Shown are the number of organoids counted after 14 days in culture.

3.3 Discussion

BCL11A has been identified as a putative oncogene of OSCC, however very little is known about the function of this gene in the oesophageal epithelium. This project sought to understand more about the function of this transcription factor in the oesophageal epithelium to gain insight into molecular mechanisms of disease. A 3D organoid system adapted from DeWard et al. was used to model *BCL11A* gene perturbations. Overall, the organoid system used here generated organoids that were poorly differentiated, with many cells identified through quantification as double positive for both basal and suprabasal markers. There were also a large proportion of cells that were negative for any markers. Although this did not preclude the derivation of meaningful results, as both knock out and over expression organoids had distinct phenotypes, it is an important consideration for future work using this model.

Organoids cultivated from *BCL11A*^{OE} cells demonstrated over 100-fold increase in *BCL11A* expression compared to control when induced *in vitro*. This induction was far higher than in the mammary gland despite using the same concentrations of 4-OHT. Indeed, *eGFP* expression was detected in these organoids when it was not in the mammary gland. This suggests that tissue specific induction protocols will be necessary for future work using these models *in vitro.* Morphologically, the over expressing organoids displayed a lack of uniformity in their outer layer, without the smooth phenotype seen in organoids of control groups. Organoids tended to be smaller with some showing little evidence of a keratinised core. *Bcl11a* knock out organoids displayed a striking blebbing phenotype, in some cases losing structural integrity entirely. Whole mount analysis revealed poorly differentiated organoids and cells with enlarged nuclei.

Importantly, these organoid studies have shown that *BCL11A* overexpression increased the colony forming ability which was in stark contrast to the phenotype of *Bcl11a* deletion. Taken together this suggests that *BCL11A* could play an important role in oesophageal progenitor cell function. These results very closely reflect those seen in the lung [3] suggesting *Bcl11a* acts in a similar capacity in multiple tissues. However, it is important to consider that these experiments represent only one independent biological replicate and future experiments should seek to confirm and strengthen these conclusions by performing multiple independent biological experiments and carrying out statistical analyses.

Building on the data presented in this chapter, future experiments performed on *Bcl11a* knock out organoids should seek to confirm the apoptosis phenotype by staining for apoptotic markers such as cleaved caspase-3 and cleaved PARP1. In addition, the staining for markers of proliferation such as ki67 in over expressing organoids as well as knock outs, could further provide insight into the question of whether *Bcl11a* behaves in a pro-proliferative manner in the oesophageal epithelium. The induction of gene perturbations upon cell seeding should also be attempted and will again, provide an insight into gene mechanisms of action. If the transcription factor is indeed behaving in a pro-proliferative manner, the over expressing cells may form bigger colonies than cells of control groups. This experiment will also determine whether *Bcl11a* is essential for organoid formation, as has been shown in the lung [3]. Another interesting avenue for experimentation could be challenge of the organoids with trans-Retinoic acid, which has been shown as a potent inducer of differentiation in the oesophageal epithelium [157]. This experiment, as a follow up to the passaging experiments shown here, could help to identify whether *Bcl11a* plays a role in maintenance of cell stemness, uncovering whether over expression of this transcription factor has a protective effect against differentiation induced by trans-Retinoic acid.

Finally, future work may want to identify potential causes of the poor differentiation identified in some organoids. Further optimisation may be necessary to ensure the culture system faithfully mimics the epithelium *in vivo* with dedicated Keratin 5 and Keratin 13 positive cells. The length of time the organoids were kept in culture for could be a contributing factor to this phenomenon, future work should try to correct for this by keeping cells in culture for shorter periods. In addition, the wholemount staining protocol should be optimised for use with these organoids. Many cells were identified within the organoids that appeared not to be positive for any markers. These cells should be Keratin 13 positive as demonstrated by DeWard et al. [127] This could be an antibody penetration issue, so further permeabilisation may be necessary with detergents such as Triton X-100.

Chapter 4: Exploring *Bcl11a* Function Using *In Vivo* Lineage Tracing

4.1 Introduction

4.1.1Exploring *Bcl11a* Gene Expression *In Vivo*

Having identified a role for *Bcl11a* in oesophageal progenitor function and differentiation using organoids, the next phase of the project sought to validate this observation and explore gene expression further *in vivo.* One way in which gene expression may be studied is through genetic lineage tracing. Lineage tracing represents a valuable *in situ* tool to gain insight into gene expression and indeed the contribution, if any, that labelled cells and their progeny provide to overall tissue maintenance. This technique was used to uncover said contributions of *Bcl11a* expressing cells to epithelial tissue maintenance in the oesophagus to further validate observations from the 3D organoid experiments. In addition to the oesophagus, genetic lineage tracing was also employed by Dr Sara Pensa to study the contribution of *Bcl11a* expressing cells to the mammary gland.

4.2 Results

4.2.1 *Bcl11a* Lineage Tracing Marks Luminal Progenitor Cells in the Mammary Gland

Lineage tracing in the mammary gland was previously employed by Dr Sara Pensa to identify the extent of *Bcl11a* expression as well as the contribution labelled cells had to the developing gland from puberty and including gestation (figure 4.1). Particularly in the case of the mammary gland, where the cyclical nature of the tissue sees much remodelling, varying lengths of chase were used to piece together information on the contribution of labelled cells to these stages. Using the *Bcl11a*^{CreERT2} mouse model, cellular progeny were tracked over time, ultimately providing insight into the long-term persistence of labelled populations of cells, as well as any particular points of expansion or regression. Cells expressing *Bcl11a* at the time of tamoxifen induction were labelled using the tdTomato reporter.

Labelled cells were identified in the luminal compartment (figure 4.1a) and seen to expand with pregnancy, seeming to contribute to alveolar lineages (figure 4.1d). Labelled cells persisted over time and could be detected after a 24-week chase. Upon pregnancy, labelled cells expanded, seeming to contribute to alveolar lineages. This was confirmed using FACS analysis (figure4.1b) The persistence of labelled cells over time and the expansion during pregnancy suggested a role for *Bcl11a* in luminal cells of the mammary gland, and more

specifically within luminal progenitors capable of giving rise to both hormone sensing and alveolar progeny. To ensure that tdTomato labelling was indeed specific to Bcl11a gene expression, mammary glands were collected from mice 24 hours post tamoxifen induction. RT-qPCR analysis confirmed the specificity of the labelling to *Bcl11a* gene expression, suggesting that the transcription factor is indeed expressed in a luminal progenitor population that expands over time.

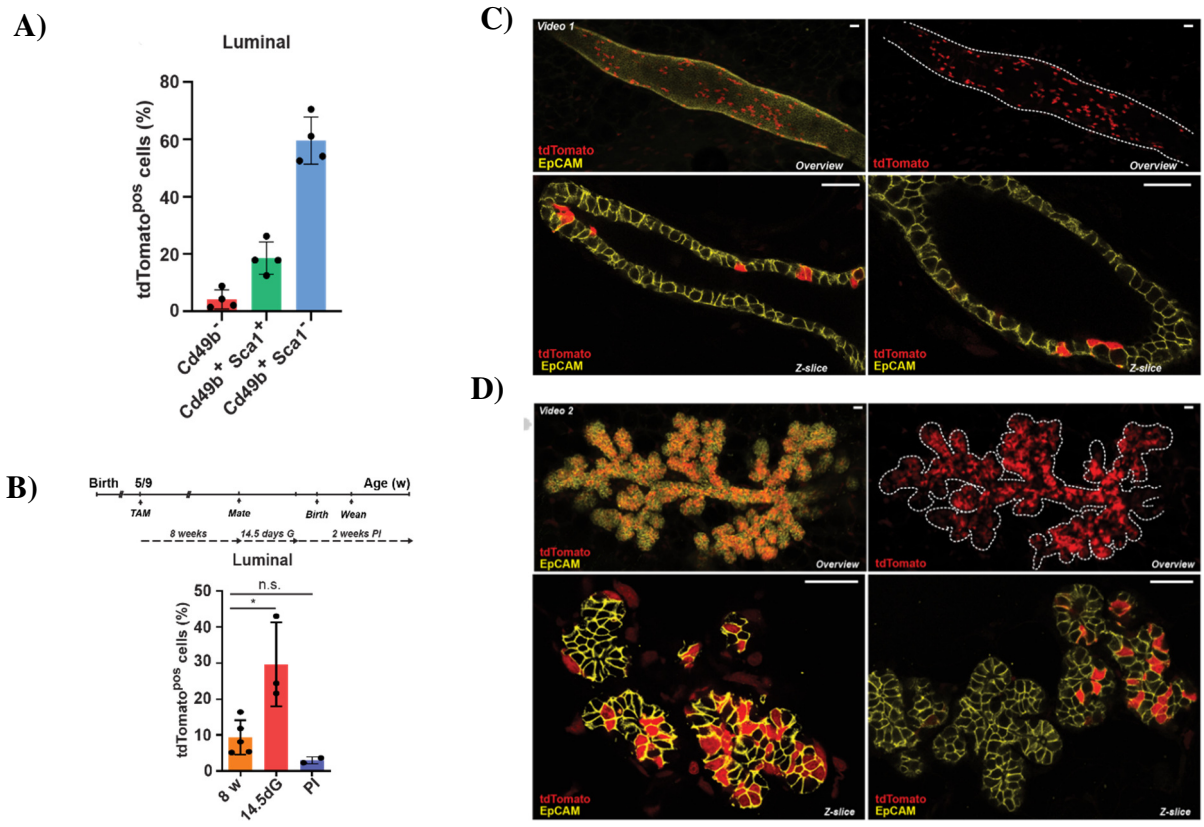


Figure 4.1: *Bcl11a* lineage tracing in the mammary gland.

All data obtained from Dr Sara Pensa. **A)** Graph depicting the percentage of tdTomato positive cells in differentiated ($CD49b^-$) or progenitor ($CD49b^+$ $Sca1^-$ $Sca1^+$) luminal cells. Populations were determined my FACS, 1 day post labelling. **B)** tdTomato labelling of *Bcl11a* expressing cells were analysed at 8 weeks, 14.5 days gestation and 2 weeks post involution. The graph depicts the percentage of tdTomato positive cells within the luminal compartment determined using FACS. **C)** Confocal imaging of tdTomato positive cells in a CUBIC cleared whole gland at 8 weeks stained with EpCam shown as an overview in the top panels and representative Z slices in the bottom panels. Scale bar is 100 µm**. D)** Confocal imaging of tdTomato positive cells in a whole CUBIC cleared gland at 14.5 days gestation and stained with EpCam. Top panels are an overview, the bottom panels are representative Z slices. Scale bar is 100 µm.

4.2.2 Lineage tracing in the Oesophageal Epithelium

The previous application of the *Bcl11a*^{CreERT2} model to the mammary gland served to preliminarily validate the use of this mouse model for the oesophageal epithelium. Lineage tracing in the oesophagus was undertaken to gain insight into whether *Bcl11a* expressing cells contributed to normal tissue maintenance. The specific aims of this phase of the project were to assess the levels of *Bcl11a* gene expression in the oesophageal epithelium, particularly the cell type to which expression was localised, as well as the persistence or expansion of these cells over time. To achieve these aims, samples were collected at varying chase lengths post tamoxifen exposure. These included 24 hours post exposure, 8 weeks post exposure and 12 weeks post exposure (figure 4.2) with two experimental mice (*Bcl11a*^{CreERT2}) per time point (labelled mouse 1 and mouse 2 for comparability) in addition to a control *Bcl11a* wild type mouse without the *CreERT2* transgene (*Bcl11a*^{WT}, for details see experimental procedures). All mice were induced with tamoxifen at either 12 or 13 weeks of age. Cell types expressing *Bcl11a* at the time of labelling were identified using a 24-hour chase and longer-term persistence or expansion of labelled cells were identified using chase lengths of 8-12 weeks.

Cellular division in the oesophageal epithelium is confined to the basal layer and these cells are thought to divide 2.9 times per week [96]. The basal layer is also thought to house stem or progenitor populations, the status of which remains unclear, however these cells are broadly described as either proliferative or differentiating [96]. Therefore, any identification of tdTomato labelled cells within the basal layer was of note, as these could imply *Bcl11a* expression in progenitor cells of the oesophagus. To better visualise the orientation of each figure, fields were captured from multiple planes shown in figure 4.3. The oesophageal epithelium was either orientated with the basal face upwards so that a plane of basal cells was visible (view 1); from the top of the basal layer downwards, so that the basal plane was visible and the tissue was stepped through to the suprabasal layer (view 2); or finally, from the suprabasal layer stepping through to the basal layer (view 3).

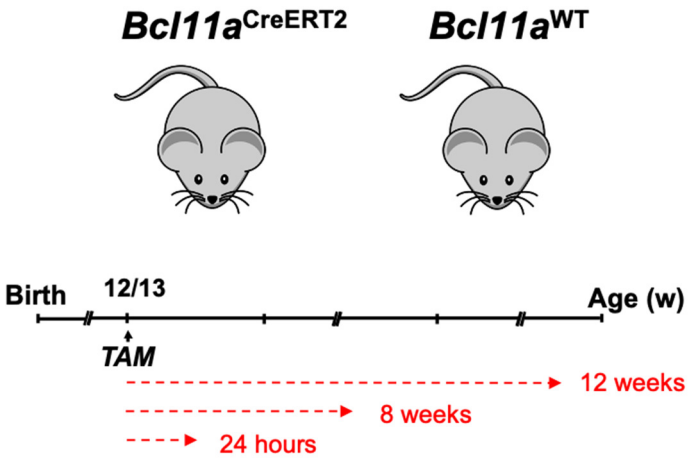


Figure 4.2: Schematic highlighting chase lengths for lineage tracing in the oesophageal epithelium.

In vitro tamoxifen was administered to two experimental *Bcl11a*^{CreERT2} mice and one *Bcl11a*^{WT} control mouse per time point. The oesophagi were collected 24 hours, 8 weeks and 12 weeks post tamoxifen exposure. Mice were induced with tamoxifen at either 12 or 13 weeks of age.

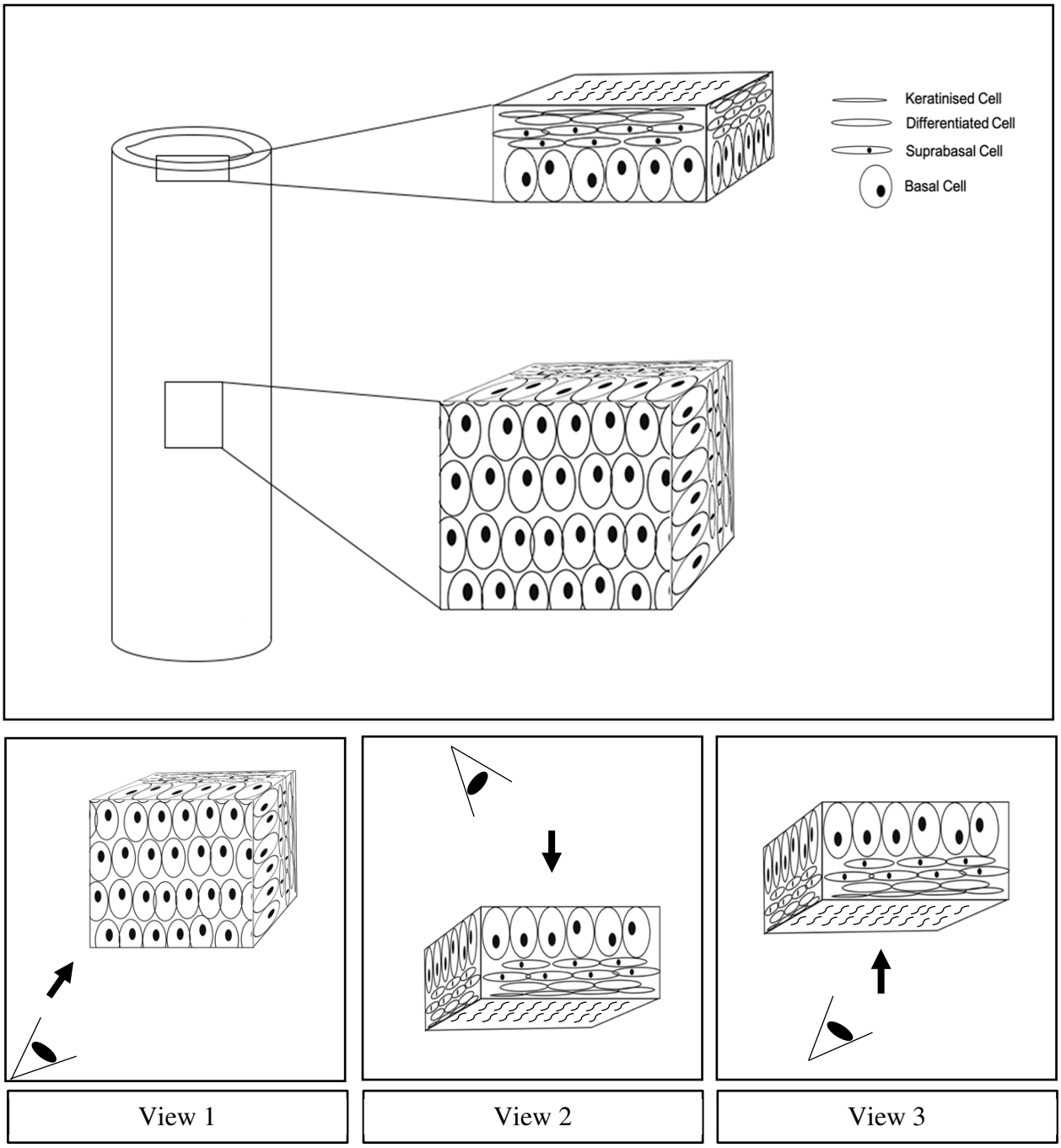


Figure 4.3: Schematic of tissue orientation from different viewpoints.

Figure 4.3 (cont.): Schematic of tissue orientation from different viewpoints.

The murine oesophagus is a stratified squamous epithelium consisting of a layer of basal cells to which proliferation is confined. As cells differentiate they move upwards becoming suprabasal before becoming keratinised and being shed off into the lumen. **A)** Orientation of squamous epithelium of intact oesophagus**. B)** Different viewpoints of tissue orientation. View 1 looks at the face of the basal layer. View 2 looks from the top of the basal layer downwards, so that the basal plane was visible and the tissue was stepped through to the suprabasal layer. View 3 looks at the tissue from the suprabasal layer stepping through to the basal layer.

4.2.3 *Bcl11a* Lineage Tracing Labels Cells in the Oesophageal Epithelium

The oesophageal epithelium was examined using whole mount analysis from two *Bcl11a*^{CreERT2} mice at 24 hours post tamoxifen exposure as well as one *Bcl11a*^{WT} control. This short chase length served to capture cells expressing *Bcl11a* at the time of labelling, identifying the cell types where expression was localised, while reducing the possibility of labelling colonies i.e. cells that have divided and differentiated multiple times since labelling. Epithelia were either cut open longitudinally or left intact for staining. Samples were stained with a basal cell surface marker (Keratin 5), to aid visualisation and to confirm the presence of tdTomato labelled cells within the basal layer. Epithelia were also stained with an RFP antibody to aid tdTomato visualisation, this antibody was previously validated for use in this capacity in the mammary gland by Dr Sara Pensa.

Analysis of the control *Bcl11a*^{WT} mouse revealed no evidence of recombination at 24 hours (figure 4.4). This suggested no expression of tdTomato indicating any labelling in other conditions was unlikely to be an artefact of the reporter construct alone. Lineage tracing in *Bcl11a*^{CreERT2} mice revealed a large amount of recombination (figure 4.5) at 24 hours. Images were analysed from multiple fields ensuring labelling was representative of the whole epithelium (figures 4.6). The extent of recombination demonstrated the potential expression of *Bcl11a* across both layers of the oesophageal epithelium. Recombination did not appear to be restricted to a particular cell type, with tdTomato expression visible across both Keratin 5 positive cells and flatter differentiated suprabasal cells in both experimental animals (figure 4.6 and 4.7). In order to better asses the level of recombination and relative location of tdTomato labelled cells, microscope imaged were quantitated (figure 4.8). Cells were scored as either Keratin 5 positive, tdTomato positive or double positive. This analysis demonstrated a proportion of tdTomato labelled cells in the basal layer (double positive), suggesting that *Bcl11a* expression was marking a subset of basal cells. There was also a small number of cells retaining the tdTomato label only, suggesting that these cells may represent suprabasal cell labelling indicating potential *Bcl11a* expression in more differentiated cells of the oesophageal epithelium.

In order to better visualise the tissue surface and relative locations (basal or suprabasal) of labelled cells, multiple planes were compiled as a 3-dimensional (3D) projection of the epithelial layer (figure 4.9 and 4.10). The 3D projection displayed clearly the contour of the

tissue surface, aiding visualisation. This analysis revealed a stronger localisation of labelled cells in the suprabasal layer, indicating a high degree of recombination in these cells. Although red cells were visible in the basal layer these were located more sparsely, perhaps indicating lower levels of recombination in these cells. This could suggest gene expression is localised to only select cells of the basal layer whilst the greater proportion of *Bcl11a* expressing cells sit in the suprabasal layer.

Closer analysis of the basal layers revealed this in greater detail (figure 4.11). In each field of vision along the basal layer a greater proportion of labelled cells resided in the suprabasal layers. Cells labelled with both Keratin 5 staining and tdTomato, were visible at varying fields along the epithelium. This indicated potential *Bcl11a* expression in some cells of the basal compartment, however these were encountered less frequently than labelled cells from the suprabasal layers indicated above. Interestingly, only select cells of each compartment were labelled indicating that *Bcl11a* may have only been expressed in a subset of cells from each layer.

In order to understand more about the role labelled cells had, the epithelia of two mice 24 hours post tamoxifen induction were taken for FACS. From the epithelium alone, tdTomato positive and negative cells were sorted and plated as organoids. Cells were plated in equal numbers and colonies were analysed after 10 days in culture (figure 4.12). Red cells generated more colonies than non-red cells (figure 4.12). This suggested that *Bcl11a* expressing cells possessed a greater colony forming capacity, potentially signifying a progenitor role in the oesophageal epithelia. To see whether this was the case and to understand whether *Bcl11a* labelled cells persisted in the oesophageal epithelium, samples were analysed from *Bcl11a*^{CreERT2} mice 8 weeks post tamoxifen induction.

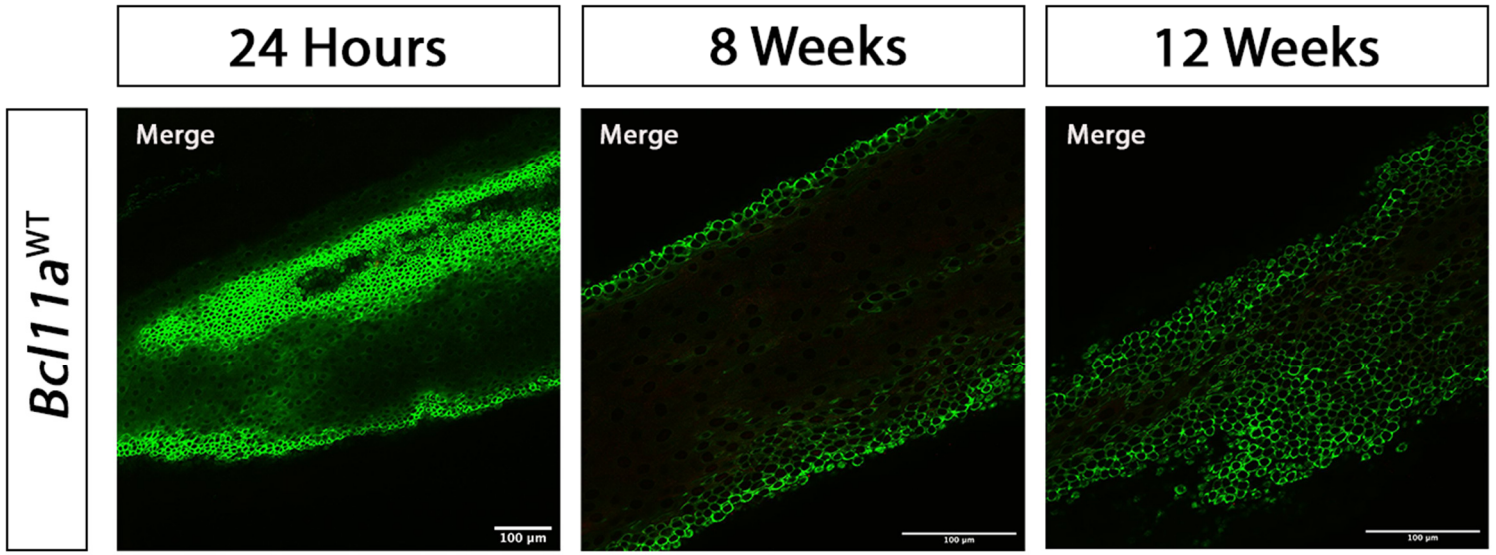


Figure 4.4: Fluorescence microscope images of cleared *Bcl11a^{WT}* oesophagi.

In vitro tamoxifen was administered to three Bcl11aWT mice and oesophagi collected 24 hours, 8 weeks and 12 weeks post exposure. The epithelia were isolated and optically cleared before staining with a marker of basal cells (Keratin 5 in green) and tdTomato (in red). Scale bar represents 100 µm.

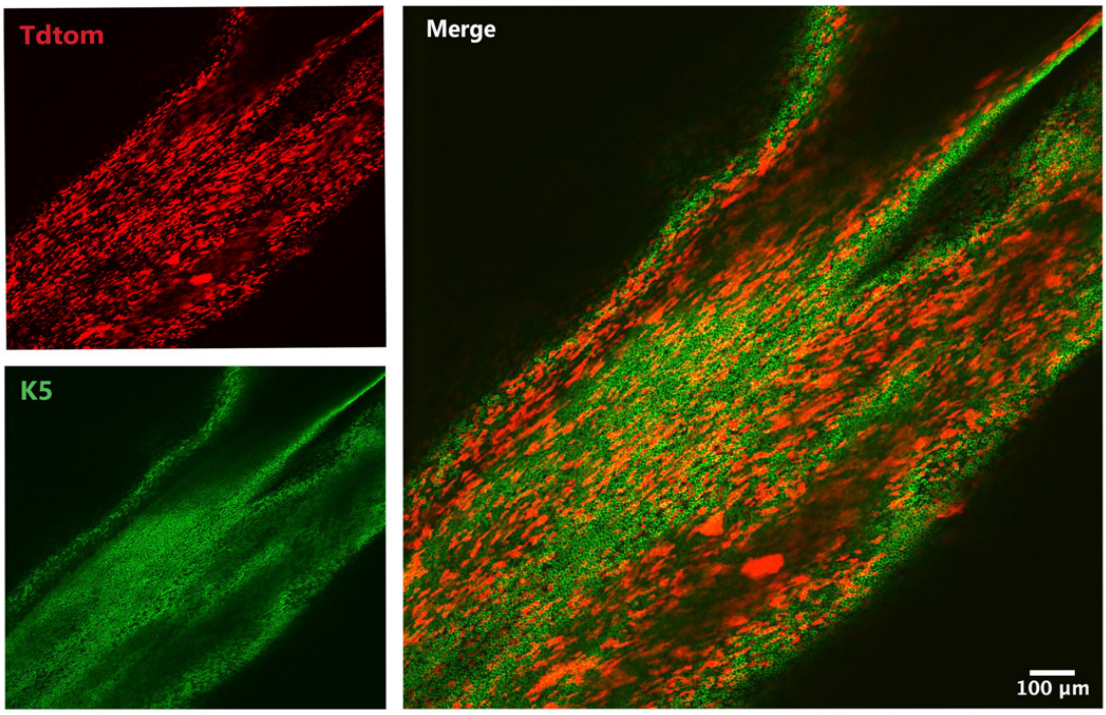


Figure 4.5: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure, global view.

24 hours post tamoxifen induction he epithelium of a *Bcl11a*^{CreERT2} mouse (mouse 1) was isolated and optically cleared before staining with a marker of basal cells (Keratin 5 in green) and tdTomato (in red). Scale bar represents 100 µm.

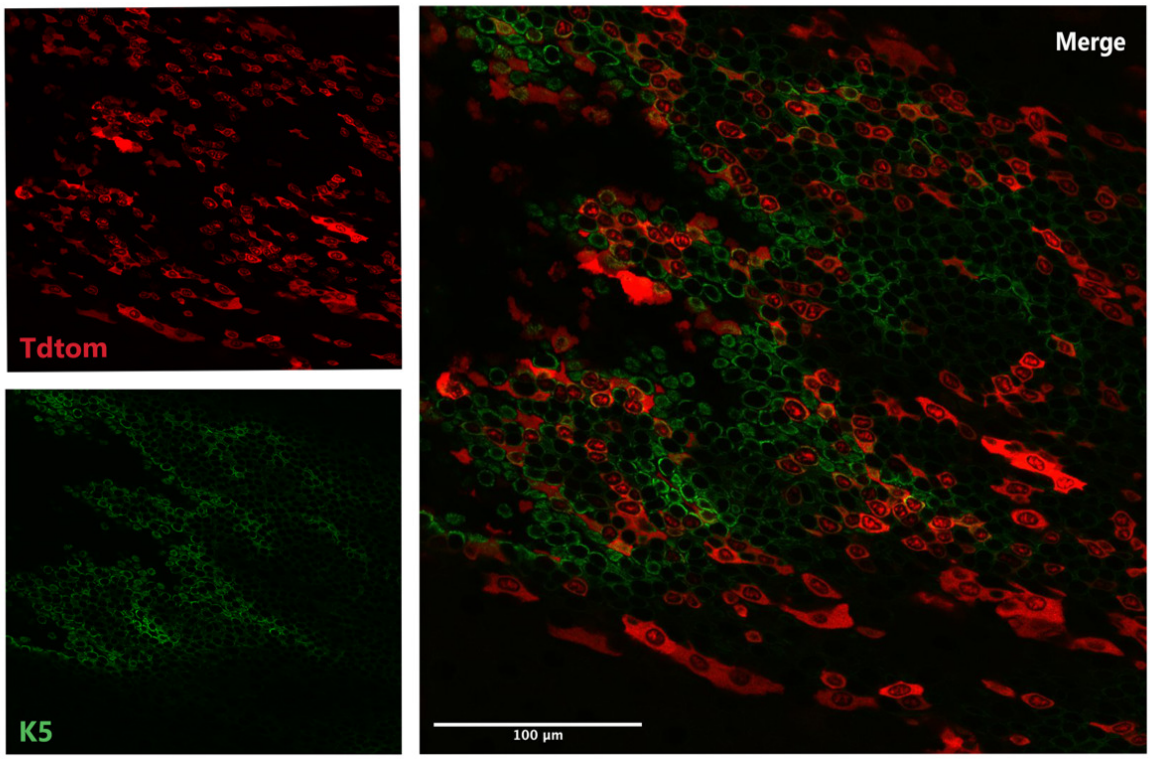


Figure 4.6: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure view 1.

The epithelia of one *Bcl11a*^{CreERT2} mouse (mouse 1) 24 hours post tamoxifen exposure was isolated and optically cleared before staining with a marker for Keratin 5 in green and tdTomato in red. Scale bar represents 100 µm.

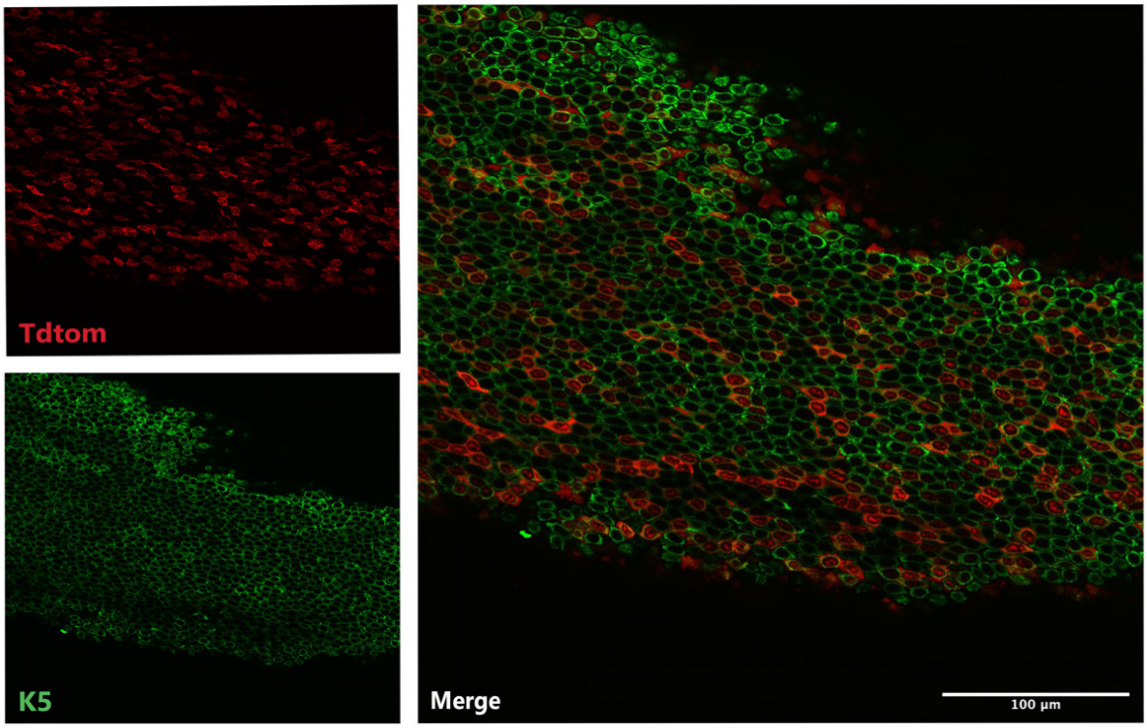


Figure 4.7: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure view 1.

The epithelia of one *Bcl11a*^{CreERT2} mouse (mouse 2) 24 hours post tamoxifen exposure was isolated and optically cleared before staining with a marker for Keratin 5 in green and tdTomato in red. Scale bar represents 100 µm.

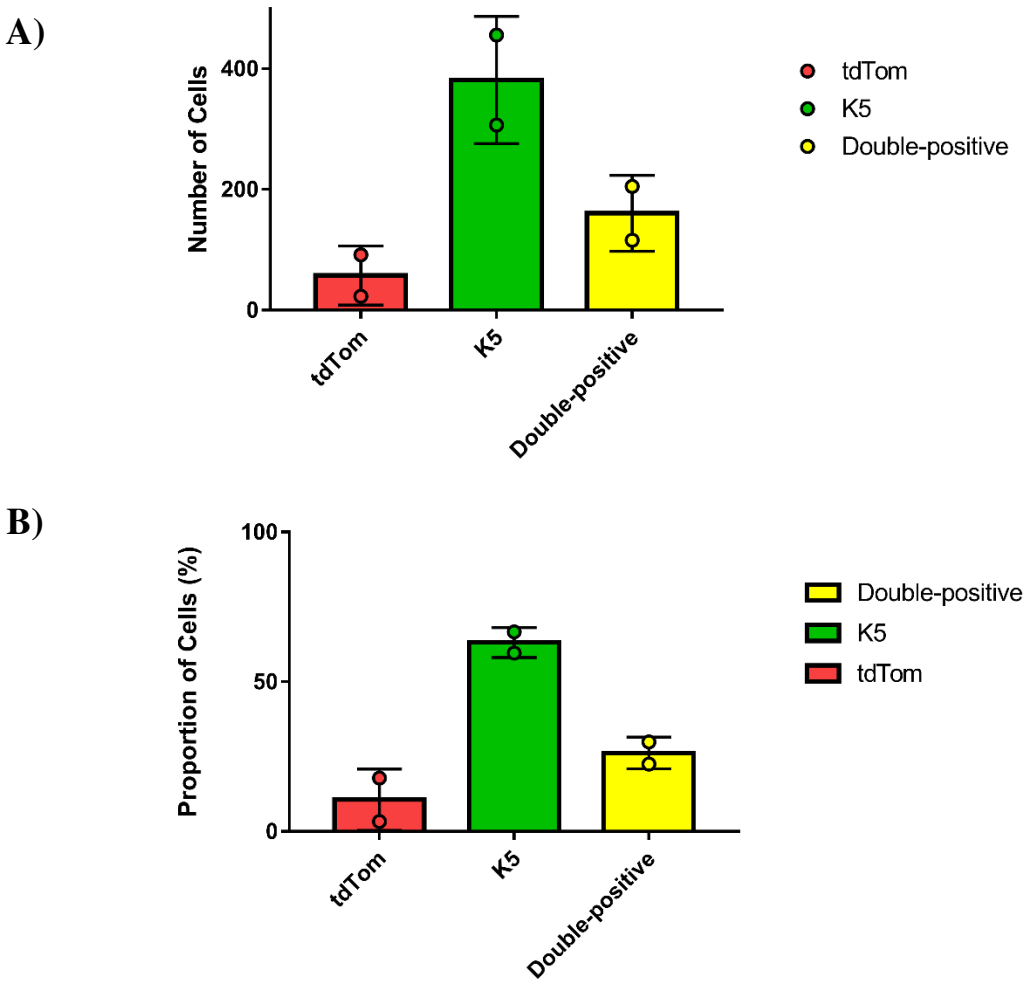


Figure 4.8: Quantification of microscope images from 2 *Bcl11a*^{CreERT2} mice 24 hours post tamoxifen exposure.

