Modelling the development of Barrett's Oesophagus

Towards better treatment of Oesophageal Adenocarcinoma

Elhadi Iich

M.Sc.

Submitted in fulfilment of the requirements of the degree of

Doctor of Philosophy

March 2020

Division of Cancer Surgery

Surgical Oncology Research Laboratory

Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

&

The University of Melbourne, Victoria, Australia

ORCID ID: 0000-0002-4194-5137

Abstract

Barrett's oesophagus (BO) is a metaplastic condition in which the normal squamous epithelium of the oesophagus is replaced by a columnar gastro-intestinal like epithelium due to repeated gastro-intestinal reflux. BO is generally accepted to be a precursor condition with the potential to develop into oesophageal adenocarcinoma.

Consensus for the cell of origin for Barrett's oesophagus is still lacking. Different sources for the cell of origin have been proposed, one of which is the submucosal gland (SMG) and duct cells of the oesophagus. This hypothesis, however, has not been properly studied due to the lack of proper model systems. In this thesis, pig and human SMGs and ducts were characterised and compared to each other to assess the suitability of pig SMGs as a substitute for human SMGs using immunohistology and multiplex immunofluorescence staining. Pig epithelial SMG cells were also further characterised using single cell RNA sequencing. Furthermore, organoid culture systems were developed to functionally assess the potential of the progenitor cells found using histologic and transcriptomic characterisation.

Pig and human SMGs show different distributions at the distal end of their respective oesophagi but are largely similar in cell type and progenitor cell marker expression as demonstrated by histologic characterisation. Both pig and human SMGs showed progenitor cell marker expression in their respective basal duct cells (CD49f and p75) and the myoepithelial cells (CD49f), suggesting that both the ductal and glandular compartments to contain their own respective pool of progenitor cells. The transcriptomic analysis of pig SMGs single cell RNA sequencing largely supports their role in maintaining oesophageal homeostasis, as previously is known for human.

Furthermore, the pseudotime trajectory inference data support the notion of the basal duct cells to be progenitor which give rise luminal duct cells. Interestingly, the data suggest the myoepithelial cells to be progenitor cells that could give rise to both basal duct cells and the gland compartment cells.

Sorting and organoid culturing of basal duct cells demonstrated their capacity to grow into squamous spheroids, supporting their role in contributing to normal oesophageal repair. This process could be inhibited by treatment with retinoic acid. Similarly, culture of gland compartment cells gave rise to spheroids with two distinct morphologies. Dense type spheroids showed a squamous morphology but also mucin production found in BO. The second type of spheroid showed a cystic morphology and similarly produces BO type mucin. Finally, viral induction of intestinalisation in submucosal basal duct cells, in the current culture system, did not show metaplastic changes similar to BO.

In summary, pig SMGs show great similarity to human SMGs. The pig SMGs contain progenitor cells in their ductal and gland compartments. These progenitors can be purified and cultured *in vitro*. The current developed protocols could be used to test the hypothesis that the SMGs contain the cell of origin of BO.

Declaration

This declaration certifies that:

I. This thesis comprises only my original work towards the PhD except where indicated.
II. Due acknowledgement has been made in the text to all other material used.
III. The thesis is fewer than 100,000 words, exclusive of tables, maps, bibliographies and appendices as approved by the Research Higher Degrees Committee.

Elhadi Iich

18 March 2020

Preface

I. I have conducted the majority of the work in this thesis and estimate my overall contribution to the thesis to be greater than 95%.
II. The histology core performed some of the histology stains as described in material and methods in Chapter 2 and single cell RNA sequencing was performed by the Single Cell Innovation Lab as described in Chapter 3.
III. No part of this thesis has been submitted for the award of any other qualification or degree.

Acknowledgements

I would like to acknowledge the people who have contributed to the current work and those who provided great moral support during the difficult times.

The labheads **Wayne Phillips** and **Nicholas Clemons** helped greatly in the writing up and presentation of this work. Also, for the freedom to pursue this project.

My mentors **Benjamin Blyth** and **Florian Wiede** who shared their great insight and expertise and helped me during really tough times when I was at a loss and losing motivation.

I would like to thank previous and current people from multiple clusters: **Keefe Chan**, **Kun-Hui Lu**, **Shuwei Liang**, **Yu-kuan Huang**, **Haoran Zhu**, **Shaun Blake**, **Jinbei Son**, **Alizee Huglo**, **Shunfei Yan**, **Jovana Gotovac**, **Glen Guerra**, **Matthew Read**, **Rosie Millen**, **Sara Roth**, **Shienny Sampurno**, **Rita Bussutil**, **Tony Tiganis and Rick Pearson** for the fun moments both inside and outside of the lab.

My friends back home in the Netherlands and Aruba **Ryan**, **Bernardo, Peter** and **Alexi** thank you for being there for me even over such a long distance. I miss you all a lot.

Thank you to my housemate who has been a joy to live with.

Abbreviations

AB	Alcian blue
ACTA2	Alpha Smooth muscle Actin 2
ALDH1A3	Aldehyde dehydrogenase 1A3
AQP	Aquaporin
BMP	Bone morphogenetic protein
BO	Barrett's Oesophagus
BrdU	Bromodeoxyuridine
Car4	Carbonic anhydrase 4
CASCADE	Cancer tissue Collection After Death
CD	Cluster of Differentiation
CDX2	caudal type homeobox 2
CF	Cystic Fibrosis
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
COX2	Cyclooxygenase 2
DAB	3’3’diaminobenzidine
EGF	Epidermal Growth Factor
EPCAM	Epithelial Cell Adhesion Molecule
EPIC	European Prospective Investigation into Cancer and Nutrition
FACS	Fluorescent Activated Cell Sorting
FFPE	Formalin Fixed Paraffin Embedded
FGF	Fibroblast Growth Factor
gam	Generalised Additive Model
GOJ	Gastro-Oesophageal Junction
GORD	Gastro-Oesophageal Reflux Disease
H&E	Haematoxylin and Eosin
H2O2	Hydrogen Peroxide
HCO3	Bicarbonate
HGD	High Grade Dysplasia
HRP	Horseradish Peroxidase
IBL	Interpapillary Basal Layer
IHC	Immunohistochemistry
IL1β	Interleukin 1β
ITGA	Integrin Alpha
ITGB	integrin Beta
ITS	Insulin Transferrin Selenium
KNG1	Bradykinin
KRT	Cytokeratin
LGD	Low grade dysplasia
LOS	Lower Oesophageal Sphincter
LP	Lamina Propria
LYZ	Lysozyme
MGST	Microsomal glutathione S-transferase
m-IF	Multiplex Immunofluorescence
MLE	Multi-layered Epithelium

MM	Muscularis Mucosa
MT	Masson's Trichrome
mTEC	Mouse Tracheal Epithelial cell
MUC	Mucin
MYL9	Myosin regulatory light polypeptide 9
MYLK	Myosin light chain kinase
NFkB	Nuclear factor kappa-light-chain-enhancer of activated B cells
OAC	Oesophageal Adenocarcinoma
ODA	Oesophagoduodenal Anastomosis
OGDA	Oesophago-gastro-duodenal Anastomosis
OJ	Oesphagojejunostomy
OSCC	Oesophageal Squamous Cell Carcinoma
p63	Tumour protein 63
p75/NGFR	Nerve Growth Factor Receptor
PAS	Periodic-Acid-Schiff
PBL	Papillary basal layer
PBS	Phosphate buffer saline
PCA	Principal Component Analysis
PF5	PBS/5%FBS
PGE	Prostaglandin E
RA	Retinoic Acid
ROI	Regions Of Interests
RT	Room Temperature
SCJ	Squamo-Columnar-Junction
Sc-RNA seq	Single cell RNA sequencing
SHH	Sonic Hedgehog
SLC	Solute Carrier
SMG	Submucosal Glands
SOX2	SRY (sex determining region Y)-box 2
SOX9	SRY (sex determining region Y)-box 9
TBST	Tris Buffer Saline Tween-20
TFF3	Trefoil Factor 3
TSA	Tyramide Signal Amplification
UEA	Ulex Europaeus Agglutinin I
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
WGA	Wheat Germ Agglutinin
Wnt	Wingless
WT	Wildtype

Table of Contents

Chapter 1

Introduction

1.1 Anatomy of the oesophagus

The oesophagus connects the oral cavity to the stomach, serving as the channel for food transport through peristaltic contractions of the muscles wrapped around it. Proximally, access to the oesophagus is controlled by the Cricopharyngeus muscle, which functions as a sphincter at the upper end. Distally, the opening to the stomach is constricted by the lower oesophageal sphincter (LOS), which protects the oesophagus from reflux of gastric and duodenal components (Figure 1.1).

The oesophageal wall is constituted of multiple layers starting with the mucosa containing a stratified squamous epithelium at the luminal surface. The role of the stratified epithelium is to protect the oesophagus against damage from gastroduodenal reflux components and passing ingested material[1].

The squamous epithelium can be further subdivided into three sublayers. The most luminal sublayer called the stratum corneum contains the squamoid pyknotic epithelial cells. The role of the stratum corneum is to protect the underlying layers from acid through their cell surface membranes and the tight junctions connecting the cells[1-3]. The middle sublayer of the squamous epithelium called the spinosum contains metabolically active cells[1,3]. The final and third sublayer, adjacent to the basement membrane, known as the stratum germinativum, contains the progenitor cells attached to the basal membrane through hemidesmosomes which give rise to the aforementioned layers[1,4]. Beneath the stratum germinativum is the lamina propria, which is composed of a layer of connective tissue, separating the squamous epithelium from the smooth muscle fibres of the muscularis mucosa[5] (Figure 1.2).

The next section of the oesophageal wall is the submucosal region, which is constituted of a matrix of connective tissue, nerves, blood and lymphatic vessels. The submucosa also contains submucosal glands (SMGs), which are connected to the lumen of the oesophagus through ducts. These structures are known play a role in secreting mucin and bicarbonate to protect the oesophagus from physical insults and refluxed gastric acid and are carried to the lumen through the ducts[6] (Chapter 2, Figure 2.1C).

Surrounding the submucosa is the muscularis propria which is composed of two layers. The inner layer is composed of a thicker circular smooth muscle while the outer layer is composed of longitudinally oriented muscle bundles[5]. Finally, the fourth layer surrounding these three oesophageal layers is an adventitial layer.

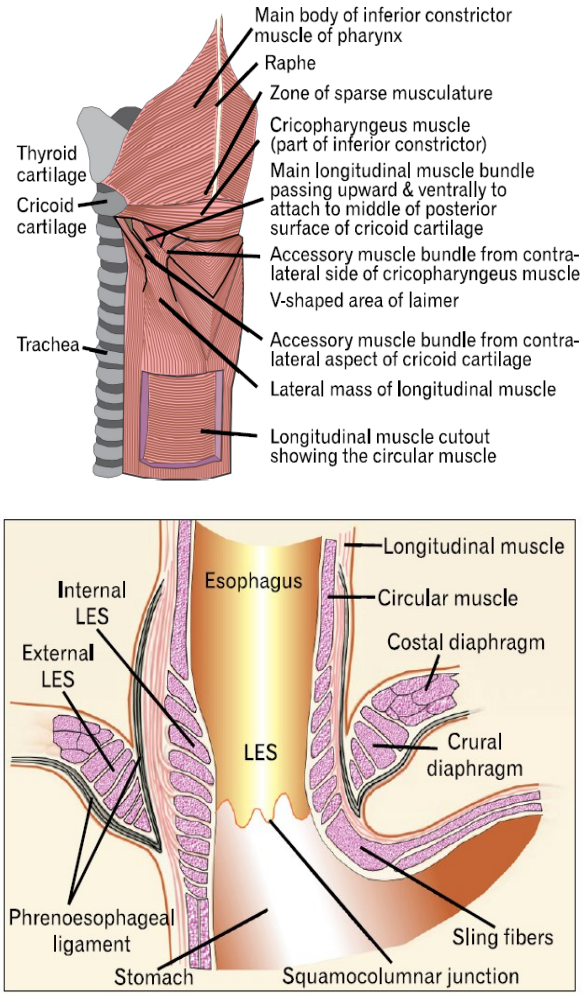


Figure 1.1: Schematic of the of the upper end (A) and lower end (B) of the human oesophagus showing the muscles responsible for peristaltic and opening and closing of the oesophagus. LES: Lower oesophageal sphincter. Adapted from Mittal *et al*.[7]

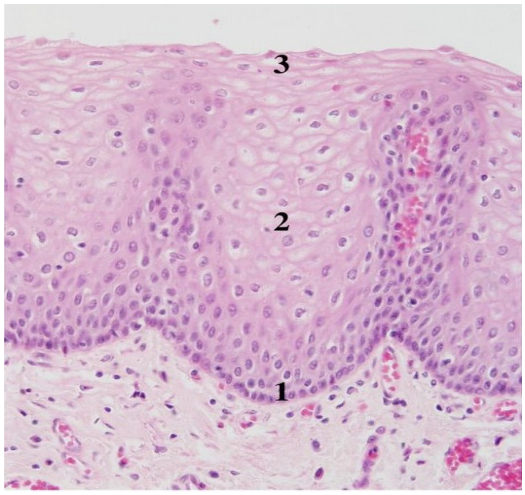


Figure 1.2: H&E staining of stratified squamous epithelium showing (1) stratum germinativum, (2) stratum spinosum and (3) stratum corneum.

1.2 Embryonic development and morphogenesis of the oesophagus

During embryonic development, the oesophagus is derived from the foregut of the primordial gut[8]. The oesophageal diverticulum grows in a caudal direction and separates from the tracheal diverticulum, with which it shares a lumen initially through septum formation[8]. This separation process happens at approximately four to six weeks of gestation in human.

Molecular studies in mice have shown the process of separation to be controlled by a variety of morphogens and transcription factors (reviewed by Zhang *et al*.[5]). Specifically, SRY (sex determining region Y)-box 2 (Sox2) expression was shown to be important for the specification of oesophageal epithelium in the dorsal foregut endoderm[9]. Conversely, Sox2 expression is repressed in the ventral foregut tracheal diverticulum by bone morphogenetic protein (Bmp) signalling[10].

After separation of the oesophagus from the trachea, it is initially lined with ciliated cells, which begin to form in the middle third of the oesophagus and expand both in rostral and caudal directions after septum formation[11]. These ciliated cells are subsequently replaced through transdifferentiation by a stratified squamous epithelium, which also starts forming in the middle third of the oesophagus[11-13].

During the conversion from a columnar (ciliated) to a squamous epithelium in the oesophagus, the basal layer cells start expressing the basal cell markers cytokeratin 5 and 14 (Krt5, Krt14)[12]. Following this conversion, the newly formed suprabasal layers start expressing the squamous markers cytokeratin 4 and 13 (Krt4, Krt13), and involucrin and loricrin at the top layers of the epithelium[12].