Microscope images from both *Bcl11a*^{CreERT2} mice (figures 4.6 and 4.7) were quantitated by scoring cells as tdTomato positive, Keratin 5 positive or double positive. **A)** Depicts the mean number of cells displaying each marker ± S.D. **B)** Numerical data in (A) presented as the mean proportion of cells positive for each marker ± S.D (n=2 biological replicates).

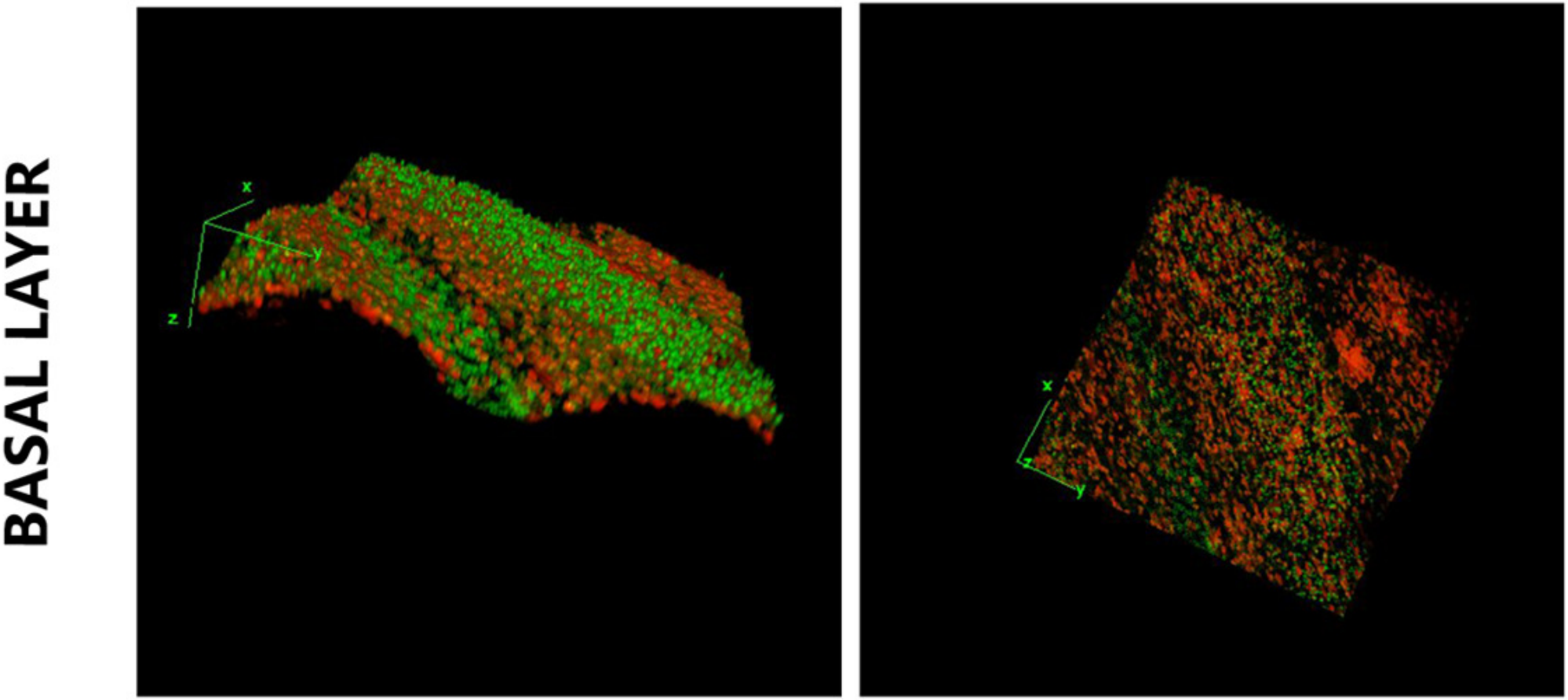


Figure 4.9: 3D maximum projection of Z stack view 2 of *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure.

The epithelium was prepared as described in figure 4.7 from mouse 2. Using a fluorescence microscope multiple planes of view were acquired in a single field and compiled as a 3D maximum projection using ImageJ software. Keratin 5 cells are marked in green and tdTomato in red.

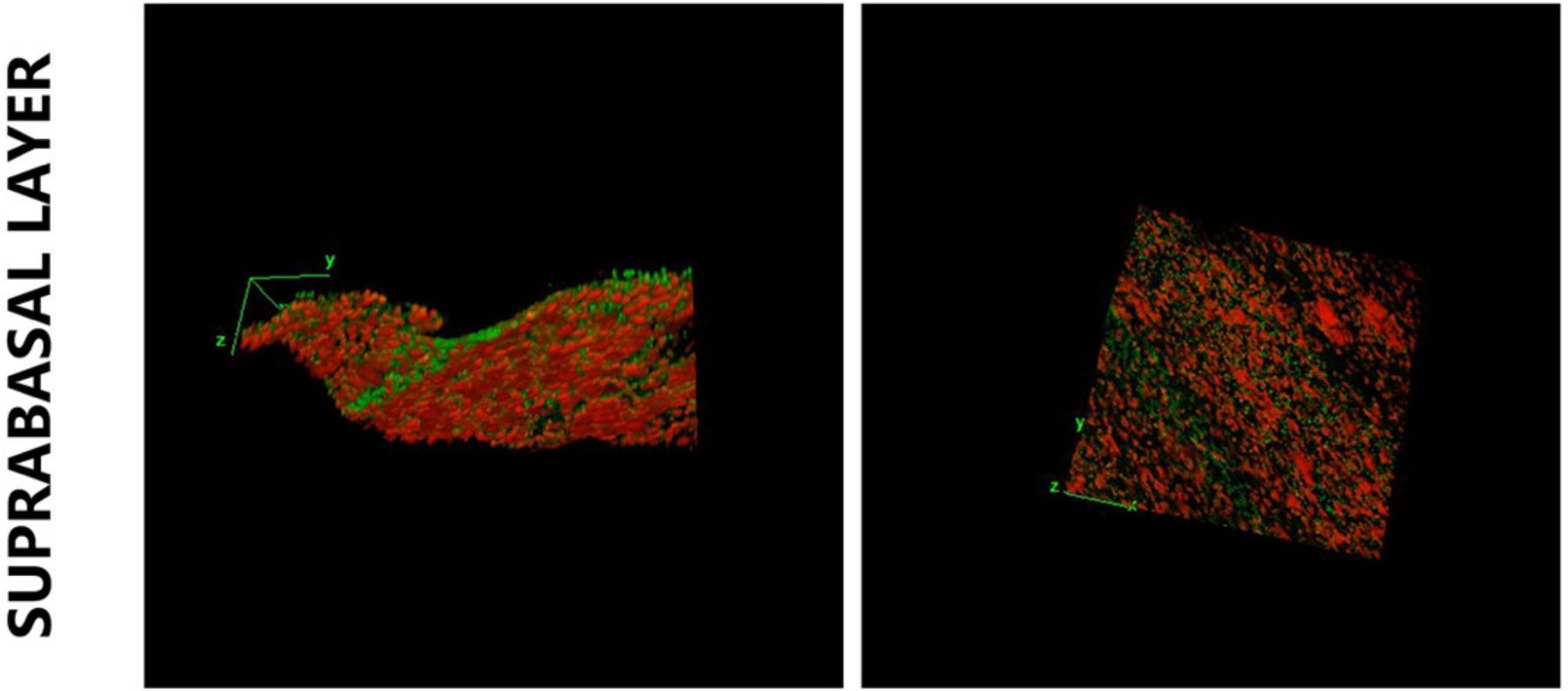


Figure 4.10: 3D maximum projection of Z stack view 3 of *Bcl11a*^{CreERT2} mouse 24 hours post tamoxifen exposure.

The epithelium was prepared as described in figure 4.7 from mouse 2. Using a fluorescence microscope multiple planes of view were acquired in a single field and compiled as a 3D maximum projection using ImageJ software. Keratin 5 cells are marked in green and tdTomato in red.

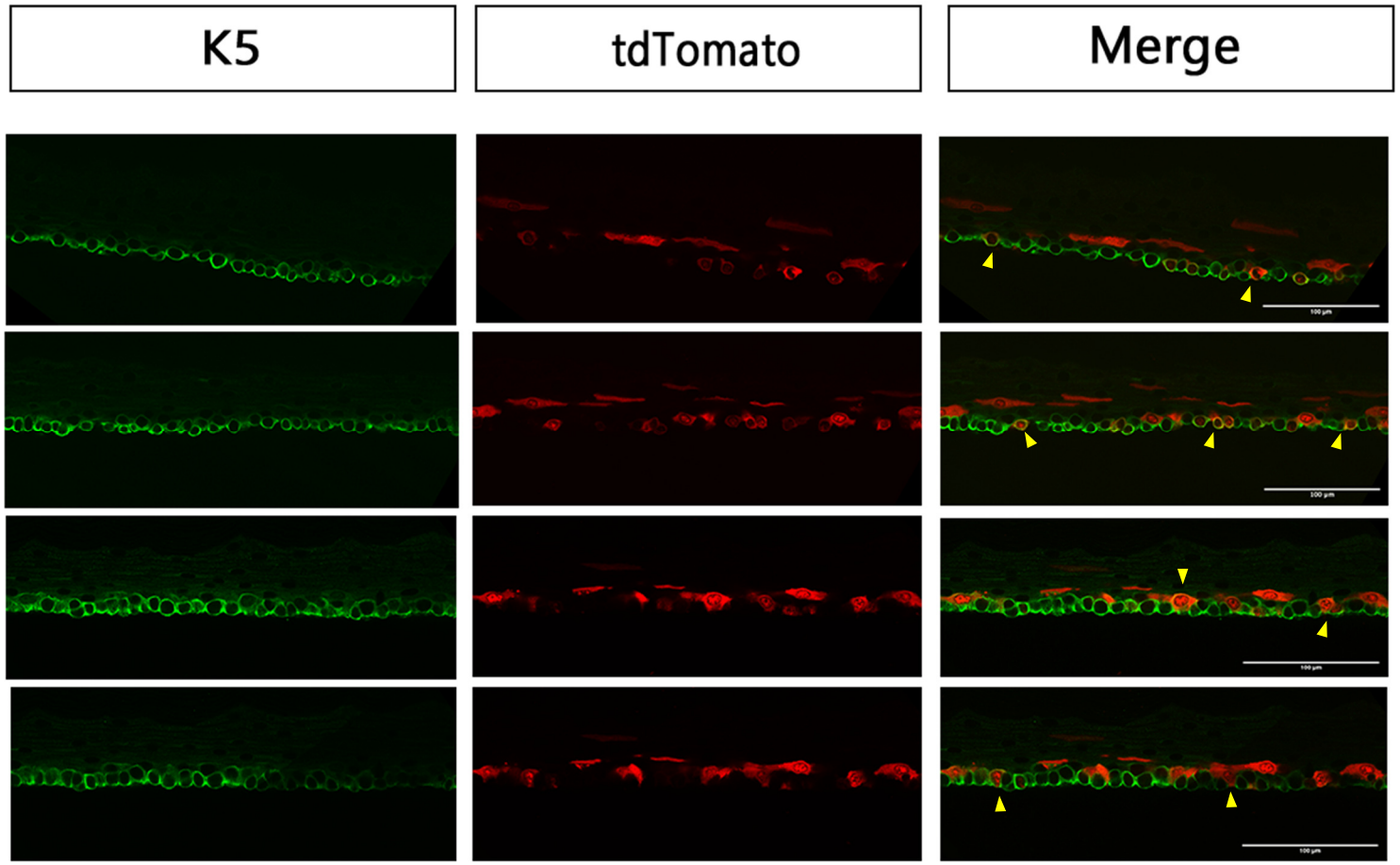


Figure 4.11: Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 24 hours post tamoxifen exposure.

Figure 4.11 (cont.): Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 24 hours post tamoxifen exposure.

Oesophageal epithelium was prepared as described in figure 4.7 from mouse 2. Arrows indicate cells that have stained positive for both Keratin 5 and Keratin 13. Scale bar represents 100 μm in the first 3 rows and 50 μm in the bottom row of images.

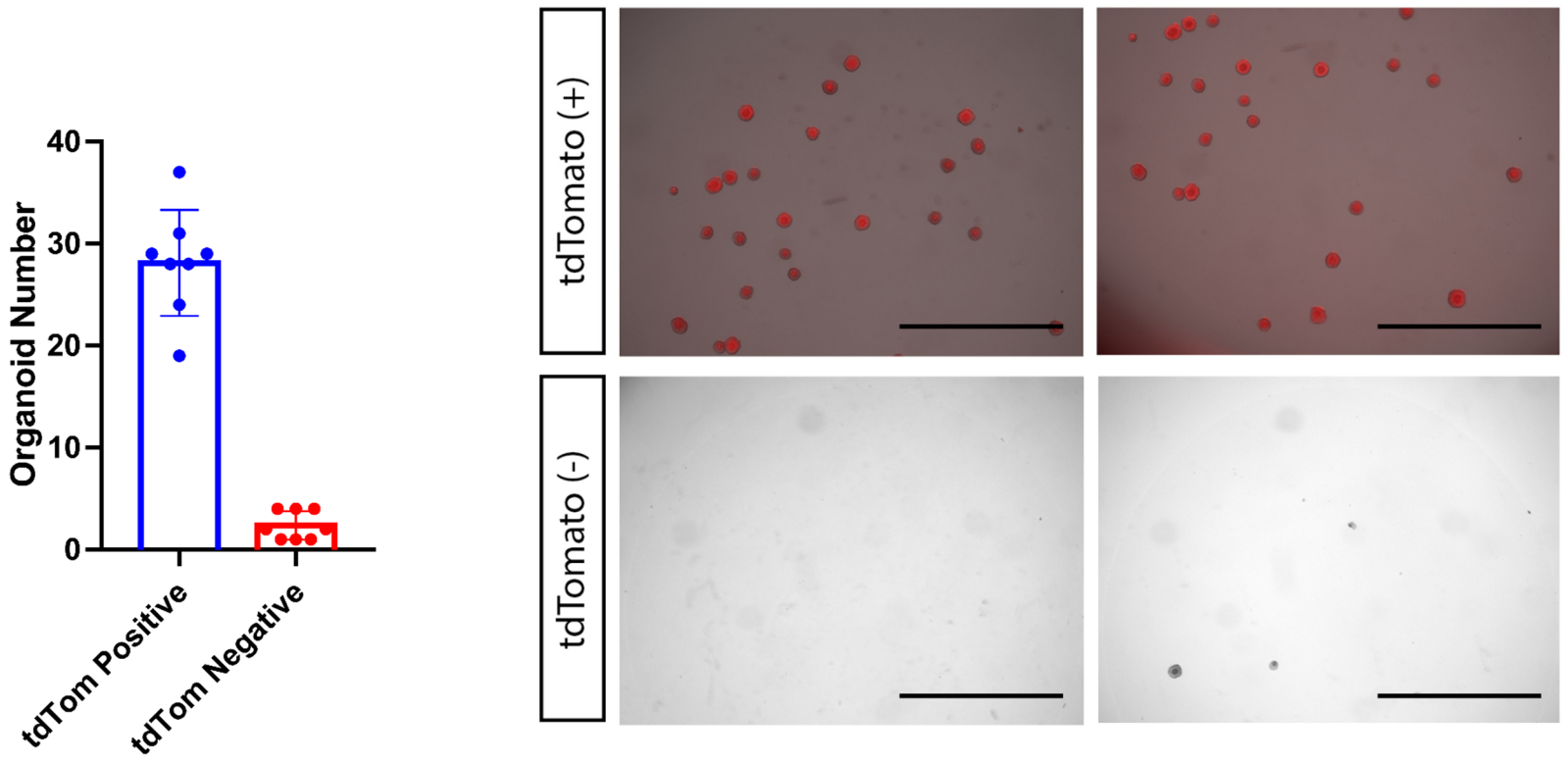


Figure 4.12: tdTomato positive and negative cells were plated as organoids.

Figure 4.12 (cont.): tdTomato positive and negative cells were plated as organoids.

Bcl11a^{CreERT2} mice were administered tamoxifen and oesophagi were collected 24 hours post exposure. The epithelial layer was dissociated and subjected to FACS based on tdTomato expression. Cells positive and negative for the marker were plated separately and growth was observed after 14 days. A) Shown are the number of organoids resulting from 1000 cells plated per condition. Data are represented as mean ± S.D. of n=8 technical replicates. Scale bar is 2000 μm.

4.2.4 *Bcl11a* Labelled Cells Persist and Expand in the Oesophageal Epithelium

At 8 weeks post tamoxifen exposure, whole mount analysis from cleared epithelia of two *Bcl11a*^{CreERT2} mice revealed a persistence of labelled cells. As cells of the basal layer divide 2.9 times per week, 8 weeks left time for roughly 24 cell divisions in each basal cell [96]. Owing to the fact that divisions give rise to proliferating or differentiating daughters with equal probability, these divisions are unlikely to represent large areas of clonal expansion in the basal layer. In addition, over 8 weeks cells marked with tdTomato expression at day one would have become keratinised and shed into the lumen [95,96] therefore any labelling of cells in the suprabasal layers seen at 8 weeks would be likely to indicate daughters of labelled cells rather than the persistence of the same labelled cells from day one. Analysis of the cleared epithelia of a *Bcl11a*^{WT} mouse 8 weeks post tamoxifen exposure revealed no evidence of recombination (figure 4.4).

The levels of recombination seen at 8 weeks post tamoxifen exposure in *Bcl11a*^{CreERT2} mice were again high, with many labelled cells visible across the epithelia of both mice (see figure 4.13 and 4.14). Similar to 24 hours post tamoxifen exposure, labelled cells were visible in both layers of the epithelium, indicating the persistence of labelled cells after requisite cell divisions. In visualising cells of the basal layer, there did appear to be a persistence of tdTomato positive cells that were also positive for Keratin 5, (see figure 4.14). Indeed, apparent colonies of labelled cells do appear at 8 weeks post tamoxifen exposure, suggesting the replicative potential of *Bcl11a* labelled cells (figure 4.14 an figure 15) further supporting a potential role for this transcription factor in progenitor cells. Similar quantitative analysis of microscope images revealed a greater proportion of tdTomato positive cells only as compared to the 24hour time point, suggesting an expansion of labelled cells into the suprabasal layers (figure 4.15). Similar proportions of double positive cells remained at 8 weeks, suggesting the persistence of labelled cells within the basal layer (figure 4.15).

In order to better visualise the relative locations of labelled cells, multiple planes of vision were combined to form a 3D projection of the epithelium. This analysis did not reveal a clear localisation of labelled cells. However, it did indicate the presence of tdTomato labelling to a particular area of the basal layer (figure 4.16). Analysis of the suprabasal layer immediately bellow these cells did suggest an expansion of labelled cells (figure 4.17). To assess whether this was the case, more planes of vision were collected to better visualise areas of clonal basal

cell division and differentiation into suprabasal cells (figure 4.18). Indeed, this analysis of multiple planes did reveal areas of apparent clonal expansion and differentiation of tdTomato labelled cells. These small clonal regions of labelled cells in the basal layer were seen to expand into larger areas of differentiated cells in the suprabasal layer, suggesting the labelling of proliferative and differentiating cells (figure 4.18). Closer inspection of multiple fields of the basal layer revealed in greater detail the extent of labelling in the basal and surrounding suprabasal layers (figure 4.19). In addition to some evidence of expansion, far more labelled cells appeared in the basal layer (figure 4.19) as compared to that observed in the epithelia 24 hours post tamoxifen exposure, suggesting the expansion of these cells. Interestingly some areas displayed clearly labelled differentiated suprabasal cells below basal cells which were not labelled (figure 4.19). This does appear to fit with the paradigm of cellular maintenance in the epithelium as differentiating cells move upwards and outwards to the suprabasal layer [96]. Together this data suggests that lineage tracing of *Bcl11a* appears to label a proliferating progenitor population in the basal layer with evidence of expansion in both basal and suprabasal layers.

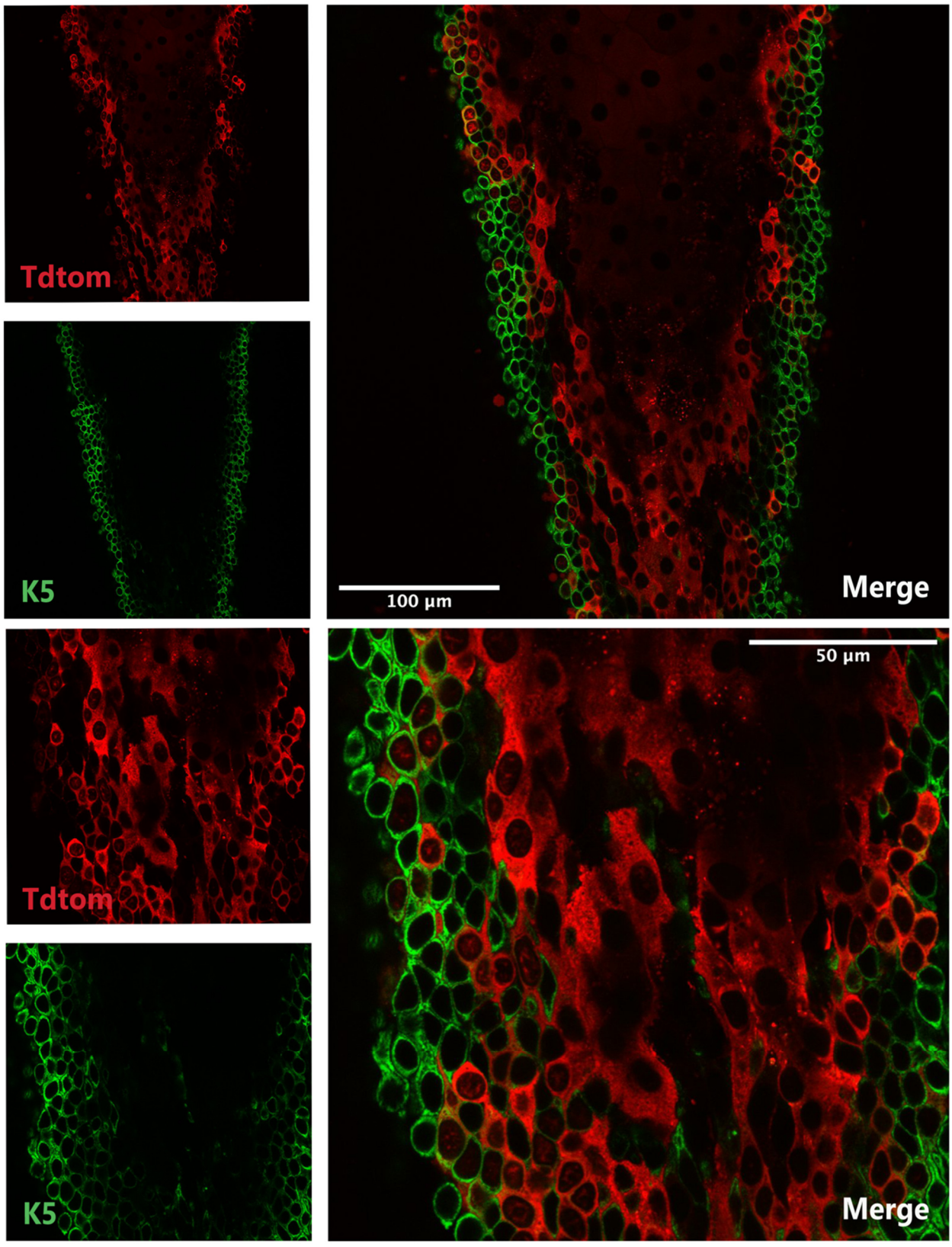


Figure 4.13: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure.

Figure 4.13 (cont.): Fluorescence microscope images from a *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure.

The oesophageal epithelium of one *Bcl11a*^{CreERT2} mouse (Mouse 1) 8 weeks post tamoxifen exposure was isolated and optically cleared before staining with a marker for Keratin 5 in green and tdTomato in red. Scale bar represents 100 μm.

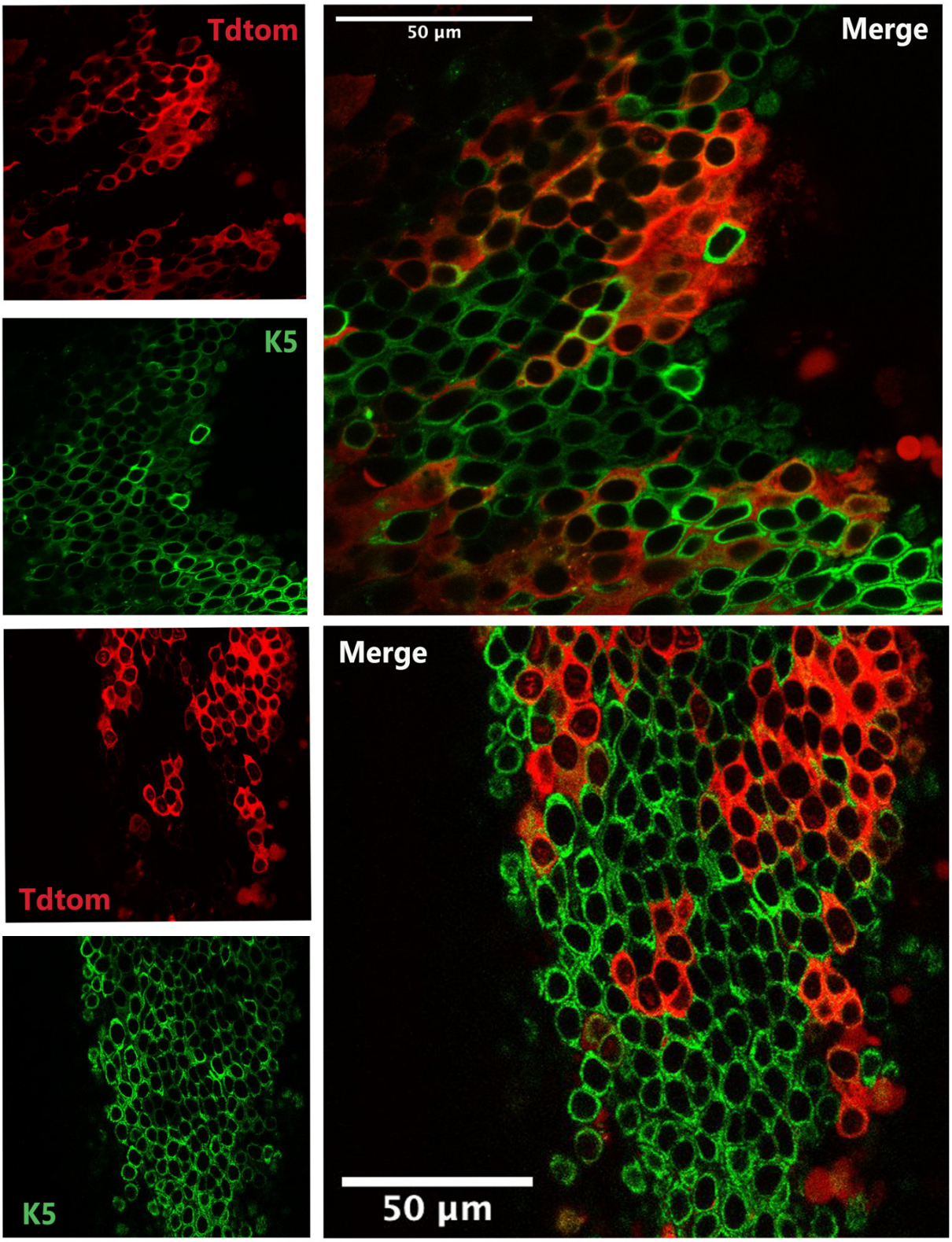


Figure 4.14: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure, view 1.

Figure 4.14 (cont.): Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure, view 1.

The oesophageal epithelium of one *Bcl11a*^{CreERT2} mouse (mouse 2) 8 weeks post tamoxifen exposure was isolated and optically cleared before staining with a marker for Keratin 5 in green and tdTomato in red. Scale bar represents 100 μm.

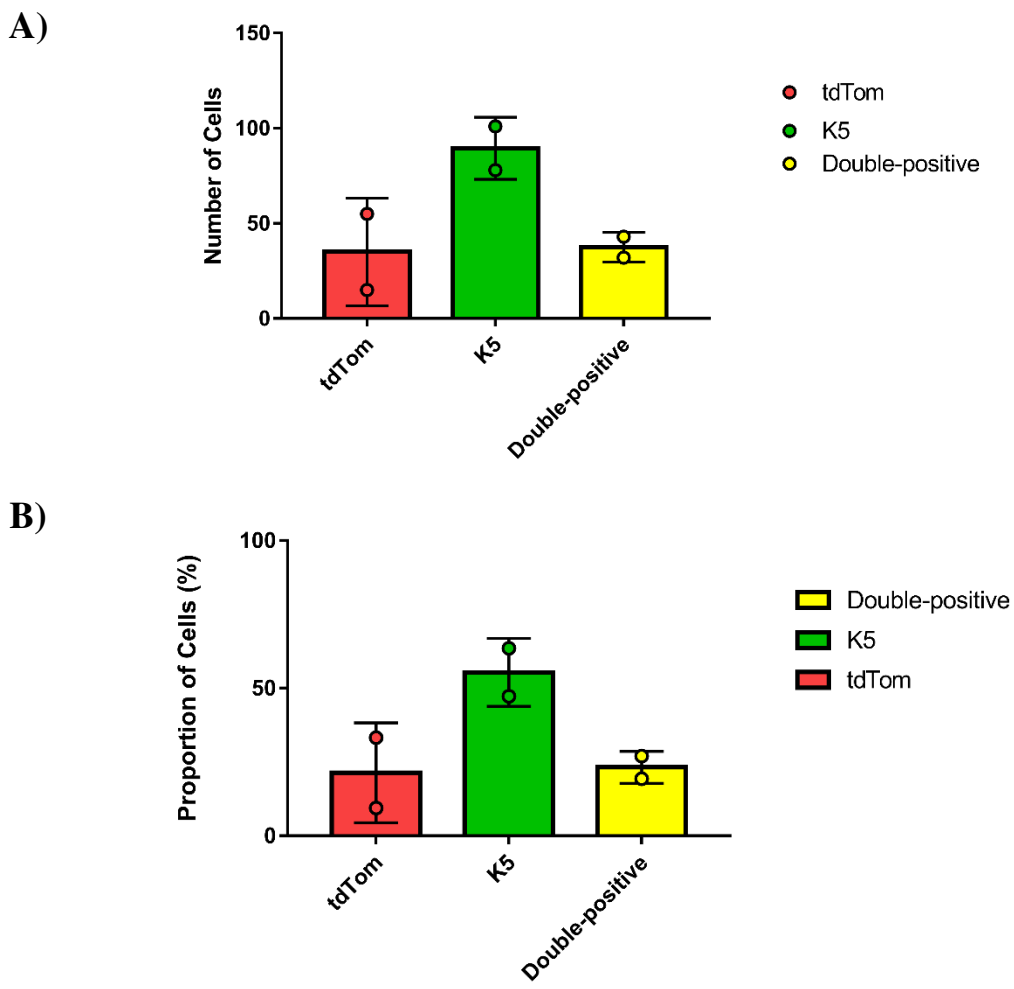


Figure 4.15: Quantification of microscope images from 2 *Bcl11a*^{CreERT2} mice 8 weeks post tamoxifen exposure.

Microscope images from both *Bcl11a*^{CreERT2} mice were quantitated by scoring cells as tdTomato positive, Keratin 5 positive or double positive. **A)** Depicts the number mean of cells displaying each marker ± S.D. **B)** Numerical data from (A) presented as the mean proportion of cells positive for each marker ± S.D (n=2 biological replicates).

.

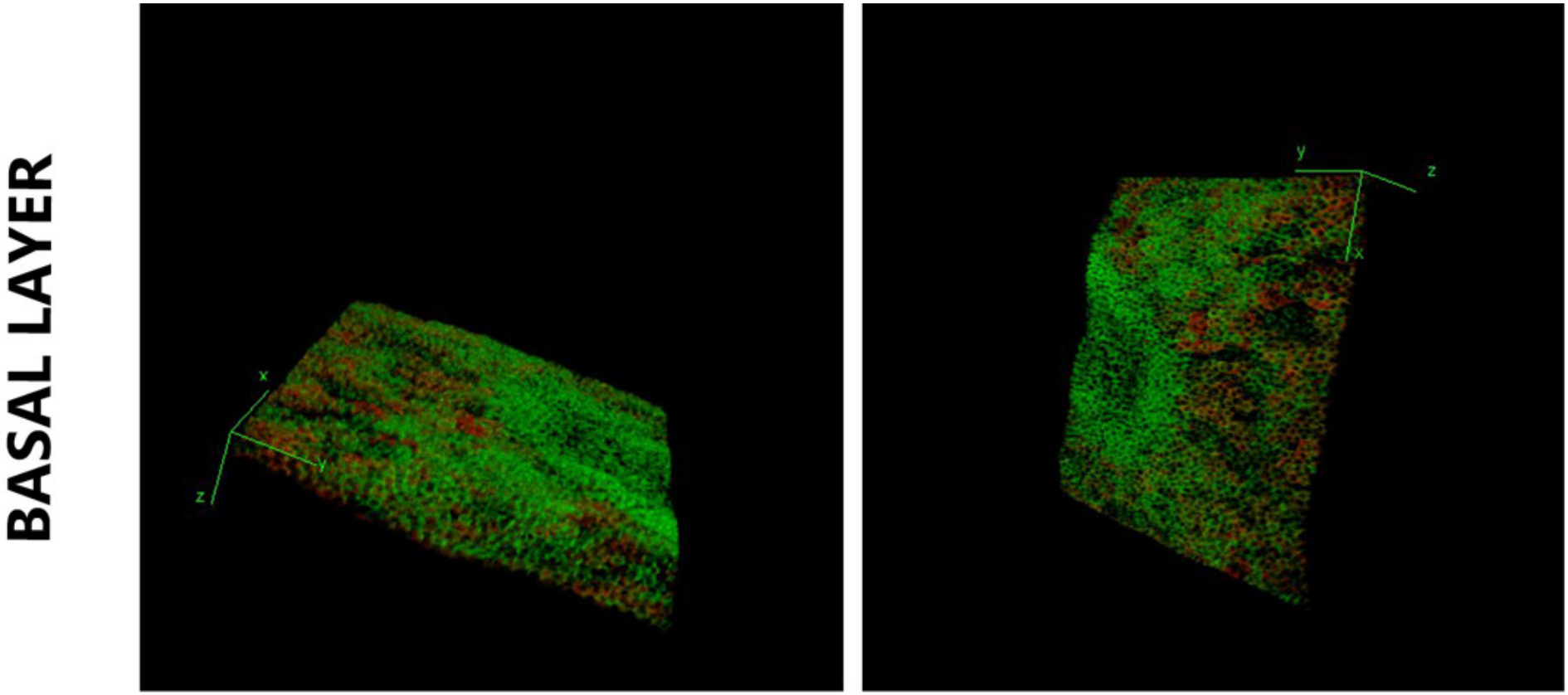


Figure 4.16: 3D maximum projection of Z stack view 2 of *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure.

The epithelium was prepared as described in figure 4.14. Using a fluorescence microscope multiple planes of view were acquired in a single field and compiled as a 3D maximum projection using ImageJ software. Keratin 5 cells are marked in green and tdTomato in red.