Multiple molecular pathways have been shown to contribute to the development and stratification of the oesophageal squamous epithelium. Bmp signalling was shown to be active in the suprabasal layer of the developing and adult stratified squamous epithelium in mice[14]. Treatment of oesophageal basal layer cells with Bmp4 was shown to induce senescence and differentiation through expression of squamous differentiation markers[15]. Conversely, treatment with Noggin, a Bmp inhibitor, promoted expansion and maintenance of basal layer oesophageal cells in a 3D culture system[16].

Expression of the transcription factor Tumour protein p63 (p63) was shown to be critical for the development of both oesophageal and tracheal epithelia[17]. Experiments using p63-null transgenic mice have shown that loss of p63 prevents the conversion of ciliated cells lining the columnar oesophageal epithelium into a squamous epithelium during development. However, lung basal layer cells, which also express p63, give rise to ciliated cells among others. This suggests that p63 alone is not the only factor governing the transition from ciliated to squamous cells in the oesophagus[17-19].

1.3 Oesophageal epithelial stem cells and homeostasis

As with other organs, the oesophageal epithelium is known to continuously regenerate. This turnover was shown to require the basal layer epithelial cells in the adult oesophagus[20]. The precise location of progenitor cells for the oesophageal epithelium is

still unknown. In humans but not in mice, the oesophageal epithelium is interrupted by papillae of stromal cells, thus giving rise to the papillary and interpapillary basal layers (PBL and IBL, respectively). Previous work by Seery *et al*., suggests the IBL contains progenitor cells, which divide asymmetrically, infrequently and have the highest clonogenic potential *in vitro*[20]. Conversely, PBL cells divided more frequently, symmetrically and showed lower clonogenic potential *in vitro*[20], leading the authors to designate PBL cells as transit-amplifying cells similar to suprabasal cells in the IBL region. These findings are not in agreement with more recent work by Barbera *et al.*[21]. In the work performed by Seery *et al*., integrin beta 1 (ITGB1) was used as a marker for labelling the basal layer progenitor cells. Barbera *et al*. resorted to the progenitor marker Cluster of differentiation 34 (Cd34), which they have shown to colocalise with ITGB1 thereby potentially labelling the same cells that Seery *et al*. was studying. Barbera *et al*. found a decreasing gradient of proliferating basal layer cells from the IBL to the PBL, thereby contradicting the findings of Seery *et al.* Additionally, they did not find any directionality in terms of division symmetry. Barbera *et al*. attributed the directionality discrepancies to the use of 2D sections in previous studies, which can suffer from sectioning artefacts.

In the mouse, multiple markers have been used to identify and characterise oesophageal progenitor cells. Kalabis *et al.* used Hoechst label extrusion and Bromodeoxyuridine (BrdU) retention as markers for potential oesophageal progenitor cells[22]. In their analyses, they were able to characterise a side population positive for the Cd34, which had enhanced self-renewal and differentiation capacity *in vitro*[22]. Similarly, Croagh *et al.* showed a small population of BrdU label retaining cells positive for integrin alpha 6 (Itga6/Cd49f) and low expression of Cd71 to represent a putative progenitor cell population in the mouse oesophagus[23]. DeWard *et al*. showed that specific basal layer cells with high Sox2, Itgb1 and Cd73 expression had higher colony forming efficiency in a 3D organoid model and were able to grow into squamous spheroids *in vitro*[16]. They also showed that basal layer cells expressing high levels of Itga6, similar to Croagh *et al*., and Itgb4 have higher spheroid forming efficiency compared to the Itga6/Itgb4 low or non-purified populations[16]. This potentially indicates that Croagh *et al* and DeWard *et al.* were purifying the same population of basal layer cells.

Recently, a subset of oesophageal basal layer cells positive for Krt15, was identified by Giroux *et al*. [24] using lineage tracing in Krt15-CrePR1/R26$^{mT/mG}$ mice to be a long-lived population able to give rise to all the layers of differentiated squamous epithelium cells[24]. Following diphtheria toxin-mediated ablation of the Krt15^{+} population, Giroux *et al*. showed the oesophageal epithelium to be considerably thinner, the basal layer to contain lower proportions of proliferative cells, and BrdU incorporation. The loss of BrdU incorporating cells potentially indicates that the Krt15^{+} might represent the same population of Cd34^{+} Kalabis *et al*.[22] described earlier. Moreover, Krt15^{+} cells were also shown to be resistant to high dose radiation and able to regenerate the squamous epithelium after radiation injury[24]. These data suggest that Krt15^{+} oesophageal basal layer cells have features of stem/progenitor cells.

Both DeWard *et al*.[16] and Croagh *et al*.[23], suggest the presence of a hierarchy within the basal layer progenitor cells of the oesophagus based on the markers they have used to characterise them. This hierarchy makes it difficult to ascertain whether the characterised populations, with the different markers, are the same. Adding to the difficulty of identifying the stem/progenitor cell(s) of the oesophagus, are the plasticity

findings by Barbera *et al*.[21], which could potentially also be extended to mice and not just humans.

The notion of a distinct progenitor population in the mouse oesophagus is contradicted by the work of Doupe *et al*.[25] where they used an inducible GFP labelled Histone-2B transgenic mice to identify label retaining cells in the mouse oesophagus. To their surprise, the label retaining cells at the basal layer of the oesophagus were negative for the stem/progenitor markers Cd34 and *Lgr5*, while the majority of cells were positive for the lymphocyte marker Cd45. This led the authors to conclude that the mouse oesophagus basal layer was homogenous and did not contain any slow cycling cells that could function as resident progenitor cells. This observation seems to hold in cases where wounding or hyperproliferation was induced through the use of retinoic acid (RA)[25]. These wounding experiments suggest that the oesophageal progenitors can switch between a homeostatic and regenerative state to maintain the epithelium.

The large disparity in the different findings could be explained by the different methods used to assess for the potential progenitor cells. Methods requiring *in vitro* culture could alter the behaviour of cells. This makes approaches using *in vivo* lineage tracing more reliable for such goals. Furthermore, a temporal aspect of the dynamic progenitor cell behaviour could also explain the disparate results found by the different groups in terms of markers expression. If one were to assume that all cells in the basal layer of the mouse oesophagus are equivalent, then a dividing cell at one time point could show a certain pattern of marker expression. This pattern could be different when looking at that same cell in a quiescent state, thereby making it difficult to identify a staining profile for a population of progenitor cells.

1.4 Oesophageal epithelial stem cells and cancer

Cancer cells are known to possess stem/progenitor-like properties such as long-term self-renewal and the capacity give rise to daughter cells of heterogenous lineages[26]. Consequently, this makes stem/progenitor cells interesting candidates as the cell of origin of cancer. Although this is not unequivocal proof that all cancers develop from stem/progenitor cells, it does however support it. Therefore, a good understanding of the resident progenitor cell(s) in the oesophagus might enable a better characterisation of the aetiology of oesophageal cancers and potential avenues for their treatment.

Oesophageal cancer is the sixth most common cause of cancer related death world-wide[27]. Oesophageal cancers are generally distinguished into two different types: oesophageal squamous cell carcinoma (OSCC) and oesophageal adenocarcinoma (OAC)[28]. Although these cancers arise in the same organ, research over the past decades on their aetiology and characterization has shed some light into the distinct processes and complexities underlying these diseases. This accumulated knowledge has led to the recognition of these two conditions as being distinct diseases[29].

OSCC accounts for the majority of oesophageal cancer cases worldwide (70%) however, its incidence is decreasing globally[30]. OSCC is mostly prevalent in developing countries such as southern and eastern Africa, Latin America and China[30]. Its pathophysiology, although an interesting topic, falls outside the scope of this thesis.

As described further in detail in the following sections, OAC comprises a lower percentage of oesophageal cancer but its incidence has been increasing mainly in western countries[31]. OAC is known to arise from a generally stable, yet phenotypically diverse metaplastic lesion known as Barrett's oesophagus (BO)[32] suggesting a multipotent cell as the origin of this disease.

1.5 Oesophageal adenocarcinoma

1.5.1 Epidemiology of oesophageal adenocarcinoma

The incidence of OAC has been increasing rapidly over the past five decades[33,34]. The rapid increase in incidence seems to have begun in the 1970s and is largely sequestered to western nations (North America, Australia, Northern and Western Europe) and higher socioeconomic classes[27]. In addition, OAC seems to affect men disproportionately with 6.2 men being affected for each woman in Australia[27]. Some studies have suggested the incidence rate to have reached a plateau[35]. However, more recent reports indicate a continuous but slower increase from the year 2000[36,37].

1.5.2 Aetiology of oesophageal adenocarcinoma

The aetiology of OAC is still under continuous investigation. The past decades have led to a better understanding of the contribution of multiple factors in the development of BO, a precursor lesion thought to culminate into OAC described further below. Importantly, the increase in OAC is concomitant with the increased prevalence of BO[38]. The factors thought to contribute to the development of BO and OAC include gastroesophageal reflux disease (GORD)[39,40], obesity[41,42], diet[41,42] and smoking[43].

1.6 Risk factors

1.6.1 Demographics

OAC shows the highest incidence in Caucasian males versus females or any other ethnicities which is comparable to the distribution of BO[38]. Caucasian males were shown to have a 13-fold higher rate of BO diagnosis than the lowest group (black females) in the US. Interestingly, a systematic review and meta-analysis of 47 studies of BO and OAC incidence showed men diagnosed with BO also had a higher rate of progressing to OAC compared to females with BO suggesting an association between the male gender and OAC prevalence[44].

1.6.2 GORD

GORD was shown to be strongly associated with the development of OAC in a pooled analysis from the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON)[45]. Furthermore, the association of reflux was also shown to increase with increasing symptom duration[45].

Gastro-oesophageal refluxate is a mixture of acid, proteases such as pepsin, pancreatic enzymes and duodenal bile acids secretions[46]. Exposure of the oesophageal epithelium to acid results in alteration of the cell-cell junctions and dilated intercellular spaces

thereby increasing epithelial permeability. The decrease in epithelial barrier was shown to be due to decreases in Claudin-3 and E-cadherin levels at the cell junctions[47,48]. Similarly, the acid activated protease pepsin secreted by chief cells in the stomach was shown to increase intercellular space dilation through its proteolytic activity. This effect seems to be abrogated when pepsin is outside its proteolytic pH range (pH >3.5) supporting the use of proton pump inhibitors in the treatment of chronic reflux[49]. Finally, bile acids refluxed from the duodenum can also be found in gastro-oesophageal refluxate. The increased permeability of the oesophageal epithelium and noxious effects of the reflux component together are known to promote inflammation and oxidative stress.

Inflammation is a known response to diverse types of insults to tissues. Damage from reflux components induces the release of cytokines and chemokines from damaged cells, which attracts inflammatory cells from the circulation. However, high influx of inflammatory cells into the damaged tissues due to chronic reflux can induce more damage. Inflammatory cells were shown to produce reactive oxygen species (ROS), which in turn promote more inflammation[50]. Interestingly, a rat reflux model showed infiltration of lymphocytes in the submucosal region of the rat oesophagus prior to observable damage to the mucosa[51]. This suggests that chronic inflammation might come prior to mucosal damage and not as a consequence. This chronic inflammation could then lead to mucosal damage and increase the inflammatory response in the tissue.

Similarly, bile acids are thought to induce oxidative stress due to their properties as detergents, leading to damage to the membrane of mitochondria and release of their content into the cytoplasm[52]. ROS released by mitochondria are known to promote DNA mutations leading to oncogenic transformation[53].

1.6.3 Obesity

Obesity alone, shows a moderate increase in risk of OAC, particularly among men. However, in combination with GORD, the risk of developing OAC increased by three-fold according to a population-based case–control study in Australia[54]. This is thought to be the result of pressure of abdominal fat on the stomach, leading to increased gastric pressure and the possible relaxation of the LOS. This in turn leads to longer exposure of the distal oesophagus to acid and bile. Alternatively, obesity might contribute to the development of OAC through the release of adipokines such as Leptin[55]. Leptin was shown to be associated with increased risk of BO in men but not in women[55]. This association remained valid even after correcting for potential confounding factors such as body mass index (BMI) or gastro-oesophageal reflux.

1.6.4 Diet

Western diet was shown to be associated with the development of OAC. Independent of the risk of developing obesity, high intake of trans-fats has been associated with increased risk of developing BO and OAC[56]. In a surgically induced reflux model in rats, the highest incidence of oesophageal tumours (both OSCC and OAC) occurred when animals were fed high fat diet in conjunction with administration of a carcinogen compared to carcinogen alone[57].

Conversely, epidemiological studies show that intake of poly-unsaturated fats, omega-3-fatty-acids, raw fruits, cruciferous and leafy vegetables rich in vitamin A, vitamin C, vitamin E and fibre have an inverse correlation with incidence of BO and OAC[58,59]. Processed meat, a known source of nitrites and nitrosamines which are known carcinogens in animals[60], also correlated positively with the incidence of OAC. This association however, did not reach statistical significance according to the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study[61]. Alcohol consumption shows no association with the development of OAC but significant association with OSCC[62-64].

1.6.5 Smoking

Smoking was shown to have a weak association with GORD[65,66]. Furthermore, a strong dose-response association was shown between smoking and developing OAC[67]. Additionally, the risk of developing OAC in men who smoke was shown to be slightly higher than in women[67].
Mechanistically, smoking is thought to promote GORD through the reduction of the LOS pressure[68,69]. This led to more frequent and prolonged exposure of the distal oesophagus to reflux during and after smoking as measured by pH at the LOS. Smoking was also shown to reduce saliva and bicarbonate secretion, thereby increasing the time needed for clearing and neutralising acid in the oesophageal lumen[70]. This was shown to be restored upon smoking cessation[70].

1.6.6 Helicobacter pylori

Helicobacter pylori (*H. pylori*) in the stomach is known to induce a prolonged inflammatory response in the infected subject. The induced inflammation does not lead to pathogen eradication but promotes tissue damage[71]. Interestingly, *H. pylori* infection was shown to be negatively correlated with GORD[72,73]. Furthermore, the negative association of *H. pylori* with BO is even stronger than with GORD[74] suggesting a protective effect. This protective effect was recently further evidenced in a meta-analysis of 72 studies showing this negative association to be independent of geographical location as baseline *H. pylori* infections are different when comparing eastern versus western countries[75]. Similarly, *H. pylori* has been negatively associated with OAC suggesting a protective effect as with BO[76]. Mechanistically, the negative association of *H. pylori* infection with GORD, BO and OAC could be due to lower acid secretion from atrophic gastric tissue. Still, conclusive evidence of a potential protective role of *H. pylori* against BO and OAC is still lacking and, in some cases, also contradicted. A 2019 study of the prevalence of BO, in a predominantly Caucasian high *H. pylori* infection area in Azerbaijan, showed no inverse association suggesting no protective effects[77].

1.7 Barrett's oesophagus

The replacement of the normal squamous epithelium in the lumen of the lower third of the oesophagus with a columnar epithelium is referred to as Barrett's metaplasia[78]. It has been hypothesized that this condition arises as an adaptation to frequent recurring

mucosal injury and inflammation from gastric acids and bile[79]. This is mainly due to its close association to GORD[80].

Subsequent studies in the 1970s have shown that the glandular metaplasia of BO can have a spectrum-like phenotype containing cells with gastric fundic (pepsin releasing), gastric cardiac (gastric mucin producing), gastric oxyntocardiac (acid producing) and intestinal phenotype[81,82]. The cellular composition of these metaplastic lesions consist of a variety of cell types found in the native tissue of intestinal and gastric tissues such as mucinous cells, goblet cells, paneth cells and to a lesser extent also endocrine cells or enterocytes[83]. Furthermore, some studies have shown that the intestinal type metaplastic BO lesions were more likely to progress towards OAC[84-86]

The development of BO is surmised to be a two-step process[87]. The first step of the metaplastic process involves the columnarisation of the squamous epithelium[88]. This process is thought to be fairly rapid. The second step is intestinalisation as evidenced by the presence of goblet cells. However, this process is thought to progress at a much slower rate[88]. Across these two steps, BO is thought to develop through a sequence from low grade to high grade dysplasia to finally adenocarcinoma.[87]

1.7.1 Columnarisation

Historically, it was believed that the lower one to two centimetres of the oesophagus was normally lined with columnar mucosa[89]. However, advances in endoscopic techniques during the 1900s have shown that patients not suffering from GORD did not have such a columnar lined mucosa. The squamous epithelium lining was found throughout the whole length of the oesophagus reaching the Rugae (epithelium folds) of the stomach.[89] This was further evidence that the specialized mucosa lining the lower oesophagus in Barrett's patients is an adaptation to the refluxed gastric juices[90].

The columnar mucosa replacing the squamous epithelium was termed cardiac mucosa, due to its histologic similarity to gastric cardia epithelium, which is lined mainly by mucin secreting cells (foveolar cells).

1.7.2 Intestinalisation

Intestinalisation can also be observed in Barrett's metaplastic tissue. This is characterised by the development of goblet cells, normally found in the intestine, within the cardiac mucosa indicating intestinalisation[80,84]. Furthermore, the caudal-related homeobox gene 2 (CDX2) gene, responsible for the differentiation and maintenance of intestinal tissue[91], was shown to not be detectable in normal squamous epithelium and lowly expressed in Barrett's glands showing no overt intestinal phenotype (goblet cells)[92,93]. Conversely, in glands showing goblet cells, expression of CDX2 was shown to be up to 16-fold higher than non-intestinalised regions of the BO lesion[93]. Another interesting observation, was the distribution of goblet cells along the proximal-distal axis of the oesophagus. Goblet cells were mainly found proximally in the Barrett's lesions[84]. The distribution of the goblet cells coincides with the level of CDX2 expression, which also seems to be highest proximally and lowest distally along the BO lesion[94].

1.8 Barrett's metaplasia-dysplasia-carcinoma sequence

The development of OAC from Barrett's metaplasia is thought to follow a sequence from non-dysplastic/indefinite, low to high grade dysplasia (LGD/HGD), finally culminating in OAC.

The classification of dysplasia remains a difficult task for pathologist due to a lack of interobserver reproducibility[95]. In an attempt to increase classification agreement, the Vienna Classification of Gastrointestinal Epithelial neoplasia was introduced to reach consensus in gastrointestinal epithelial neoplasia nomenclature[96,97]. In this agreement, five categories for stratifying gastrointestinal epithelial neoplastic lesions were defined which increased the interobserver classification agreement:

Category 1: negative for neoplasia/dysplasia
Category 2: indefinite for neoplasia/dysplasia
Category 3: non-invasive low-grade adenoma/dysplasia
Category 4: non-invasive high-grade adenoma/dysplasia
Category 5: invasive neoplasia

Dysplasia is assessed by looking at the number and structure of glands and their folding, the number of cells (density) and their localisation within the lesion. It can also be assessed at the cell level by assessing subcellular abnormalities such as nuclei shape, polarity and nuclear staining pattern[98].

LGD is characterized by low or absent abnormality in the architecture of the Barrett's glands. Cells have eosinophilic cytoplasm with enlarged, elongated nuclei showing hyperchromasia upon staining[83]. Cell proliferation is increased in the gland where the dysplasia originates and less at the surface of the epithelium[83]. At the base of the glands, nuclei can be enlarged, pleomorphic, and show loss of polarity[83].

On the other hand, HGD is characterized by greater number of cellular abnormalities compared to LGD. The glands are more numerous, irregular in shape, appear more dilated (larger lumen), and contain necrotic cell debris. Cell nuclei are more enlarged than in LGD, show more pleomorphism and loss of polarity throughout the glands unlike LGD where this mainly occurs at the base of the glands. Cell proliferation is still increased in HGD but is found throughout the glands and at the surface epithelium, not just at the gland base [83].

1.9 Barrett's oesophagus screening and surveillance

BO is an asymptomatic condition that largely goes undetected. This most likely leads to an underestimation of its prevalence. As described in section 1.5.1, epidemiological studies suggest BO to be mostly prevalent in middle aged Caucasian males with a history GORD[99]. A systematic review and meta-analysis of the prevalence of BO in patients with known risk factors such as GORD, obesity, and smoking compared to patients without, showed the later to have an overall pooled prevalence of 0.8% compared to 3% in the former[100]. A recent study showed the proportion of BO in older patients (70 years and older) has decreased while the opposite has happened in younger age groups with 30 and 39 years showing the highest increase[101]. This is surprising as currently GORD patients are typically diagnosed with BO at a mean age around 50-59 years old[32].

Based on these known risks, The American College of Gastroenterology has recommended screening males with weekly GORD symptoms for at least five years and two or more risk factors (over the age of fifty, high waist-to-hip ratio (> 0.9), currently or with a history of smoking, familial history of BO)[102]. Similarly, The American Society for Gastrointestinal Endoscopy recommends screening Caucasian males over the age of fifty with increased GORD symptom duration, obese and smokers[103]. The British Society of Gastroenterology on the other hand, supports lowering the threshold of admission for screening to fewer than three risk factors, provided a family history of BO or OAC has already been documented in a 1st degree relative of the patient[104]. Still, the utility of screening is being questioned as most of these recommendations are based on retrospective cohorts from non-randomized studies[105].

Patients diagnosed early with BO are usually advised to participate in surveillance, which was shown in several studies to lead to improved survival from OAC[106,107]. However, other retrospective studies have shown no added benefits from surveillance[108,109].

To conclusively determine the need, cost-effectiveness, and efficacy for surveillance, Old *et al*. designed an 'ongoing' ten year follow up trial of randomized multicentre BO endoscopic surveillance versus endoscopy 'at need'[105]. Once completed, the results of this trial may also enable stratification of patients who would benefit the most from targeted surveillance.

Additionally, the development of cost-effective and minimally invasive alternatives to endoscopy based screening for BO, such as The Cytosponge™, could further incentivise more widespread screening and lower the burden on patients[110]. The Cytosponge™ is a sponge mesh attached to a string contained in a capsule, which can be swallowed by the patient. Once the capsule reaches the stomach, it dissolves and the sponge is subsequently retrieved through the oesophagus with all collected cells from the gastric cardia, gastro-oesophageal junction and oesophagus, which are subsequently analysed for BO markers (mainly TFF3)[111]. In a case control study with 1110 patients, the Cytosponge™ could detect BO of any length at a sensitivity of 79.9%. This sensitivity increased to 87.2% for segments longer than three centimetres of circumferential BO. The specificity of the Cytosponge™ for detecting BO was 92.4% compared to the current standard of endoscopy and biopsy. The predicted cost reduction in a population of patients at risk of developing BO was between 25 and 27% compared to standard endoscopy[112].

1.10 Driving factors of metaplasia of Barrett's oesophagus

Reflux, reflux components and the resulting inflammatory response thereof seem to play an important role in driving the metaplastic process. Acid alone has been shown to induce Barrett's metaplasia with a gastric phenotype[113,114].

Sonic Hedgehog (SHH) was shown to be important for the development of gastric tissues[115]. The expression of SHH was confirmed in a study comparing its expression levels in long segment Barrett's lesions[94]. Higher SHH expression was found in the non-intestinalised columnar epithelium (as indicated by the absence of goblet cells) compared to the proximal region containing goblet cells[94]. Downstream of SHH, BMP4[116] from stromal cells and the transcription factor SOX9[117] have been shown to be important in inducing columnarisation in the squamous epithelium in mice.

Treatment of primary rat keratinocytes with bile acids alone has been shown induce expression of the homeobox protein, CDX2[118], suggestive of a role in intestinalisation of Barrett's metaplasia. Exposure of mouse or human oesophageal epithelial cells to acid and bile has also been shown to induce expression of CDX2, along with other intestinal genes such as villin and MUC2[119,120]. This increase in CDX2 was abrogated through inhibition of nuclear factor-kappa B (NFκB)[121], thereby supporting a role for chronic inflammation in the development of intestinal metaplasia. NFκB expression was found in up to 60% of Barrett's metaplasia biopsies and this percentage increased to 80% in OAC, further cementing its role in the development of BO as a consequence of reflux[122,123].

RA is another component of reflux that has been shown to have a higher activity in the oesophagus of BO patients compared to a non-metaplastic group[124]. Interestingly, RA is important for squamous differentiation during embryonic development as it antagonises Hedgehog and BMP4 signalling[125]. This seems to contradict its role in the development of Barrett's[124]. However, RA has also been shown to be able to downregulate the squamous differentiation marker p63 in primary nasopharyngeal cells supporting its role in promoting columnar metaplasia[126].

1.11 Cell of origin controversy

BO is a stable metaplastic condition that generally does not progress rapidly if at all towards malignancy. This reflects the need for a progenitor cell that has undergone metaplastic changes (ie. reprogramming) presumably due to signalling induced by chronic reflux and inflammation, and is able to maintain the metaplastic tissue over a long period of time. Consequently, it is plausible that OAC, which is known to develop from Barrett's metaplasia, develops from the same metaplastic progenitor cell. However, the origin of this progenitor cell is still debated with different models suggesting different sources. The opinions on the cell of origin of BO can generally be subdivided into two groups: intrinsic or extrinsic to the oesophagus.

1.11.1 Extrinsic to the oesophagus:

Due to its continuity with the gastric mucosa, early theories proposed that BO could potentially develop through migration of gastric cells into the oesophagus to replace the squamous epithelium damaged due to reflux. Support for this migration process came from work by Bremner *et al*. using dogs that had their normal oesophageal mucosa surgically stripped[127]. Complete or partial normal squamous re-epithelialization was observed in dogs with a functioning LOS. Conversely, dogs suffering from reflux showed total or nearly total columnar re-epithelialisation with a gastric phenotype. This led the authors to conclude that a migration of gastric cells is the source of BO in dogs[127]. Furthermore, in keeping with this theory, Barrett's metaplasia contains cells with a gastric phenotype and is known to express markers also found in the stomach such as MUC5AC and TFF2[128].

Supporting this hypothesis is the work performed by Quante *et al*., who used *Lgr5-Cre-ERT/Rosa-LacZ* reporter mice to demonstrate the migration of LacZ labelled $Lgr5^{+}$ gastric progenitor cells. This effect was mediated by inducing chronic inflammation in the oesophagus and mouse forestomach[128]. In this model, Quante *et al*. used an Epstein-

Barr virus driven promoter to overexpress Interleukin 1β (IL1β) to phenocopy chronic inflammation seen in humans. IL1β has previously been shown to be increased in BO along with other inflammatory cytokines such as IL6 and IL8[129]. These mice were shown to develop oesophagitis at six months, which further progressed to columnar metaplasia at the squamo-columnar junction (SCJ) reminiscent of BO in human[128]. As a result of the chronic inflammation, migration of LacZ labelled stomach derived cells into the inflamed squamous epithelial region at the SCJ was observed. At 20-22 months, these mice developed HGD and intramucosal adenocarcinoma at the distal end of the mouse oesophagus. Furthermore, these metaplastic changes were shown to occur earlier when the mice were also treated with bile acids or a carcinogen.

A similar proposition but different cell of origin was proposed by Wang *et al.* using a p63 null mouse model[130]. As described previously, p63 is important for the replacement of the columnar epithelium in the oesophagus with a squamous epithelium during embryonic development[17]. Wang *et al*. demonstrated the columnar epithelium in p63 null mice to be positive for multiple known BO markers. In their retrospective tracing of the metaplasia in p63 null mice compared to wildtype (WT) mice, Wang *et al*. followed carbonic anhydrase 4 (Car4) expression in the embryonic columnar oesophageal monolayer epithelium of both genotypes. At embryonic day 14, Wang *et al*. showed an anterior to posterior gradient replacement of $Car4^{+}/Krt7^{+}$cells in the WT mice by p63 positive cells, but not in p63 null mice. Furthermore, they showed a small number of $Car4^{+}$ cells remained at the SCJ of WT adult mice and humans, which they defined as residual embryonic cells. Finally, they used expression of diphtheria toxin driven by Krt14 promoter to specifically ablate squamous basal layer oesophageal cells in adult mice, and showed that these residual embryonic cells were capable of expanding proximally to replace the damaged squamous epithelium[130]. Interestingly, using their ablation model, Wang *et al.* also found $Car4^{-}/Krt7^{+}$ positive cells expanding proximally into the ablated squamous epithelium[130]. This suggests the presence of another cell type at the SCJ, which does not fit the residual embryonic cell narrative, capable of giving rise to a columnar epithelium.

Jiang *et al.*[131] made a similar observation in their work. Jiang *et al*. identified cells sharing both basal/squamous markers (Krt5 and p63) and the columnar marker cytokeratin 7 (Krt7) at the SCJ. Through lineage tracing and overexpression of the basal marker Sox2, they were able to show expansion of both squamous and a transitional columnar epithelium using Krt5-Cre-ERT/Rosa-Sox2–GFP transgenic mice but not in the adjacent gastric epithelium. Purification and culture of basal layer cells from squamous epithelium ($Krt5^{+}/p63^{+}/Krt7^{-}$) at air-liquid interface and 3D organoid cultures showed that these cells only gave rise to squamous epithelial cultures. Conversely, cells from the transitional epithelium ($Krt5^{+}/p63^{+}/Krt7^{+}$) generated non keratinized epithelium. Jiang *et al*. further showed ectopic expression of CDX2 in the transitional epithelium to induce Barrett's like metaplasia by inducing expression of intestinal markers such as Mucin 2 (Muc2) and Trefoil factor 3 (Tff3).

An alternative theory for the development of BO outside the oesophagus was observed in rat oesophagojejunostomy models where irradiated female rats received bone marrow cells from male rats[132]. After, eight weeks of reflux exposure, rat oesophagi showed ulcerative oesophagitis and intestinal metaplasia. Interestingly, using fluorescent in situ hybridization, the authors showed cells within the normal squamous

and columnar oesophagus possessing a Y-chromosome in the female rats. This led the authors to conclude that circulating multi-potent progenitor cells from the bone marrow could contribute to the development of BO[132]. This however most likely reflects solitary inflammatory cells in the inflamed tissue from donor bone marrow cells as there was no clonal expansion of the donor cells in the metaplastic tissue, which would have been expected if these were progenitor cells for the columnar epithelium.