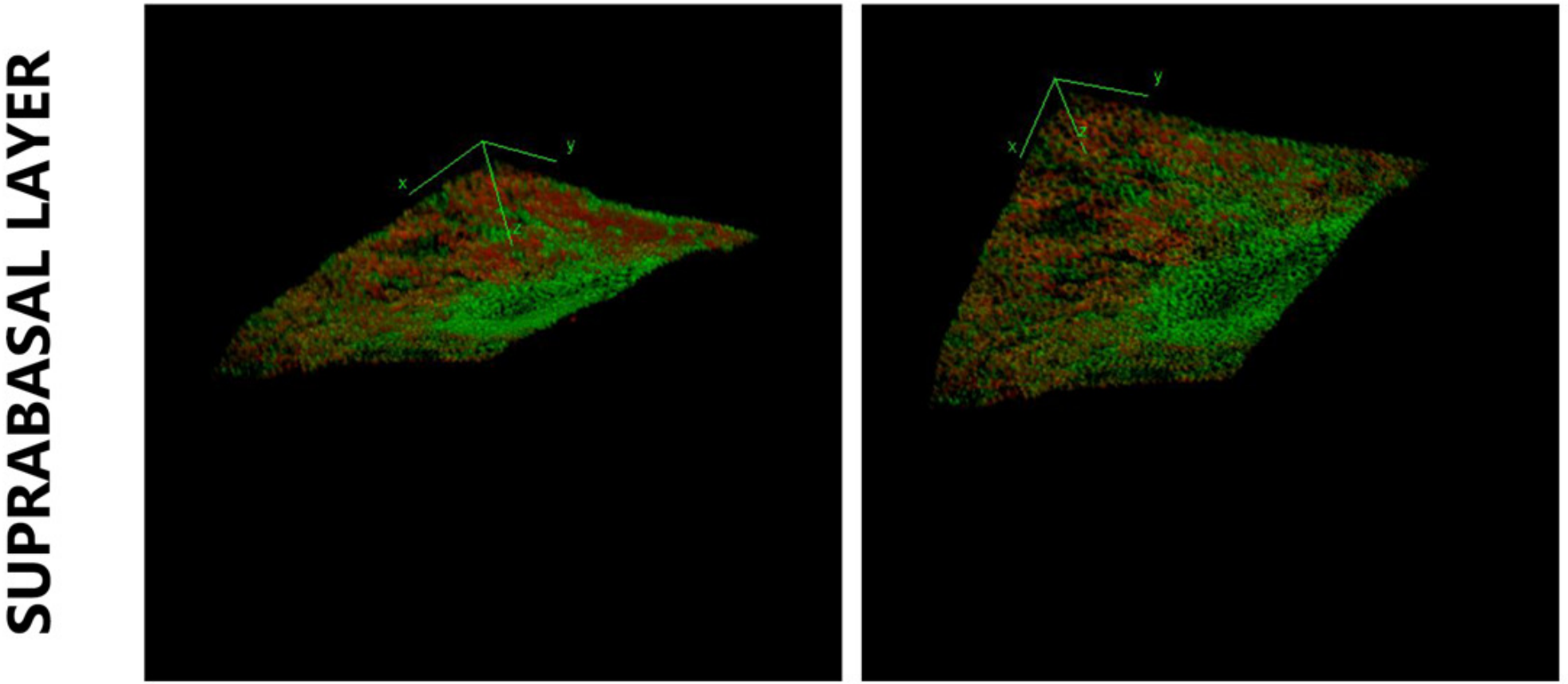


Figure 4.17: A 3D maximum projection of Z stack view 3 of *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure.

The epithelium was prepared as described in figure 4.14. Using a fluorescence microscope multiple planes of view were acquired in a single field and compiled as a 3D maximum projection using ImageJ software. Keratin 5 cells are marked in green and tdTomato in red.

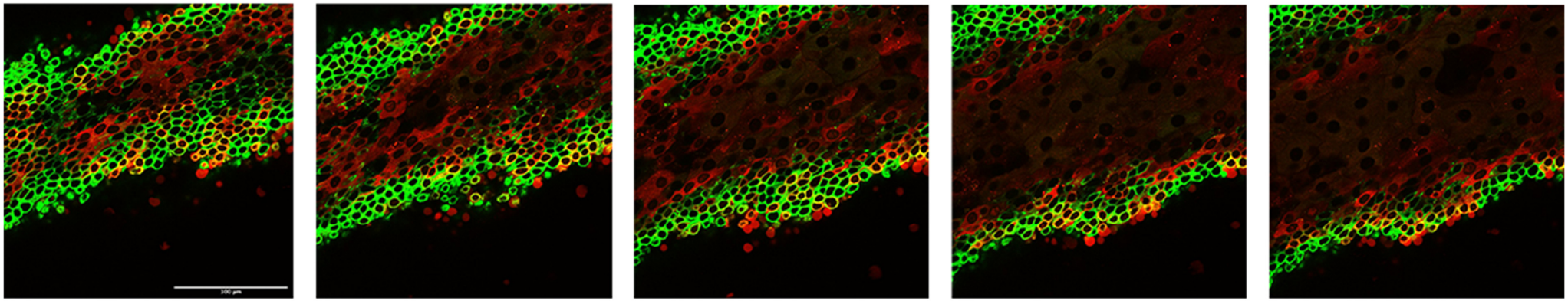


Figure 4.18: Fluorescence microscope step through of *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure.

Multiple Z slices through the oesophageal epithelium isolated from a *Bcl11a*^{CreERT2} mouse 8 weeks post tamoxifen exposure. Keratin 5 staining is shown in green and tdTomato is shown in red. Scale bar represents 100 µm.

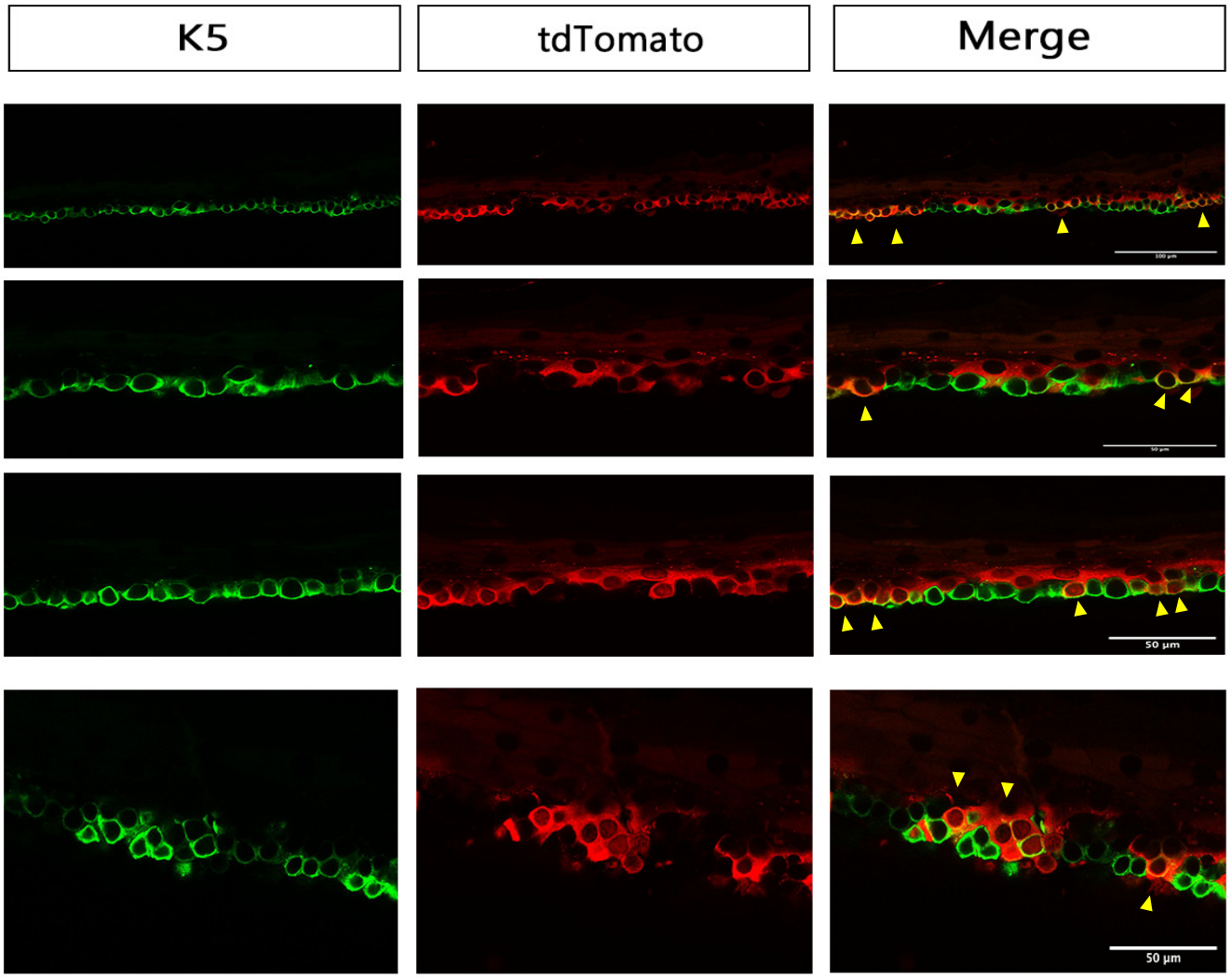


Figure 4.19: Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 8 weeks post tamoxifen exposure.

Figure 4.19 (Cont.): Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 8 weeks post tamoxifen exposure.

Oesophageal epithelium was prepared as described in figure 4.14. Arrows indicate cells or groups of cells that have stained positive for both Keratin 5 and Keratin 13. Scale bar represents 100 μm in the first 2 rows and 50 μm in the last 2 rows of images.

4.2.5 *Bcl11a* Marks a Long-Lived Population in the Oesophageal Epithelium

To assess the longer-term persistence of *Bcl11a* labelled cells, epithelia were collected and clarified from *Bcl11a*^{CreERT2} mice 12 weeks post exposure to tamoxifen. Whole mount analysis from these tissues revealed widespread recombination (see figure 4.20 and 4.21). A greater proportion of cells at this time point appear to be yellow, indicating a possible expansion of labelled cells in the basal layer. The analysis of a *Bcl11a*^{WT} epithelium 12 weeks after tamoxifen exposure did not reveal the presence of any tdTomato labelling when stained with the RFP antibody (figure 4.4).

In lineage tracing studies of the oesophagus, such high recombination rates were not reflected in long chase experiments. Clonal lineage tracing [95] identified regions of roughly 4 or 5 clones over 20 weeks. Here, much of the epithelium is labelled with large clonal areas visible. Such a high degree of recombination could be due to a high tamoxifen dose. Although tamoxifen titration has not been attempted here, it will be important for future work to ensure cells are labelled in isolation.

To understand more about the relative proportions of labelled cells, microscope images were quantitated (figure 4.22). Interestingly, the relative proportion of double positive cells increased after 12 weeks, suggesting the persistence, beyond multiple divisions, of labelled cells in the basal layer. To better visualise the relative localisation of labelled cells, multiple planes were combined to form a 3D projection of the epithelial layer (figure 4.23 and 4.24). Interestingly, much of the basal layer did appear to show some level of tdTomato expression (figure 4.23). Similar to previous time points, the majority of labelled cells appeared to be localised to the suprabasal layer (figure 4.24). In comparison to the 3D projections of the other epithelia, the basal layer at 12 weeks post tamoxifen induction did appear to house more labelled cells than at 8 weeks. To assess any areas of apparent clonality, multiple planes were analysed for any local expansion of labelled cells in the basal layer as well as evidence of differentiation into the suprabasal layers (figure 4.25). Indeed, clonal expansion did seem to occur in labelled cells of the basal compartment; this was seen to expand into the more differentiated cells of the suprabasal layer (figure 4.25). This local expansion of labelled cells could signify *Bcl11a* expression in a progenitor population of the oesophageal epithelium. Further analysis of the basal layer identified more regions of expansion in labelled cells (Figure 4.26). Clonal expansion was visible in both cells of the basal layer extending to the

differentiated layers. Again, the greatest proportion of labelled cells resided in the suprabasal layers (figure 4.26). This did seem to represent an expansion of labelled cells in this compartment when compared to the epithelia of the previous time points, further reinforcing a progenitor role for *Bcl11a* in the oesophageal epithelium.

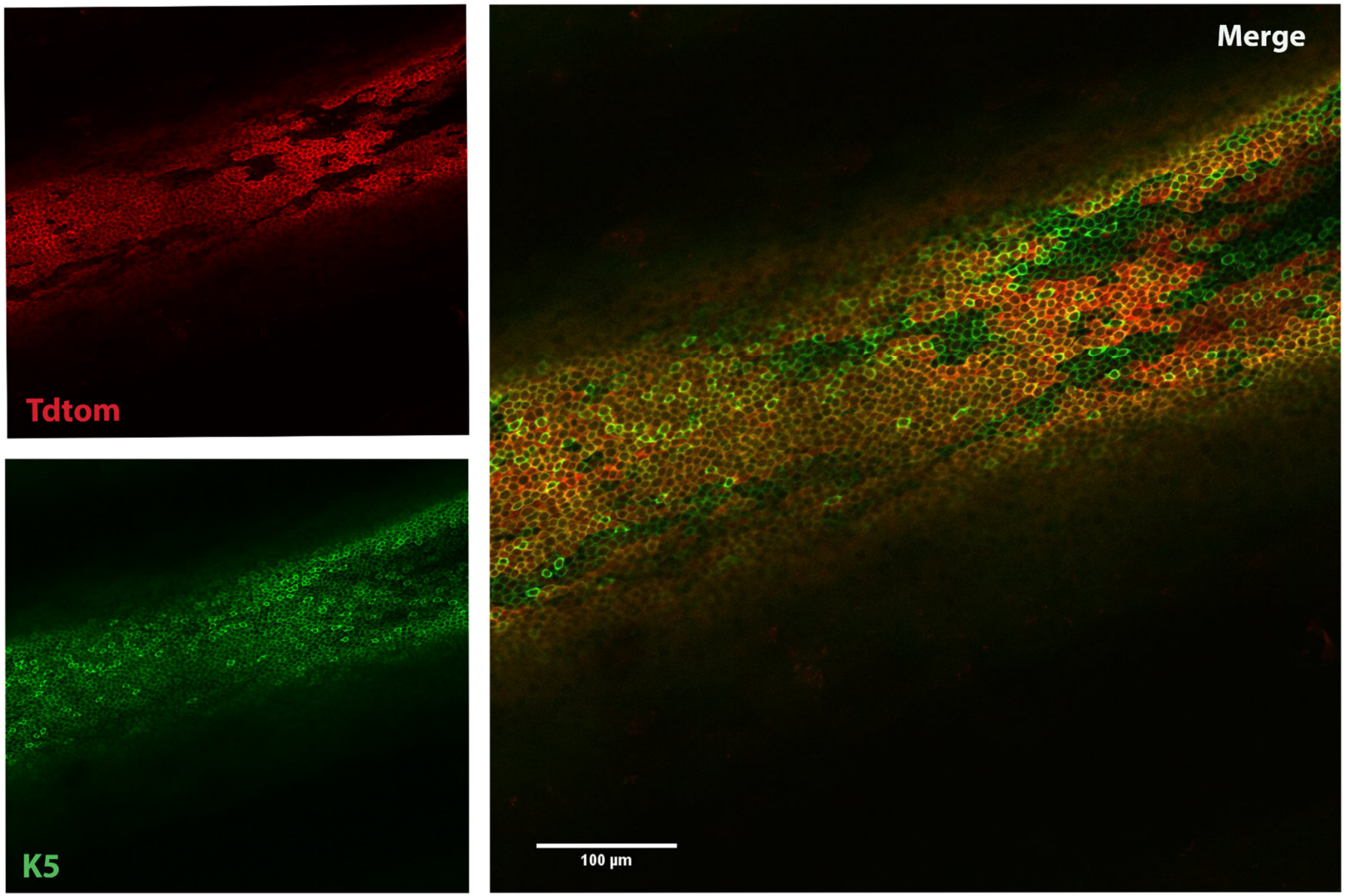


Figure 4.20: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure, global view.

Figure 4.20 (cont.): Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure, global view

The epithelium from one *Bcl11a*^{CreERT2} mouse (mouse 1) was isolated and optically cleared 12 weeks post tamoxifen exposure. The epithelium was stained with a marker of basal cells: Keratin 5 in green and tdTomato in red. Scale bar represents 100 μm.

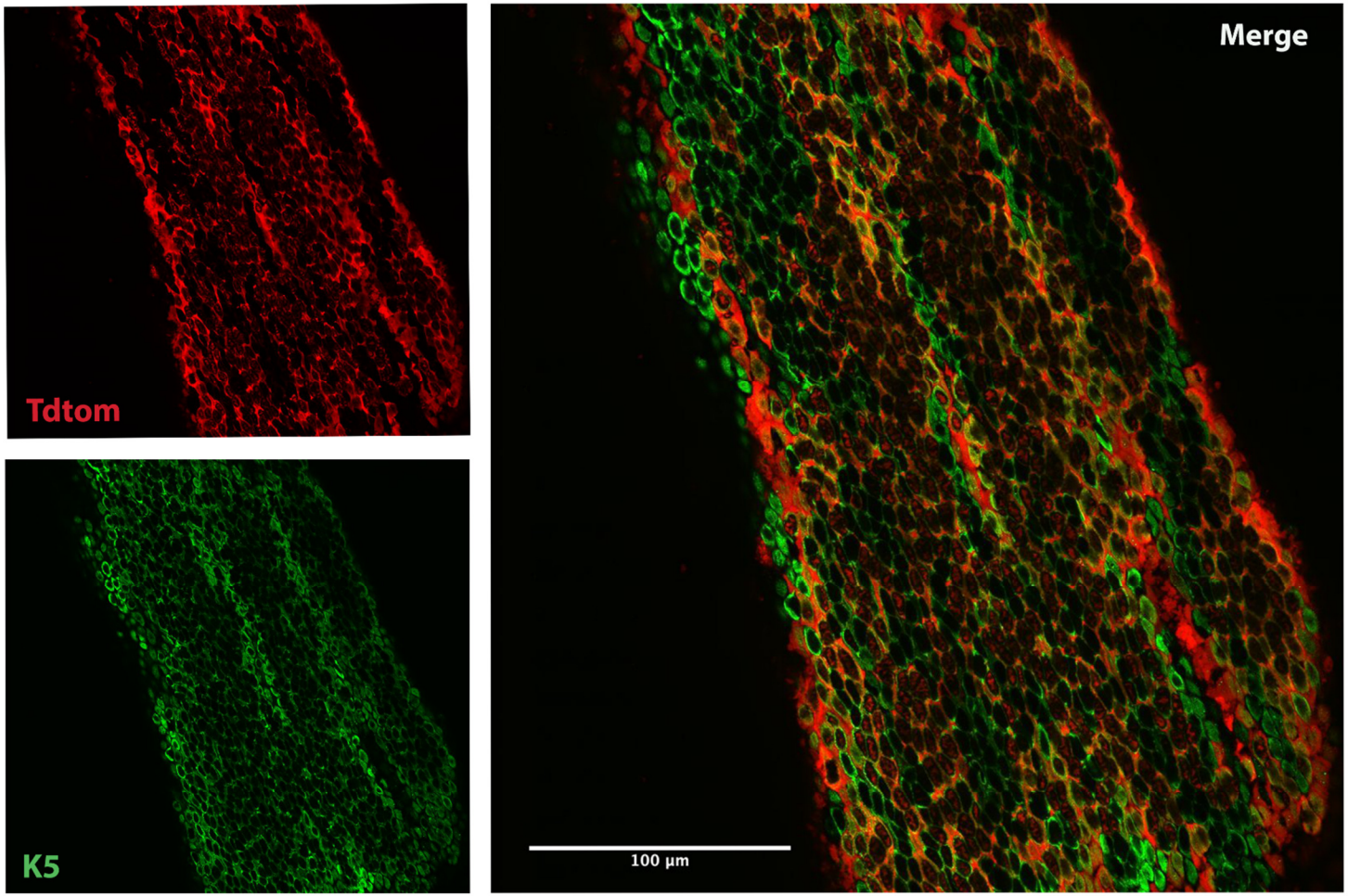


Figure 4.21: Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure, global view.

Figure 4.21 (cont.): Fluorescence microscope images from *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure, global view.

The epithelium from one *Bcl11a*^{CreERT2} mouse (mouse 2) was isolated and optically cleared 12 weeks post tamoxifen exposure. The epithelium was stained with a marker of basal cells: Keratin 5 in green and tdTomato in red. Scale bar represents 100 μm.

S

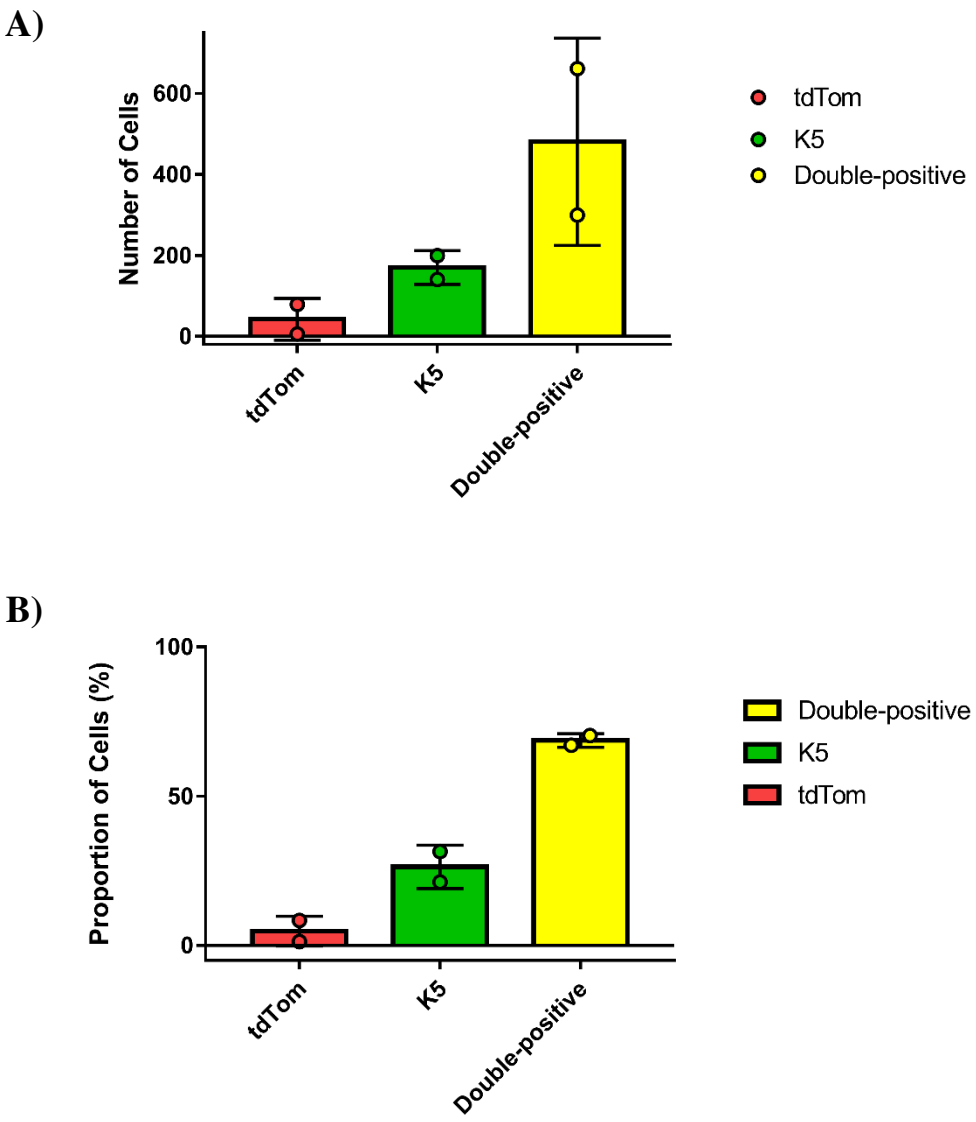


Figure 4.22: Quantification of microscope images from 2 *Bcl11a*^{CreERT2} mice 12 weeks post tamoxifen exposure.

Microscope images from both *Bcl11a*^{CreERT2} mice (figures 4.20 and 4.21) were quantitated by scoring cells as tdTomato positive, Keratin 5 positive or double positive. **A)** Depicts the number of cells displaying each marker ± S.D n=2 biological replicates. **B)** Numerical data from (A) presented as a proportion of cells positive for each marker ± S.D.

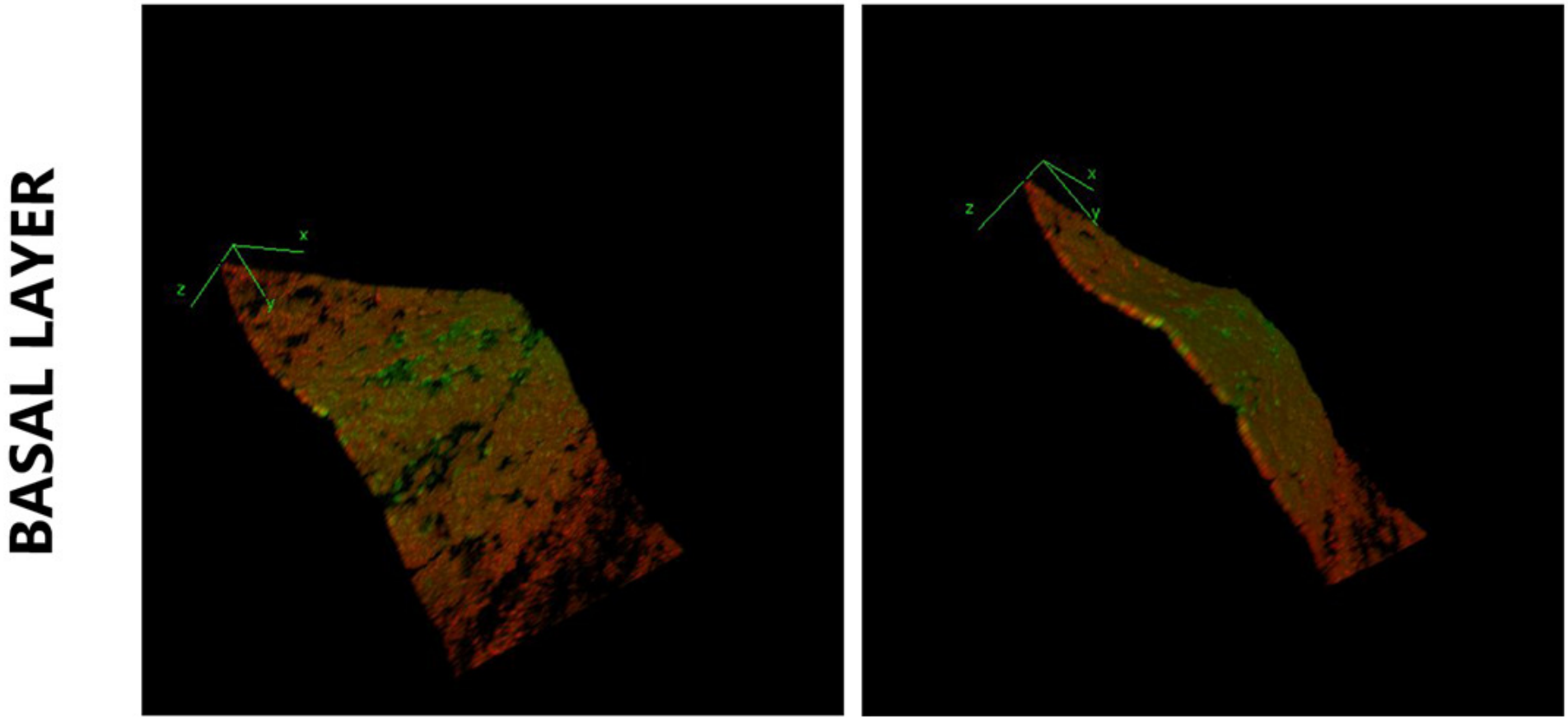


Figure 4.23: A 3D maximum projection of z stack view 2 of *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure.

The epithelium was prepared as described in figure 4.21. Using a fluorescence microscope multiple planes of view were acquired in a single field and compiled as a 3D maximum projection using ImageJ software. Keratin 5 cells are marked in green and tdTomato in red.

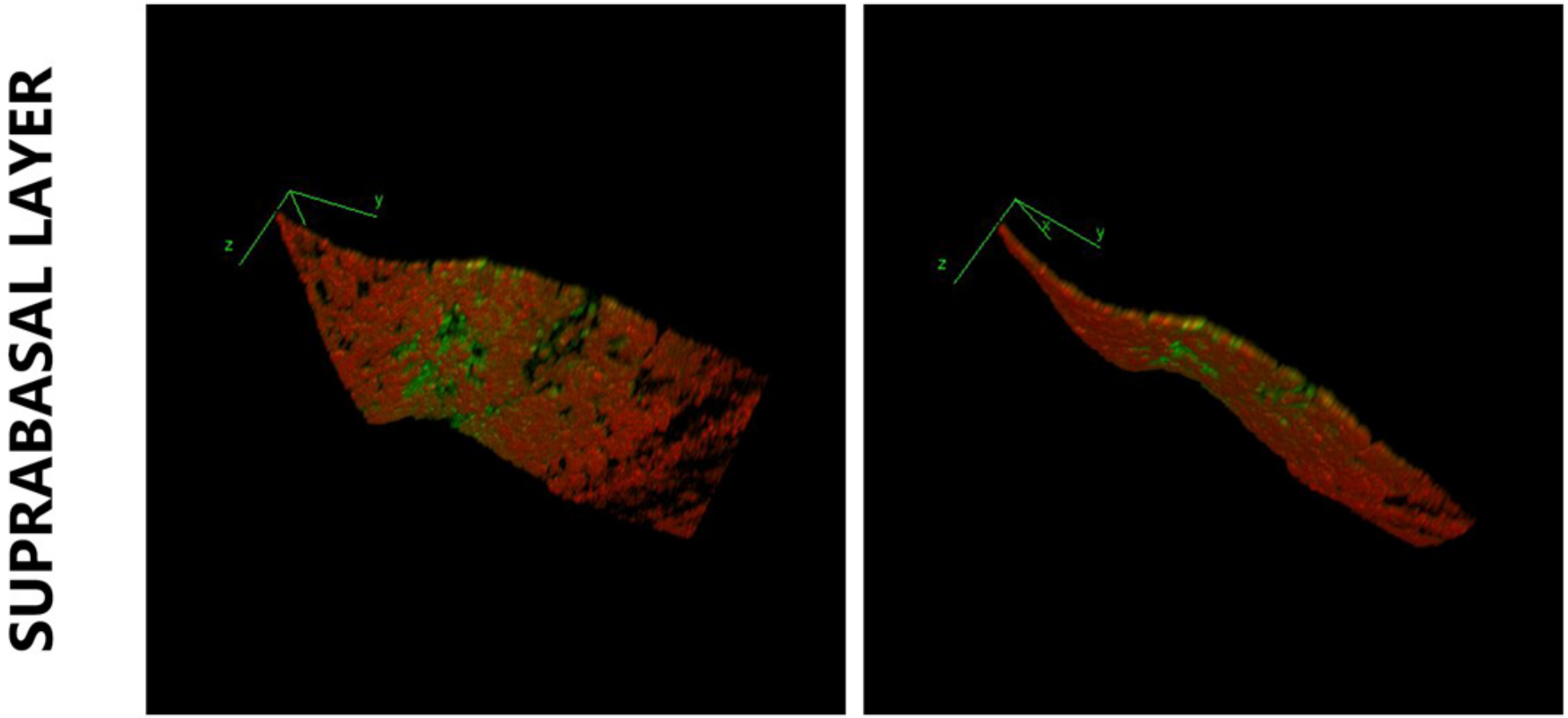


Figure 4.24: A 3D maximum projection of z stack view 3 of *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure.

The epithelium was prepared as described in figure 4.21. Using a fluorescence microscope multiple planes of view were acquired in a single field and compiled as a 3D maximum projection using ImageJ software. Keratin 5 cells are marked in green and tdTomato in red.

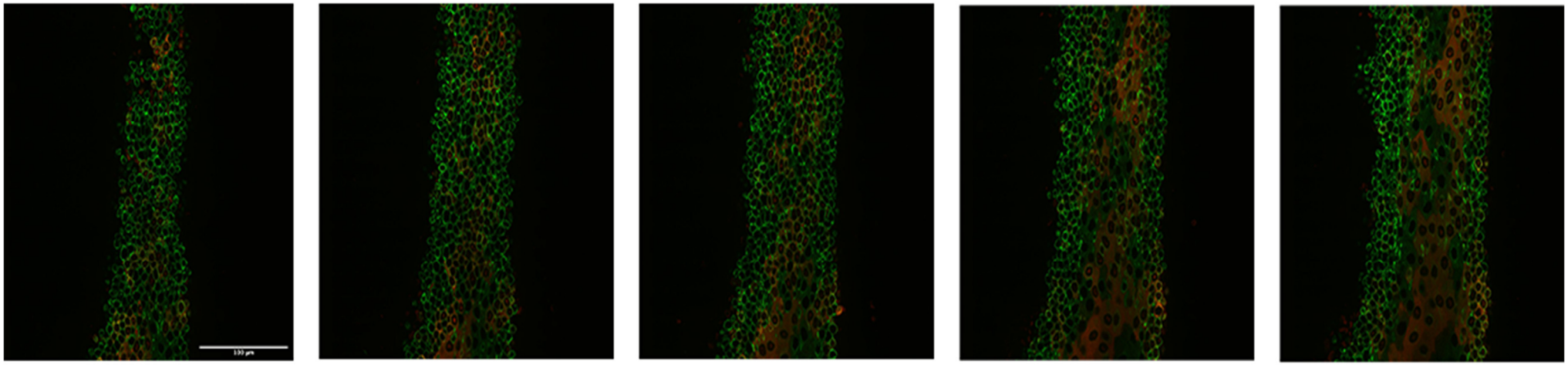


Figure 4.25: Fluorescence microscope step through of the oesophageal epithelium of a *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure.

Multiple Z slices through the oesophageal epithelium isolated from a *Bcl11a*^{CreERT2} mouse 12 weeks post tamoxifen exposure. Keratin 5 staining is shown in green and tdTomato is shown in red. Scale bar represents 100 µm.

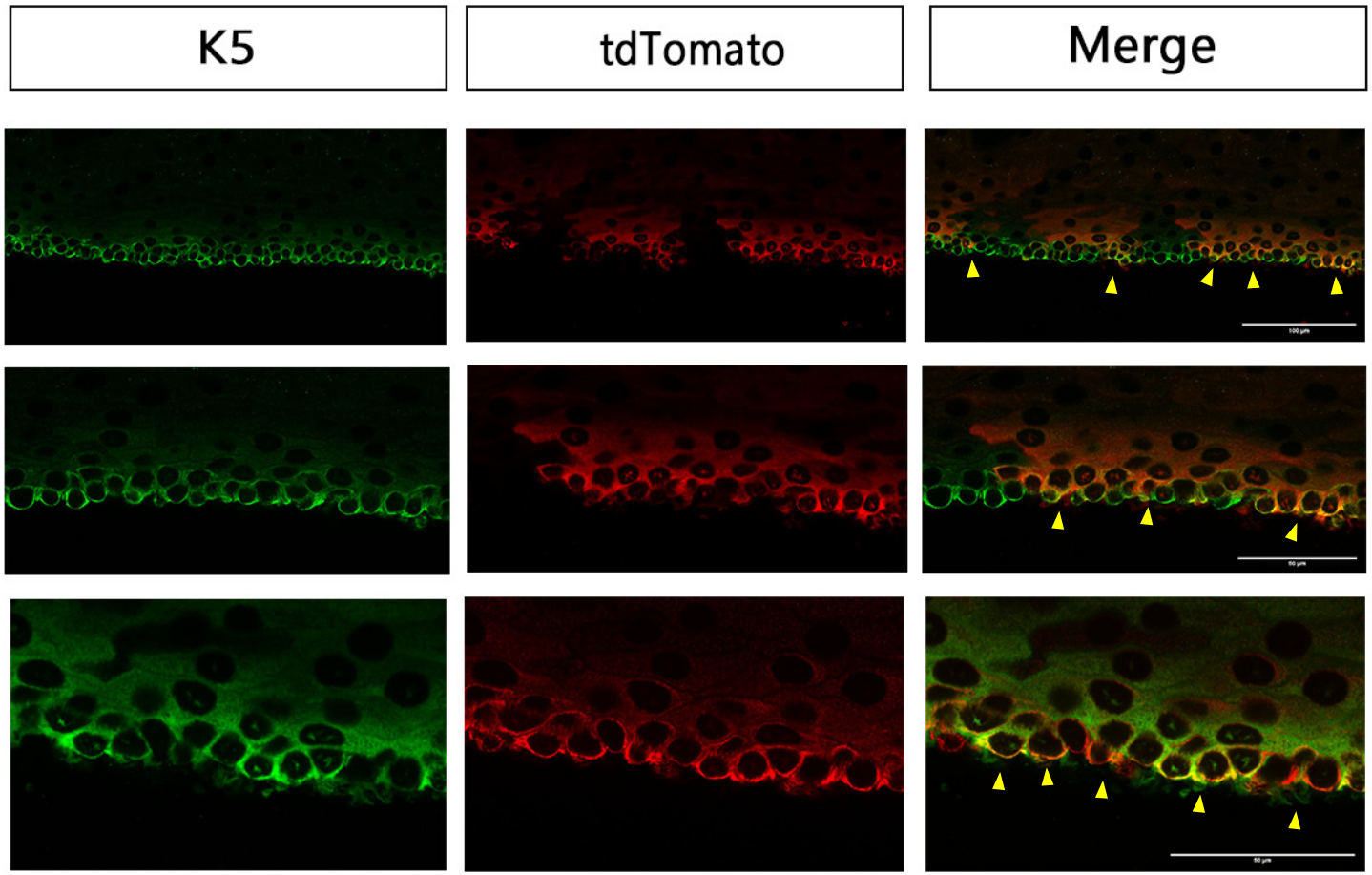


Figure 4.26: Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 12 weeks post tamoxifen exposure.