1.11.2 Intrinsic to the oesophagus:

Within the oesophagus a likely source of Barrett's metaplasia could be the squamous epithelial cells, which could directly convert into columnar cells through transdifferentiation[133]. Evidence supporting this mechanism come from scanning electron microscopy data, showing cells at the junction of Barrett's and normal squamous epithelium with both features of squamous and columnar cells[134]. This however does not concur with data showing clonality within Barrett's lesion indicating a common progenitor, as transdifferentiation is a process which occurs with no or very little proliferation[135,136].

Alternatively, molecular reprogramming of a progenitor cell in the oesophagus could give rise to BO. An example of such a phenomenon would be the multi-layered epithelium (MLE) found in BO[137]. MLE is a type of proliferating epithelium found generally near or at the neo-squamous junction[138]. This epithelium shows expression of both squamous and columnar epithelia markers and is considered a transitional stage in the development of BO[137,138]. In MLE, the columnar cells are found overlaying the squamous cells, suggesting a conversion of basal layer squamous cells due to exposure to reflux components leading to their phenotype change[138]. Studies using rat models of reflux via oesophagogastroduodenal anastomosis (OGDA) to induce bile exposure in the oesophagus [139], have shown MLE with columnar cells overlaying squamous cells similar to what is found in humans[139]. However, upon immune-histologic characterisation, rat MLE was shown to still harbour expression of basal/squamous markers p63 and Krt14, which is not seen in human MLE. The authors suggested that in their rat model, the basal layer MLE cells maintains a high degree of squamous phenotype compared to humans.

Studies done in the late 1980's by Gillen *et al.* (building on the work performed by Bremner *et al*., section 1.11.1), used dogs as an *in vivo* model system to study the contribution of gastric cardia to the development of Barrett's [140]. Gillen *et al.* performed a cardioplasty with mucosal stripping and induction of acid reflux alone or in combination with bile. To determine whether gastric cardia is the source of Barrett's, this group excised circumferentially the mucosa directly above the SCJ, leaving a two-centimetre-wide ring of intact squamous epithelium as barrier to prevent any potential SCJ or gastric cell migration, while removing another two-centimetre-wide ring above the intact mucosa. After induction of reflux, all dogs showed columnar epithelium in the lower region of excised epithelium. However, two of the five dogs also showed columnar epithelium above the intact squamous ring in the acid alone group. Importantly, from their histology, they saw what was described as a submucosal gland duct in continuity with the Barrett's epithelium[140]. These observations led the authors to conclude that the metaplastic epithelium arises from a cell originating from the submucosal gland or duct.

In humans, Coad *et al*. studied serial sections of BO tissue resections where they observed continuity between submucosal gland ducts and the overlaying Barrett's epithelium and the distinct regions of normal squamous epithelium within the BO lesions known as squamous islands[141]. These findings were further supported by Lorinc *et al*.[142] who studied the distribution of SMGs in relation to the type of mucosal epithelium overlying them. In their work, they found an accumulation of SMGs underneath the squamous islands in the BO lesions, and a direct connection between the submucosal ducts to the squamous islands[142].

Using microdissection and genetic analyses to look at mutation with BO lesions at the crypt level. Leedham *et al.*[143] showed a *p16* point mutation in submucosal duct cells and the neighbouring metaplastic BO crypt thereby demonstrating clonality between submucosal ducts and BO.

Further evidence for a potential role of the submucosal glands in the development of Barrett's has been demonstrated by Chang *et al.* [124]. Using short term *ex vivo* culture of normal oesophageal biopsy specimens, they demonstrated the role of retinoic acid (RA) in driving columnarisation of submucosal gland acini. Interestingly upon reaching the surface, the glandular acini opened up and gave rise to a glandular epithelium reminiscent of BO and showed expression of the columnar marker KRT7[124]. However, no mucin production was observed *in vitro,* possibly due to the limited culture period.

Finally, a cross-sectional study by Bartel *et al*.[144], where 515 normal squamous epithelial biopsies were examined from 106 treatment naive BO patients found a high percentage (39% and 21%) of sub-squamous intestinal metaplastic glands at 5 and 10 mm from SCJ respectively. Although not conclusive, these data also suggest BO could arise from the submucosal region.

1.12 Animal models for Barrett's oesophagus

Attempts at establishing a more fundamental understanding of the aetiology of BO, using either tissue culture or organotypic models have been insufficient. This is due to the fact that these models cannot fully recapitulate the genetic or clonal diversity of BO, or any kinds of stromal or immune interactions potentially involved in disease development or progression[145]. To date, animal models still seem the best option for studying the aetiology of BO provided they have a similar gastro-oesophageal anatomy and gastric refluxate composition. Another key advantage of animal models is the ability to assess environmental factors such as diet and exposure to certain carcinogens in the development and further progression of BO to OAC[146-148].

1.12.1 Rats

Rats have been successfully used for the study of Barrett's due to the convenient application as surgical models[57,149]. Pera *et al.*[149] developed a rat surgical model for the development of OAC through oesphagojejunostomy (OJ) with gastric preservation to induce chronic oesophagitis. Their work showed that rats receiving the combination of OJ and a low dose subcutaneous administration of a carcinogen developed adenocarcinoma[149]

Using a similar approach, Attwood *et al*. [150] performed oesophagoduodenal anastomosis (ODA). In their work they were able to show a higher rate of adenocarcinoma formation in combination with a carcinogen. Interestingly, this was not the case for the rats that underwent oesophagogastroplasty indicating that duodenal reflux components and not only gastric acid reflux are important for the development of OAC in this model. The OAC lesions however were of mixed phenotype showing cells either producing keratin or mucin[150]. Clark *et al*. further expanded on the work done by Attwood by also performing ODA and feeding rats a high-fat diet[57,150]. These rats showed a higher incidence of OAC compared to control rats. A high fat diet intake was also shown to change the composition of bile acids in the rat reflux model by increasing the concentration of taurine conjugated bile acids[146]. This led to a higher incidence of BO and OAC when compared to the control group[146], suggesting that not only the amount but also the type of fats consumed can be a significant contributing factor.

The rat OJ model was also used to support the role of cyclooxygenase 2 (COX2) in the development of OAC. Cyclooxygenase 2 (COX2) is induced by acid and bile in human intestinal and BO tissue *ex vivo* culture system[151]. Furthermore, COX2 inhibition in OAC cell lines *in vitro* reduced cell growth and increased apoptosis[152]. Inhibition of COX2 in the rat OJ model, using a selective inhibitor, reduced the rate of progression of BO to OAC, but did not affect the rate of BO development[153].

1.12.2 Mice

In addition to rats, mice have similarly been employed to study BO. Surgical approaches are more difficult in mice compared to rats due to their smaller size leading to high mortality rates post operation[154,155]. Furthermore, mice (and rats) have a different oesophago-gastric anatomy where the squamocolumnar junction extends into the mid stomach[130]. Yet, despite constant exposure to stomach acid, columnar epithelium does not naturally develop. Mice (and rats) also do not possess submucosal glands or ducts in their oesophagus and the luminal side of their mucosa is keratinised. The advantage of mice models lies in the ability to genetically manipulate them. This makes mice prime models for assessing the effect of gene gain or loss, either on their own[9], in combination with exposure to carcinogens[156] or reflux inducing surgery[157,158] in the development of BO.

Using transgenic mouse models, several genes and pathways potentially associated with the development of BO and OAC have been identified. As described previously, p63 null mice have been used as a model for studying the columnar to squamous transition of the oesophagus in mice during embryonic development[130]. Similarly, Sox2 knockout mice have been shown to develop columnar epithelium, partly due to the reduced levels of p63[9]. As mentioned in section 1.11.1, Quante *et al.* developed a mouse model of IL1β overexpression which resulted in columnar metaplasia at the SCJ[128], demonstrating the effect of chronic inflammation to drive the resulting columnar metaplasia. Work on developing intestinal metaplasia in the mouse oesophageal basal layer was performed through overexpression of CDX2 using a Krt14 driven promoter[159]. This work although not leading to an intestinal phenotype in the mouse oesophagus, showed the effect of CDX2 on decreasing cell adhesion between cells resulting in more leaky epithelium through which reflux could easily diffuse[159].

1.12.3 Dogs

Dogs are an attractive model for the study of BO as they can develop BO and OAC spontaneously similar to humans[160,161] In addition, dogs have a very similar gastroesophageal structure to human. Furthermore, as with humans, dogs contain submucosal glands in their oesophagus, which rodents also lack, suffer from reflux and develop BO and OAC[162]. Aside from difficulties in laboratory handling of dogs, dog models for Barrett's require a long time for diseases development and progression. Dog reflux models can take up to one year at the earliest to show Barrett's mucosa, provided their normal oesophageal squamous mucosa was surgically removed prior or at the time of reflux induction[127]. If the mucosa was kept this could take up to three years[163]As described in section 1.11.1, Bremner *et al*.[127] used oesophageal mucosal stripping and increased gastric reflux to demonstrate that the ensuing columnar epithelium in the distal oesophagus was acquired upon squamous epithelium destruction. In their work, Bremner *et al*. surmised that the columnar lined epithelium was derived from junctional or gastric epithelia through creeping substitution[127]. However, this was contradicted by the work performed by Gillen *et al*. as described in section 1.11.2. Similar work by Li *et al*. showed that acid suppression during the healing process of the stripped mucosa promoted squamous epithelium regeneration in addition to columnar epithelium compared to only columnar epithelium in the non-acid supressed control group[164].

1.12.4 Pigs

Similar to dogs, pigs are close to human biologically but are difficult to handle in a laboratory setting. Pigs share the same properties as humans and dogs in their oesophagi. They possess a non-keratinised epithelium; and submucosal glands[165]. Pigs have recently been used in surgical ablation models to show a contribution of SMGs in the repair of the normal oesophageal epithelium[166]. This was done in the absence of induced reflux so no reports are available on the capacity of pigs to develop Barrett's or their suitability for such a task.

1.13 Organoid cultures as *in vitro* system for modelling healthy and diseased tissues

The past decade has seen great advances in the development of organoid culture system as a way to isolate, model, and better characterise the role of multiple signalling pathways in tissue development and disease. Organoids can be defined as, a 3D cell construct approximating several features of the organ the stem/progenitor cells originate from[167].

Recent advances in understanding the composition and dynamics of diverse stem/progenitor cell niches have allowed for the generation of organoids from diverse tissue types (primarily of epithelial origin)[168]. This is achieved by taking advantage of the self-organizing capacity of stem/progenitor cells, which is allowed to take place in a laminin-rich extracellular matrix, and with the necessary morphogens and growth factors provided.

Some of the features of organoids are: composition of multiple cell types characteristic of the organ being modelled derived from the stem/progenitor cells, similar function, and cellular organisation specific to the organ of origin[167].

Organoid cultures from normal adult stem cells have been popularised by the seminal work of Sato *et al*.[169]. Using whole mouse intestinal crypts or single purified Lgr5 intestinal crypts cells, Sato *et al*. were able to grow single layered spheroids with a central lumen showing a highly polarized epithelium. These spheroids show crypt like structures expanding in outward direction. In these structures, the enterocyte brush borders are located at the luminal side as are Paneth and goblet cell secretions. Conversely, the basal side of the spheroids is exposed to the Matrigel providing the necessary signalling to maintain the crypt stem cells alongside the growth factors in the culture media[169]. These data also shed light on the importance of the Wingless (Wnt) and Lgr5 signalling in the maintenance of intestinal stem cells in the crypts[170].

This culture system has since been adapted to other organs such as the stomach where Lgr5 positive stem cells were also found at the base of the pyloric glands in the mouse stomach[171]. Interestingly, gastric organoids could also be grown from Troy positive chief cells in the gastric corpus which were shown to dedifferentiate upon damage *in vivo* and give rise to other lineages in the stomach[172,173].

In the liver, Lgr5 positive cells appear in bile duct cells upon liver damage. Similar to the intestine and stomach, culture of these cells in a modified organoid media allowed for their expansion *in vitro*[174]. Interestingly, these organoids were shown to express both liver and bile duct markers and could be pushed into a more differentiated liver phenotype by changing the composition of the culture media[175].

Organoid cultures were also developed for upper gastro-intestinal tract tissues such as the salivary glands and the oesophagus. Nanduri *et al*.[176] demonstrated the potential use of salivary stem cells as a potential avenue for curing Xerostomia due to hyposalivation. However, the culture media used by Nanduri *et al*. was different from the organoid culture media used for the intestine, stomach and liver organoid cultures. This was addressed in a follow-up study by Maimets *et al*.[177] where robust Wnt activation was shown to allow for the growth of salivary gland organoids. These organoids were shown in mice models of Xerostomia to be able to restore saliva secretion.
Mouse oesophageal basal cells were shown to grow into a pseudostratified multilayer spheroids with basal cells in contact with the Matrigel, squamoid cells at the apical layer followed by a keratinized core at the centre of the spheroids[16]. These spheroids were cultured in similar conditions as the colon or stomach derived organoids. Spheroids with no keratinisation were developed from human oesophageal tissue by Jeong *et al*.[178]. In their work, Jeong *et al*. showed that a population of CD49f high and CD24 low expressing cells is enriched for high spheroid forming potential. In their work, Jeong *et al*. also showed that deletion or knockdown of the transcription factor p63 reduced spheroid formation and expression of the basal markers KRT5 and KRT14.

Collectively, these studies demonstrate the robustness of organoids as a prime tool for modelling and studying tissue homeostasis, and how disruption thereof could promote disease development. For example, cancer can be modelled using organoids derived

from wildtype stem cells. Using human intestinal organoids Drost *et al*.[179] developed CRISPR/Cas9-based mutation strategy of multiple genes involved in intestinal cancer development. In their work, they were able to show progressive independence of the spheroids from the growth factors supplied in the culture media upon mutation of a number of tumour suppressors[179]. Aside from cancer, hereditary diseases could also be modelled *in vitro*. Cystic fibrosis (CF), an epithelial disease where mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) chloride channel, results in reduced fluid transport and decreased mucus solubility in organs such as the intestine, colon and the lungs[180]. Intestinal organoids derived from CF patients did not respond to fluid transport inducing drugs *in vitro*[181]. Using CRISPR technology and homologous recombination, defects in the CFTR gene could be repaired thereby restoring normal function[182].

Altogether, these data show the potential of organoids in modelling both homeostatic and disease conditions *in vitro*. They also show the potential of using organoids as tool for translational research to develop and test therapeutic avenues for curing a variety of diseases.

1.14 Rationale

Overall, there is still a lack of consensus on the cell of origin of BO. This is important as identification of the source would allow for more targeted and effective treatments and prevent recurrence of BO and/or OAC. Most of the current hypotheses on the origin of BO are based on mice and rat models, which as described in sections 1.12, have a different anatomy compared to humans. Furthermore, development of BO in the former only happens through some effort either through genetic or surgical interventions. Surgical models are arguably the closest to a normal situation to rapidly promote the development of BO. However, the technical difficulties, the postoperative state and non-physiologic levels of acid and bile exposures created in these models represent additional points of contention.