Figure 4.26 (cont): Multiple fields of view of oesophageal epithelia from *Bcl11a*^{CreERT2} mouse, 12 weeks post tamoxifen exposure

Oesophageal epithelium was isolated from a *Bcl11a*^{CreERT2} mouse 12 weeks after tamoxifen exposure. The epithelium was stained with a marker for Keratin 5, shown in green and tdTomato, shown in red. Arrows indicate cells or groups of cells that have stained positive for both Keratin 5 and tdTomato. Scale bar represents 100 μm in the first row and 50 μm in the last 2 rows of images.

4.3. Discussion

Genetic lineage tracing was undertaken in the oesophageal epithelium to understand the contribution of *Bcl11a* expressing cells to the overall tissue maintenance. Using multiple chase lengths, lineage tracing revealed a high degree of labelling across all cells of the epithelium. The labelled cells persisted over time, suggesting *Bcl11a* marked a progenitor population. TdTomato positive and negative populations were sorted on the basis of this fluorescence, 24 hours post tamoxifen exposure and plated as organoids. Red cells gave rise to more colonies than non-red cells, suggesting that *Bcl11a* expression promotes colony formation *in vitro*. In the later time points, areas of clonal expansion in labelled cells were visible as well as a clear increase in Keratin 5 and Keratin 13 double-positive cells, further highlighting the proliferative capacity of labelled cells. In the final 12 week chase much of the oesophageal epithelium appeared to be recombined. Taken together this suggests that *Bcl11a* marks a progenitor population in the oesophageal epithelium.

Despite this evidence however, important caveats remain to be considered. Tamoxifen dosage was not titrated for these lineage tracing experiments which could account for the extremely high level of recombination observed. The induction protocol was adapted from the mammary gland where the levels of recombination at this dosage were far less prominent. Unfortunately, the COVID-19 restrictions prevented these experiments from being performed in time. However, it is an important consideration for future work to ensure that cells are labelled in isolation, which can be achieved by ensuring the tamoxifen dose is low enough to label single cells. In addition, despite work in the mammary gland proving label specificity, tdTomato has not been confirmed as specific to *Bcl11a* expression using an independent method in the oesophageal epithelium. This is important to ensure that labelling is representative of the real expression both in terms of location and kinetics.

Future lineage tracing studies should also be undertaken to analyse cell division and differentiation kinetics to confirm the exact identity of *Bcl11a* expressing cells. For example, Edu incorporation could be performed alongside the lineage tracing. This will provide a greater insight into how these cells aid tissue maintenance. Many important questions remain about the nature of these cells, in particular in response to tissue injury. It would be interesting to see whether *Bcl11a* marks a population responsible for tissue repair, confirming the importance of gene expression in progenitor cells of the epithelium. By the same token, the role of labelled

cells in tumour formation is another interesting future perspective for this work. Given the putative oncogene status of *BCL11A* in OSCC, it is possible that *Bcl11a* expressing cells could contribute to much of the tumour mass during the development of such a malignancy in the oesophagus.

In a similar vein, *Bcl11a* lineage tracing in the mammary gland has many important avenues for further work in addition to the oesophagus. *Bcl11a* lineage tracing was undertaken in the mammary gland by Dr Sara Pensa with the aim of understanding the contribution *Bcl11a* expressing cells had to the tissue. *Bcl11a* was found to mark a luminal progenitor population which was seen to expand during gestation giving rise to alveolar cells. Similar to data obtained for the oesophagus, further characterisation of these labelled cells is necessary to deduce the exact position of *Bcl11a* expressing cells within the wider cellular hierarchy of the mammary gland. The status of these cells in response to tumour formation is also an important avenue for further research.

Chapter 5: Exploring *BCL11A* Function in Mammary Organoids

5.1 Introduction

5.1.1 *BCL11A* Gene Function in the Mammary Gland

The status of *BCL11A* as a TNBC gene [2] in the mammary gland makes this transcription factor a valuable candidate for further study. Key questions remain around the function of this gene in both the resting and diseased gland. In particular; whether gene expression is required for normal tissue development; or indeed whether gene expression is associated with stemness as has been described in some tissues [14,26]. Questions also remain about the effect of gene perturbations on cellular and molecular dynamics, to model TNBC and gain insight into molecular mechanisms of disease.

To perform experiments to alter *BCL11A* gene expression, faithful tissue modelling of the mammary gland was required, for which a 3D organoid system was selected. This culture system represented the *in vitro* model used for gene function study and offered several key advantages over traditional cell culture methods. For the purposes of this work, a key feature of the technique was the ability to cultivate structures derived from single cells. It was important to maintain a handle on the cell type plated to observe any potential plasticity induced by alterations in *BCL11A* expression. Based on the recent understanding that basal-like breast cancers (BLBCs) arise in cells of luminal origin [117] each experiment was started with a homogenous population of luminal progenitor cells. This population has been identified as $EpCam^{high}$ $CD49f^{low}$, $Sca1^{neg}$ $CD49b^{high}$ cells within the mammary gland using FACS [102]. The sorting strategy was adapted from Shehata et al.: a representative plot can be seen in figure 5.1. The derivation of cultures from this population in $BCL11A^{OE}$ and $Bcl11a^{KO}$ mice therefore gave an insight into cellular dynamics and the molecular mechanisms of disease. For the purposes of final analyses, it was important to cultivate as many organoids as possible to have enough material for RNA extraction, to analyse gene expression, as well as to perform immunofluorescence staining to assess the cellular composition of organoids and whether gene expression had an influence on this.

Therefore, for accurate modelling, the organoid system required a robust protocol. Although some publications did claim to have made the organoid culture system possible for the mammary gland, it does remain a sensitive and difficult technique to master, particularly in the context of generating organoids from single cells. Therefore, the first aim was to compare 3D organoid culture methods to decide on the most robust system to take forward. The second aim

was to assess the biological function of *BCL11A* in the mammary gland using the luminal progenitor cells from *BCL11A*^{OE} and *Bcl11a*^{KO} mice, cultivated in 3D, to model the effects of these changes in gene expression.

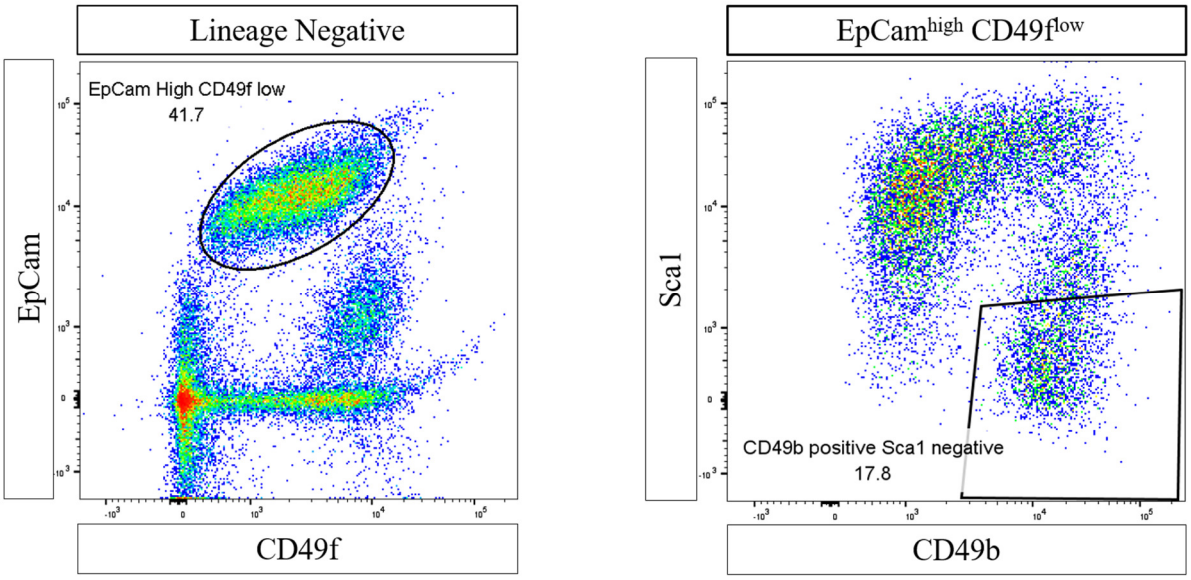


Figure 5.1: Representative FACS plots for the selection of luminal progenitor cells.

For the culture of 3D organoids, cells of the luminal progenitor compartment were seeded as single cells. To visualise the gross population of luminal cells, lineage negative cells were separated based on EpCam and CD49f expression (left). EpCam^{high} CD49f^{low} cells were gated (ellipse) and then visualied based on CD49b and Sca1 expression. CD49b^{positive} Sca1^{negative} cells (square) were sorted for plating as organoids (right).

5.2 Results

5.2.1 Comparing Organoid Models of the Mammary Gland

Before gene function could be explored in the mammary gland, published models of 3D culture were compared. Organoid culture systems were judged on technical aspects such as colony forming capacity in each setting, as well as cellular composition using whole mount analysis. Organoids were collected for whole mount analysis and stained with basal and luminal markers (Keratin 5 and Keratin 8) to characterise the cellular composition of luminally derived organoids in each system.

The first organoid method characterised represented a newly defined protocol adopted from Jardé et al, and has been named 'medium formulation 1' for the purposes of this work [148]. Single sorted luminal progenitor cells from WT (C57BL/6) mice were seeded in matrigel at a density 3000 cells per well to firstly assess growth. Organoids were cultivated using the medium in table 2, a formulation with a few key growth factors such as Nrg1, R-spondin-1 and Noggin. The luminal progenitor cells generated very small structures after seven days, hereafter referred to as luminal organoids (figure 5.2). These were then cultivated for a further 7 days to assess growth and ultimately cultivated for a total of 21 days, during which organoids demonstrated modest growth in size.

Organoids were defined as structures 50 µm or larger, giving an average OFU frequency of 2% for cells grown in these conditions. Interestingly, the spherical morphology was retained for the duration of the 21 days, with no budding visible. This morphology, described as a 'simple cyst' could have been a result of a lack of branching factors in the media [145]. Although Jardé et al. do report the appearance of lobular structures, these organoids were cultured for a longer period (30 days) whilst also being derived from tissue fragments and not from single cells.

Upon closer inspection using a bright field microscope luminal organoids did appear to display a degree of heterogeneity. Some organoids contained areas of lower density within the structures as indicated by the arrows in figure 5.3. These could be indicative of inner lumens. Organoids cultivated by Jardé et al. demonstrated the presence of lumens within the lobules of the organoids. This suggests that the phenomenon of lobule formation does not require the presence of heterogeneous cell populations and can be driven by luminal cells alone. To see

whether these areas were indeed lumens and to gain a greater insight into cellular composition, organoids cultivated for 21 days were collected for whole mount analysis.

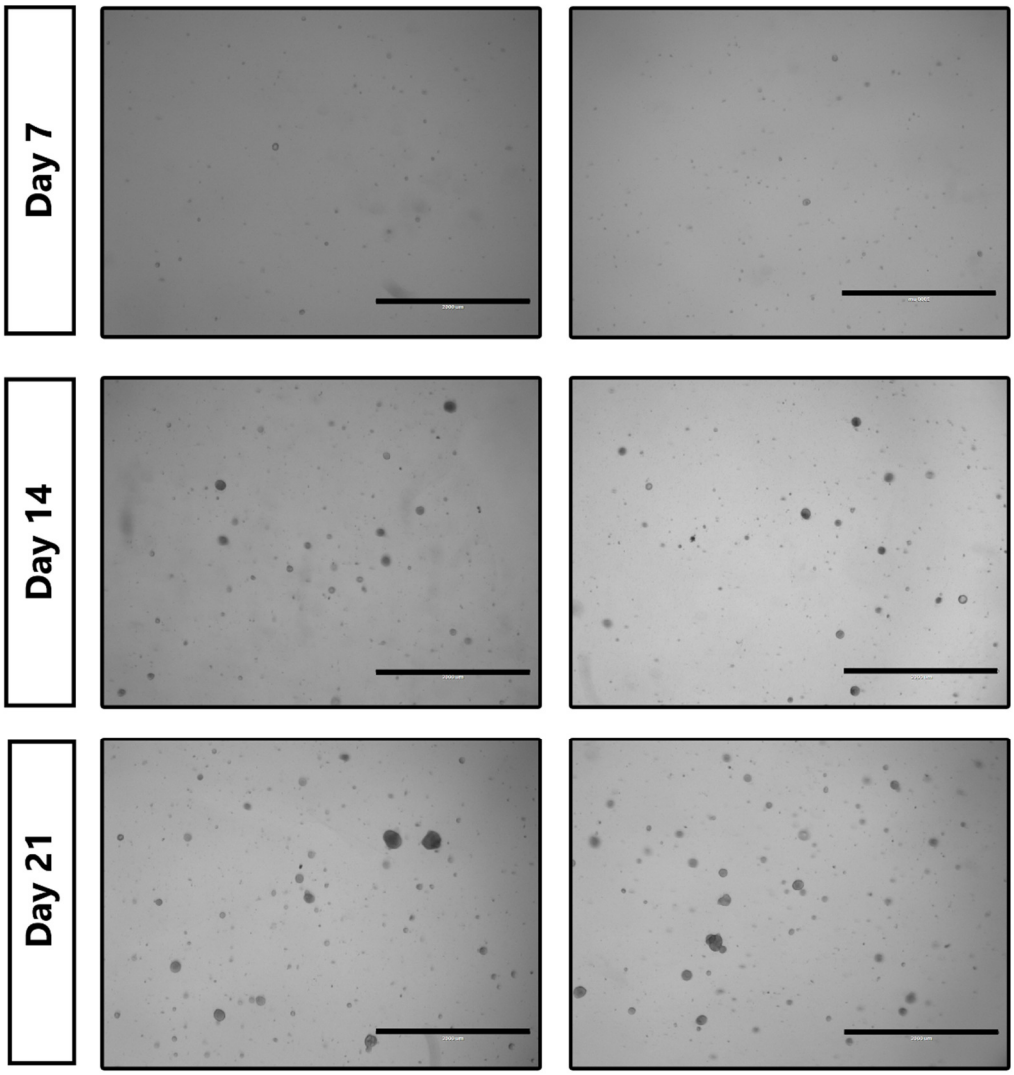


Figure 5.2: Luminal organoids cultivated using medium formulation 1.

Luminal progenitor cells from WT mice were cultivated as organoids for a period of 21 days. Scale bar indicates 2000 μm. Images are representative of wells where 3000 cells were plated per well.

The progenitor subset comprises luminal cells that are not hormone sensing and give rise to colonies that are hormone receptor negative but that are predominantly luminal [102]. More recently, the capacity of these cells to give rise to a few Keratin 5 positive cells within an organoid *in vitro* has been demonstrated [147]. Therefore, organoids derived from luminal progenitor cells would be expected to generate structures composed largely of Keratin 8 positive cells with perhaps some evidence of Keratin 5 positive cells around the periphery. Indeed, luminal organoids cultivated using this method were found to broadly display Keratin 8 positive cells with some organoids demonstrating the presence of Keratin 5 positive cells (figure 5.4). Although the quality of the staining of these organoids did appear to be poor, some evidence of Keratin 5 positive cells were found arranged around the outer layer of the organoids. This polarity is related to the bilayered structure found *in vivo*. Structures retained their spherical morphology and did indeed contain (in some cases) multiple lumens per organoid as indicated in figure 5.4. Overall staining towards the middle of the organoids appeared to be weaker, potentially indicating poor antibody penetration. It is therefore not possible to comment on the extent of Keratin positivity in these inner cells.

Whole mount analysis did not appear to reflect the heterogeneity in size of the organoids observed in the brightfield view. In fact, the organoids that were finally analysed were mostly only those large enough to be visible by eye. This meant that the majority of the organoids were lost during the sample preparation stages. Despite cultivating organoids for 21 days, many structures remained too small to view by eye and therefore inadequate for use in this analysis. In addition to this, organoids displayed a OFU frequency so low that many mice were required for each preparation of the culture. This was unsustainable for use in the final phase of the project as the *BCL11A*^{OE} and *Bcl11a*^{KO} mouse colony sizes were substantially smaller than normal, due to the COVID-19 pandemic. The organoids using medium formulation 1 did not display the branching morphogenesis expected in 3D culture of the mammary gland. This phenomenon was not representative of the *in vivo* gland and, coupled with the technical short falls, this method of culture was ultimately not taken forward for use in the project and a different media formulation was sought.

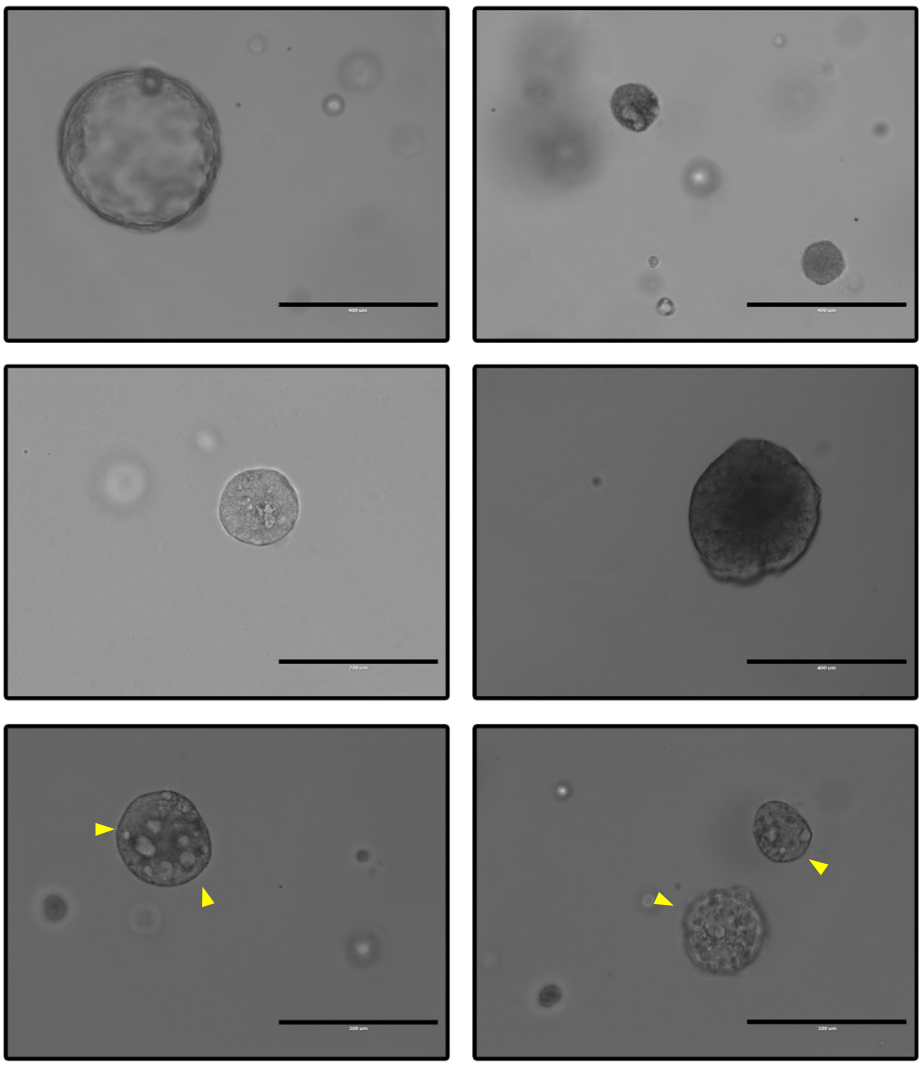


Figure 5.3: Individual luminal organoids cultivated using medium formulation 1.

Luminal progenitor cells from WT mice were cultivated as organoids for a period of 21 days using medium formulation 1. Images are representative of wells where 3000 cells were plated per well. Scale bar represents either 200 µm in the bottom two images or 400 µm in the top four images. Arrows indicate areas of lower density within the centre of the organoids.

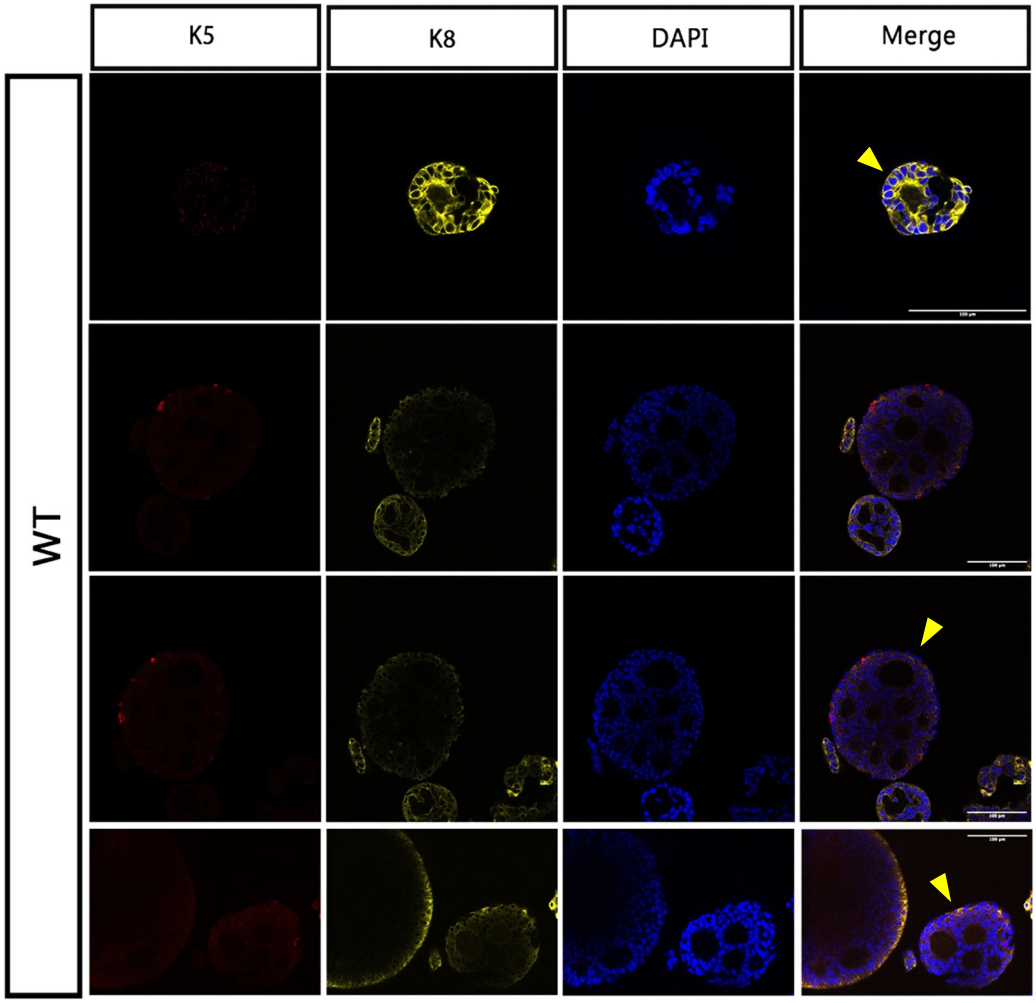


Figure 5.4: Fluorescence microscopy of whole mount WT mammary luminal organoids cultivated in medium formulation 1.

Luminal mammary organoids were subjected to whole mount analysis after 21 days in culture**.** Organoids were stained with a luminal marker (Keratin 8 in yellow), a basal marker (Keratin 5 in red) and DAPI (in blue) for nuclear visualisation. Arrows indicate organoid lumens**.** The scale bar indicates 100 µm. Images were obtained using a confocal microscope with one optical z plane displayed.

To explore further media formulations, a new recipe was adopted from Jamieson et al. and for the purposes of this work it was named 'medium formulation 2'. Due to the retained use of matrigel ECM, the plating protocol was similar to that used in the previous system. This medium was more extensive in supplementation (see table 2), including many factors previously not added to the culture medium. As such, it was more closely related to the organoid culture medium used for other organs [127,155]. There was a more extensive FGF supplementation, including FGF2 and FGF10 as well as heparin, added to initiate branching in the organoids [158]. In addition, both Wnt3A and R-Spondin-2 were used to stimulate the Wnt pathway required in mammary gland development [147].

Luminal progenitor cells from WT mice were seeded as single cells at a density of 3000 cells per well and cultured in a similar manner to before. Within 7 days of seeding, colonies began to emerge as small structures, demonstrating an average OFU frequency of 8.7%, higher than the previous media formulation. After 21 days in culture luminally derived organoids remained small with a spherical morphology (figure 5.5). In order to assess whether this was an intrinsic property of the cells, or perhaps a case of contact inhibition due to high colony numbers, fewer cells were seeded. Cells were seeded at varying dilutions, after which a density of 2000 cells per well was selected. Cells cultivated in this manner gave rise to larger organoids with budding structures even after 14 days (figure 5.6). The presence of budding in the organoids is considered to precede ductal morphogenesis where organoids display (to a lesser extent) the branching reminiscent of the ductal tree in puberty [145]. Closer analysis revealed some heterogeneity in the structures with some organoids demonstrating a variable degree of budding (figure 5.7).

Organoids were collected for whole mount analysis after 14 days in culture. This was due to the superior size and OFU frequency exhibited by these organoids, meaning they were suitable for whole mount analysis at this stage. This offered an advantage over the previous culture method due to improved efficiency in both organoid size and number, whilst the shorter time in culture served to minimise the likelihood of potential tissue culture-induced artefacts in the organoids.

To analyse the cellular composition, organoids were stained with markers of basal and luminal cells. Similar to the organoids cultured using medium formulation 1, the structures were broadly composed of Keratin 8 positive cells with a few cells demonstrating Keratin 5

positivity (figure 5.8). Similar to before, these cells were located on the outer areas of the organoids, displaying a polarity and bilayered structure reminiscent of the *in vivo* gland. However, again similar to previous preparations organoid staining appeared to be sub optimal, with weaker signal from the antibodies towards the centre of the organoids as well as some distortion around the structures. Organoids generally displayed elaborate budding morphogenesis, identified using whole mount analysis, without demonstrating any branching morphology. It is possible that poor antibody penetration may have been responsible for sub optimal staining due to the high number of cells at the centre of the organoids.

It has been suggested by Nguyen-Ngoc et al. that the organoid matrix is responsible for restricting the branching and lobular development of organoids compared to the extent observed in the *in vivo* gland [159]. In an attempt to further optimise the organoid system, and to understand more about the effects of ECM on branching, luminal progenitor cells were seeded in a mix of matrigel and collagen. Due to the collagen rich ECM in the mouse mammary gland, this composition was thought to represent a more physiologically accurate model. Adapted from Nguyen-Ngoc et al. cells were seeded in the matrigel collagen mix and cultured in medium formulation 2. After 7 days organoids began to form, continuing growth and elongation to become extensively branched structures by day 21 (figure 5.9). Such morphology was not observed in the culture of cells in matrigel alone (figure 5.10). This suggests that extensive branching in this manner could be inhibited in the matrigel as organoid branches did not invade the ECM, remaining compact for the duration of cell culture. Branching morphology appears to more closely resemble that of the mammary gland *in vivo.* This finding was consistent with that reported by Nguyen-Ngoc et al. and suggests that the collagen-matrigel ECM composition may represent a more physiologically relevant model of the mammary gland *in vitro.*

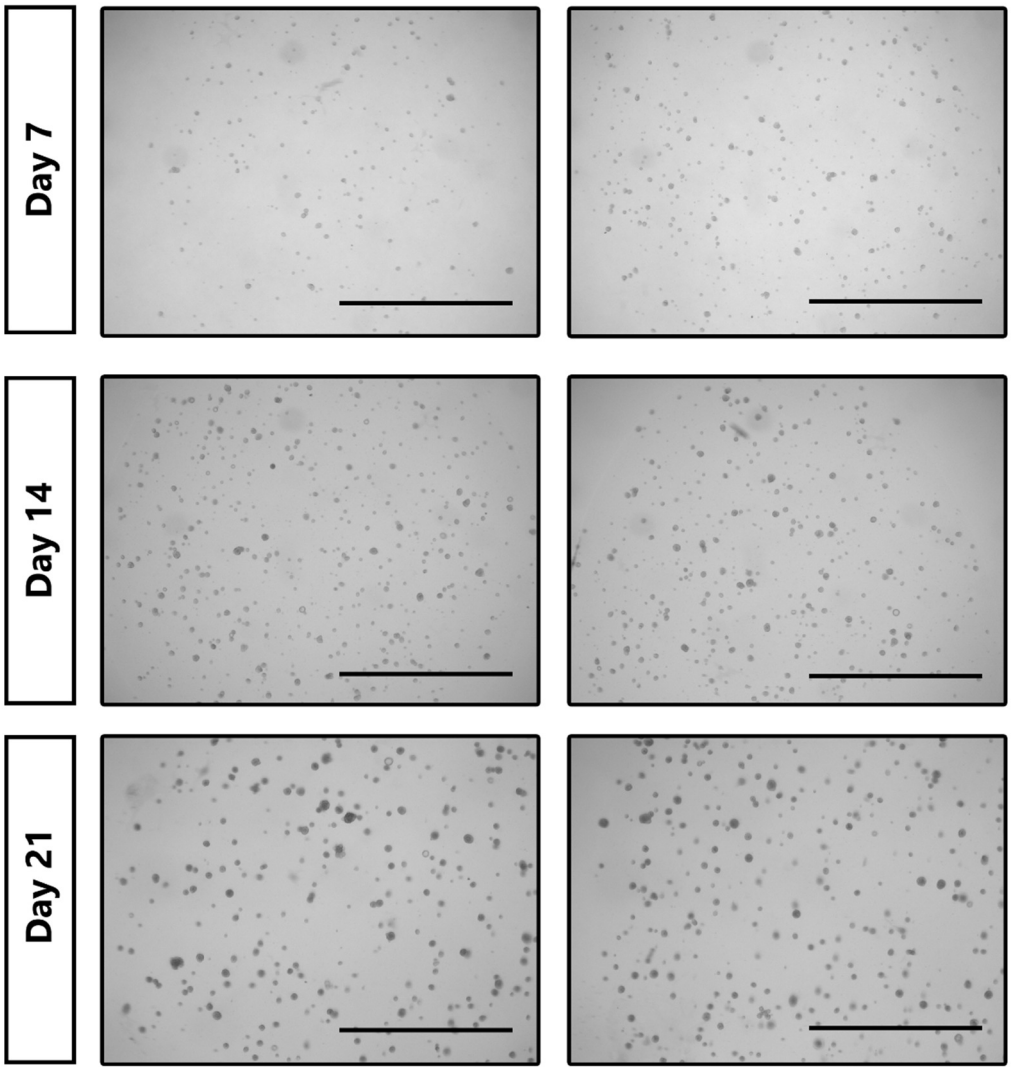


Figure 5.5: Luminal organoids cultivated using medium formulation 2.

Luminal progenitor cells from WT mice were cultivated as organoids for a period of 21 days. Scale bar indicates 2000 μm. Images are representative wells where 3000 cells were plated per well.

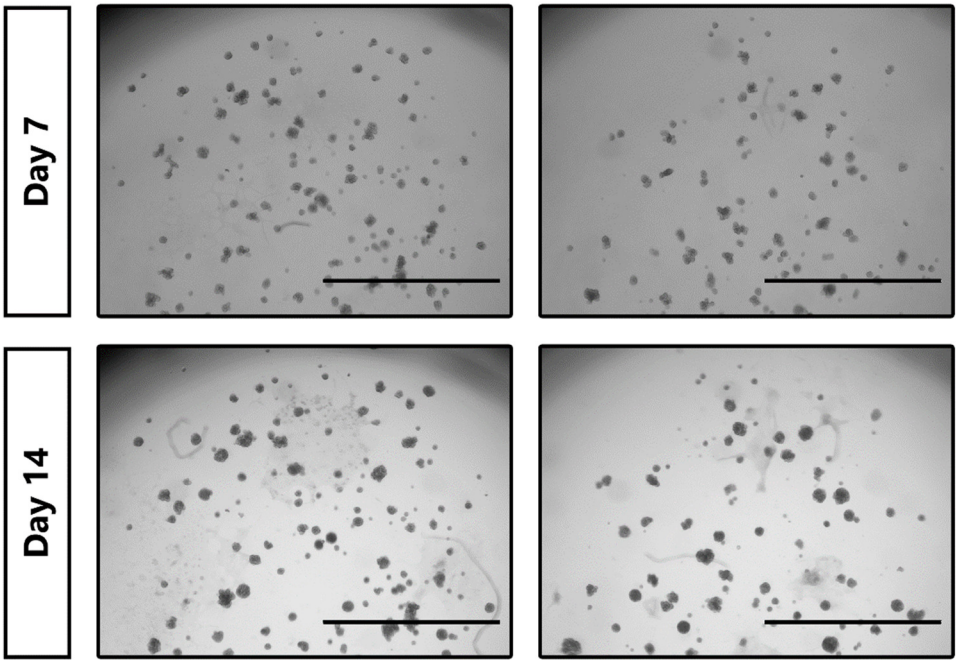


Figure 5.6: Luminal organoids cultivated using medium formulation 2 seeded at a lower density.

Luminal progenitor cells from WT mice were seeded at a density of 2000 cells per well and cultivated as organoids for a period of 14 days. Scale bar indicates 2000 μm, images are representative.

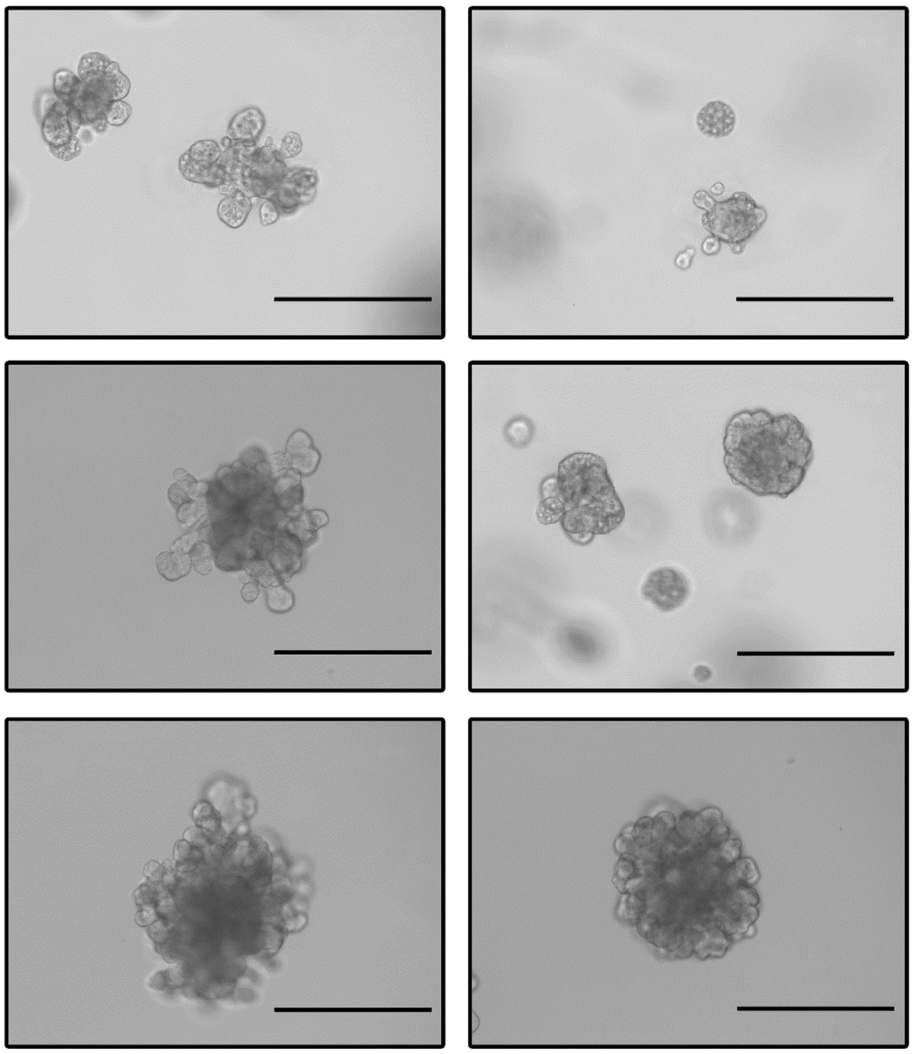


Figure 5.7: Individual luminal organoids cultivated using medium formulation 2.

Luminal progenitor cells from WT mice were cultivated as organoids for a period of 14 days. Cells were seeded at a density of 2000 cells per well. Scale bar represents 400 µm.

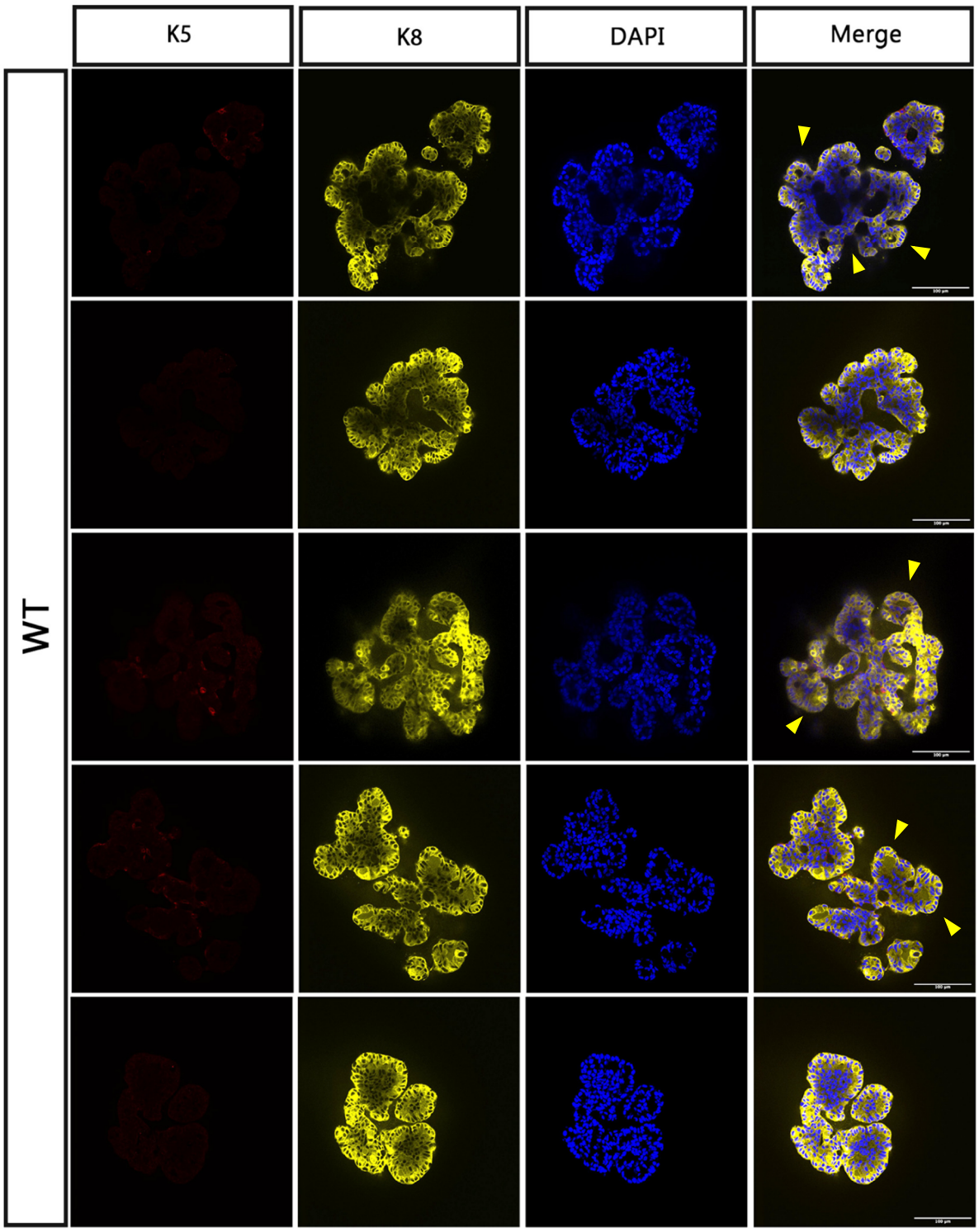


Figure 5.8: Fluorescence microscopy of whole mount WT mammary luminal organoids cultivated using medium formulation 2.

Figure 5.8 (cont.): Fluorescence microscopy of whole mount WT mammary luminal organoids cultivated using medium formulation 2.

Luminal organoids were cultivated for a period of 14 days and subjected to whole mount analysis. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Arrows indicate luminal regions. Scale bar indicates 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

To assess the cellular composition of these structures, whole mount staining was attempted using immunofluorescence markers. Due to the presence of collagen in the ECM, the usual methods of matrigel dissociation were not possible. Authors Nguyen-Ngoc et al. suggested to stain organoids whilst encased in the ECM, without liberation. When this was attempted, however, antibodies were unable to penetrate the matrigel and collagen mix. This meant that organoids were not sufficiently stained for analysis and subsequently could not be used. As mentioned previously, this method of analysis was integral to downstream experiments; without verifying the cytokeratin expression of these cells it was impossible to characterise these organoids. This culture condition was therefore unsuitable for further use in the project. Therefore, despite efforts to optimise the organoid culture system further, the matrigel collagen ECM could not be taken forward. Ultimately, medium formulation 2 was selected for use in the project, cultivating organoids derived from *BCL11A*^{OE} and *Bcl11a*^{KO} mice in Matrigel ECM alone.

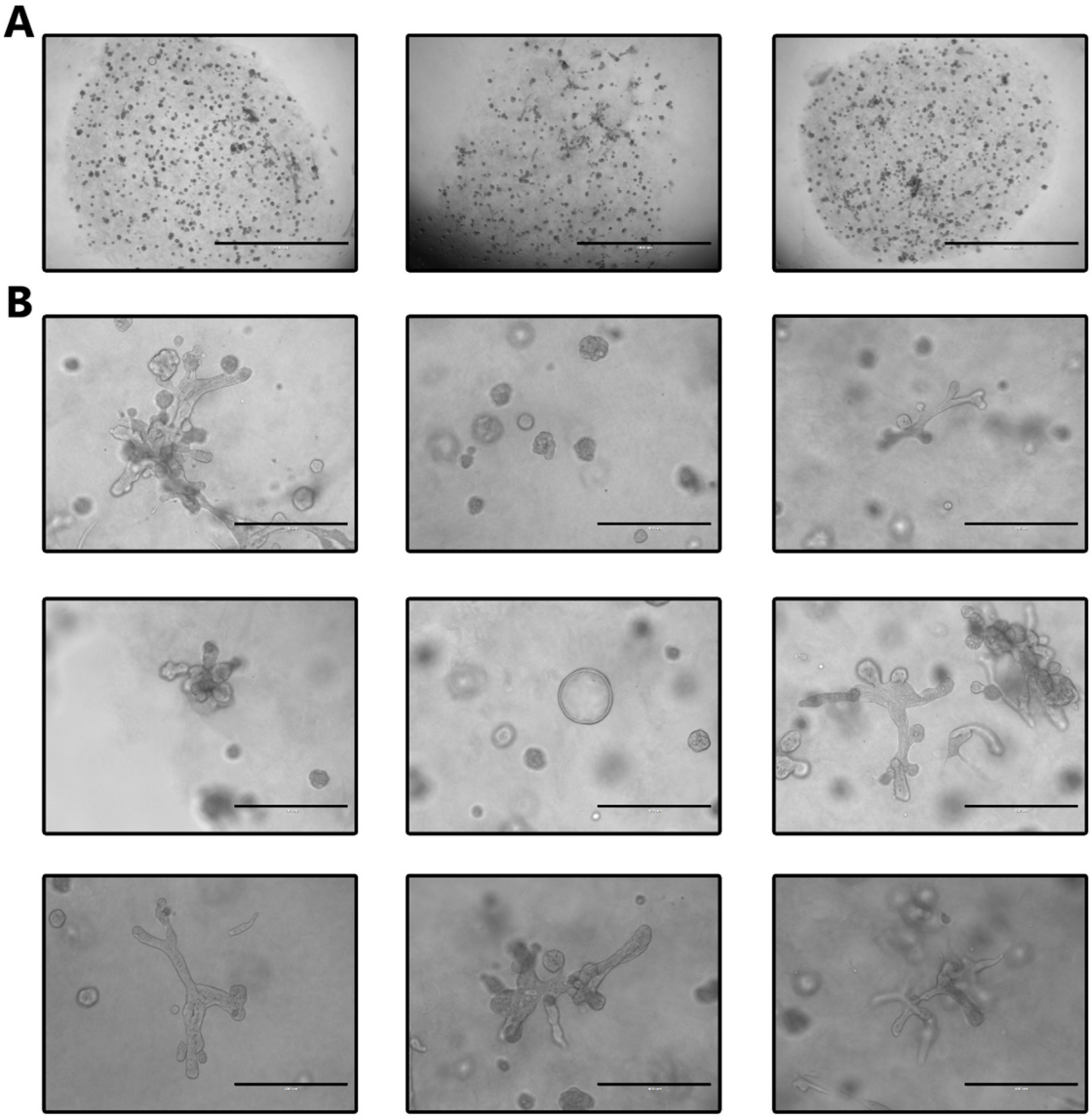


Figure 5.9: Luminal mammary organoids cultivated in a collagen and matrigel ECM mix using medium formulation 2.

Luminal organoids were cultivated for a period of 14 days in an ECM composed of collagen and matrigel. **A)** Representative wells of luminal organoids seeded at a density of 3000 cells per well. Scale bar is 2000 µm. **B)** Individual luminal organoids. Scale bar is 400 µm.

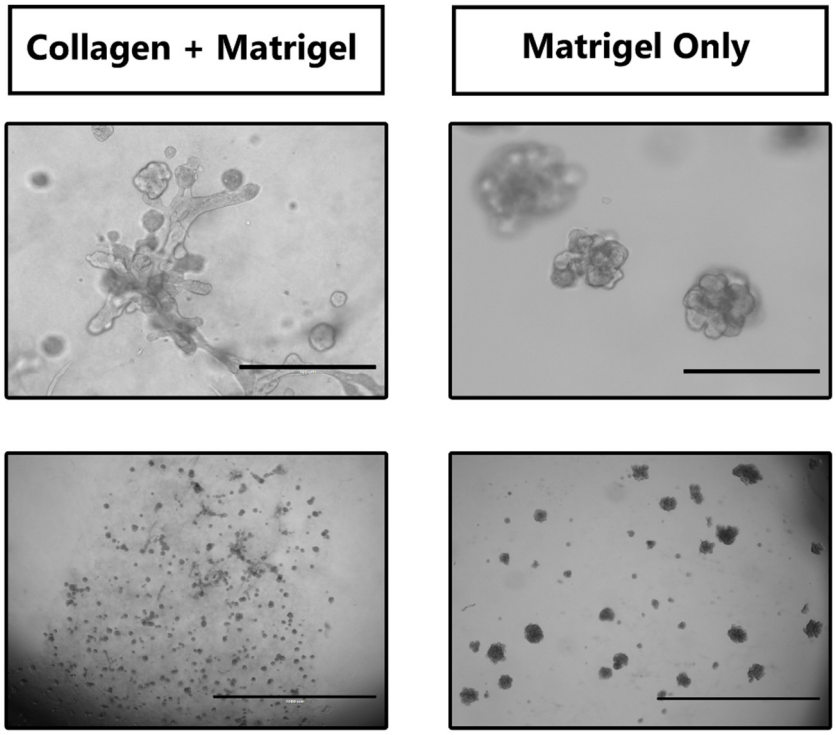


Figure 5.10: Morphological comparison of organoids cultivated using different ECM compositions.

Luminal organoids from WT mice were cultivated for 21 days in either collagen and matrigel matrix or matrigel matrix alone and supplemented using medium 2. Top panel, scale bar is 400 µm, lower panel scale bar is 2000 µm.

5.2.2 Modelling *BCL11A* Overexpression in Mammary Organoids

To model the function of *BCL11A* in the mammary gland, organoids were generated from the luminal progenitor cells of *BCL11A*^{OE} mice. Organoids were cultivated for a period of 7 days, after which induction of the transgene was mediated through 4-OHT exposure *in vitro* for a period of 24 hours (see experimental procedures for details). Organoids were treated with either 4-OHT, vehicle control or left untreated. The organoids were then cultivated for a further 7 days, after which they were collected for whole mount analysis and RT-qPCR analysis for *BCL11A* gene expression. The size and number of organoids within each condition was also analysed.

In order to first assess the effect 4-OHT or vehicle treatment had on cells, WT organoids were cultivated and exposed to these reagents. After 14 days in culture, organoids appeared to show little difference in appearance between the conditions. The structures in treated conditions displayed budding morphology consistent with control organoids, suggesting that exposure to these reagents did not induce abnormal morphology (figure 5.11). When subjected to size-based analysis and colony number, organoids did not tend to be variable in size between conditions although those treated with 4-OHT and vehicle did tend to be larger in size than untreated (figure 5.11). Organoids were subjected to whole mount analysis to identify whether exposure to these reagents had induced any abnormalities in cellular conformation or differentiation. As seen in figure 5.12, organoids appear to display broadly similar morphology and cellular composition. Organoids were largely composed of Keratin 8 labelled cells, with some evidence of Keratin 5 positive cells, in many of the structures. Similar to previous preparations, staining appeared to be weaker towards the centre of the organoids with distorted regions visible around the outside of the structures. In order to quantitate the cellular composition of these organoids further, cells were counted and scored as either Keratin 5 positive, Keratin 8 positive or double positive (figure 5.13). Cells with no discernible markers were scored as ‘none’. There appeared to be little difference between the treatment groups in terms of cellular composition of the organoids. The vehicle control treatment did give rise to organoids with more Keratin 5 positive cells however, this was only a small number of cells.

Organoids exposed to 4-OHT appeared to display similar morphology and cellular composition to the vehicle group, suggesting that 4-OHT exposure alone had little effect on

these parameters (figures 5.14 and 5.15). Taken together this suggests that 4-OHT and ethanol exposure alone had little effect on organoid size or cellular composition.

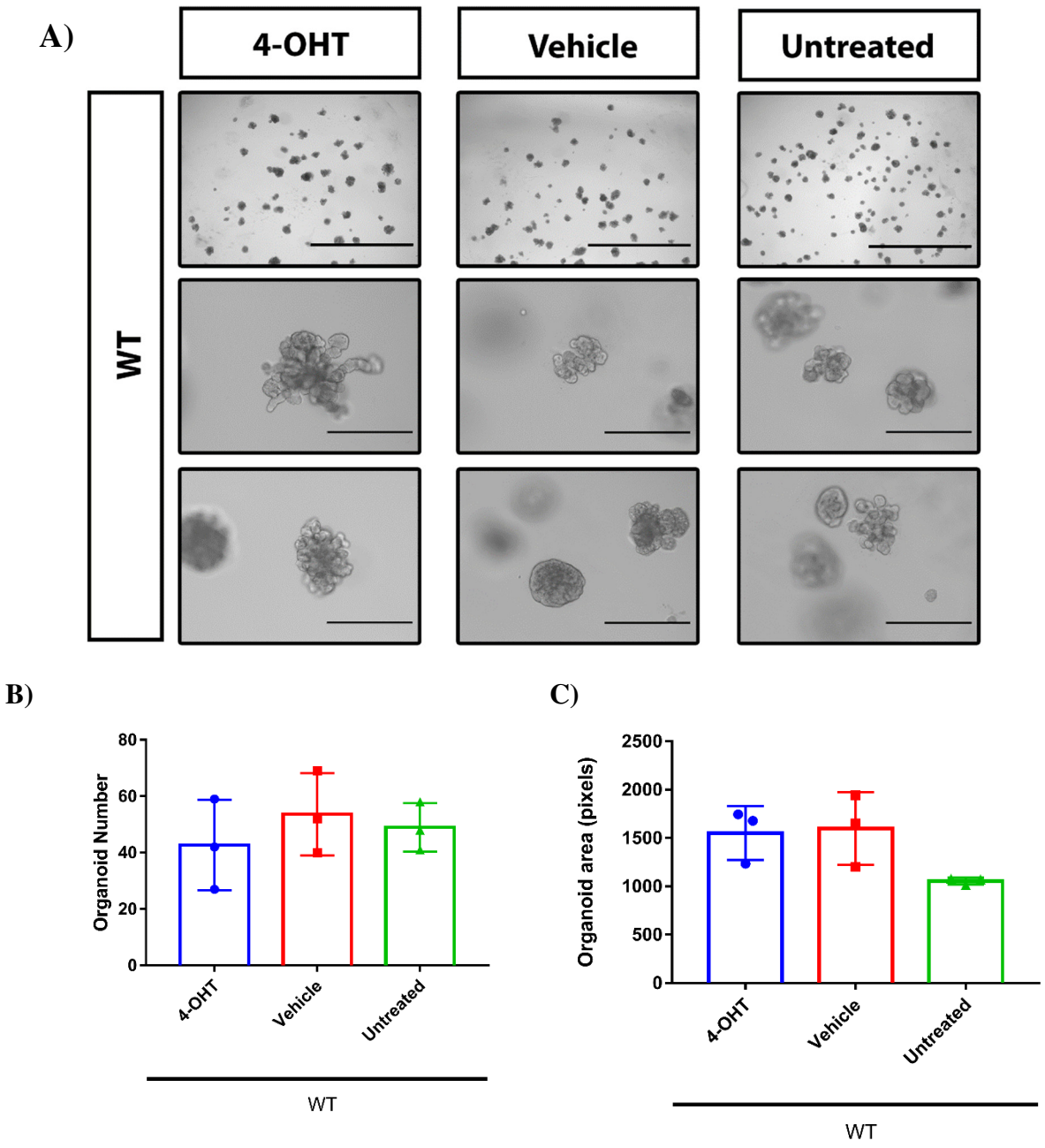


Figure 5.11: Quantification of WT luminal mammary organoid number and size.

Figure 5.11 (cont.): Quantification of WT luminal mammary organoid number and size.

Luminal mammary organoids were cultivated from WT mice for 7 days after which they were exposed to either 4-OHT or ethanol vehicle control or left untreated and cultivated for a further 7 days. Data are indicative of one biological replicate. **A)** Organoids were cultivated for a total of 14 days and subjected to different treatment. Scale bar represents either 2000 μm for the lower magnification (top row) and 400 μm at the higher magnification (lower two rows). **B)** Shows organoid number cultivated from 2000 cells after 14 days in culture for the varying conditions. **C)** shows the mean size of organoids per well in pixels. Data are represented as mean ± S.D. of n=3 technical replicates

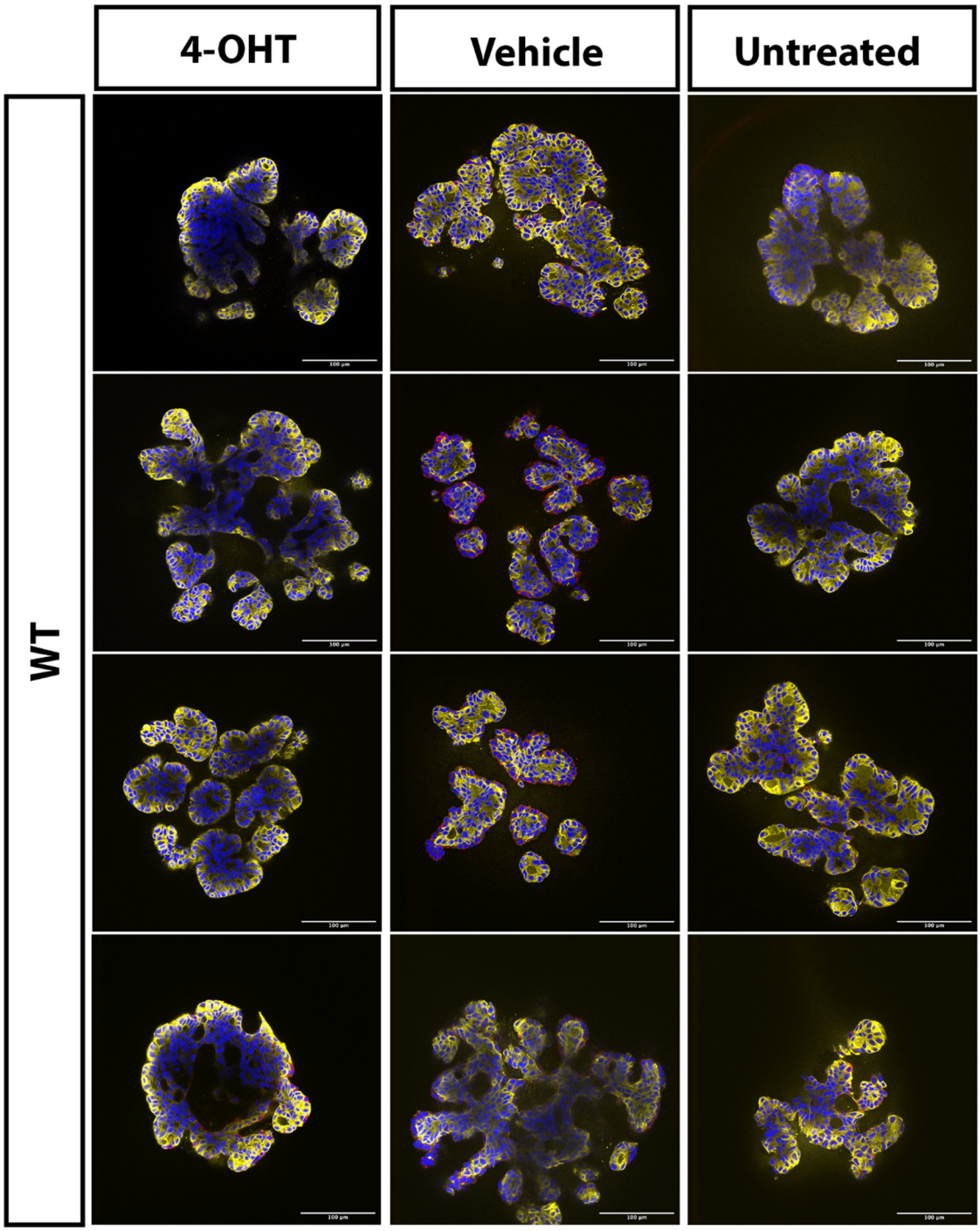


Figure 5.12: Fluorescence microscopy of whole mount WT mammary luminal organoids exposed to 4-OHT vehicle control and left untreated.

Organoids cultivated from luminal progenitor cells of WT mice exposed to 4-OHT or ethanol vehicle control after 7 days in culture and then cultivated for a further 7 days or left untreated for 14 days. Organoids were stained with luminal marker (Keratin 8 in yellow) a basal marker (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

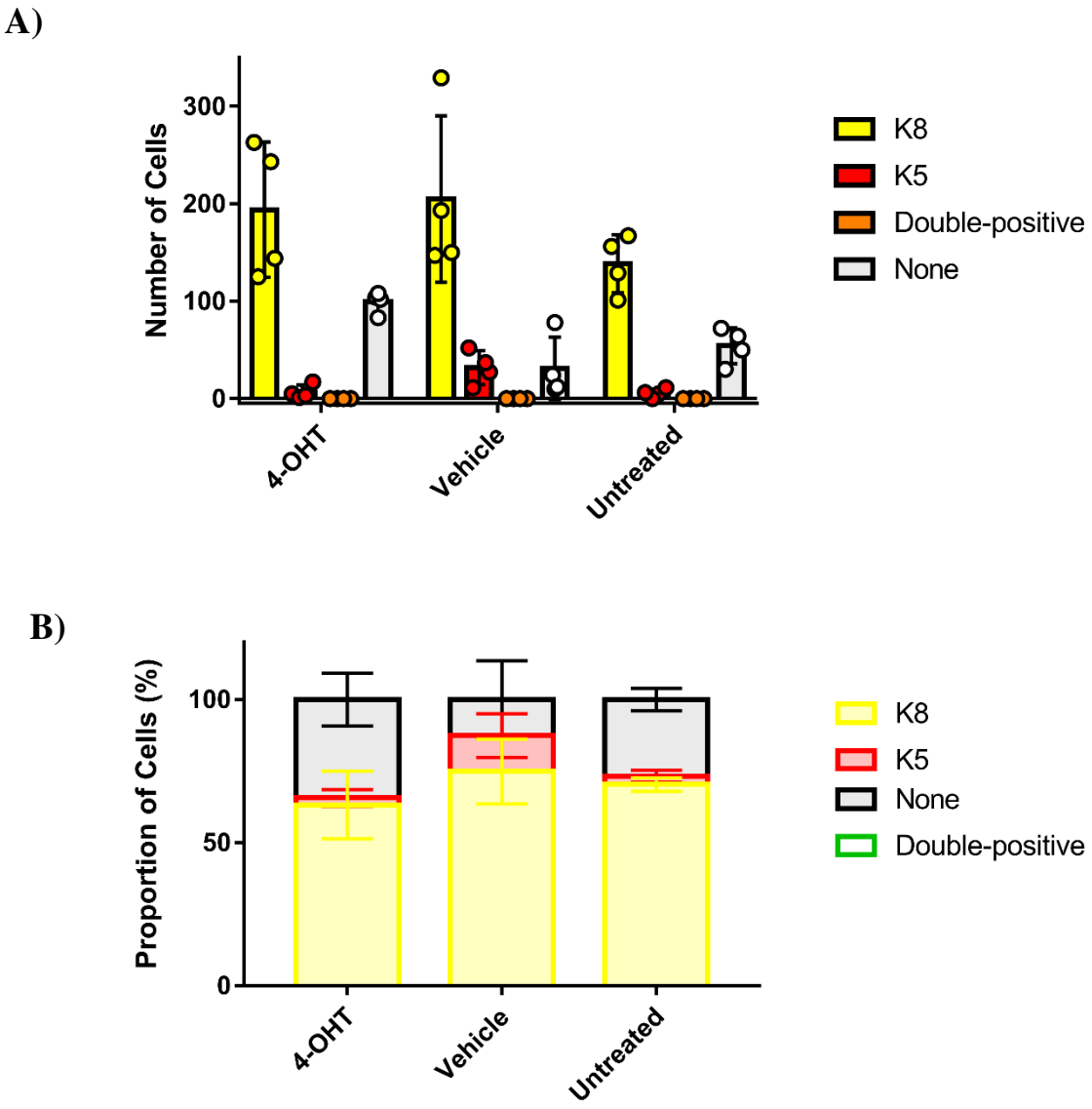


Figure 5.13: Quantification of WT mammary organoid microscope images.

Cells from the organoids displayed in figure 5.12 were counted and scored as either Keratin 8 or Keratin 5 positive. Cells with no marker or indistinguishable in their identity were scored under 'none'. **A)** Depicts the mean number of cells displaying each marker ± S.D. of n=4 technical replicates. **B)** Data from (A) presented as the mean proportion of cells positive for each marker in the organoids from each treatment group ± S.D. of n=4 technical replicates.

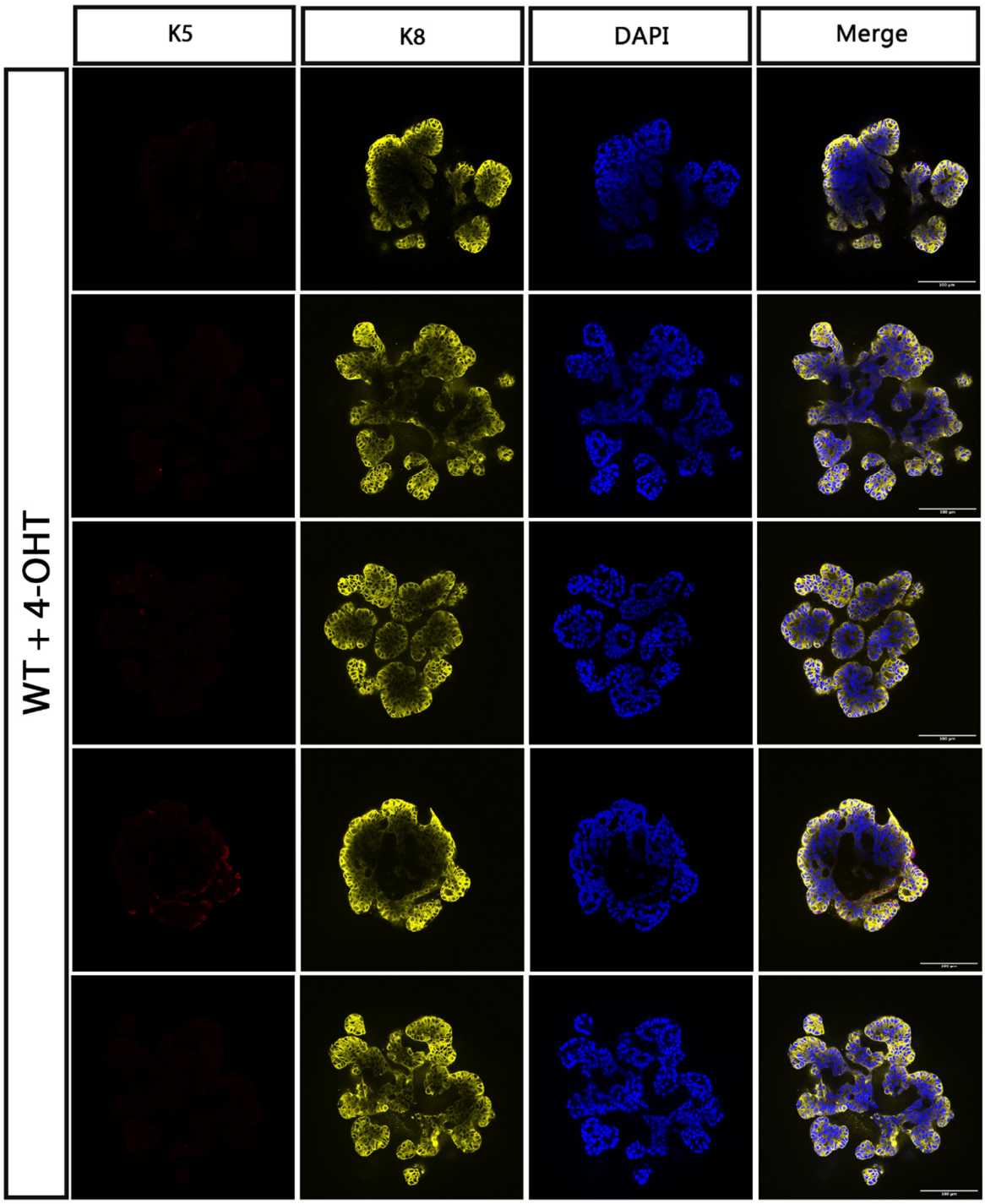


Figure 5.14: Fluorescence microscope images of whole mount WT organoids treated with 4-OHT.

Organoids were treated with 4-OHT (see figure 5.12 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 µm. Images were obtained using a confocal microscope with one optical z plane displayed.

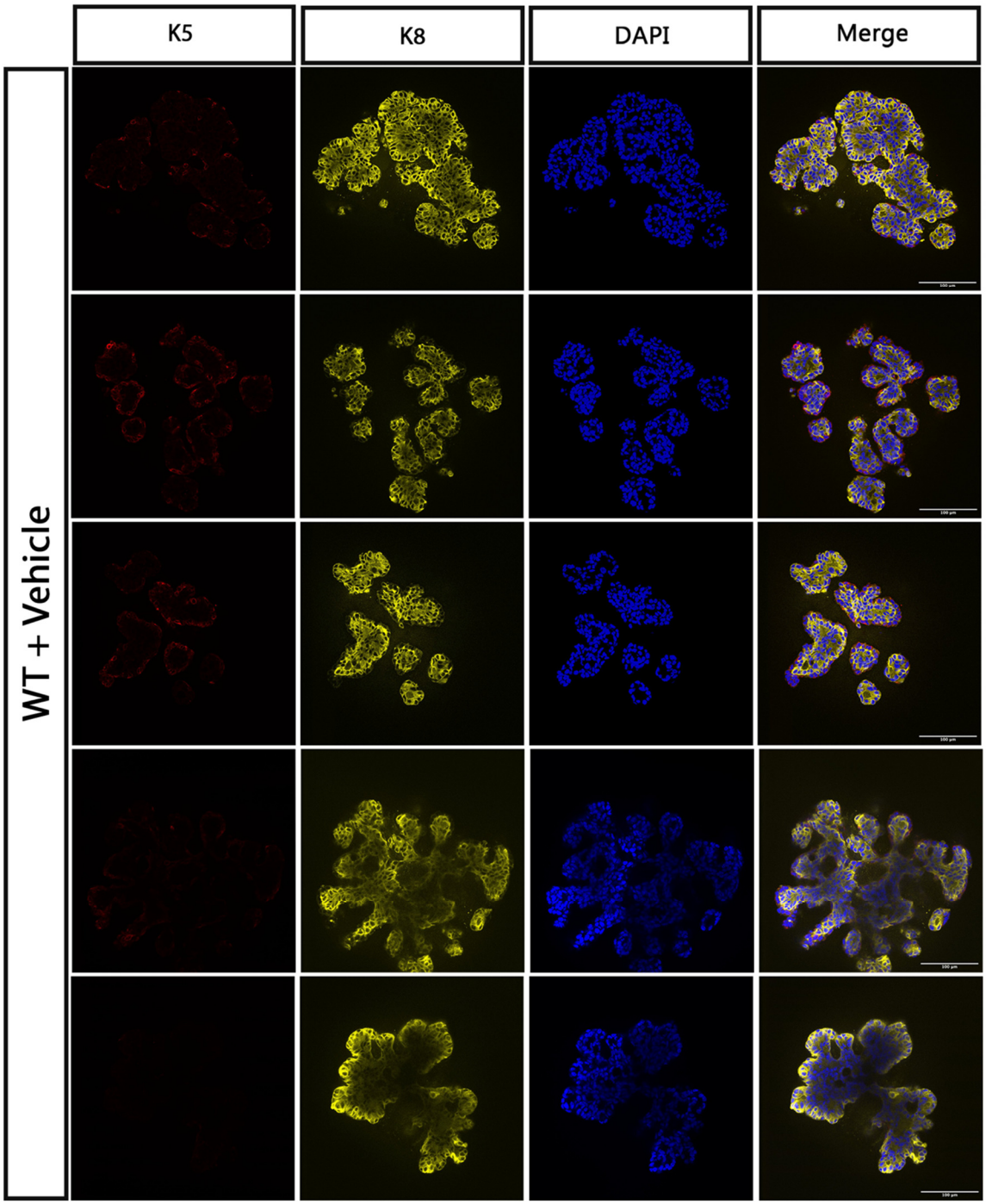


Figure 5.15: Fluorescence microscope images of whole mount WT organoids treated with vehicle.

Figure 5.15 (cont): Fluorescence microscope images of whole mount WT organoids treated with vehicle.

Organoids were treated with ethanol vehicle control (see figure 5.12 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

Next, luminal progenitor cells were cultivated from *BCL11A*^{OE} mice. After 7 days in culture, organoids were subjected to the same induction protocol: treated with 4-OHT, vehicle control or left untreated. At 7 days post induction organoids were analysed using bright field microscopy to identify any immediately visible morphological differences. Broadly, organoids did not appear to display large differences between conditions, as shown in representative wells (figure 5.16). Next, the size of organoids was quantified across all wells and each condition. The average organoid size did not tend to be different between conditions (figure 5.16b), suggesting that *BCL11A* overexpression has no effect on the size of the organoids. Closer inspection of the organoids confirmed that structures were consistent between conditions with any variations in budding morphology in line with the normal heterogeneity observed in mammary organoids (figure 5.17a). Organoids across all conditions display normal budding morphology consistent with control, suggesting that 4-OHT induction had little effect on the organoids.

Part of the *BCL11A*^{OE} construct contains *e*GFP protein under an *IRES* (see figure 2.1). Recombined cells were expected to express the *eGFP* fluorescence marker. To assess whether this was the case, organoids were analysed for this marker, however no fluorescence was visible (figure 5.17b). This was an unexpected result that suggested 4-OHT induction was apparently insufficient to induce recombination, evident when compared to the extent of *e*GFP expression observed in the same experiment performed with organoids of oesophageal origin (See figure 3.13). To assess whether this was the case, and to confirm whether *BCL11A* was over expressed, gene expression was assessed using RT-qPCR analysis. Analysis, when normalising to an internal control (*Gapdh* expression) revealed an increase in *BCL11A* expression (figure 5.18). Surprisingly the *eGFP* expression is visible in cells from other tissues (see chapter 3). This suggests that the *IRES* mediated *eGFP* expression is tissue specific or that the 4-OHT mediated induction in the mammary cells need to be optimised. Coupled with a lack of difference in organoid size between conditions, this could suggest that the overexpression of *BCL11A* to the extent demonstrated using RT-qPCR may not have an effect on the organoids. However, the lack of *eGFP* expression in these cells did suggest that expression had perhaps not reached a threshold level to induce phenotypic changes. Given that over expression only demonstrated a 13-fold increase here as compared to over 100-fold in the oesophageal organoids it suggests that tissue specific induction protocols are a necessary consideration for future work. Nonetheless, organoids were collected and subjected to whole

mount analysis to understand whether this level of overexpression had any effect on cellular composition or differentiation.