To date, arguably dogs represent the closest model to recapitulate the aetiology of BO in human. Based on this model, the likely cell of origin of BO inferred from the dog models seems to be the oesophageal submucosal glands as alluded to by the work of Gillen *et al*. Additionally, the presence of sub-squamous intestinal metaplasia in treatment naive BO patients[144], and the BO markers found in SMGs[180], support the hypothesis of BO emerging from either the submucosal gland acini or ducts.

The potential role of the SMGs in the development of BO has not been appreciably explored due to the lack of suitable animal or *in vitro* models. In order to properly explore this possibility, it is necessary to have a reliable source of tissue to procure sufficient amounts for experimentation. Unfortunately, human oesophageal tissue is not readily available and therefore alternatives are necessary. Similar to human, dog oesophageal tissue is also difficult to procure. To circumvent these issues, pig oesophagi, which can be easily acquired in abundant amounts from abattoirs, and seem structurally similar to human (and dog), could be used as an alternative. Still, the available data are insufficient to determine whether pig SMGs can be used as a substitute for human SMGs. Importantly, a proper characterisation of the potential progenitor cells present in the SMGs of both organisms is lacking.

The goals of this thesis were: to more thoroughly characterise and compare human and pig submucosal glands acini and ducts histologically; use the newly acquired histologic characterisation, identify and isolate progenitor cell(s) present in the tissue of both species; and, finally, develop a 3D culture system that allows for the *in vitro* maintenance of isolated progenitor cell(s) and test their capacity to undergo Barrett's-like metaplastic changes.

1.15 Hypothesis

The submucosal gland cells of the oesophagus contain progenitor cell(s) capable of developing into Barrett's epithelium.

1.16 Aims

- Determine the suitability of pig SMGs as a substitute for human SMGs through characterisation of both and assessing the degree of similarity.
- Develop a fluorescent activated cell sorting strategy for sorting and purifying potential progenitor cell(s) from the SMGs to culture and characterise *in vitro*.
- Develop an *in vitro* culture system to grow organoids from SMG cells and to test factors potentially involved in the development of Barrett' metaplasia.

Chapter 2

Histologic characterisation and comparison of human and porcine oesophageal submucosal glands

2.1 Introduction

Human submucosal ducts have been shown to be associated with squamous islands found within BO lesions[142]. Additionally, submucosal ducts and oesophageal squamous epithelium were shown to share basal/squamous markers p63[166,183] and KRT5[183]. Furthermore, the BO marker TFF3 can be found in human SMGs[111]. These associations, along with seminal work performed in dog models of reflux[140] among others[142,184], and the long term stable nature of BO, strongly suggest the SMGs and submucosal ducts to contain 'progenitor cells', which may play a role in the development of BO and repair of normal squamous epithelium.

The absence of oesophageal SMGs in standard laboratory animal models has greatly impeded progress in addressing the hypothesis of the SMGs containing the cell of origin of BO. This prompted the search for alternative models to be used in the laboratory. Aside from human, only dogs are known to spontaneously develop BO, which further progresses to oesophageal adenocarcinoma (OAC)[160,161]. However, both human and dog tissue are not readily available in abundant amounts to use for testing the hypothesis of SMG cells being the source of BO.

Pigs are among the few mammals known to possess SMGs in their oesophagus, however unlike humans and dogs, pigs have so far not been reported to develop BO. Furthermore, multiple other factors contributing to the development of BO, such as diet and frequency of reflux are expected to be different between pigs and humans or dogs.

To determine if pig SMGs are comparable to human SMGs and can be used as a substitute model, histologic analyses and comparisons of the SMGs from both human and pig should give more insights into these differences. Furthermore, staining for progenitor cell markers, based on known stem/progenitor markers from structurally and functionally similar tissues, might provide further supporting evidence for the presence of progenitor cell(s) within the submucosal glands and/or ducts. Additionally, this would further support the rationale of a progenitor cell's role in the development Barrett's metaplasia.

2.2 Material and methods

2.2.1 Tissues

Human distal oesophageal resections were acquired from post-mortem rapid autopsies from the Cancer tissue Collection After Death (CASCADE)[185] programme of consented patients with no prior diagnosis of BO or OAC. Resections were transported in RPMI-1640 (11875093; GIBCO; Thermo Fischer Scientific; MA USA) media with no additives on ice and stored at 4˚C until the moment of dissection. All work with human tissue was approved by the Human Research Ethics Committee of the Peter MacCallum Cancer Centre (project 18/211).

Full length oesophagi of six months old pigs were acquired from a local abattoir. Oesophagi were transported in phosphate buffer saline pH7.4 (PBS) on ice and stored at 4˚C until the moment of dissection.

2.2.2 Histology stains

Histology stains were performed by the histology and microscopy core at Peter MacCallum Cancer Centre.

For Alcian blue staining, FFPE blocks were cut at 4 µm thickness, dewaxed and rehydrated using standard protocols. Sections were incubated for 30 min in Alcian blue pH2.5 at room temperature, washed for 2 min under running tap water then counterstained with 0.5% Nuclear Fast Red solution for 5 to 10 min at room temperature. Finally, slides were washed for 1 min under running tap water, dehydrated and coverslipped.

Masson Trichrome staining was performed on 4 µm sections. After dewaxing and rehydration, slides were incubated for 60 min in Bouin's fixative at 60˚C then washed for 10 min under running tap water. Following the fixation, the sections were incubated in Weigert haematoxylin for 2 min, washed in water for 2 min, incubated in 1% Ponceau 2R/1% Acid fuchsin solution for 5 min then washed again in water. The slides were then incubated in 1% Phosphomolybdic acid solution for 3 min, washed in water followed by a 5 min incubation in 2% Light green solution then washed for 1 min in 1% acetic acid. Finally, slides were dehydrated and coverslipped.

Periodic acid-Schiff staining was performed on 4 µm dewaxed and rehydrated sections. Sections were incubated for 5 min in 1% Periodic acid followed by a wash in water for 1 min. Slides were placed in Schiff reagent for 10 min then washed for 10 min under running tap water. Finally, slides were counterstained with Haematoxylin, dehydrated and coverslipped.

2.2.3 Immunohistochemistry

FFPE blocks were cut at 4 µm, dewaxed and rehydrated. Antigen retrieval was performed in citrate buffer (10mM sodium citrate, pH 6) in a pressure cooker (125˚C, 3 min). Slides were allowed to cool at room temperature before proceeding with the staining procedure. Slides were washed once with PBS for 3 min, incubated for 10 min in 3% hydrogen peroxide (H_2O_2, 1.07209 ,Sigma-Aldrich, St. Louis, MO), washed for 3 min with Tris buffer saline + 0.5% Tween-20 (85114, Thermo Fischer Scientific)(TBST) then blocked for 30 min with blocking buffer: 5% goat serum (G9023, Sigma-Aldrich, St. Louis, MO), 1% bovine serum albumin (BSA, BSAS-AU, VIC, Australia), and 0.2% Triton-X100 (28313, Thermo Fischer Scientific) in PBS at room temperature (RT). Slides were incubated with primary antibodies (Table 2.1) overnight at 4˚C in blocking buffer followed by 3 x 3 min washes with TBST. Slides were subsequently incubated for 30 min with the appropriate horseradish peroxidase (HRP) secondary antibody (Rabbit, K4003 or. Mouse, K4001 EnVision+ Single Reagents HRP (DAKO, Santa Clara, CA) or

STREPTAVIDIN-HRP 1:1000 (ab7403, Abcam) at RT followed by 3 x 3 min washes with TBST. Slides were incubated with 3'3'diaminobenzidine (DAB, K346811-2, DAKO), washed for 2 min with water followed by haematoxylin counterstain, dehydration and coverslipping.

Table 2.1: Primary and secondary antibodies used for IHC staining

Marker	Primary antibody / dilution	Catalogue number	Manufacturer	Secondary antibody / dilution
WGA	Lectin from Triticum vulgaris (wheat) / 1:1000	L5142	Sigma-Aldrich	STREPTAVIDIN-HRP / 1:1000
UEA	UEA I Lectin (Biotin) / 1:1000	GTX01511	Genetex Hsinchu City, Taiwan	STREPTAVIDIN-HRP / 1:1000
ACTA2	Anti-Actin, α-Smooth Muscle antibody, Mouse monoclonal clone 1A4 / 1:4000	A5228	Sigma-Aldrich	EnVision+ Single Reagents HRP. Mouse / undiluted
p63	Anti-p63 antibody [4A4] / 1:200	ab735	Abcam	EnVision+ Single Reagents HRP. Mouse / undiluted
KRT5	Anti-Keratin 5 / 1:2000-1:3000	PRB-160P	Covance, Burlington, NC	EnVision+ Single Reagents HRP. Rabbit / undiluted
KRT13	Human KRT13 (clone Ks13.1) / 1:1000	CBL176	Millipore, Burlington, NC	EnVision+ Single Reagents HRP. Mouse / undiluted
KRT7	Anti-Keratin 7 / 1:4000	M7018	DAKO	EnVision+ Single Reagents HRP. Mouse / undiluted
SOX9	Sox-9 (H-90)/ 1:200	sc-20095	Santa Cruz, Dallas, TX	EnVision+ Single Reagents HRP. Rabbit / undiluted
MUC5AC	MUC5AC / 1:4000	MA 1-38221	Thermo Scientific, Waltham, MA	EnVision+ Single Reagents HRP. Mouse / undiluted
MUC2	Mucin 2 (H-300) / 1:2000	sc-15334	Santa Cruz	EnVision+ Single Reagents HRP. Rabbit / undiluted
CDX2	CDX-2 (EPR2764Y) / 1:1000	235R-15	Cell Marque, Rocklin, CA	EnVision+ Single Reagents HRP. Rabbit / undiluted
EPCAM (pig)	EPCAM Antibody (aa25-265) / 1:500	LS-C374109	LSBio, Seattle, WA	EnVision+ Single Reagents HRP. Rabbit / undiluted

EPCAM (human)	Purified anti-human CD326 (EpCAM) Antibody clone 9C4 / 1:200	324202	Biolegend, San Diego, CA	EnVision+ Single Reagents HRP. Mouse / undiluted
p75/NGFR	Recombinant Anti-p75 NGF Receptor antibody [EP1039Y] / 1:200	Ab52987	Abcam	EnVision+ Single Reagents HRP. Rabbit / undiluted
ITGA6	Integrin α6 Antibody clone F-6 / 1:200	sc-374057	Santa Cruz	EnVision+ Single Reagents HRP. Mouse / undiluted
MUC1	Anti-MUC1 antibody [115D8]	ab36690	Abcam	EnVision+ Single Reagents HRP. Mouse / undiluted
MUC6	Novus Gastric Mucin Antibody (CLH5), 0.02 mg /	NBP2-44374	Novus Biological, Centennial, CO	EnVision+ Single Reagents HRP. Mouse / undiluted

2.2.4 Immunofluorescence staining

Frozen sections cut at 8 µm were fixed in -20˚C chilled 100% methanol for 10 min, washed in PBS for 3 min, blocked for 30 min with blocking buffer (described above) followed by overnight incubation with primary antibodies (Table 2.2). After primary antibody incubation, slides were washed 3 x 3 min with TBST, followed by incubation with secondary antibody when applicable in secondary antibody buffer (1% BSA, 0.2% Triton-X100 in PBS) for 30 min at room temperature. Slides were washed 3 x 3 min in TBST, counterstained with DAPI (D9542, Sigma-Aldrich), mounted in mounting media (VECTASHIELD mounting media, VEH1000, Vectorlabs, Burlingame, CA) and coverslipped.

Table 2.2: Primary and secondary antibodies used for IF staining

Marker	**Primary antibody / dilution**	**Catalogue number**	**Manufacturer**	**Secondary antibody / dilution**
EPCAM (pig)	EPCAM Antibody (aa25-265) / 1:500	LS-C374109	LSBio	Goat anti Rabbit Alexa Fluor 488 (A-11034, Thermo Fischer Scientific)/ 1:500
CD49f-AF647	Alexa Fluor® 647 anti-human/mouse CD49f Antibody / 1:200	313609	Biolegend	NA
CD271-PE	CD271 (LNGFR)-PE, human / 1:200	130-112-790	Miltenyi Biotec, Bergisch Gladbach, Germany	NA

NA: Not applicable

2.2.5 Multiplex immunofluorescence staining

FFPE sections were prepared as described previously (section 2.2.3). Blocking was performed for 10 min at RT. Primary antibody (Table 2.3) incubation was performed for one hour at RT or overnight at 4˚C depending on the antibody (Table 2.3). Slides were washed 3 x 3 min with TBST, incubated with the appropriate HRP secondary antibody (Table 2.3) for 10 min at RT, washed again 3 x 3 min with TBST then incubated for 10 min with the appropriate Tyramide Signal Amplification reagent (TSA; Opal 7-Colour Manual IHC Kit, NEL811001KT, PerkinElmer, San Jose, CA). Slides were washed 5 x 3 min with TBST then antibody stripping was performed in citrate buffer in a microwave set at 100 mW for 15 min after which the slides were left to cool on ice for 15 min, washed 3 x 3 min with TBST then the previous steps were repeated starting from the blocking step. After the final round of staining and antigen retrieval, slides were cooled, washed as described, counterstained with DAPI, mounted (VECTASHIELD) and coverslipped for imaging.

Table 2.3: Primary and secondary antibodies used for IHC staining

Marker	**Primary antibody / dilution**	**Incubation**	**Catalogue number**	**Manufacturer**	**Secondary antibody / dilution**
EPCAM (pig)	EPCAM Antibody (aa25-265) / 1:200	1hr at RT	LS-C374109	LSBio	EnVision+ Single Reagents HRP. Rabbit / undiluted
EPCAM (human)	Purified anti-human CD326 (EpCAM) Antibody clone 9C4 / 1:200	1hr at RT	324202	Biolegend	EnVision+ Single Reagents HRP. Mouse / undiluted
p75/NGFR	Recombinant Anti-p75 NGF Receptor antibody [EP1039Y] / 1:200	1hr at RT	Ab52987	Abcam	EnVision+ Single Reagents HRP. Rabbit / undiluted
SOX9	Sox-9 (H-90)/ 1:200	Overnight at 4˚C	sc-20095	Santa Cruz	EnVision+ Single Reagents HRP. Rabbit / undiluted
KRT5	Anti-Kerating 5 / 1:2000	1hr at RT	PRB-160P	Covance	EnVision+ Single Reagents HRP. Rabbit / undiluted
p63	Anti-p63 antibody [4A4] / 1:200	Overnight at 4˚C	ab735	Abcam	EnVision+ Single Reagents HRP. Mouse / undiluted
KRT7	Anti-Keratin 7 / 1:1000	1hr at RT	M7018	DAKO	EnVision+ Single Reagents HRP. Mouse / undiluted

WGA	Lectin from Triticum vulgaris (wheat) / 1:200	1hr at RT	L5142	Sigma-Aldrich	STREPTAVIDIN-HRP / 1:1000
UEA	UEA I Lectin (Biotin) / 1:100	1hr at RT	GTX01511	Genetex	STREPTAVIDIN-HRP / 1:1000
ACTA2	Anti-Actin, α-Smooth Muscle antibody, Mouse monoclonal clone 1A4 / 1:100	Overnight at 4˚C	A5228	Sigma-Aldrich	EnVision+ Single Reagents HRP. Mouse / undiluted
ITGA6	Integrin α6 Antibody (F-6) / 1:100	Overnight at 4˚C	sc-374057	Santa Cruz	EnVision+ Single Reagents HRP. Mouse / undiluted

2.2.6 Imaging and processing

Brightfield images of IHC staining were taken using a VS120 OLYMPUS microscope. Overview images were taken using 2X objective after which areas for high magnification imaging were selected and imaged using 40X objective.

Immunofluorescence images were taken using a FV1000 OLYMPUS confocal microscope for confocal images or a BX51 microscope for immunofluorescence images. Merged images and confocal images Z-stacks were generated using FIJI (ImageJ) software.

Multiplex immunofluorescence images were taken using a Vectra Multispectral Tissue Imaging System. An overview scan was taken using the 10X objective. Regions of interests were selected using Phenochart (PerkinElmer) and imaged at 40X magnification. Vectra images were processed using InForm analysis software (PerkinElmer) to generate component tiff-files for analysis using the HALO analysis software (Indica Labs, Albuquerque, NM). HALO results were analysed and graphed using R programming software.

2.3 Results

2.3.1 Abundance, localisation and distribution of SMGs in human and pig oesophagus

BO is known to extend proximally from the SCJ, this should coincide with the presence and distribution of human SMGs if they are the source of Barrett's metaplasia.

Figure 2.1A shows a longitudinal section of normal human oesophagus including the SCJ, marked as the Z-line in Figure 2.1B, indicating the end of the oesophagus and beginning of the gastric cardia. Figure 2.1C shows a normal SMG in the submucosal region with submucosal duct underneath.

To assess whether normal human SMGs fit the historical localisation of BO and whether there are any correlations, SMGs were dissected from normal oesophagi resected post-

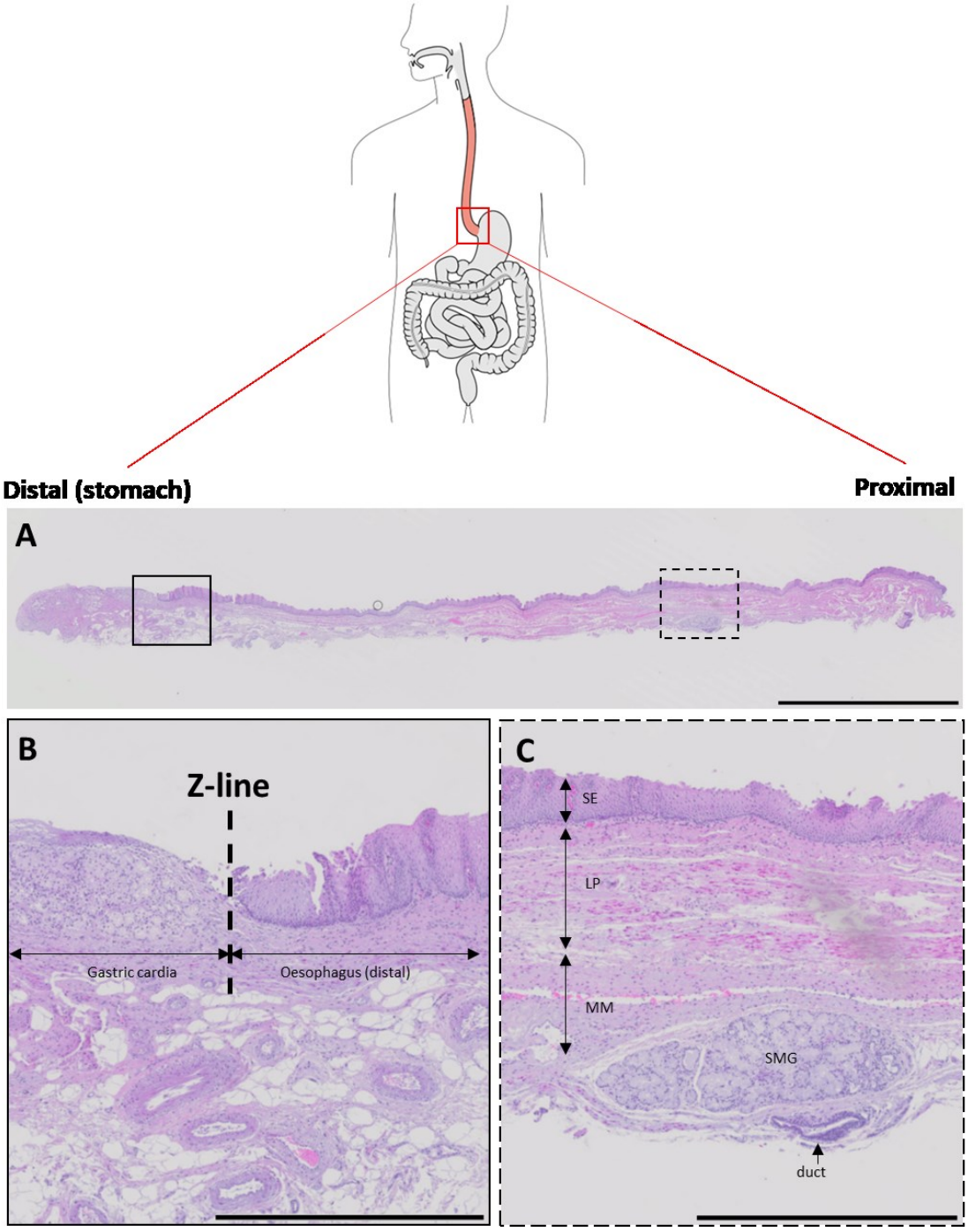


Figure 2.1: Resection of human distal normal oesophagus and proximal gastric cardia. H&E stained sections of distal normal human oesophageal tissue (A), inset of intersection (z-line) between oesophagus and gastric cardia (B), and oesophageal SMG approximately 15 mm from z-line (C). Scalebars= (A) 5 mm, (B) and (C) 1 mm. SE: Squamous epithelium, LP: Lamina propria, MM: Muscularis mucosa, SMG: submucosal gland.

mortem and counted at one-centimetre intervals relative to the SCJ (Figure 2.2A). The quantifications show a trend towards a decreasing number of SMGs from distal to proximal direction except in one case (19.002.62) where the trend is reversed. This prevalence of SMGs above the z-line coincides with the region where BO is diagnosed. There was a clear difference in number of SMGs found in the different donor samples when normalised to resection length (Figure 2.2B), indicating high inter-person variation in the density of SMGs at the distal end of the oesophagus.

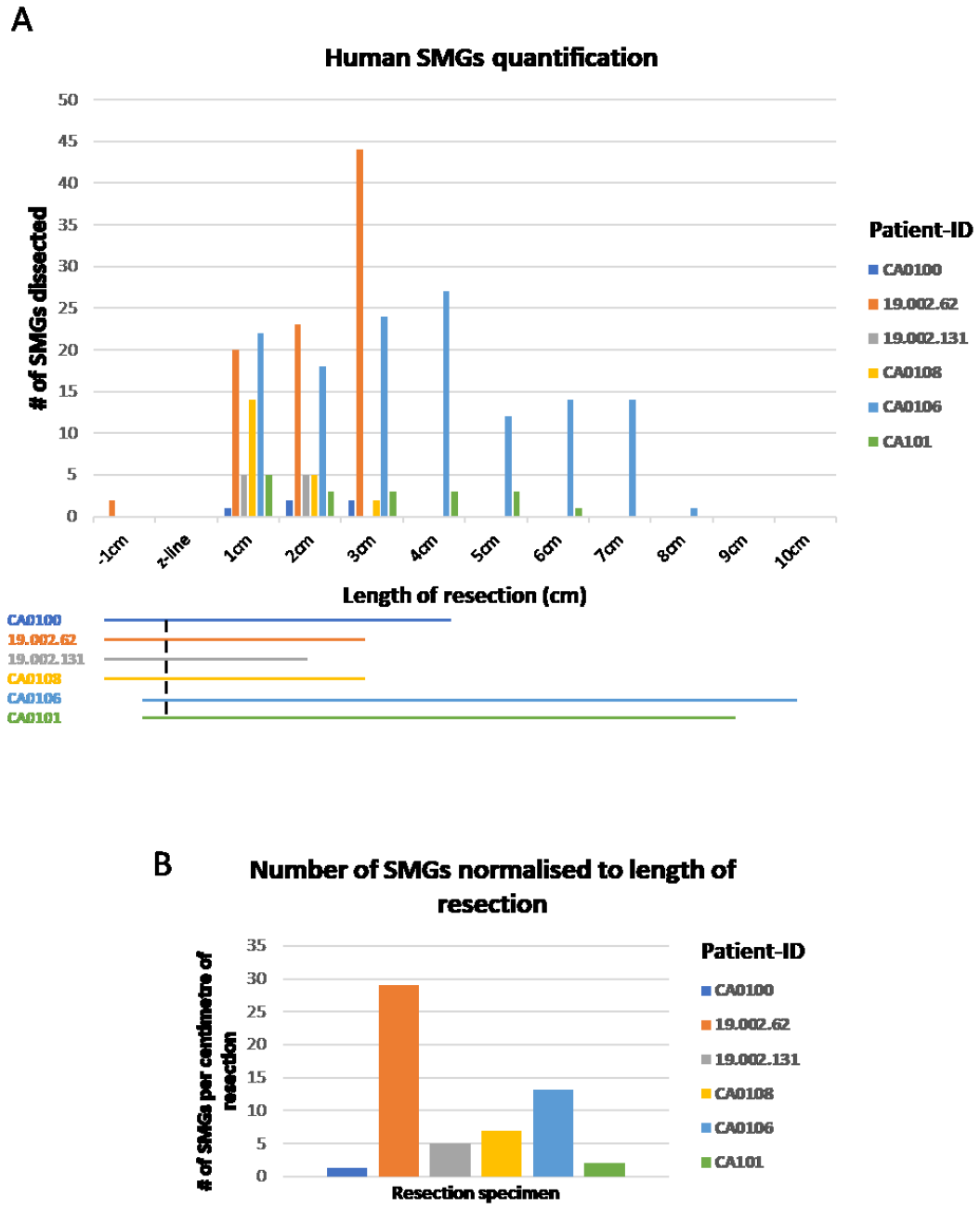


Figure 2.2: Distribution of SMGs in human oesophageal resections of CASCADE patients. (A) Quantification of the number of SMGs dissected from oesophageal resections from CASCADE patients. X-axis of graph indicates the distance from distal to proximal region in 1 cm bands. Y-axis indicates the number of SMGs found at each band. Lines below the graph indicate the length of each resection specimen relative to the Z-line. (B) Bar graph shows the number of SMGs normalised to resection length.

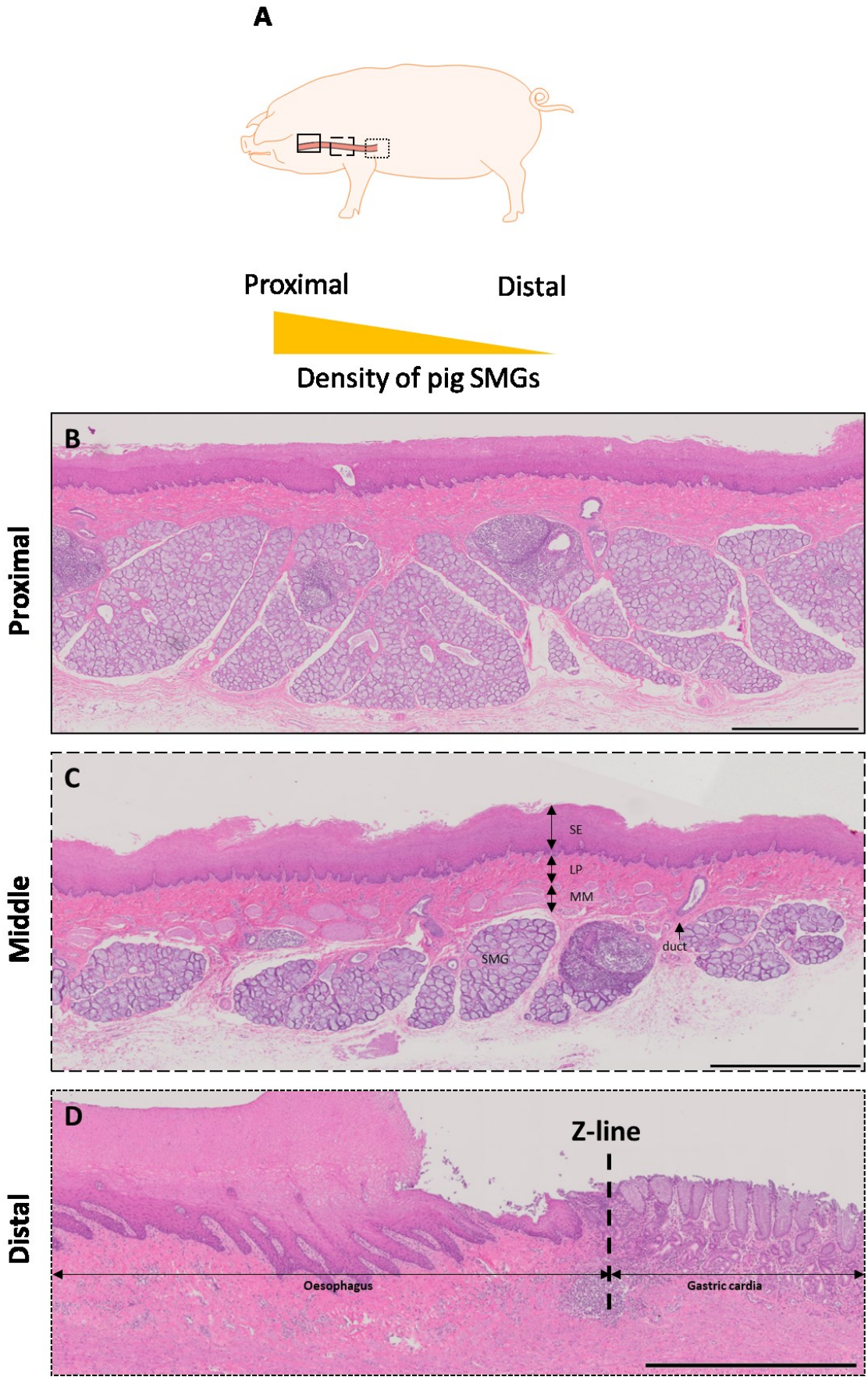

A
Proximal
Distal
Density of pig SMGs
Proximal
B
Middle
C
SE
LP
MM
SMG
duct
Distal
D
Z-line
Oesophagus
Gastric cardia

Figure 2.3: Distribution of SMGs in pig oesophagus. Schematic of SMG density distribution in pig oesophagus (A). H&E stained sections of resected mucosal and submucosal regions of (B) proximal, (C) middle and (D) distal pig oesophagus, respectively. Z-line in panel (D) indicates the intersection between the distal oesophagus and gastric cardia. SE: Squamous epithelium, LP: Lamina propria, MM: Muscularis mucosa, SMG: submucosal gland. Scalebars=1 mm.