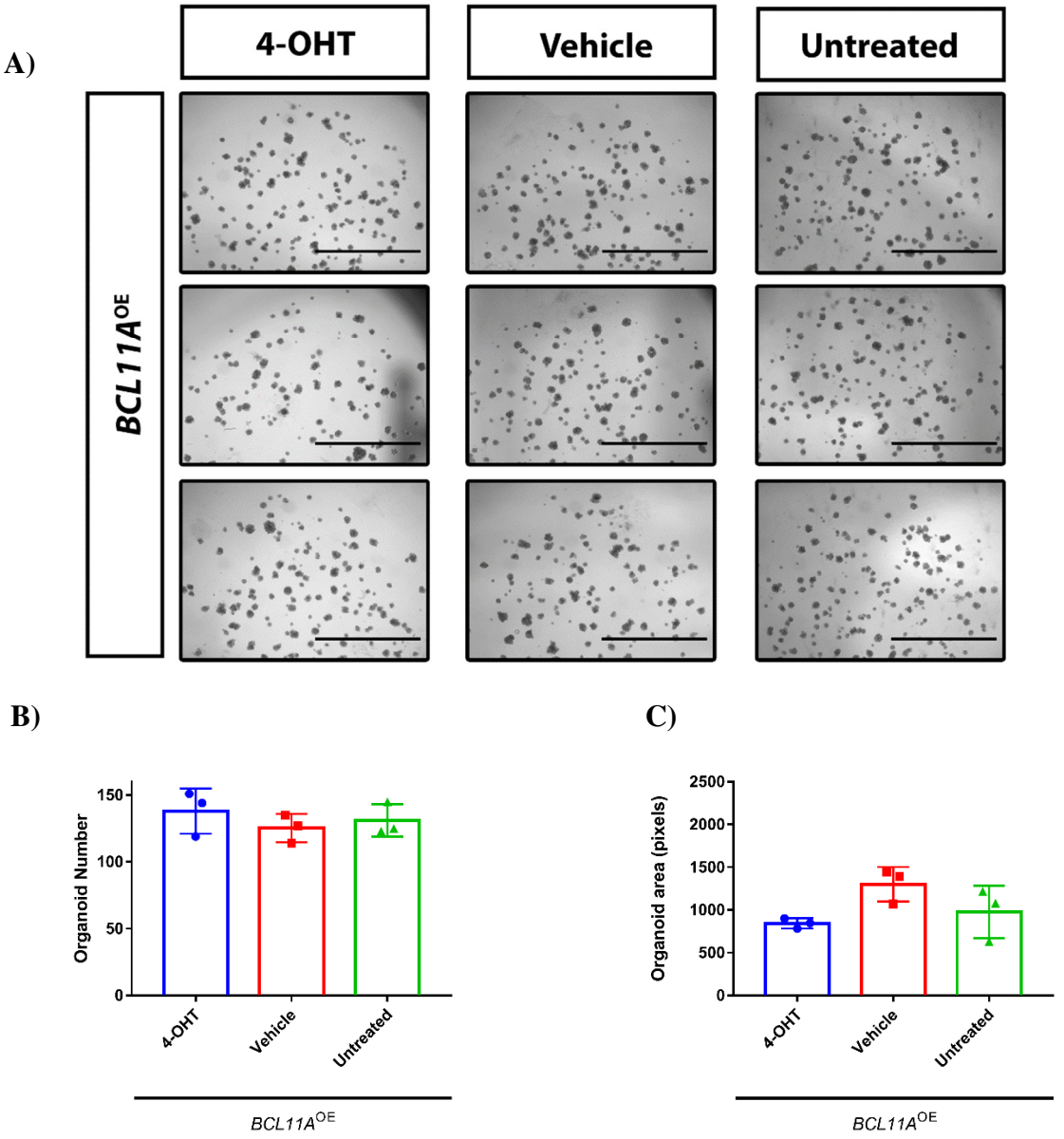


Figure 5.16: Brightfield view of *BCL11A*^{OE} mammary luminal organoids and quantification of organoid numnber an size after 14 days in culture.

Data shown are representative wells of one biological experiment. Organoids were cultivated from luminal progenitor cells of *BCL11A*^{OE} mice for 7 days. Organoids were then exposed to 4-OHT or ethanol vehicle control and cultivated for a further 7 days or left untreated for the duration. **A)** Data shows representative wells of organoids from each condition. Scale bars represent 2000 µm. **B)** shows the number of organoids derived from 2000 cells per well. **C)** shows the mean size of organoids per well in pixels. Data are represented as mean ± S.D. of n=3 technical replicates.

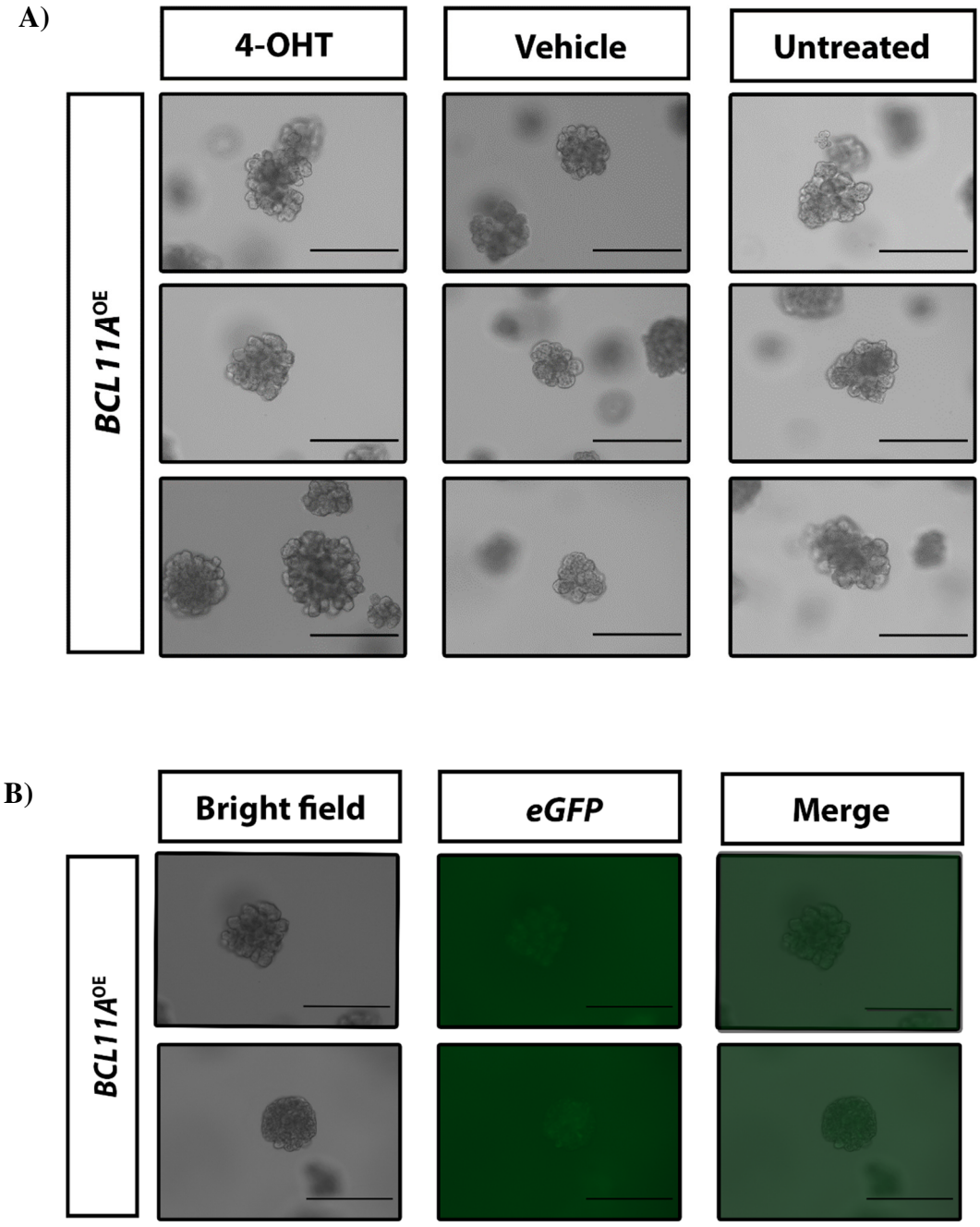


Figure 5.17: Brightfield view of individual *BCL11A*^{OE} luminal mammary organoids cultivated for 14 days.

A) Individual organoids cultivated from luminal progenitor cells for 7 days and exposed to 4-OHT, ethanol vehicle control or left untreated and cultivated for a further 7 days. Scale bar represents 200 µm **B)** Representative images of organoids treated with 4-OHT and visualised through an GFP channel. Scale bar represents 200 µm.

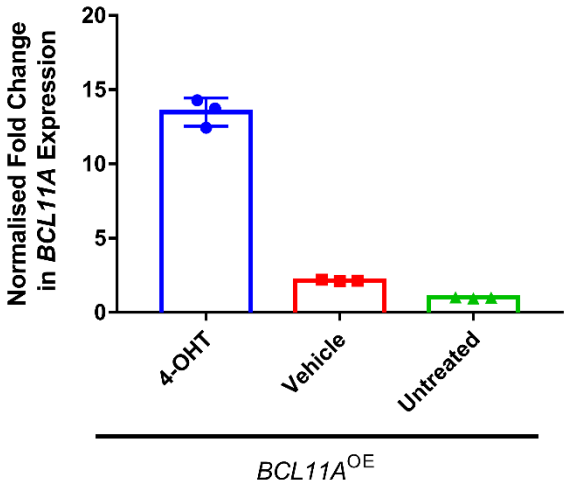


Figure 5.18: Analysis of *BCL11A* expression in luminal mammary organoids cultivated from *BCL11A*^{OE} mice.

After 14 days, organoids were collected and RNA isolated. After reverse transcription, the relative abundance of cDNA for the *BCL11A* gene was quantified using qPCR. Shown is the expression of *BCL11A* relative to that of *Gapdh* for organoids treated with 4-OHT, ethanol vehicle control or left untreated (by the same protocol as in figure 5.16). Data represent multiple technical replicates per treatment condition.

Whole mount analysis of organoids cultivated from *BCL11A*^{OE} mice did not reveal any distinct differences in cellular composition between the organoids of the experimental group versus controls (figure 5.19). Between the groups organoids appeared to be largely similar in cellular composition and broader appearance. Structures were predominantly composed of Keratin 8 positive cells with some Keratin 5 positive cells visible. Interestingly, multiple lumens per organoid were visible, with some organoids displaying intricate internal structures (figure 5.19). Similar to previous staining, mammary organoids appeared to have a weaker antibody signal towards the centre as well as distorted regions around the outside of the structures. Analysis of the cellular composition of these organoids from microscope pictures did not reveal any large differences across conditions (figure 5.20). Organoids were fairly uniformly composed mainly of Keratin 8 positive cells as well as cells that were not positive for any markers (figure 5.20). Closer analysis of the 4-OHT treated organoids revealed this in greater detail (figure 5.21), organoids contained very few Keratin 5 positive cells as well as multiple lumens. This was similar to that observed in the control groups (figure 5.22 and figure 5.23). Taken together, this suggested that the level of gene induction did not have any effect on the phenotypes or cellular composition of luminal organoids. *BCL11A* overexpression after 7 days of organoid culture therefore had little effect on either the size or the cellular composition of luminal mammary organoids. It remains to be seen if *BCL11A* overexpression at the point of seeding would have an impact on organoid phenotypes.

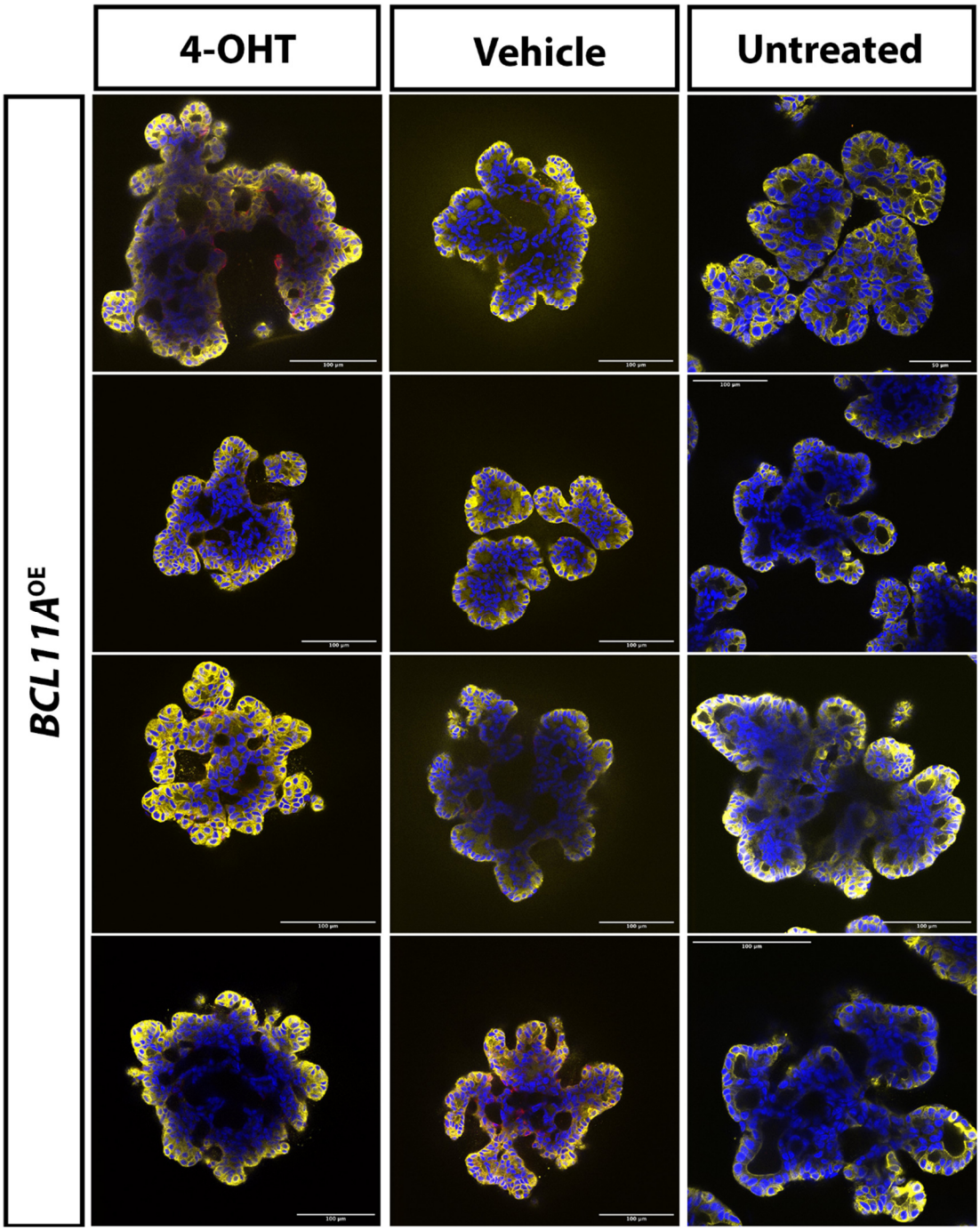


Figure 5.19: Fluorescence microscopy of whole mount *BCL11A*^{OE} luminal mammary organoids.

Figure 5.19 (cont.): Fluorescence microscopy of whole mount *BCL11A*^{OE} luminal mammary organoids

Data representative of one biological experiment. Luminal mammary organoids cultivated from *BCL11A*^{OE} mice were cultured for 7 days and treated with 4-OHT, vehicle control or left untreated. Organoids were cultured for a further 7 days and taken for whole mount analysis. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

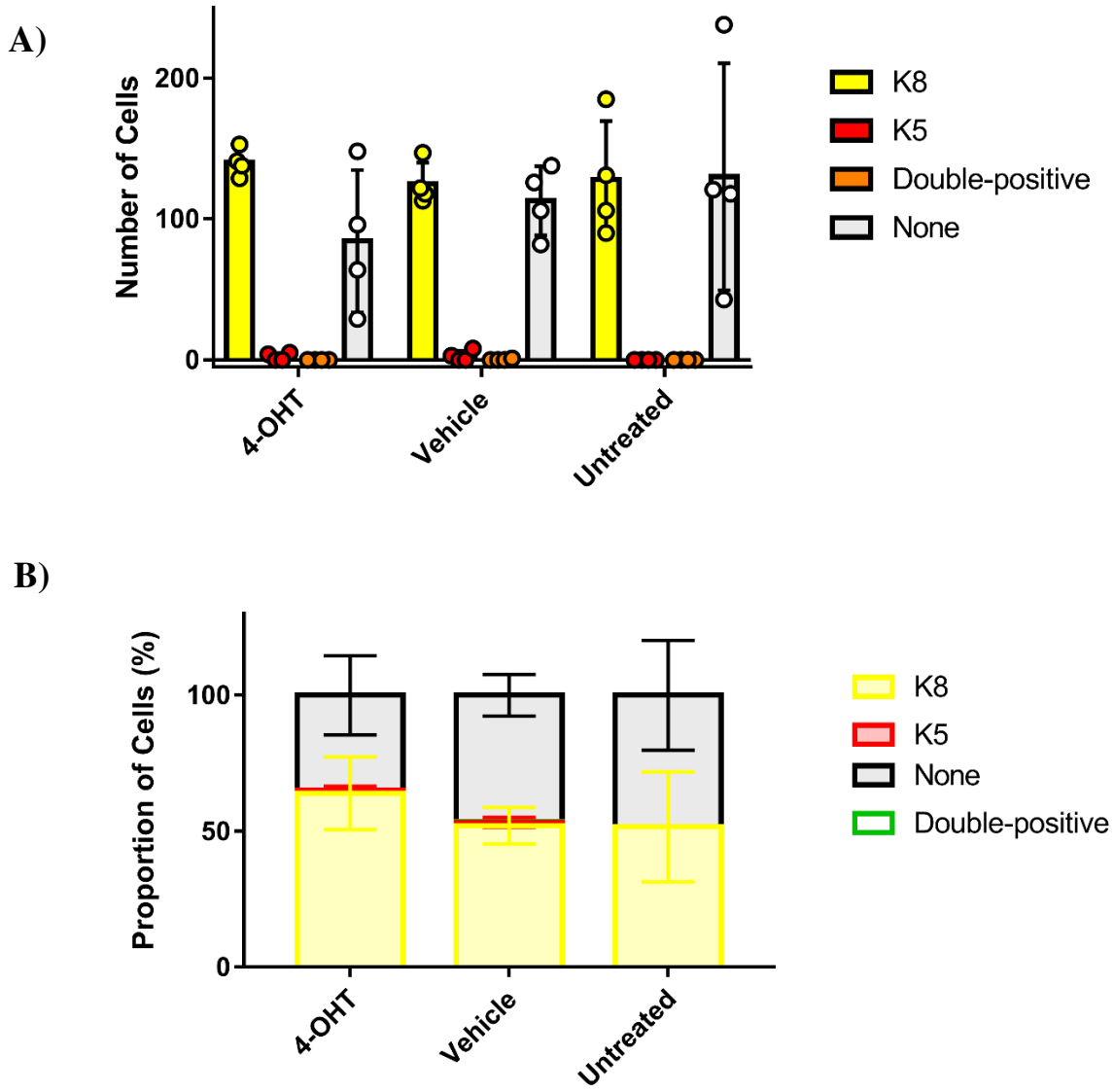


Figure 5.20: Quantification of *BCL11A*^{OE} mammary organoid microscope images.

Cells from the organoids displayed in figure 5.19 were counted and scored as either Keratin 8 or Keratin 5 positive. Cells with no marker or indistinguishable in their identity were scored under 'none'. **A)** Depicts the number of cells displaying each marker ± S.D. of n=4 technical replicates. **B)** Data from (A) presented as a proportion of cells positive for each marker in the organoids from each treatment group ± S.D. of n=4 technical replicates.

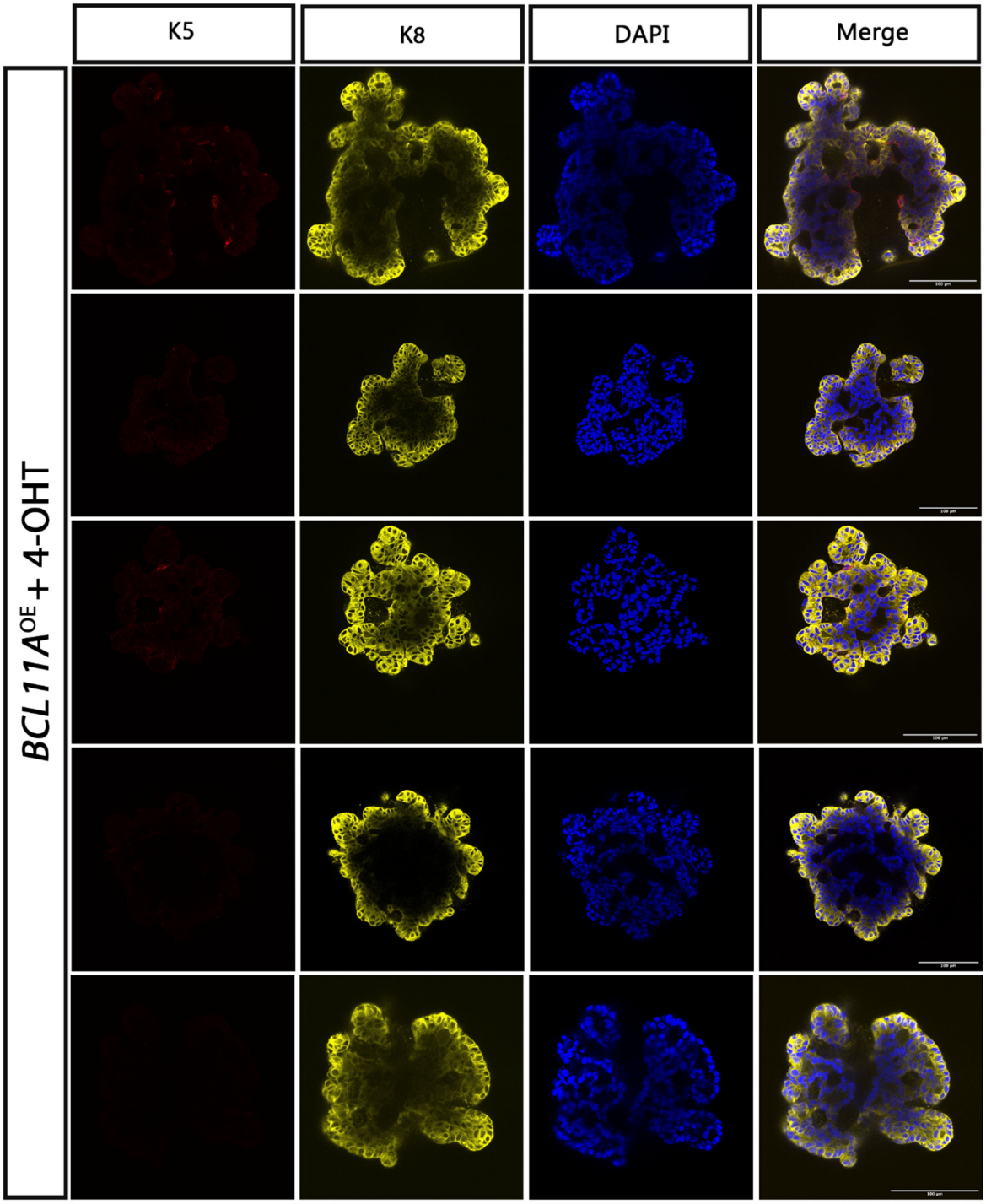


Figure 5.21: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with 4-OHT.

Figure 5.21 (cont): Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with 4-OHT.

Organoids were treated with 4-OHT (see figure 5.16 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

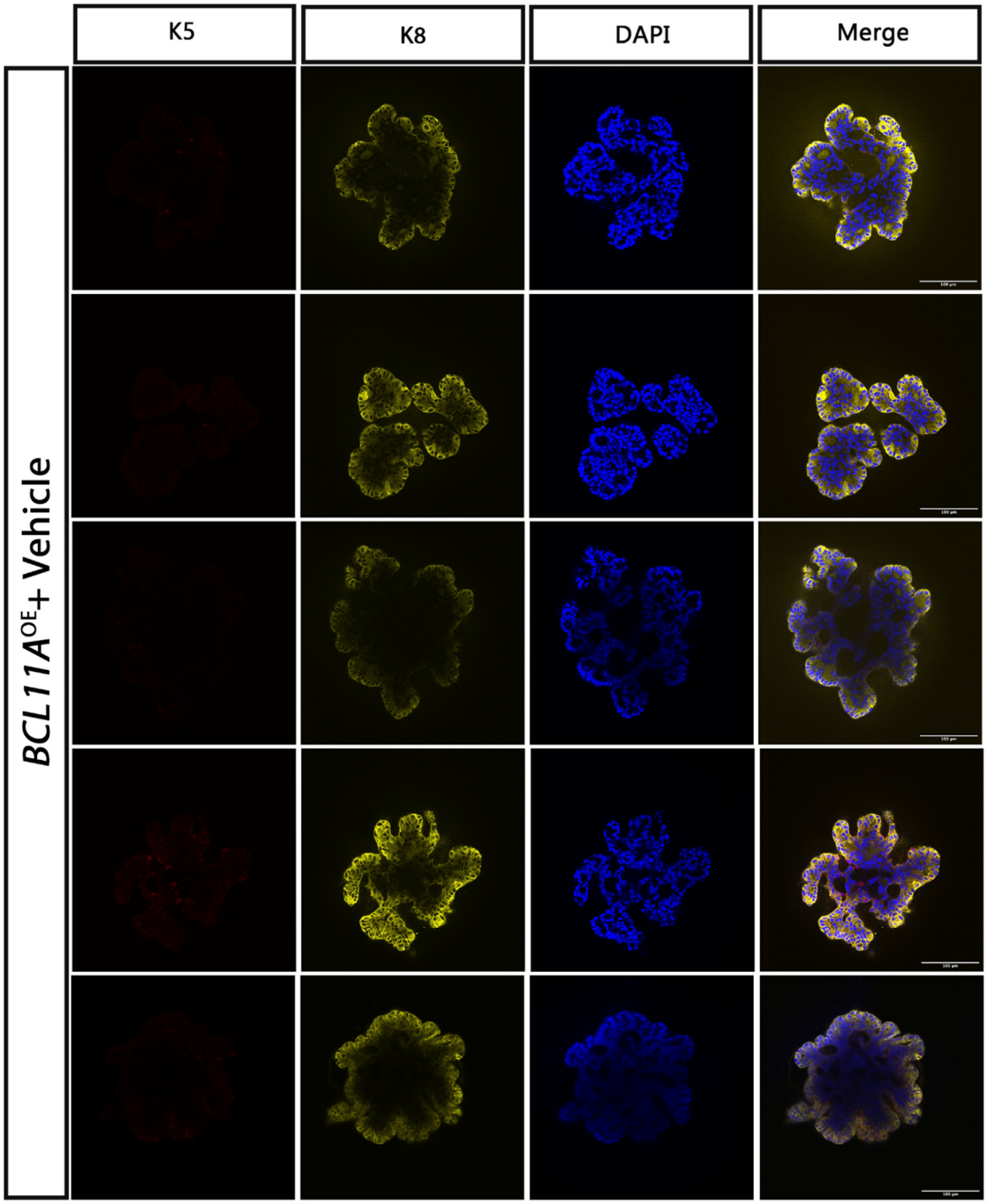


Figure 5.22: Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with vehicle.

Figure 5.22 (cont): Fluorescence microscope images of whole mount *BCL11A*^{OE} organoids treated with vehicle.

Organoids were treated with vehicle control (see figure 5.16 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

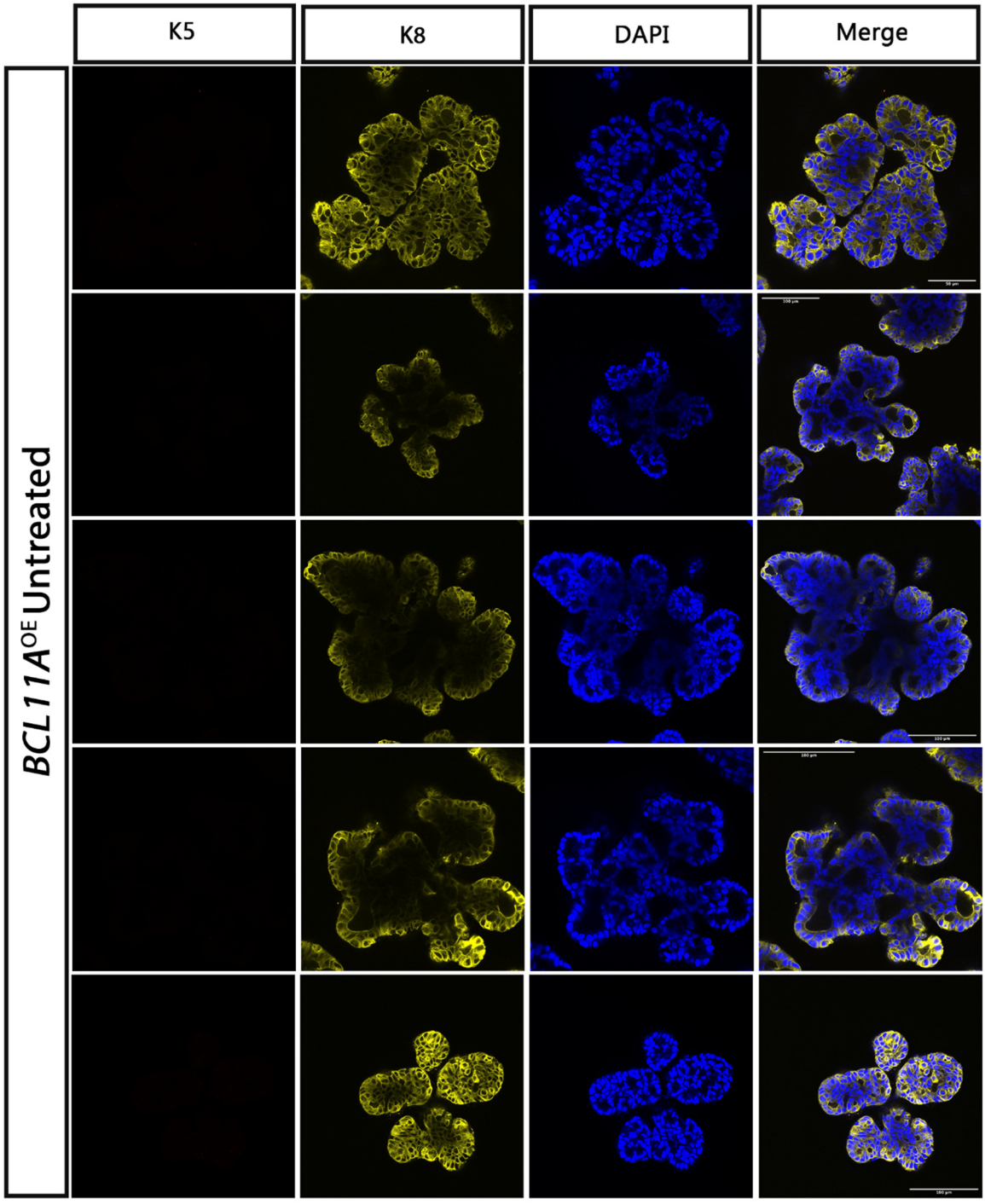


Figure 5.23: Fluorescence microscope images of untreated whole mount *BCL11A*^{OE} organoids.

Figure 5.23 (cont): Fluorescence microscope images of untreated whole mount *BCL11A*^{OE} organoids.

Organoids were left untreated and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

5.2.3 Modelling *Bcl11a* deletion in Mammary Organoids

To assess the effect of *Bcl11a* deletion on luminal organoids from the mammary gland, organoids were cultivated from luminal progenitor cells of *Bcl11a*^{KO} mice. Similar to the previous condition, knock out of *Bcl11a* was induced through *in vitro* 4-OHT exposure after 7 days in culture. Organoids were treated with either 4-OHT or ethanol or left untreated and then cultivated for a further 7 days. After 14 days in culture organoids were analysed for phenotypic differences including colony number and sized based analysis. Analysis did not reveal differences in organoid size or the number of colonies in each condition (figure 5.24). Closer analysis of 4-OHT treated organoids did not reveal differences in the organoids cultivated in the different conditions (figure 5.24 and 5.25). Organoids appeared to be morphologically similar between conditions, each demonstrating similar size and branching to control organoids (figure 5.24b and c). This suggested that *Bcl11a* deletion had no effect on the overall structural integrity of luminal mammary organoids. To ensure that recombination had occurred and *Bcl11a* had indeed been deleted, RNA was collected for gene expression analysis using RT-qPCR. Analysis revealed a low level of *Bcl11a* expression in both 4-OHT treated and also in the vehicle control group. Such a result would suggest that the qPCR had failed, potentially due to the low quality and concentration of RNA extracted from the mammary organoids (data not shown). Due to a greater level of relative abundance in the cell, *Gapdh* levels may have artificially appeared higher than *Bcl11a* cDNA levels. This could generate a false knock down effect when the data was normalised to this internal control, thus proving that the data was inconclusive.

Despite this result however, organoids were collected for confocal analysis to assess whether *Bcl11a* deletion induced any alteration in cellular composition. Organoids appeared to display similar morphology across all the conditions (figure 5.26), again demonstrating structures composed mainly of Keratin 8 positive cells with some evidence of Keratin 5 positive cells located around the outermost layers of the organoids. Organoids within the control groups did not display this phenotype (figure 5.29 and 5.30). Quantification of the microscope images did identify a small number of Keratin 5 cells in the 4-OHT treated organoids (figure 5.27). This could suggest that *Bcl11a* knock down led to an expansion of the basal population within the organoids, suggesting that this transcription factor could have a role in cellular differentiation in the mammary gland. Taken together with work from Khaled et al, to show the expression of *Bcl11a* in luminal cells of the mammary gland, this result could suggest that this transcription

factor is responsible for the maintenance of the luminal compartment as its knock down can be seen to lead to an expansion of Keratin 5 positive cells. However, this experiment represents only one biological replicate and a knock out could not be confirmed via qPCR, so these experiments should be repeated to make more robust conclusions.