2.3.2 Histologic comparison of human and pig SMGs shows numerous similarities

In terms of cellular composition, mucinous cells are the major cell type within the human oesophageal SMGs (Figure 2.4A and C-dashed arrow). These cells are gathered into lobules surrounded by smaller elongated cells (Figure 2.4C-solid arrow). In addition to these two cell types another cell type referred to as oncocytes[183] can been observed in certain dissected glands at varying densities (Figure 2.4A and 4B-arrow). Interestingly, oncocytes have a centrally located nucleus and a mostly cuboidal strongly eosinophilic cytoplasm.

Similarly, mucinous cells compose the majority of the pig glands (Figure 2.4F and G-dashed arrow). In addition, these mucinous cells form sacciform lobules gathering gland secretions, which are then released into the lumen upon stimulation. Similar to human, the secretory lobules in pig are surrounded by smaller elongated cells (Figure 2.4G-solid arrow).

The submucosal ducts compartment of the SMGs, in both human and pig, shows a pseudostratified two cell layer epithelium. The basal layer consists of pyramid-like shaped basal cells (Figures 2.4E and H-dashed arrow), while the second layer of the ductal epithelium is constituted of cuboidal cells lining the luminal side of the ducts (Figure 2.4E and H solid arrow).

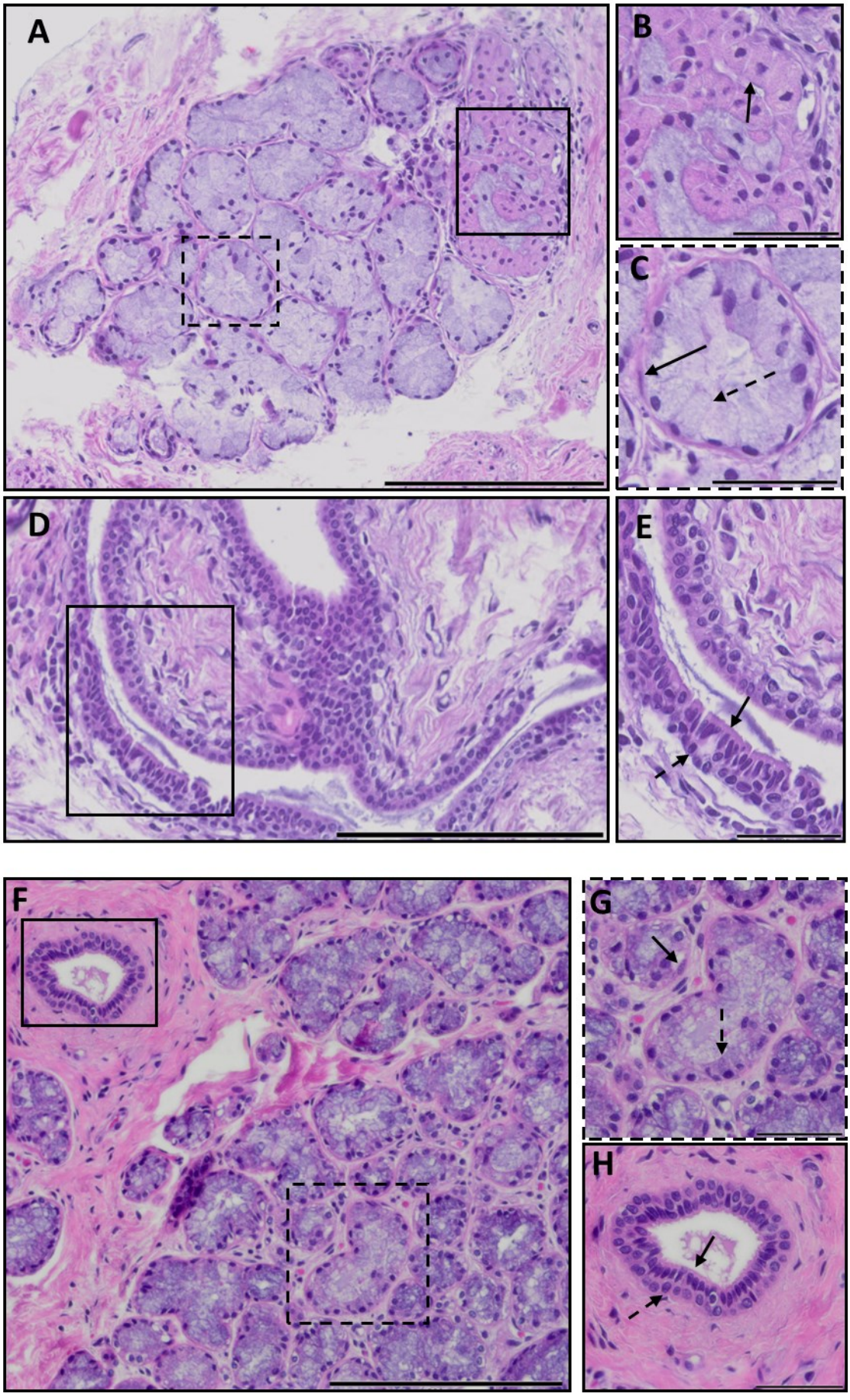

A
B
C
D
E
F
G
H

Figure 2.4: Histology of dissected SMGs from the normal human and pig oesophagus. H&E section (A) of human SMG containing oncocytes (inset B, solid arrow) and an acinar gland (inset C) shows myoepithelial (solid arrow) and mucinous cells (dashed arrow), respectively. H&E of human submucosal duct (D). Inset of submucosal duct (E) shows luminal (solid arrow) and basal duct cells (dashed arrow). H&E section (F) of pig SMG containing acinar glands. Inset of acinar glands (G) shows myoepithelial (solid arrow) and mucinous cells (dashed arrow), and pig submucosal duct (H) shows luminal (solid arrow) and basal duct cells (dashed arrow). Scalebars= (A) 200 µm, (B) 50 µm, (C) 50 µm, (D) 200 µm, (E) 50 µm, (F) 100 µm, (G) 50 µm, (H) 50 µm.

Masson's trichrome (MT) allows for distinction between smooth muscle (red) and collagen (blue) fibres[186]. MT staining shows clear differences in the size of the cytoplasm of elongated cells surrounding the gland's acinar cells in pig versus human SMGs (Figures 2.5A, A', as marked by the light red stained area and black arrows). Conversely, oncocytes in human SMGs, which are absent in pig SMGs, also stain in a similar fashion to the elongated cells (Figure 2.5A', yellow arrow). Areas surrounding the glands of both organisms stain positive for collagen (blue). Ductal cells of both human and pig do not show differences in staining with MT (Figure 2.6A, A').

Periodic Acid-Schiff (PAS) staining is used for the detection of mucin, glycogen and basement membrane[187]. PAS staining shows no obvious difference between human and pig mucinous cells potentially indicating similar types of secretions (Figure 2.5B, B'). Oncocytes in human SMGs however, are generally devoid of staining, indicating that these cells do not produce or secrete mucin (Figure 2.5B', yellow arrow). However, the lumens surrounded by oncocytes are positive for PAS staining, indicating a potential role in transport of secretions. Similarly, comparing ductal cells of both organisms, no obvious differences can be distinguished (Figure 2.6B, B').

Finally, in concordance with PAS staining, mucin staining with Alcian Blue (AB) shows positive staining in all mucinous cells for both human and pig SMGs (Figure 2.5C, C'). Oncocytes do not show positive staining for AB, although luminal content does, similar to PAS staining (Figure 2.5C', yellow arrow). Ductal cells are negative for AB in both organisms (Figure 2.6C, C').

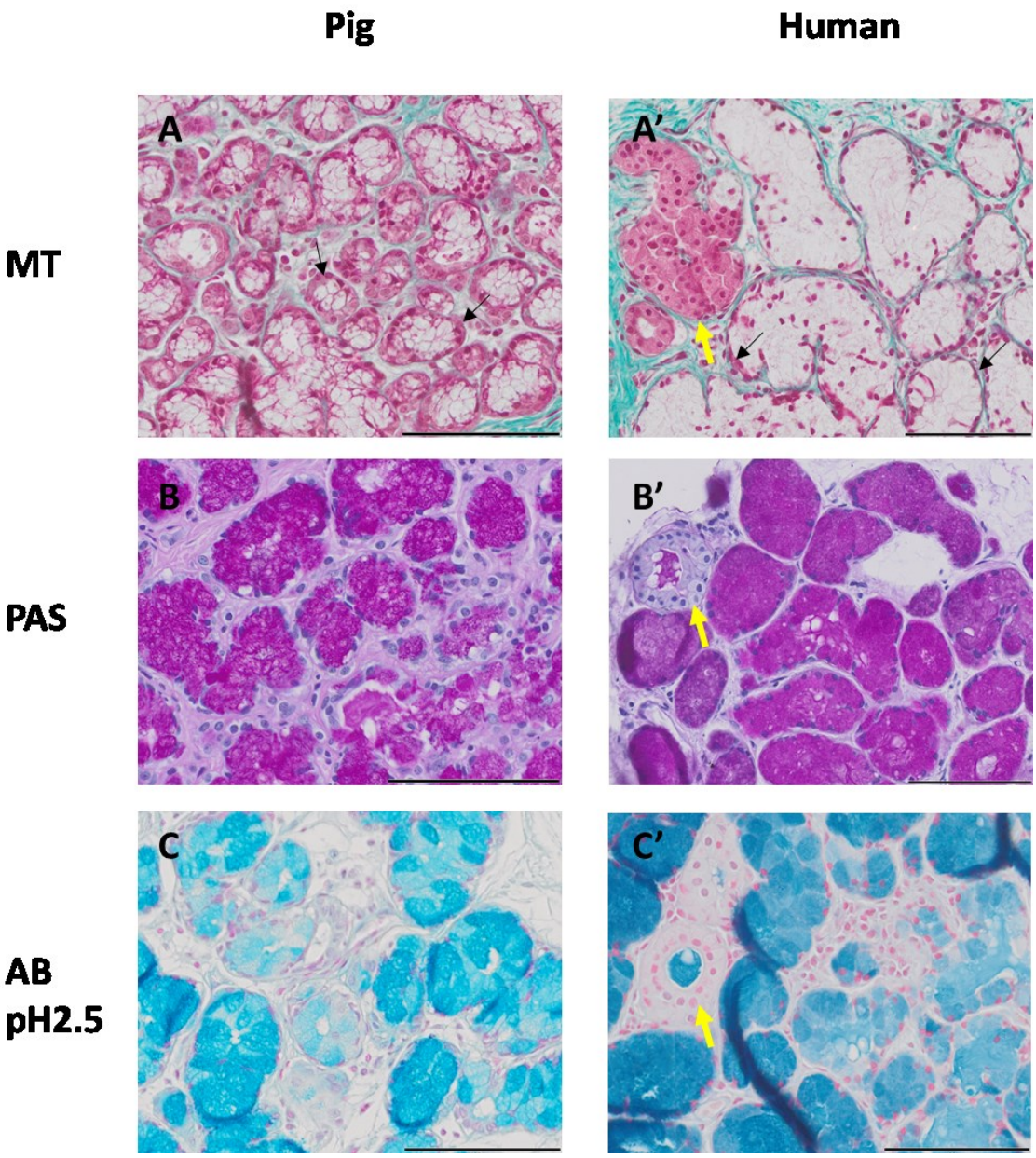


Figure 2.5: Histologic comparison of SMG compartment from pig and human oesophagi. Representative images of MT (A-A', pig n=2, human n=2), PAS (B-B', n=2), and AB pH2.5 staining (C-C, pig n=3, human n=2) of pig and human SMG acini, respectively. Yellow arrow indicates oncocytes. Black arrows in (A-A') indicate elongated cells surrounding the glands. Scalebars=100 µm.

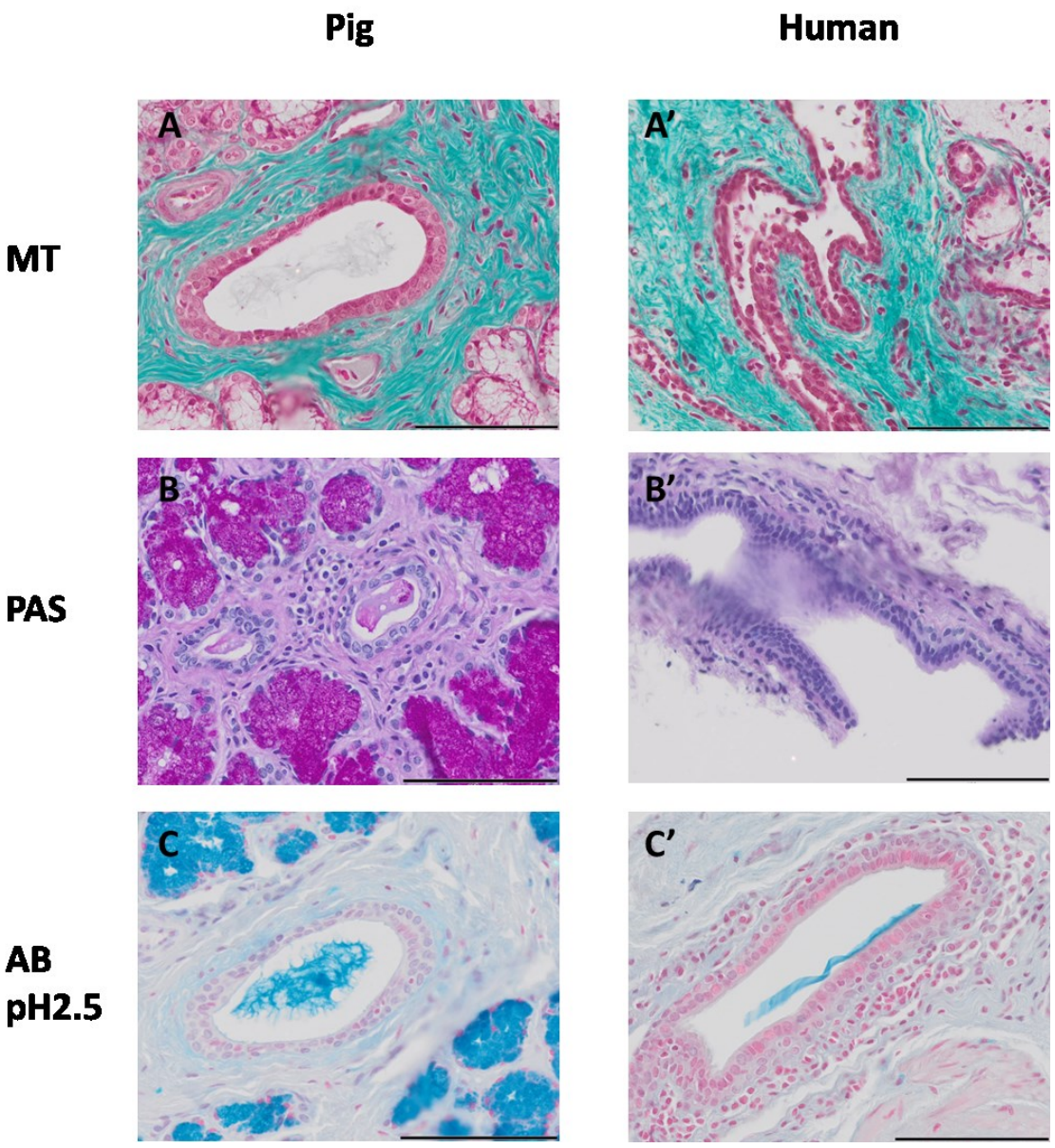


Figure 2.6: Histologic comparison of submucosal ducts from pig and human oesophagi. Representative images of MT (A-A', pig n=2, human n=2), PAS (B-B', n=2), and AB pH2.5 staining (C-C, pig n=3, human n=2) of pig and human submucosal ducts, respectively. Scalebars=50 µm.