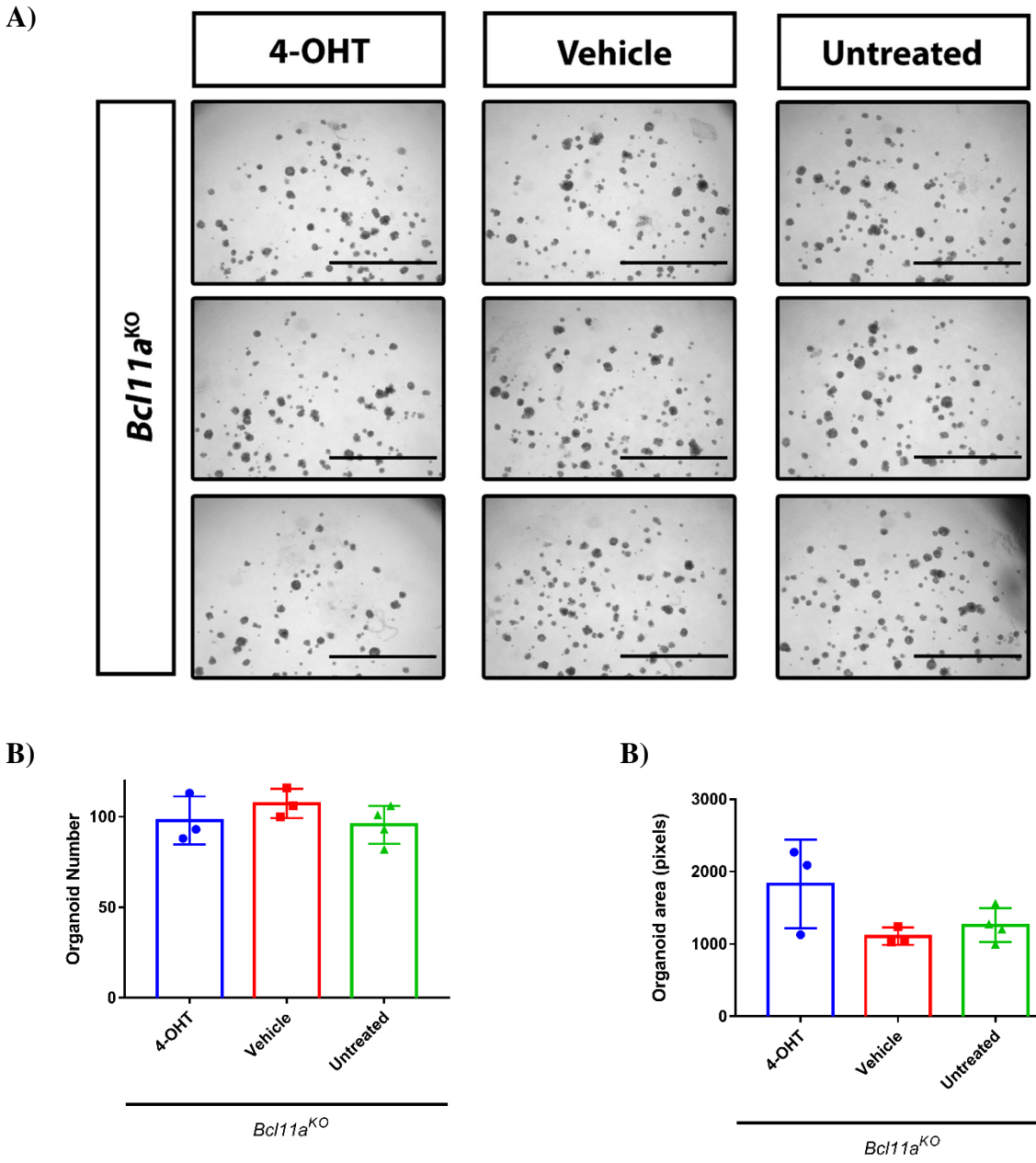


Figure 5.24:Representative brightfield view of *Bcl11a*^{KO} mammary organoids and quantification of organoid number and size after 14 days in culture.

Luminal mammary organoids were cultivated from *Bcl11a*^{KO} mice for 7 days and exposed to either 4-OHT or ethanol vehicle control, or left untreated and cultivated for a further 7 days. Data are representative of one biological replicate. **A)** Representative wells from each treatment condition Scale bar indicates 2000 µm. **B)** shows the number of organoids derived from 2000 cells per well. **C)** shows the mean size of organoids in pixels per well. Data are represented as mean ± S.D. of n=3-4 technical replicates.

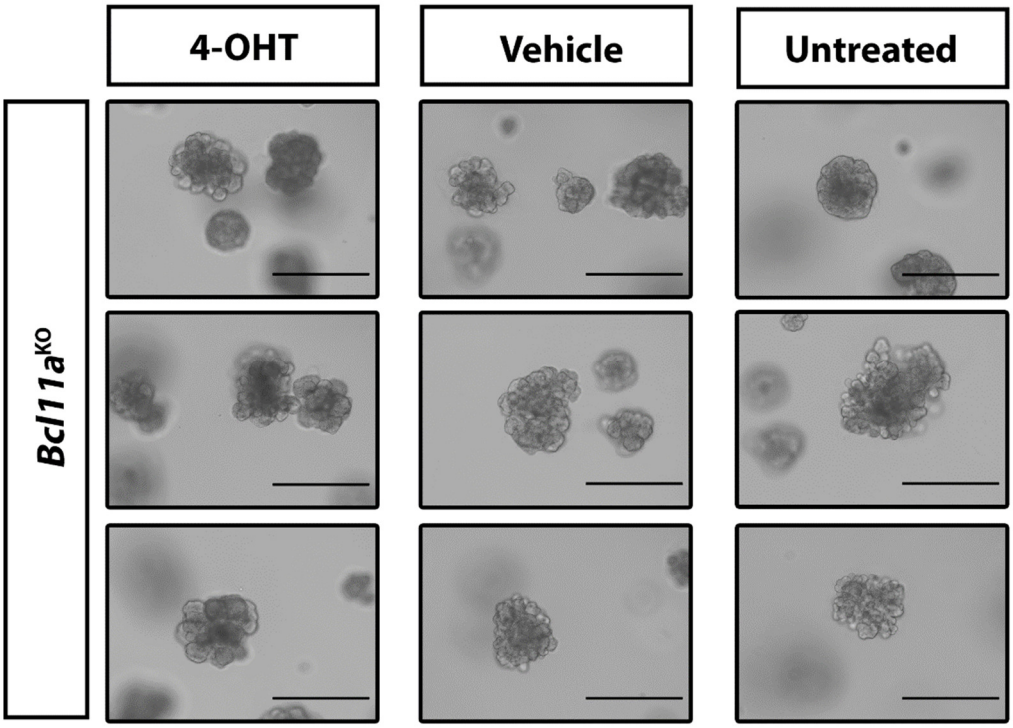


Figure 5.25: individual *Bcl11a*^{KO} mammary luminal organoids after 14 days in culture.

Organoids were cultivated from luminal progenitor cells of *Bcl11a*^{KO} mice for 14 days (see figure 5.24 for details). Scale bar indicates 400 µm.

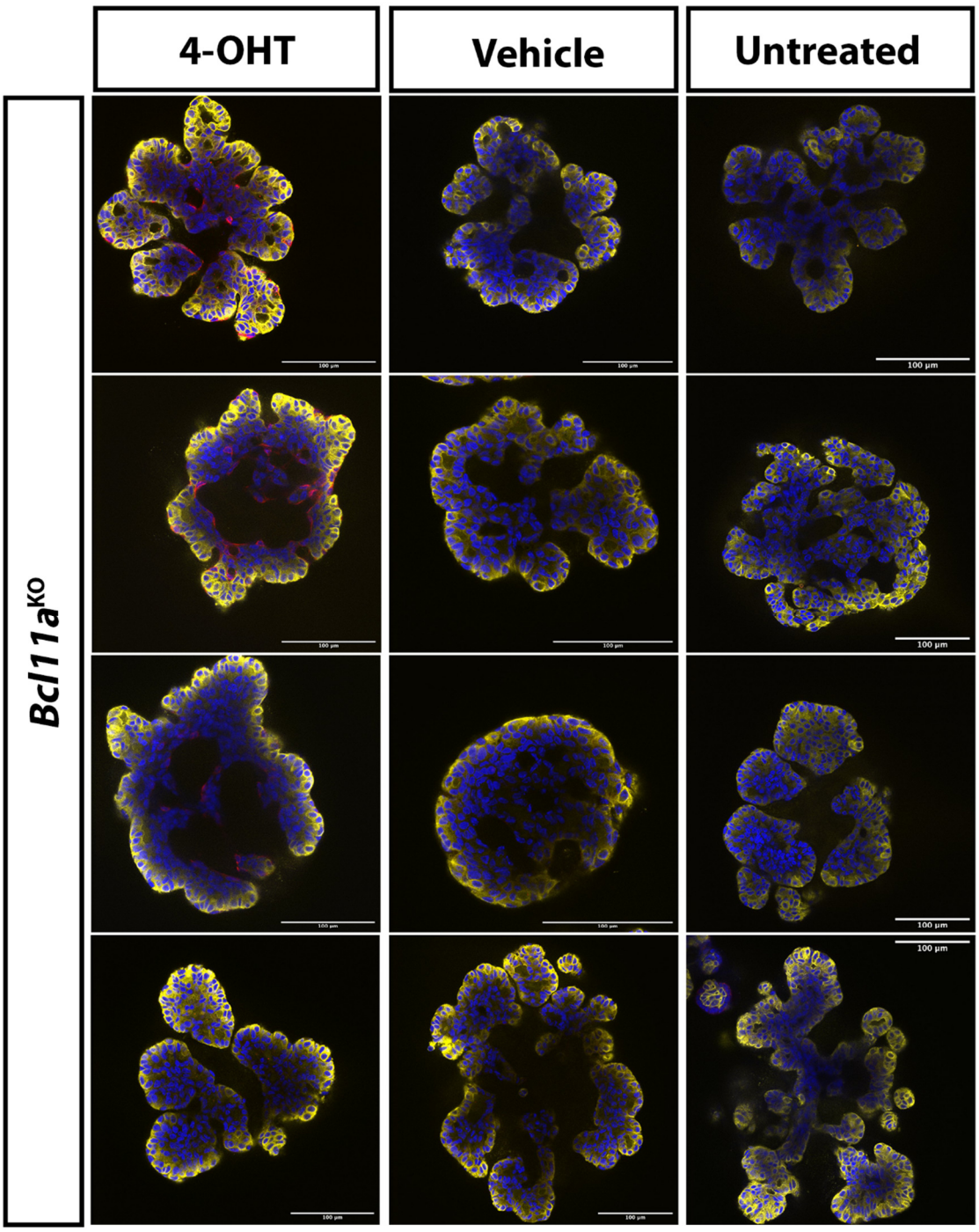
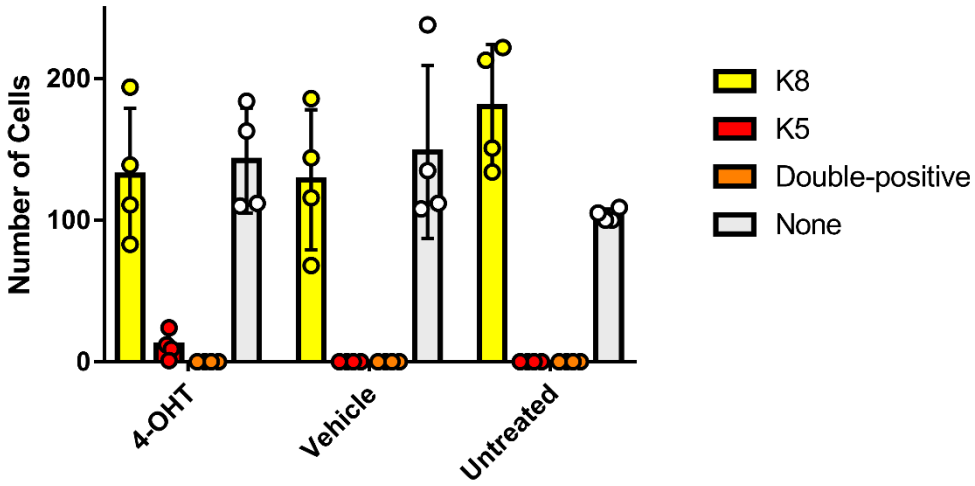


Figure 5.26: Fluorescence microscopy of whole mount *Bcl11a*^{KO} organoids.

Figure 5.26 (cont): Fluorescence microscopy of whole mount *Bcl11a*^{KO} organoids.

Luminal mammary organoids cultivated from *Bcl11a*^{KO} mice were cultured for 7 days and treated with 4-OHT or ethanol vehicle control or left untreated. Organoids were cultured for a further 7 days and taken for whole mount analysis. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 μm. Images were obtained using a confocal microscope with one optical z plane displayed.

A)



B)

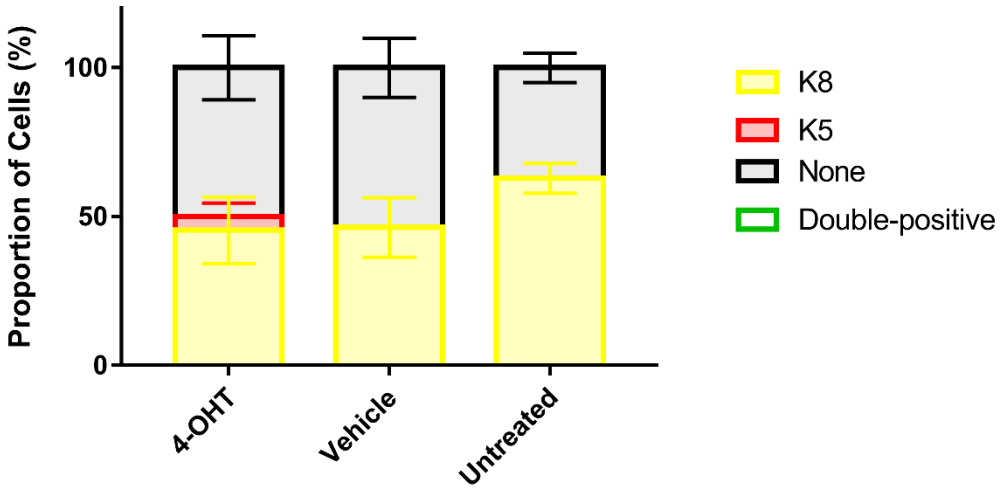


Figure 5.27: Quantification of *Bcl11a*^{KO} mammary organoid microscope images.

Cells from the organoids displayed in figure 5.19 were counted and scored as either Keratin 8 or Keratin 5 positive. Cells with no marker or indistinguishable in their identity were scored under 'none'. **A)** Depicts the number of cells displaying each marker ± S.D. of n=4 technical replicates. **B)** Data from (A) presented as a proportion of cells positive for each marker in the organoids from each treatment group ± S.D. of n=4 technical replicates.

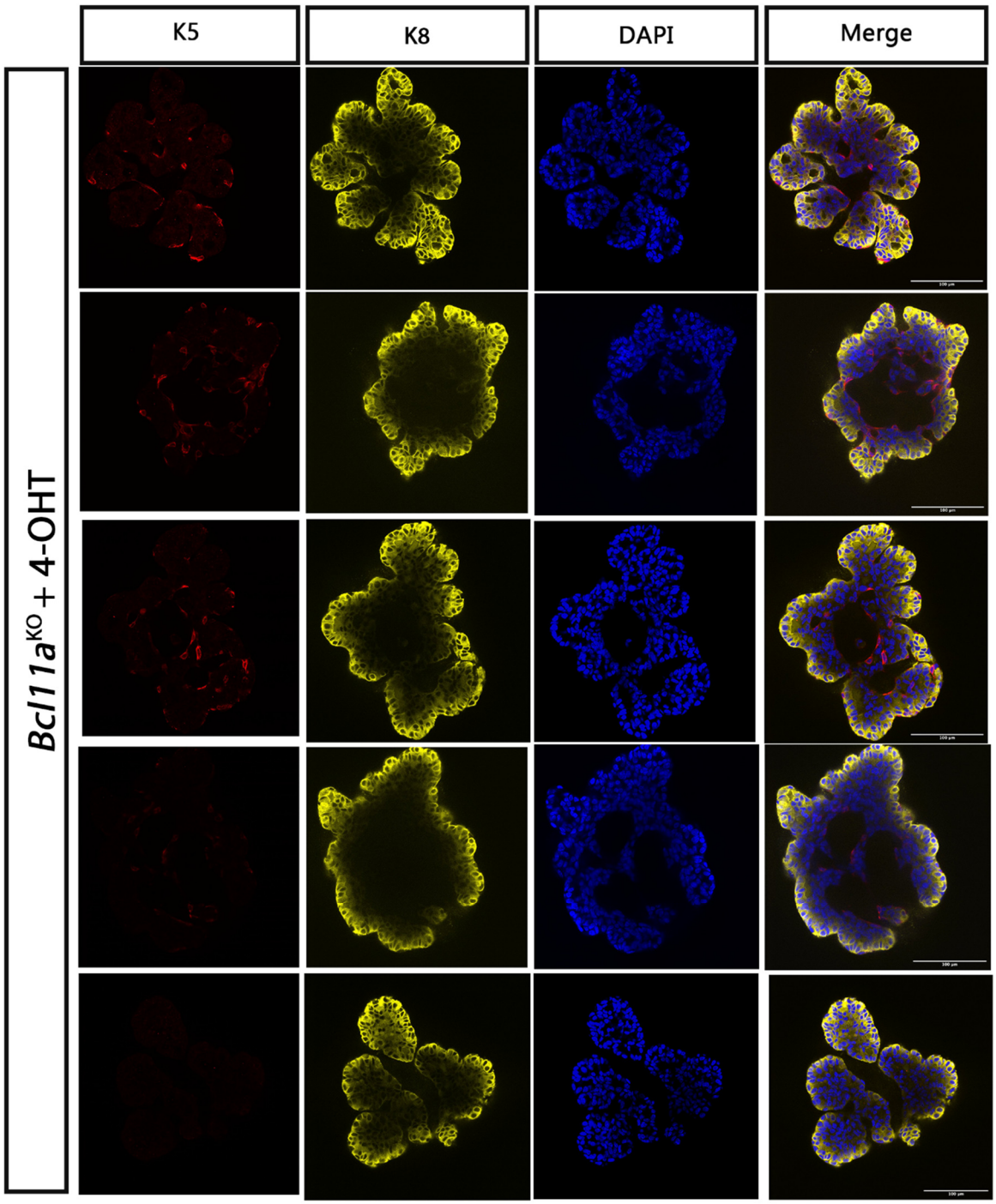


Figure 5.28: Fluorescence microscope images of whole mount *Bcl11a*^{KO} organoids treated with 4-OHT.

Organoids were treated with 4-OHT (see figure 5.23 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 µm. Images were obtained using a confocal microscope with one optical z plane displayed.

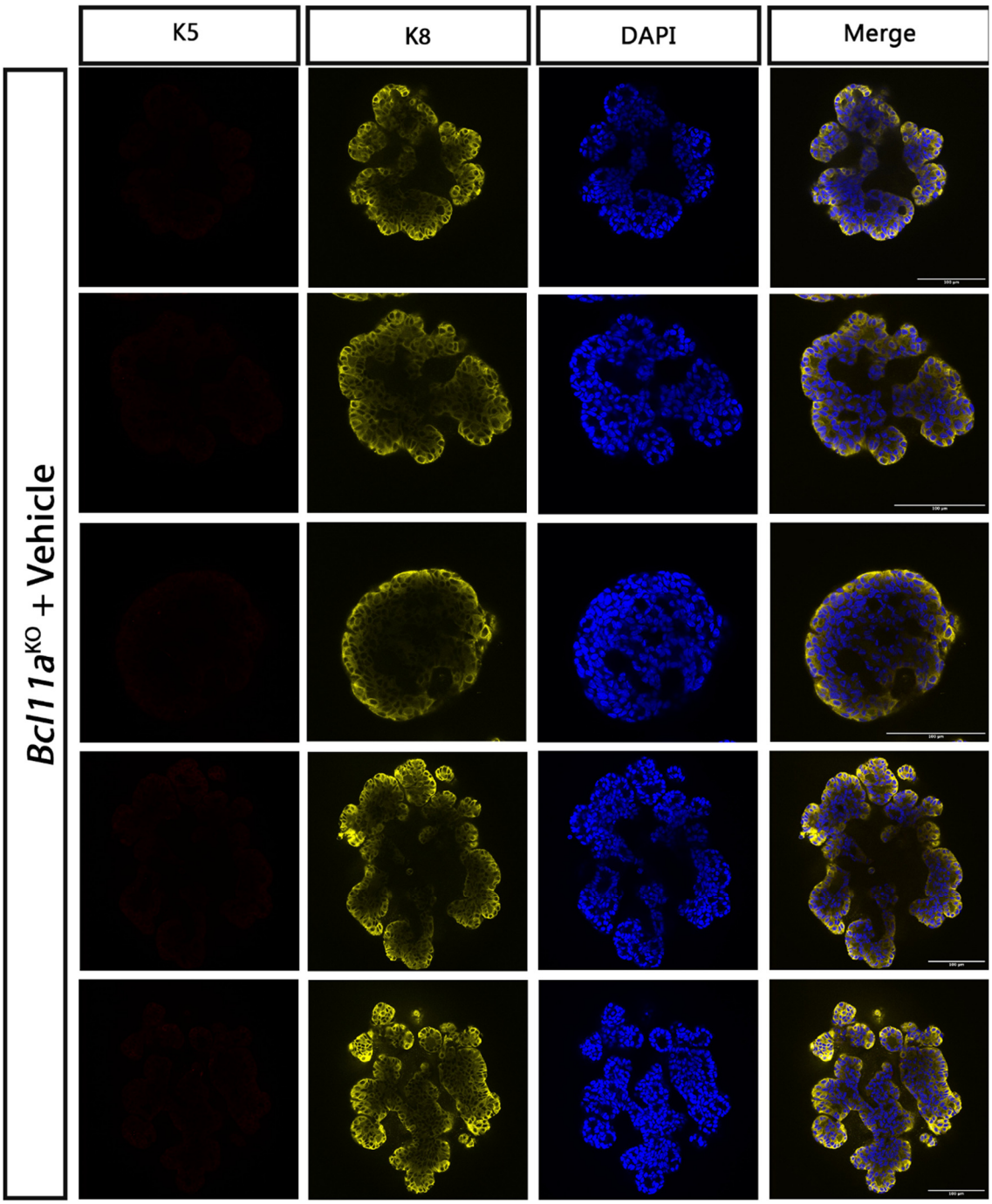


Figure 5.29: Fluorescence microscope images of whole mount *Bcl11a*^{KO} organoids treated with vehicle.

Organoids were treated with ethanol vehicle control (see figure 23 for details) and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain. Scale bar represents 100 µm. Images were obtained using a confocal microscope with one optical z plane displayed.

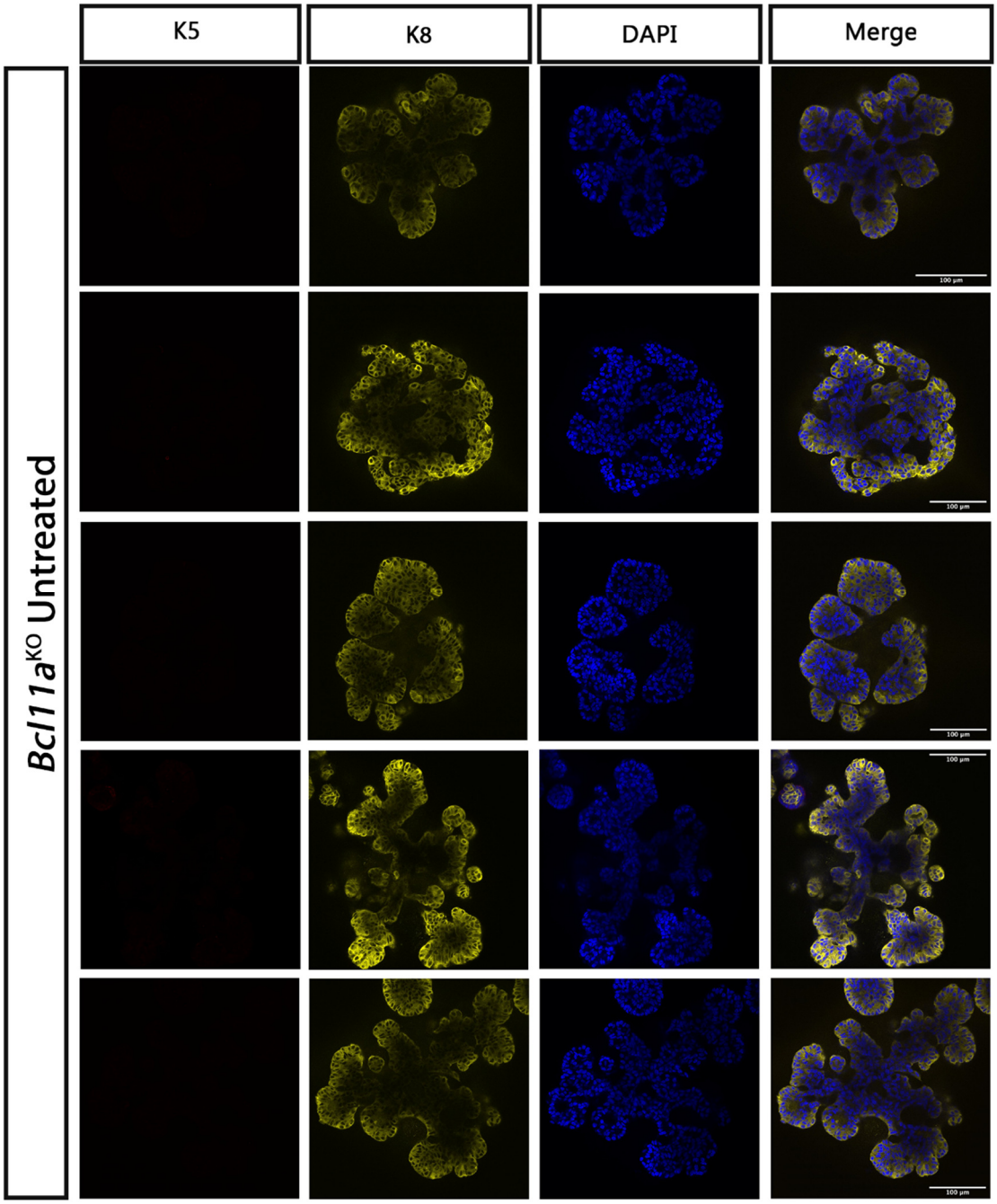


Figure 5.30: Fluorescence microscope images of untreated whole mount *Bcl11a*^{KO} organoids.

Organoids were left untreated and cultivated for a period of 14 days. Organoids were stained with a marker of luminal cells (Keratin 8 in yellow), a marker of basal cells (Keratin 5 in red) and DAPI nuclear stain in blue. Scale bar represents 100 µm. Images were obtained using a confocal microscope with one optical z plane displayed.

5.3 Discussion

Despite the prominence of *BCL11A* as a marker of TNBC, little is known about its wider function in the mammary gland. This project sought to understand more about *BCL11A* gene function by using a 3D organoid system to model gene expression changes. The first aim was to compare published organoid models of the mammary gland and select one for use in the project. Media compositions were adapted from Jardé et al. .and Jamieson et al. (medium formulation 1 and 2, respectively). This work in comparing organoid models highlights the variability between mammary organoid media compositions and their effect on luminal progenitor cells when cultured *in vitro*. Medium formulation 2 was ultimately used due to the superior colony forming capacity (8.7%) of luminal progenitor cells when cultivated in this medium. Organoids cultivated in this manner displayed budding morphology and, when subjected to whole mount analysis were seen to be composed mainly of Keratin 8 positive cells as well as multiple lumens.

Luminal progenitor cells were cultivated in this manner from *BCL11A*^{OE} and *Bcl11a*^{KO} mouse models, with the aim of exploring gene function. *BCL11A* overexpression appeared to have no effect on organoid size or number, although RT-qPCR analysis revealed an increase in gene expression. Interestingly, cells also did not appear to display *eGFP* fluorescence indicating a lack of expression. It could however be possible that this level of induction was insufficient to induce high enough levels of recombination that *eGFP* would be visible in the mammary organoids. Interestingly this concentration of 4-OHT was sufficient to induce *e*GFP expression in the same experiment on organoids of oesophageal origin. More importantly, the qPCR results from the oesophageal experiment demonstrated superior *BCL11A* over expression levels (over 100-fold, in. comparison to only round 13-fold in the mammary organoids). This highlights the need for further optimisation to the 4-OHT induction protocol that is specific to the needs of the mammary organoid culture. *Bcl11a* knock out organoids did appear to generate a small population of Keratin 5 positive cells. Such a result suggests there could be a role for the transcription factor in the maintenance of the luminal compartment of the mammary gland. However, the knock out could not be confirmed via qPCR, so it remains unclear whether such perturbations in *Bcl11a* gene expression have any phenotypic effects. In addition, single optical z planes were analysed to generate this data. Future experiments should acquire multiple z planes for each organoid to more comprehensively asses the proportions of Keratin 5 and Keratin 8 positive cells.

Taken together, this evidence suggests that further optimisation of the 4-OHT induction protocol for mammary organoids is necessary in order to reach a more conclusive result. Certainly, 4-OHT titration should be attempted to identify the concentration necessary for recombination in this setting. Indeed, it may have been the case here that only very few cells were targeted and so only some cells were recombined. Homogenisation across 4-OHT treated organoids for RNA extraction may therefore have diluted levels of expression and any phenotypes.

In addition, future work in this area should seek to perform experiments with a high number of biological replicates providing greater statistical power helping to provide more conclusive results. Multiple preparations of cells from different mice will be required to ensure any results are consistent and biological. In addition to replicates, future experiments, should seek to understand whether gene overexpression or knock out has a bearing on colony forming capacity of these cells as well as an expansion of basal cell population. Indeed, it could be that although overexpression does not induce molecular changes at late-stage induction, it may have an effect on the colony formation rate as demonstrated in tracheal organoids[3]. These experiments will also help to identify what role if any *Bcl11a* may have in the maintenance of the luminal compartment, particularly luminal progenitors, as taken together with lineage tracing experiments in the mammary gland, this work could suggest it plays a role in cellular identity.

More generally on the point of the mammary organoid culture, this work highlights some of the intrinsic variability of culturing these cells. Varying colony numbers and relatively low organoid forming efficiencies highlight the sensitivity these cells possess in contrast to highly adapted and robust organoid systems such as the intestine or oesophagus. The 3D culture of mammary organoids from single cells therefore still requires optimisation to be considered as robust as the 3D culture of other epithelial cells.

Chapter 6: Discussion

6.0 Discussion

One of the major hurdles remaining for cancer treatment is the lack of targeted therapy. The advent of NGS technologies have provided a wealth of knowledge to the cancer community, making available gene expression data for many cancer subtypes. Such data allows the identification of target markers to be studied and validated for therapeutic potential. What has also been notable from these efforts is the degree of genetic similarities between some epithelial cancers, particularly the clustering of squamous carcinomas of the lung and head and neck, with basal like breast cancers [1]. *BCL11A* has been identified as a triple negative breast cancer gene and a driver of lung squamous carcinoma [2,3] as well as a driver of laryngeal squamous carcinoma, a cancer of the neck. In this thesis, data presented indicate that *BCL11A* is a putative oncogene in oesophageal squamous carcinoma, further elevating its status as a prominent marker of squamous epithelial malignancies.

Key questions remain around the function of this transcription factor in the mammary gland and the oesophagus, indeed whether this gene has any homeostatic role in these tissues is of particular interest. Therefore, this project aimed to explore *BCL11A* gene function in these tissues using a combination of *in vitro* and *in vivo* techniques. Using a reverse genetics approach, gene function was explored using cells cultivated from models of BCL11A overexpression and knock out ($BCL11A^{OE}$ and $Bcl11a^{KO}$ respectively). Due to its high physiological relevance, and given its relative success [3] a 3D organoid approach was selected to model phenotypic changes induced by gene perturbations. In addition, *in vivo* studies of *Bcl11a* were achieved using a lineage tracing mouse model $Bcl11a^{CreERT2}$ in the oesophagus and in the mammary gland by Dr Sara Pensa.

6.1 *Bcl11a* Has a Homeostatic Role in the Oesophagus.

Organoid systems of the oesophagus displayed a high degree of variability between published models. These experiments demonstrate the importance of comparing models to ensure the most suitable system is used. Despite the initial variability however, the oesophageal organoid system ultimately proved to be efficient and reproducible across multiple biological replicates. A key advantage of the oesophageal organoid system is the clear display of a differentiation gradient. The organoid system chosen to be used in the project was that adapted from DeWard et al. [127]. Despite this however, there were some notable differences between the organoids

from the publication and those generated for this work. Firstly, some organoids failed to demonstrate dedicated Keratin 5 and Keratin 13 cells, instead cells were double positive for these markers. In addition, many cells remained unstained for either marker, where organoids cultivated by authors DeWard et al. did not display this phenomenon [127]. However, these characteristics did not preclude the use of the organoid system to generate useful data. They should instead be points for future experiments to address.

Organoids were cultivated from *BCL11A*^{OE} and *Bcl11a*^{KO} models and treated with 4-OHT to induce changes in gene expression. *BCL11A*^{OE} organoids demonstrated an induction of gene expression nearly ten times that of the mammary gland, as well as clearly displaying *eGFP* positivity. Over expressing organoids were found to be smaller than control groups and when analysed via whole mount immunofluorescence displayed striking phenotypes. Over expressing organoids demonstrated a clear lack of the differentiation seen in control groups. Some organoids did not contain a keratin core, the hallmark of cellular differentiation in the murine oesophagus. Passaging experiments revealed an increased colony forming capacity in cells over expressing *BCL11A*, similar to the findings reported in the lung [3]. Perhaps even more striking was the phenotype observed in *Bcl11a*^{KO} organoids treated with 4-OHT. These organoids demonstrated severe blebbing, in some cases losing structural integrity entirely. Organoids were smaller than control, consistent with an arrest in growth. Immunofluorescence analysis revealed enlarged nuclei in the cells of knock out organoids, consistent with an apoptotic phenotype. Passage of these cells revealed a diminished colony forming capacity, suggesting *Bcl11a* was essential for organoid initiation and colony formation. Although these experiments represented only single independent biological repeats, taken together this data suggests a progenitor role for *BCL11A* in the oesophageal epithelium. Indeed, this function does seem to fit with the wider biology and role of this transcription factor in other tissues [14,26]. It remains to be seen whether cells induced at the time of seeding display phenotypic outcomes as seen in the formation of lung organoids [3].

To validate this data and to understand whether *Bcl11a* did indeed have a progenitor role in the oesophageal epithelium, a lineage tracing model was used. Varying lengths of chase were implemented post tamoxifen exposure to uncover more about the dynamics of labelled cells. Labelled cells at 24 hours post exposure demonstrated improved colony forming capacity compared to non-labelled cells. These cells were seen to persist and expand in both layers of the oesophageal epithelium over a period of 12 weeks, culminating in a high degree of cells

positive for both Keratin 5 and tdTomato. This suggested that *Bcl11a* did indeed mark a progenitor population important for tissue maintenance, demonstrating a function for this gene in the homeostatic oesophagus with a progenitor function.

Taken in context with the current understanding of tissue maintenance within the field, these lineage tracing results do fit with the observation of expansion over time [96]. However, the extent of recombination observed in this study does not fit with that observed at lower doses of tamoxifen in other studies [95]. This may suggest that the tamoxifen dose is too high to label cells individually and will require titration for a more consistent approach.

6.2 Future Perspectives for Exploring Gene Function in the Oesophagus

This thesis has clearly demonstrated the use of organoids to study novel biology in the oesophageal epithelium, with many prospective avenues for future research. Following on from the data presented in this thesis, future experiments may want to focus on the optimisation of the organoid protocol to eliminate the presence of Keratin 5 and Keratin 13 double-positive cells. These cells were likely an artefact of culture, points to consider may be the length of time cells are maintained to ensure this is as short as possible. Further optimisation of the wholemount immunofluorescence protocol may also be necessary to address issues with antibody penetrance and unstained cells. Longer incubation with detergents such as Triton X-100 may be necessary to increase permeability.

Future experiments may want to repeat organoid seeding experiments with 4-OHT induction upon seeding. This experiment would provide insight into the necessity of *Bcl11a* in organoid establishment and whether over expression of this gene causes an increase in organoid size. Organoids derived from $BCL11A^{OE}$ and $Bcl11a^{KO}$ mouse models may also be stained with markers of apoptosis and proliferation to provide insight into gene function. The challenge of organoids derived from both models with trans-Retinoic acid is another experiment which builds on this work and could determine whether Bcl11a has a role in progenitor cell maintenance.

Quantitative lineage tracing in the manor described by the authors Doupé et al and Piedrafita et al [95,96] is key to answering remaining questions about the identity of *Bcl11a* expressing cells, thereby contributing to the growing body of work on cellular hierarchy and tissue maintenance in the oesophageal epithelium. It will be important to understand whether *Bcl11a* lineage tracing fits the paradigm of tissue maintenance recently described whereby a single progenitor population maintains the oesophageal epithelium [96]. In addition to this analysis, future experimental set up should also seek to include a further control animal. Collecting the oesophageal epithelium from a *Bcl11a*^{CreERT2} animal without induction would help ensure transgene specificity to tamoxifen induction in future lineage tracing experiments.

Tissue maintenance may also be studied using injury experiments in *Bcl11a*^{CreERT2} mice as exemplified by Doupé et al. The response of *Bcl11a* expressing cells to wound healing is vital to identifying its status as a marker of progenitor cells. As described by Doupé et al. progenitor cells quickly expand to heal the injury and restore equilibrium. If *Bcl11a* expression is indicative of a progenitor population then this should be represented by a high degree of label retaining daughter cells at the point of injury. Given the biological similarities between the oesophageal epithelium and the skin epithelium [96,133] it would be interesting to see whether *Bcl11a* plays a similar role in the skin. Using the same *Bcl11a*^{CreERT2} model, lineage tracing in the skin epithelium could easily address such questions.

Still questions remain around carcinogenesis in the oesophageal epithelium. To address the questions of gene function in the oesophageal epithelium a tumour model of oesophageal squamous carcinoma may be used [160]. These involve the administration of a chemical agent. when used in conjunction with the *Bcl11a*^{CreERT2} model will allow the study of the contribution of *Bcl11a* expressing cells to the tumour mass, helping to shed more light on the cellular origins of oesophageal squamous carcinoma. It would also be interesting to explore the *Bcl11a* Sox2 axis in this setting. Given the status of *SOX2* as an oncogene in oesophageal squamous carcinomas it could be that these proteins interact to drive tumour formation as has been identified in LUSC [3,43]. Such genetic similarities between these tumours suggests that mechanistic properties could be similar. Lazarus et al. explore regulation of *SETD8* by both *BCL11A* and *SOX2*, observing the sensitivity of LUSC cell lines to SETD8 inhibitors. Therefore, the inhibition of SETD8 in OSCC lines is an experiment that would easily address questions on shared molecular mechanisms. The implications of such a discovery would be

important for the biological understanding of squamous carcinoma tumour formation as well as for therapeutic intervention in patients with these tumours.

Indeed, the identification of *BCL11A* as a marker of carcinomas opens many questions about the mechanistic properties of this protein. *BCL11A* has been identified as a transcriptional repressor, that co localises with large co repressor and chromatin remodelling complexes such as the NuRD and SWI/SNF [161,162]. Such complexes enable the assembly of multi protein complexes containing structural chromatin binding and DNA histone modifying enzymes that are able to both supress transcription and modify chromatin. The function of these complexes is both context and subunit dependant suggesting a high level of promiscuity. This ultimately enables the intricate modulation to needed suit the transcriptional needs of complex organisms. Understanding associating proteins and the context specificity of will be greatly important in deducing the mechanistic properties of this protein. One way that associated proteins may be discovered is using Rapid immunoprecipitation of endogenous proteins (RIME) [163]. This method allows the pull down of a bait protein that is sub-linked to associated proteins. Proteins are then identified using mas spec to define interactors. This technique could be useful to identify *Bcl11a* interactors, ultimately shedding light on the mechanisms that lead malignancy in these sides [2,3].

6.3 *BCL11A* in the Mammary Gland

Initial work in comparing published models for the culture of mammary epithelial cells in 3D, highlighted the high degree of variability in these methods. Despite authors [147,148] demonstrating the cultivation of structures from single cells, in practice cultivation of single luminal progenitor cells in this manner did not prove robust. Using different media formulations, highly variable and inconsistent organoid forming efficiencies were observed across multiple biological replicates. This works to highlight the variability and sensitivity of luminal progenitor cells in culture, particularly when cultivated from single cells. This would suggest that there is still work to be done in technically optimising the 3D culture of mammary cells if they are to be considered as robust as the culture of other epithelial cells (eg. Intesitine, oesophagus and lung). Indeed, many of the more recent examples of this technique describe the derivation of cultures from tissue fragments rather than single cells. This method of culture appears more widely adopted and may therefore be a more robust approach to adopt for future experiments.

In this thesis, 3D organoid experiments were attempted to further understand the role of *BCL11A* in TNBC. Luminal progenitor cells isolated form the *BCL11A*^{OE} mice were seeded in vitro to establish organoids before being treated with 4-OHT to induce transgene expression. Quantitative analysis of organoid size did not reveal differences between conditions. In addition, no differences in cellular composition were observed between organoids using whole mount immunofluorescence staining. Despite RT-qPCR data showing an increase in BCL11A expression, *eGFP* was not visible in the organoids. Organoids cultivated in the same manner from luminal progenitor cells of *Bcl11a*^{KO} mice, showed some degree of variation in cellular composition, with a small expansion of Keratin 5 positive cells. This could suggest a role for *Bcl11a* in luminal cell maintenance. However, an inconclusive qPCR result meant that the knock out could not be confirmed, so taken together these findings remain inconclusive on the exact role of *Bcl11a* in the mammary gland. It is unclear whether the findings are truly biological or whether technical limitations such as 4-OHT delivery were responsible for insufficient recombination. These finding are inconsistent with the 2D culture of luminal progenitor cells and fat pad transplantation experiments suggesting technical aspects of the experiments were at fault [2].

Given the observations in the *BCL11A*^{OE}experiment, it may have been possible that the level of induction seen in these organoids was not sufficient to induce transgene expression. Such a lack of *eGFP* expression suggests that this was a factor. As such, *BCL11A* gene expression may not have reached the threshold required to induce phenotypic changes in the cells. In addition, it is important to consider the method of RNA preparation, where organoids were pooled across common wells, potentially diluting overall expression levels. Very few cells may have displayed high levels of recombination whilst other cells may not have recombined at all, therefore the effect of the RT-qPCR analysis may have been to show overall low levels of expression that were not representative. Future experiments should therefore consider titrating the 4-OHT to ensure optimal delivery to the culture. Similarly, experiments should seek to induce cells at an earlier time point, to uncover whether over expressing cells at the point of seeding do indeed demonstrate divergent phenotypes as has been demonstrated in the lung [3].

In addition to the induction protocol, aspects of the experimental design may also have confounded the results. The lack of biological replicates undoubtedly contributed to the low statistical power and overall lack of significance after the analysis of organoid size. The

visibility of overall trends suggests that organoid size may indeed vary between experimental groups, however due to a lack of significance it is not possible to reach a conclusion.

Bcl11a lineage tracing in the mammary gland did reveal gene expression in luminal cells that was seen to expand during pregnancy, giving rise to alveolar lineages. This work was undertaken by Dr Sara Pensa and demonstrates a role for *Bcl11a* in the functional mammary gland. The position of the organoid data in this wider context together with work from Khaled et al. could suggest that *Bcl11a* function is dependent on the developmental stage of the mammary gland [2]. Given the cyclical nature of the tissue there is a possibility that phenotypic changes may only be evident at certain time points, for example during puberty or pregnancy. Here, only cells of the adult gland were used which could represent an oversimplification of the tissue and therefore a suboptimal model on which to study mammary gland gene function. Therefore, taken in context with published work and data from lineage tracing studies, the function of *BCL11A* in the mammary gland warrants further study to fully elucidate mechanistic properties, whether specific to developmental time points or otherwise.

6.4 Future Perspectives for Exploring Gene Function in the Mammary Gland

In the mammary gland, this work has revealed the intrinsic sensitivity of mammary epithelial cells when cultured from single cells. Although the data remains inconclusive, there are still many possibilities for future work. Given the functional characterisation of *BCL11A* as a TNBC gene [2] coupled with promising lineage tracing data, evidence would suggest there is a role for this transcription factor in the mammary gland. Future work could involve the culture of mammary organoids from tissue fragments to address the technical questions arising from this study. There remain questions about the timing of 4-OHT induction for cells of the $BCL11A^{OE}$ and $Bcl11a^{KO}$ models. Given the striking effects the induction of these perturbations upon seeding had in lung organoids [3], it remains to be seen whether there is such an effect in cells of the mammary gland.

Future work should also focus on the optimisation of the wholemount immunofluorescence staining protocol. Demonstrated in the data presented in this thesis, there were many cells within the organoids which did not stain positive for either of the basal or luminal markers.

This could have been an antibody penetrance issue, so future work should focus on the permeabilisation of the organoids by incubating them for longer periods with a detergent. In addition, future work should seek to confirm the phenotype observed with the *Bcl11a* knock out organoids of an expansion in Keratin 5 positive cells. For a more robust quantification, multiple z planes should be obtained for each organoid, allowing the overall number of Keratin 5 positive cells to be counted.

Another possible avenue for exploration could be the DMBA/MPA tumour model used in conjunction with the *Bcl11a*^{CreERT2} lineage tracing mouse, to identify whether an expansion of *Bcl11a* expressing cells contribute to the tumour mass. This would aid understanding about TNBC tumour initiation in the mammary gland as well as contributing to the growing body of work aimed at identifying initiating events and cell of origin theories in breast cancer. Indeed, the identification of this gene as a marker of TNBC does suggest *Bcl11a* expressing cells might contribute to tumour formation in some way. To supplement this study, more work is required to fully characterise any developmental functions of this transcription factor in the mammary gland. Evidence would suggest *Bcl11a* is expressed in the luminal progenitor population which undergo an expansion in response to pregnancy. Might an amplification in the genetic locus of *Bcl11a* within these cells contribute to TNBC tumour initiation? Identifying the mechanistic properties of physiological gene function are integral to understanding driver mutational events in malignancy, and would go hand in hand with future work on *Bcl11a* in TNBC.

6.5 Conclusion

This thesis has demonstrated the application of 3D organoid models to explore novel gene function, validating and expanding on these findings, using *in vivo* lineage tracing. A novel function for *Bcl11a* has been demonstrated, identifying it as a transcription factor with a potential role in progenitor cells of the oesophageal epithelium. In addition to these findings, this thesis also works to highlight some of the difficulty still faced in the culture of single mammary cells as organoids. It is clear that different tissues have different requirements in cell culture and that there is still optimisation required for the mammary system before it can be considered as robust as the 3D culture of other epithelial cells.

7 Bibliography

1. Xia, Y., Fan, C., Hoadley, K. A., Parker, J. S. & Perou, C. M. Genetic determinants of the molecular portraits of epithelial cancers. *Nat. Commun.* **10**, 5666 (2019).

2. Khaled, W. T. *et al.* BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nat. Commun.* **6**, 5987 (2015).

3. Lazarus, K. A. *et al.* BCL11A interacts with SOX2 to control the expression of epigenetic regulators in lung squamous carcinoma. *Nat. Commun.* **9**, 3327 (2018).

4. Lee, Y. T., Tan, Y. J. & Oon, C. E. Molecular targeted therapy: Treating cancer with specificity. *Eur. J. Pharmacol.* **834**, 188–196 (2018).

5. Hinck, L. & Näthke, I. Changes in cell and tissue organization in cancer of the breast and colon. *Curr. Opin. Cell Biol.* **26**, 87–95 (2014).

6. Potten, C. S. & Loeffler, M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110**, 1001–1020 (1990).

7. Chang, K. *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).

8. Yin, J., Xie, X., Ye, Y., Wang, L. & Che, F. BCL11A: a potential diagnostic biomarker and therapeutic target in human diseases. *Biosci. Rep.* **39**, BSR20190604 (2019).

9. Nakamura, T. *et al.* Evi9 encodes a novel zinc finger protein that physically interacts with BCL6, a known human B-cell proto-oncogene product. *Mol. Cell. Biol.* **20**, 3178–3186 (2000).

10. Satterwhite, E. *et al.* The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. *Blood* **98**, 3413–3420 (2001).

11. Masliah-Planchon, J., Bièche, I., Guinebretière, J.-M., Bourdeaut, F. & Delattre, O. SWI/SNF Chromatin Remodeling and Human Malignancies. *Annu. Rev. Pathol. Mech. Dis.* **10**, 145–171 (2015).

12. Tang, L., Nogales, E. & Ciferri, C. Structure and Function of SWI/SNF Chromatin Remodeling Complexes and Mechanistic Implications for Transcription. *Prog. Biophys. Mol. Biol.* **102**, 122–128 (2010).

13. Saiki, Y., Yamazaki, Y., Yoshida, M., Katoh, O. & Nakamura, T. Human EVI9, a

homologue of the mouse myeloid leukemia gene, is expressed in the hematopoietic progenitors and down-regulated during myeloid differentiation of HL60 cells. *Genomics* **70**, 387–391 (2000).

14. Luc, S. *et al.* Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. *Cell Rep.* **16**, 3181–3194 (2016).

15. Ippolito, G. C. *et al.* Dendritic cell fate is determined by BCL11A. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E998-1006 (2014).

16. Liu, P. *et al.* Bcl11a is essential for normal lymphoid development. *Nat. Immunol.* **4**, 525–532 (2003).

17. Powers, A. N. & Satija, R. Single-cell analysis reveals key roles for Bcl11a in regulating stem cell fate decisions. *Genome Biol.* **16**, 199 (2015).

18. Yu, Y. *et al.* Bcl11a is essential for lymphoid development and negatively regulates p53. *J. Exp. Med.* **209**, 2467–2483 (2012).

19. Sankaran, V. G. *et al.* Developmental and species-divergent globin switching are driven by BCL11A. *Nature* **460**, 1093–1097 (2009).

20. Xu, J. *et al.* Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6518–6523 (2013).

21. Kuo, T.-Y., Hong, C.-J., Chien, H.-L. & Hsueh, Y.-P. X-linked mental retardation gene CASK interacts with Bcl11A/CTIP1 and regulates axon branching and outgrowth. *J. Neurosci. Res.* **88**, 2364–2373 (2010).

22. Kuo, T.-Y., Hong, C.-J. & Hsueh, Y.-P. Bcl11A/CTIP1 regulates expression of DCC and MAP1b in control of axon branching and dendrite outgrowth. *Mol. Cell. Neurosci.* **42**, 195–207 (2009).

23. Greig, L. C., Woodworth, M. B., Greppi, C. & Macklis, J. D. Ctip1 Controls Acquisition of Sensory Area Identity and Establishment of Sensory Input Fields in the Developing Neocortex. *Neuron* **90**, 261–277 (2016).

24. Cánovas, J. *et al.* The Specification of Cortical Subcerebral Projection Neurons Depends on the Direct Repression of TBR1 by CTIP1/BCL11a. *J. Neurosci.* **35**, 7552–7564 (2015).

25. Li, S. *et al.* Transcription Factor CTIP1/ BCL11A Regulates Epidermal Differentiation and Lipid Metabolism During Skin Development. *Sci. Rep.* **7**, 13427 (2017).

26. Zhou, J. *et al.* BCL11A Promotes the Progression of Laryngeal Squamous Cell Carcinoma. *Front. Oncol.* **10**, 375 (2020).

27. Cai, Y.-D. *et al.* Identification of the Gene Expression Rules That Define the Subtypes in Glioma. *J. Clin. Med.* **7**, 350 (2018).

28. Li, S.-H., Li, J.-P., Chen, L. & Liu, J.-L. miR-146a induces apoptosis in neuroblastoma cells by targeting BCL11A. *Med. Hypotheses* **117**, 21–27 (2018).

29. Zhou, J. *et al.* Genetic polymorphisms and plasma levels of BCL11A contribute to the development of laryngeal squamous cell carcinoma. *PLoS One* **12**, e0171116 (2017).

30. Kasaian, K. *et al.* The genomic and transcriptomic landscape of anaplastic thyroid cancer: implications for therapy. *BMC Cancer* **15**, 984 (2015).

31. Criscitiello, C., Azim, H. a., Schouten, P. C., Linn, S. C. & Sotiriou, C. Understanding the biology of triple-negative breast cancer. *Ann. Oncol.* **23**, vi13-8 (2012).

32. Eroles, P., Bosch, A., Pérez-Fidalgo, J. A. & Lluch, A. Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. *Cancer Treat. Rev.* **38**, 698–707 (2012).

33. Weigelt, B. & Reis-Filho, J. S. Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nat Rev Clin Oncol* **6**, 718–730 (2009).

34. Koboldt, D. C. *et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

35. Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast* **24**, S26–S35 (2015).

36. Ahmad, A. *et al.* Endoxifen, a New Cornerstone of Breast Cancer Therapy: Demonstration of Safety, Tolerability, and Systemic Bioavailability in Healthy Human Subjects. *Clin. Pharmacol. Ther.* **88**, 814–817 (2010).

37. Nahta, R. & Esteva, F. J. Herceptin: mechanisms of action and resistance. *Cancer Lett.* **232**, 123–138 (2006).

38. Prat, A. *et al.* Molecular characterization of basal-like and non-basal-like triple-negative

breast cancer. *Oncologist* **18**, 123–33 (2013).

39. Jézéquel, P. *et al.* Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response. *Breast Cancer Res.* **17**, 1–16 (2015).

40. Lehmann, B. D. *et al.* Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One* **11**, e0157368 (2016).

41. Prat, A. *et al.* Genomic Analyses across Six Cancer Types Identify Basal-like Breast Cancer as a Unique Molecular Entity. *Sci. Rep.* **3**, 3544 (2013).

42. Yan, W., Wistuba, I. I., Emmert-Buck, M. R. & Erickson, H. S. Squamous Cell Carcinoma - Similarities and Differences among Anatomical Sites. *Am. J. Cancer Res.* **1**, 275–300 (2011).

43. Bass, A. J. *et al.* SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat. Genet.* **41**, 1238–1242 (2009).

44. Kim, J. *et al.* Integrated genomic characterization of oesophageal carcinoma. *Nature* (2017) doi:10.1038/nature20805.

45. Lin, Y. *et al.* Epidemiology of esophageal cancer in Japan and China. *J. Epidemiol.* **23**, 233–242 (2013).

46. Doll, R., Peto, R., Wheatley, K., Gray, R. & Sutherland, I. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ* **309**, 901–911 (1994).

47. Anderson, L. A. *et al.* Risk factors for Barrett's oesophagus and oesophageal adenocarcinoma: results from the FINBAR study. *World J. Gastroenterol.* **13**, 1585–1594 (2007).

48. Smyth, E. C. *et al.* Oesophageal cancer. *Nat. Rev. Dis. Prim.* **3**, 17048 (2017).

49. Quante, M. *et al.* Bile acid and inflammation activate gastric cardia stem cells in a mouse model of Barrett-like metaplasia. *Cancer Cell* **21**, 36–51 (2012).

50. Galipeau, P. C., Prevo, L. J., Sanchez, C. A., Longton, G. M. & Reid, B. J. Clonal expansion and loss of heterozygosity at chromosomes 9p and 17p in premalignant esophageal (Barrett's) tissue. *J. Natl. Cancer Inst.* **91**, 2087–2095 (1999).

51. Reid, B. J. *et al.* Predictors of progression in Barrett's esophagus II: baseline 17p (p53)

loss of heterozygosity identifies a patient subset at increased risk for neoplastic progression. *Am. J. Gastroenterol.* **96**, 2839–2848 (2001).

52. Stachler, M. D. *et al.* Paired exome analysis of Barrett's esophagus and adenocarcinoma. *Nat. Genet.* **47**, 1047–1055 (2015).

53. Duits, L. C. *et al.* Barrett's oesophagus patients with low-grade dysplasia can be accurately risk-stratified after histological review by an expert pathology panel. *Gut* **64**, 700–706 (2015).

54. Freedman, N. D. *et al.* Fruit and vegetable intake and head and neck cancer risk in a large United States prospective cohort study. *Int. J. cancer* **122**, 2330–2336 (2008).

55. Cooper, S. C. *et al.* The influence of deprivation and ethnicity on the incidence of esophageal cancer in England. *Cancer Causes Control* **20**, 1459–1467 (2009).

56. Liu, X. *et al.* Genetic Alterations in Esophageal Tissues From Squamous Dysplasia to Carcinoma. *Gastroenterology* **153**, 166–177 (2017).

57. Fagundes, R. B., Melo, C. R., Pütten, A. C. K., Moreira, L. F. & de Barros, S. G. S. p53 immunoexpression: an aid to conventional methods in the screening of precursor lesions of squamous esophageal cancer in patients at high-risk? *Cancer Detect. Prev.* **29**, 227–232 (2005).

58. Couch, G. *et al.* The Discovery and Validation of Biomarkers for the Diagnosis of Esophageal Squamous Dysplasia and Squamous Cell Carcinoma. *Cancer Prev. Res. (Phila).* **9**, 558–566 (2016).

59. Song, Y. *et al.* Identification of genomic alterations in oesophageal squamous cell cancer. *Nature* **509**, 91–95 (2014).

60. Gen, Y., Yasui, K., Nishikawa, T. & Yoshikawa, T. SOX2 promotes tumor growth of esophageal squamous cell carcinoma through the AKT/mammalian target of rapamycin complex 1 signaling pathway. *Cancer Sci.* **104**, 810–816 (2013).

61. Hussenet, T. *et al.* SOX2 Is an Oncogene Activated by Recurrent 3q26.3 Amplifications in Human Lung Squamous Cell Carcinomas. *PLoS One* **5**, e8960 (2010).

62. Daniel, C. W. & Smith, G. H. The mammary gland: a model for development. *J. Mammary Gland Biol. Neoplasia* **4**, 3–8 (1999).

63. Watson, C. J. & Khaled, W. T. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development* **135**, 995–1003 (2008).

64. Sakakura, T., Suzuki, Y. & Shiurba, R. Mammary stroma in development and carcinogenesis. *J. Mammary Gland Biol. Neoplasia* **18**, 189–197 (2013).

65. Hovey, R. C., McFadden, T. B. & Akers, R. M. Regulation of mammary gland growth and morphogenesis by the mammary fat pad: a species comparison. *J. Mammary Gland Biol. Neoplasia* **4**, 53–68 (1999).

66. Williams, J. M. & Daniel, C. W. Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol.* **97**, 274–290 (1983).

67. Russo, I. H. & Russo, J. Developmental stage of the rat mammary gland as determinant of its susceptibility to 7,12-dimethylbenz[a]anthracene. *J. Natl. Cancer Inst.* **61**, 1439–1449 (1978).

68. Inman, J. L., Robertson, C., Mott, J. D. & Bissell, M. J. Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development* **142**, 1028–1042 (2015).

69. Hens, J. R. & Wysolmerski, J. J. Key stages of mammary gland development: Molecular mechanisms involved in the formation of the embryonic mammary gland. *Breast Cancer Res.* **7**, 220–224 (2005).

70. Bresciani, F. Effect of ovarian hormones on duration of DNA synthesis in cells of the C3H mouse mammary gland. *Exp. Cell Res.* **38**, 13–32 (1965).

71. Huebner, R. J., Lechler, T. & Ewald, A. J. Developmental stratification of the mammary epithelium occurs through symmetry-breaking vertical divisions of apically positioned luminal cells. *Development* **141**, 1085–1094 (2014).

72. Humphreys, R. C. *et al.* Apoptosis in the terminal endbud of the murine mammary gland: a mechanism of ductal morphogenesis. *Development* **122**, 4013–4022 (1996).

73. Neumann, N. M. *et al.* Coordination of Receptor Tyrosine Kinase Signaling and Interfacial Tension Dynamics Drives Radial Intercalation and Tube Elongation. *Dev. Cell* **45**, 67-82.e6 (2018).

74. Daniel, C. & Silberstein, G. Postnatal Development of the Rodent Mammary Gland. *Mammary Gland Dev. Regul. Funct.* 3–36 (1987).

75. Couldrey, C. *et al.* Adipose tissue: A vital in vivo role in mammary gland development but not differentiation. *Dev. Dyn.* **223**, 459–468 (2002).

76. Schroeder, J. A. & Lee, D. C. Dynamic expression and activation of ERBB receptors in the developing mouse mammary gland. *Cell growth & Differ. Mol. Biol. J. Am. Assoc. Cancer Res.* **9**, 451—464 (1998).

77. Wang, Z. Transactivation of Epidermal Growth Factor Receptor by G Protein-Coupled Receptors: Recent Progress, Challenges and Future Research. *Int. J. Mol. Sci.* **17**, 95–106 (2016).

78. Coleman, S., Silberstein, G. B. & Daniel, C. W. Ductal morphogenesis in the mouse mammary gland: evidence supporting a role for epidermal growth factor. *Dev. Biol.* **127**, 304–315 (1988).

79. Sternlicht, M. D. Key stages in mammary gland development: The cues that regulate ductal branching morphogenesis. *Breast Cancer Res.* **8**, 201 (2005).

80. Pond, A. C. *et al.* Fibroblast growth factor receptor signaling is essential for normal mammary gland development and stem cell function. *Stem Cells* **31**, 178–189 (2013).

81. Ornitz, D. M. & Itoh, N. The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 215–266 (2015).

82. Hynes, N. E. & Watson, C. J. Mammary gland growth factors: roles in normal development and in cancer. *Cold Spring Harb. Perspect. Biol.* **2**, a003186 (2010).

83. Cui, Y. & Li, Q. Expression and functions of fibroblast growth factor 10 in the mouse mammary gland. *Int. J. Mol. Sci.* **14**, 4094–4105 (2013).

84. Parsa, S. *et al.* Terminal end bud maintenance in mammary gland is dependent upon FGFR2b signaling. *Dev. Biol.* **317**, 121–131 (2008).

85. Moses, H. & Barcellos-Hoff, M. H. TGF-beta biology in mammary development and breast cancer. *Cold Spring Harb. Perspect. Biol.* **3**, a003277–a003277 (2011).

86. Soriano, J. V, Pepper, M. S., Orci, L. & Montesano, R. Roles of hepatocyte growth factor/scatter factor and transforming growth factor-beta1 in mammary gland ductal

morphogenesis. *J. Mammary Gland Biol. Neoplasia* **3**, 133–150 (1998).

87. Pierce, D. F. J. *et al.* Inhibition of mammary duct development but not alveolar outgrowth during pregnancy in transgenic mice expressing active TGF-beta 1. *Genes Dev.* **7**, 2308–2317 (1993).

88. Lane, T. F. & Leder, P. Wnt-10b directs hypermorphic development and transformation in mammary glands of male and female mice. *Oncogene* **15**, 2133–2144 (1997).

89. Lindvall, C. *et al.* The Wnt co-receptor Lrp6 is required for normal mouse mammary gland development. *PLoS One* **4**, e5813 (2009).

90. van Amerongen, R., Bowman, A. N. & Nusse, R. Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* **11**, 387–400 (2012).

91. Rajaram, R. D. *et al.* Progesterone and Wnt4 control mammary stem cells via myoepithelial crosstalk. *EMBO J.* **34**, 641–652 (2015).

92. Roarty, K. & Serra, R. Wnt5a is required for proper mammary gland development and TGF-beta-mediated inhibition of ductal growth. *Development* **134**, 3929–3939 (2007).

93. Wagner, D. E. & Klein, A. M. Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet.* **21**, 410–427 (2020).

94. Tang, L. Integrating lineage tracing and single-cell analysis. *Nat. Methods* **17**, 359 (2020).

95. Doupé, D. P. *et al.* A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science* **337**, 1091–1093 (2012).

96. Piedrafita, G. *et al.* A single-progenitor model as the unifying paradigm of epidermal and esophageal epithelial maintenance in mice. *Nat. Commun.* **11**, 1429 (2020).

97. Kolios, G. & Moodley, Y. Introduction to stem cells and regenerative medicine. *Respiration.* **85**, 3–10 (2013).

98. DeOme, K. B., Faulkin, L. J. J., Bern, H. A. & Blair, P. B. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.* **19**, 515–520 (1959).

99. Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells.

Nature **439**, 993–997 (2006).

100. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).

101. Regan, J. & Smalley, M. Prospective isolation and functional analysis of stem and differentiated cells from the mouse mammary gland. *Stem Cell Rev.* **3**, 124–136 (2007).

102. Shehata, M. *et al.* Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. *Breast Cancer Res.* **14**, R134 (2012).

103. Watson, C. J. & Khaled, W. T. Mammary development in the embryo and adult: new insights into the journey of morphogenesis and commitment. *Development* **147**, dev169862 (2020).

104. Fu, N. Y., Nolan, E., Lindeman, G. J. & Visvader, J. E. Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. *Physiol. Rev.* **100**, 489–523 (2019).

105. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189–193 (2011).

106. Davis, F. M. *et al.* Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat. Commun.* **7**, 13053 (2016).

107. Lloyd-Lewis, B., Harris, O. B., Watson, C. J. & Davis, F. M. Mammary Stem Cells: Premise, Properties, and Perspectives. *Trends Cell Biol.* (2017) doi:10.1016/j.tcb.2017.04.001.

108. Ahmed, M. I., Elias, S., Mould, A. W., Bikoff, E. K. & Robertson, E. J. The transcriptional repressor Blimp1 is expressed in rare luminal progenitors and is essential for mammary gland development. *Development* **143**, 1663–1673 (2016).

109. Elias, S., Morgan, M. A., Bikoff, E. K. & Robertson, E. J. Long-lived unipotent Blimp1-positive luminal stem cells drive mammary gland organogenesis throughout adult life. *Nat. Commun.* **8**, 1714 (2017).

110. Van Keymeulen, A. *et al.* Lineage-Restricted Mammary Stem Cells Sustain the Development, Homeostasis, and Regeneration of the Estrogen Receptor Positive Lineage. *Cell Rep.* **20**, 1525–1532 (2017).

111. Wang, C., Christin, J. R., Oktay, M. H. & Guo, W. Lineage-Biased Stem Cells Maintain Estrogen-Receptor-Positive and -Negative Mouse Mammary Luminal Lineages. *Cell Rep.* **18**, 2825–2835 (2017).

112. Centonze, A. *et al.* Heterotypic cell-cell communication regulates glandular stem cell multipotency. *Nature* **584**, 608–613 (2020).

113. Pal, B. *et al.* Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat. Commun.* **8**, 1627 (2017).

114. Bach, K. *et al.* Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat. Commun.* **8**, 2128 (2017).

115. Giraddi, R. R. *et al.* Single-Cell Transcriptomes Distinguish Stem Cell State Changes and Lineage Specification Programs in Early Mammary Gland Development. *Cell Rep.* **24**, 1653-1666.e7 (2018).

116. Molyneux, G. *et al.* BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* **7**, 403–417 (2010).

117. Koren, S. *et al.* PIK3CAH1047R induces multipotency and multi-lineage mammary tumours. *Nature* **525**, 114–118 (2015).

118. Nguyen, Q. H. *et al.* Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nat. Commun.* **9**, 2028 (2018).

119. Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. In situ identification of bipotent stem cells in the mammary gland. *Nature* **506**, 322–327 (2014).

120. Rosekrans, S. L., Baan, B., Muncan, V. & van den Brink, G. R. Esophageal development and epithelial homeostasis. *Am. J. Physiol. Liver Physiol.* **309**, G216–G228 (2015).

121. Kuo. B, and Urma.D, M. D. Esophagus - anatomy and development. *GI Motil. online* (2006).

122. Zhang, Y. *et al.* Development and stem cells of the esophagus. *Semin. Cell Dev. Biol.* **66**, 25–35 (2017).

123. Domyan, E. T. *et al.* Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. *Development* **138**, 971–981 (2011).

124. Lazzaro, D., Price, M., de Felice, M. & Di Lauro, R. The transcription factor TTF-1 is

expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. *Development* **113**, 1093–1104 (1991).

125. Ioannides, A. S. *et al.* Foregut separation and tracheo-oesophageal malformations: the role of tracheal outgrowth, dorso-ventral patterning and programmed cell death. *Dev. Biol.* **337**, 351–362 (2010).

126. Jacobs, I. J., Ku, W.-Y. & Que, J. Genetic and cellular mechanisms regulating anterior foregut and esophageal development. *Dev. Biol.* **369**, 54–64 (2012).

127. DeWard, A. D., Cramer, J. & Lagasse, E. Cellular heterogeneity in the mouse esophagus implicates the presence of a nonquiescent epithelial stem cell population. *Cell Rep.* **9**, 701–711 (2014).

128. Morrisey, E. E. & Rustgi, A. K. The Lung and Esophagus: Developmental and Regenerative Overlap. *Trends Cell Biol.* **28**, 738–748 (2018).

129. Rodriguez, P. *et al.* BMP signaling in the development of the mouse esophagus and forestomach. *Development* **137**, 4171–4176 (2010).

130. Raymond, C., Anne, V. & Millane, G. Development of esophageal epithelium in the fetal and neonatal mouse. *Anat. Rec.* **230**, 225–234 (1991).

131. Yu, W.-Y., Slack, J. M. W. & Tosh, D. Conversion of columnar to stratified squamous epithelium in the developing mouse oesophagus. *Dev. Biol.* **284**, 157–170 (2005).

132. Alcolea, M. P. & Jones, P. H. Cell competition: Winning out by losing notch. *Cell Cycle* **14**, 9–17 (2015).

133. Alcolea, M. P. & Jones, P. H. Lineage analysis of epidermal stem cells. *Cold Spring Harb. Perspect. Med.* **4**, a015206–a015206 (2014).

134. Marques-Pereira, J. P. & Leblond, C. P. Mitosis and differentiation in the stratified squamous epithelium of the rat esophagus. *Am. J. Anat.* **117**, 73–87 (1965).

135. Potten, C. S. & Booth, C. Keratinocyte stem cells: a commentary. *J. Invest. Dermatol.* **119**, 888–899 (2002).

136. Croagh, D., Phillips, W. A., Redvers, R., Thomas, R. J. S. & Kaur, P. Identification of candidate murine esophageal stem cells using a combination of cell kinetic studies and cell surface markers. *Stem Cells* **25**, 313–318 (2007).

137. Li, A., Simmons, P. J. & Kaur, P. Identification and isolation of candidate human keratinocyte stem cells based on cell surface phenotype. *Proc. Natl. Acad. Sci.* **95**, 3902–3907 (1998).

138. Tumbar, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004).

139. Lampreht Tratar, U., Horvat, S. & Cemazar, M. Transgenic Mouse Models in Cancer Research. *Front. Oncol.* **8**, 268 (2018).

140. Breslin, S. & O'Driscoll, L. Three-dimensional cell culture: the missing link in drug discovery. *Drug Discov. Today* **18**, 240–249 (2013).

141. Simian, M. & Bissell, M. J. Organoids: A historical perspective of thinking in three dimensions. *J. Cell Biol.* **216**, 31–40 (2017).

142. Orkin, R. W. *et al.* A murine tumor producing a matrix of basement membrane. *J. Exp. Med.* **145**, 204–220 (1977).

143. Emerman, J. T. & Pitelka, D. R. Maintenance and induction of morphological differentiation in dissociated mammary epithelium on floating collagen membranes. *In Vitro* **13**, 316–328 (1977).

144. Barcellos-Hoff, M. H., Aggeler, J., Ram, T. G. & Bissell, M. J. Functional differentiation and alveolar morphogenesis of primary mammary cultures on reconstituted basement membrane. *Development* **105**, 223–235 (1989).

145. Nguyen-Ngoc, K.-V. *et al.* 3D culture assays of murine mammary branching morphogenesis and epithelial invasion. *Methods Mol. Biol.* **1189**, 135–162 (2015).

146. Murphy, S. V & Atala, A. 3D bioprinting of tissues and organs. *Nat. Biotechnol.* **32**, 773–785 (2014).

147. Jamieson, P. R. *et al.* Derivation of a robust mouse mammary organoid system for studying tissue dynamics. *Development* **144**, 1065–1071 (2017).

148. Jardé, T. *et al.* Wnt and Neuregulin1/ErbB signalling extends 3D culture of hormone responsive mammary organoids. *Nat. Commun.* **7**, 13207 (2016).

149. Sumbal, J., Chiche, A., Charifou, E., Koledova, Z. & Li, H. Primary mammary organoid model of lactation and involution. *bioRxiv* 859645 (2019) doi:10.1101/859645.

150. Li, M. L. *et al.* Influence of a reconstituted basement membrane and its components on casein gene expression and secretion in mouse mammary epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 136–140 (1987).

151. Andl, C. D. *et al.* Coordinated functions of E-cadherin and transforming growth factor beta receptor II in vitro and in vivo. *Cancer Res.* **66**, 9878–9885 (2006).

152. Reynolds, B. A. & Weiss, S. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* **255**, 1707–1710 (1992).

153. Kasagi, Y. *et al.* The Esophageal Organoid System Reveals Functional Interplay Between Notch and Cytokines in Reactive Epithelial Changes. *Cell. Mol. Gastroenterol. Hepatol.* **5**, 333–352 (2018).

154. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

155. O'Rourke, K. P., Ackerman, S., Dow, L. E. & Lowe, S. W. Isolation, Culture, and Maintenance of Mouse Intestinal Stem Cells. *Bio-protocol* **6**, e1733 (2016).

156. Ramesh, S., Wildey, G. M. & Howe, P. H. Transforming growth factor beta (TGFbeta)-induced apoptosis: the rise & fall of Bim. *Cell Cycle* **8**, 11–17 (2009).

157. Chang, C.-L. *et al.* Retinoic acid-induced glandular differentiation of the oesophagus. *Gut* **56**, 906–917 (2007).

158. Garner, O. B. *et al.* Stage-dependent regulation of mammary ductal branching by heparan sulfate and HGF-cMet signaling. *Dev. Biol.* **355**, 394–403 (2011).

159. NGUYEN-NGOC, K.-V. & EWALD, A. J. Mammary ductal elongation and myoepithelial migration are regulated by the composition of the extracellular matrix. *J. Microsc.* **251**, 212–223 (2013).

160. Nair, D. V. & Reddy, A. G. Laboratory animal models for esophageal cancer. *Vet. world* **9**, 1229–1232 (2016).

161. Zhang, Z. *et al.* Architecture of SWI/SNF chromatin remodeling complex. *Protein Cell* (2018) doi:10.1007/s13238-018-0524-9.

162. Moody, R. R. *et al.* Probing the interaction between the histone methyltransferase/deacetylase subunit RBBP4/7 and the transcription factor BCL11A

in epigenetic complexes. *J. Biol. Chem.* **293**, 2125–2136 (2018).

163. Mohammed, H. *et al.* Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326 (2016).