2.3.3 Lectin immunohistochemistry shows profound differences between human and pig mucinous but not ductal cells

Staining using lectins specific to different types of sugars, which are part of mucins, was used to provide more clues as to whether human and pig SMGs secrete similar kinds of mucins. Wheat germ agglutinin (WGA) is specific to N-acetyl glucosamine and sialic acid residues while Ulex Europeaeus agglutinin (UEA) is specific for α-L-fucose residue.

WGA staining in pig SMGs is patchy, marking only a subset of mucinous cells (Figure 2.7A). Interestingly, WGA seems to also stain the elongated cells surrounding the acini (Figure 2.7A, arrows). In contrast, human SMGs show ubiquitous staining of all acinar cells (Figure 2.7A'). Ductal cells show no staining in either species, aside from the lumen where the mucin gets transported (Figures 2.8A, A').

Similar to WGA, UEA staining in pig SMGs seems to be present only in a subset of mucinous cells, but unlike WGA does not stain the elongated cells surrounding the acini (Figure 2.7B). Human acini show ubiquitous staining for UEA similar to WGA (Figure 2.7B'). These results indicate potential specialisation of mucinous cells in secreting different types of mucins in pig compared to human.

Finally, similar to WGA, both human and pig ductal cells show no staining for UEA (Figures 2.8B, B') consistent with the lack of PAS and AB staining (Figure 2.6).

2.3.4 ACTA2 marks myoepithelial cells in both human and pig and eludes to the cell of origin of oncocytes

Acinar units of secretory tissues, such as the salivary and mammary glands, are known to have myoepithelial cells that play a number of roles ranging from homeostasis to control of mucin secretion[188,189]. These organs share numerous structural and functional similarities with the oesophageal SMGs such as their glandular morphology and secretory function. Staining for α-smooth muscle actin (ACTA2), a marker of myoepithelial cells, was performed to assess whether SMG acini also contain myoepithelial cells potentially fulfilling a similar role.

Strong ACTA2 staining can be seen in the elongated cells surrounding the mucinous cells of the acini, suggesting that these cells are most likely of myoepithelial origin (Figure 2.7C and C'). Interestingly, oncocytes in human SMGs show diffuse staining for ACTA2 (Figure 2.7C', arrow), which could indicate that oncocytes have a myoepithelial origin. Ductal cells show no staining for ACTA2; however, some weak nuclear staining can be noticed in the pig ductal cells (Figure 2.8C, C').

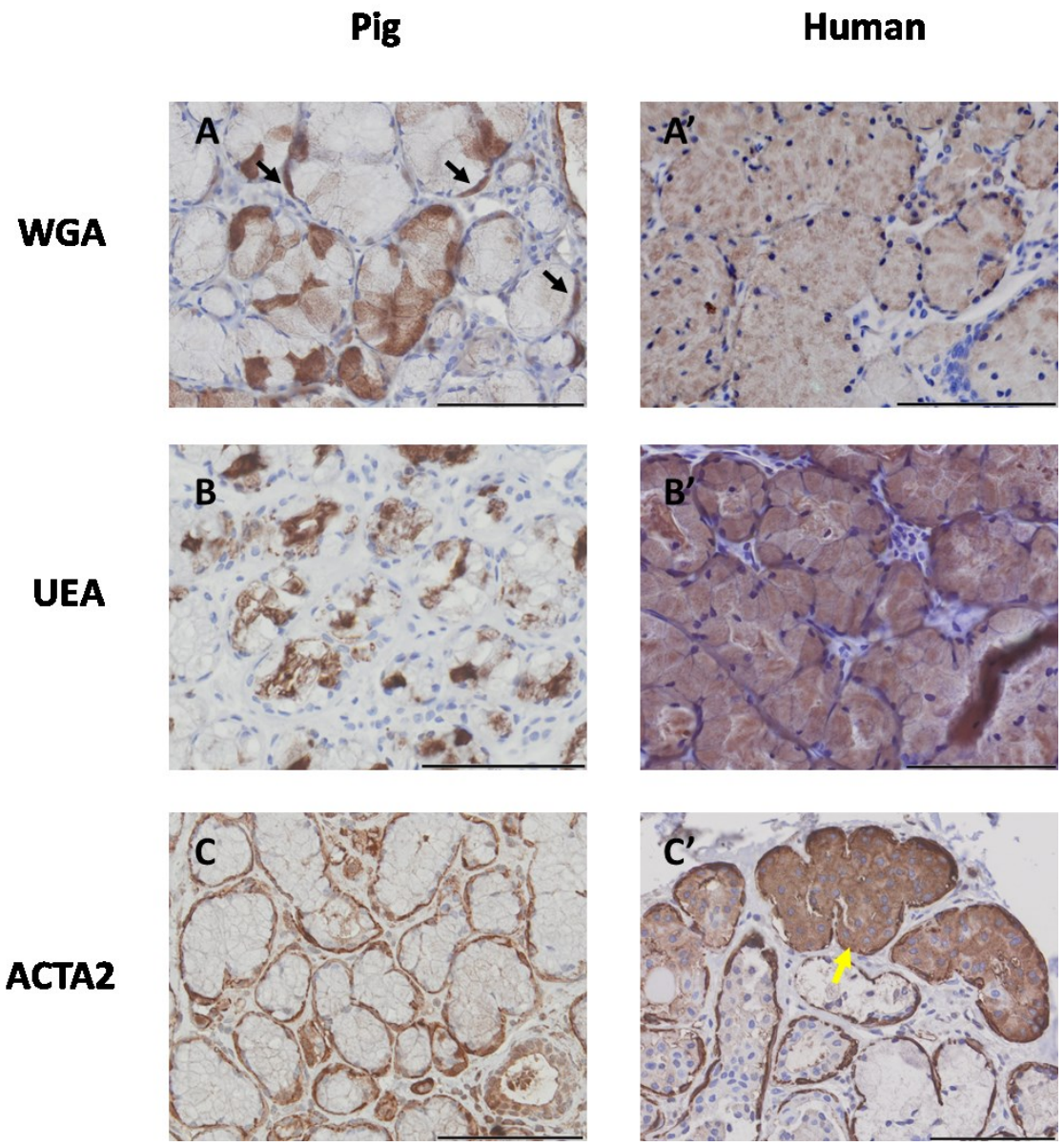


Figure 2.7: Immunohistologic comparison of SMG acini dissected from pig and human oesophagi. Representative images of WGA (A-A', n=3), UEA (B-B', n=3), and ACTA2 (C-C', n=3) staining of pig and human SMG acini, respectively. Yellow arrow indicates oncocytes, black arrows indicate myoepithelial cells positive for WGA. Scalebars=100 µm.

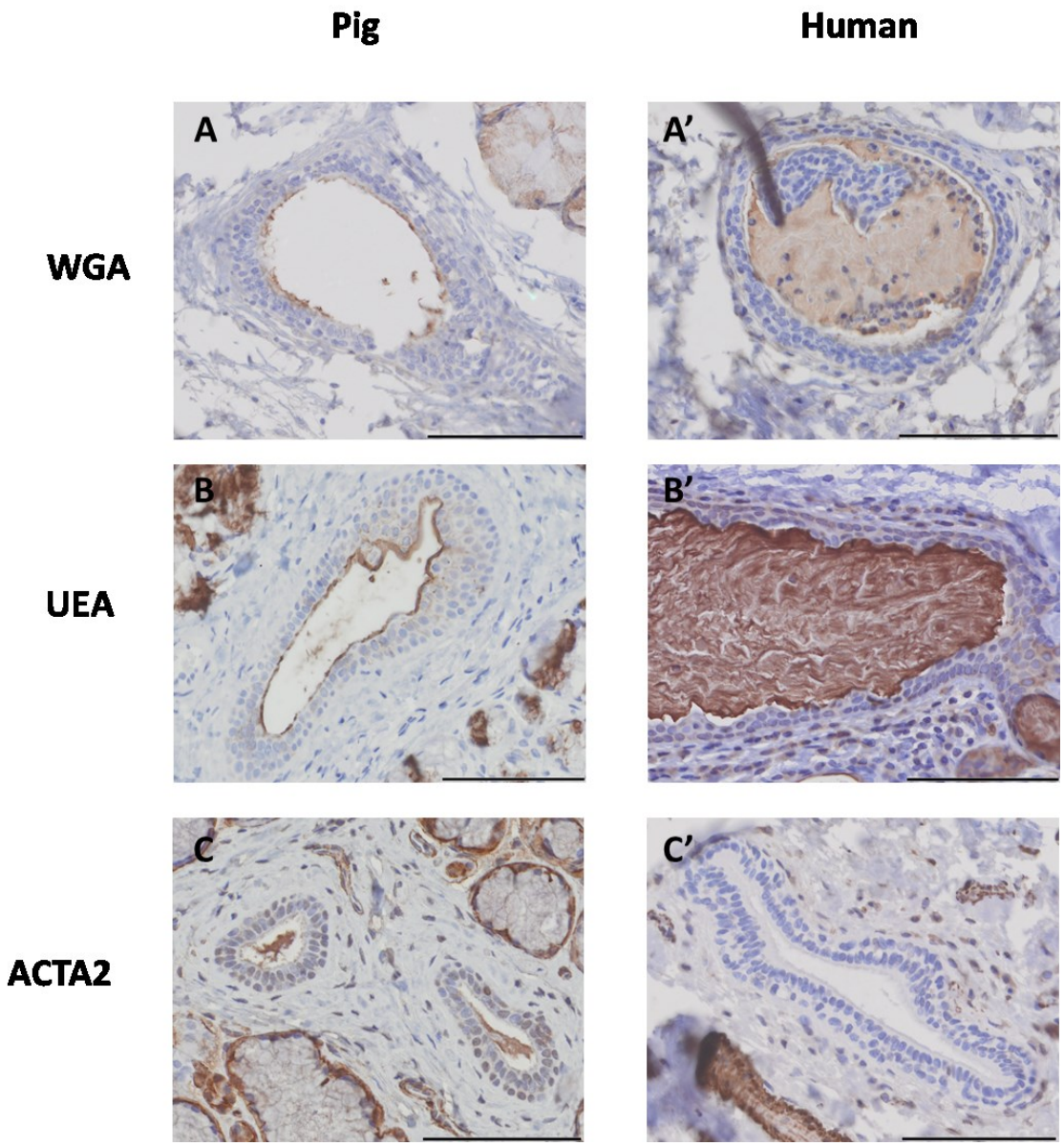


Figure 2.8: Immunohistologic comparison of submucosal ducts dissected from pig and human oesophagi. Representative images of WGA (A-A', n=3), UEA (B-B', n=3), and ACTA2 (C-C', n=3) staining of pig and human submucosal ducts, respectively. Scalebars=100 µm.

2.3.5 Basal and squamous markers can be found on the same cells of human and pig SMGs

Immunohistochemical staining for the basal and squamous marker Tumour Protein 63 (p63) shows weak nuclear staining in a subset of myoepithelial cells in both pig and human SMGs (Figures 2.9A, A' black arrows). In human SMG sections, oncocyte nuclei are negative for p63 (Figure 2.9A', yellow arrow). The majority of submucosal duct basal cells show clear p63 staining whereas luminal duct cells are negative in both pig and human tissue (Figures 2.10A, A').

Cytokeratin 5 (KRT5) another known basal and squamous marker shows weak staining in the myoepithelial cells of both pig and human SMGs (Figures 2.9B, B', black arrows). Interestingly, oncocytes in human SMGs show diffuse staining for KRT5 (Figure 2.9B', yellow arrow) similar to ACTA2 (Figure 2.7C'), supporting a potential myoepithelial origin. In addition, KRT5 staining in oncocytes seems to be predominantly sequestered to cell membranes rather than the cytoplasm. KRT5 staining in the submucosal ducts is localised to the basal duct cells, similar to p63 (Figures 2.10B, B'). Luminal duct cells show weaker and diffuse KRT5 staining in pig but are still prominent in human.

The squamous marker cytokeratin 13 (KRT13) is absent in both human and pig SMG acini (Figures 2.9C, C'). Similarly, oncocytes are negative for KRT13 (Figure 2.9C', yellow arrow). Pig luminal ductal cells show clear KRT13 staining (Figure 2.10C). Human ductal cells are also KRT13 positive, however, the intensity of staining is much weaker in comparison (Figure 2.10C'). Basal duct cells of both organisms show no staining for KRT13.

2.3.6 Submucosal gland acini but not ducts show slight differences in columnar marker distribution

Pig and human SMGs and ducts were assessed for the presence of two markers known to be present in columnar tissues. Clear SRY (Sex-Determining Region Y)-Box 9 (SOX9) nuclear staining is found in a subset of pig (Figure 2.9D, black arrows) and human myoepithelial cells (Figure 2.9D', black arrows). Oncocytes show no clear nuclear SOX9 staining but do show some weak cytoplasmic staining (Figure 2.9D', yellow arrow). Ductal cells on the other hand show clear SOX9 staining in the majority of luminal duct cells and a portion of basal duct cells (Figure 2.10D, D').

Cytokeratin 7 (KRT7) shows the biggest difference between human and pig SMG acini. Pig SMGs are largely negative for KRT7 staining except for rare distinct spindle shaped cells (Figure 2.9E, black arrows), whereas human acini and oncocytes are positive (Figure 2.9E'). Conversely, both human and pig luminal ductal cells show clear staining for KRT7, while basal duct cells are mostly negative (Figures 2.10E, E'). These findings recapitulate the results found by Krüger *et al.*[166]

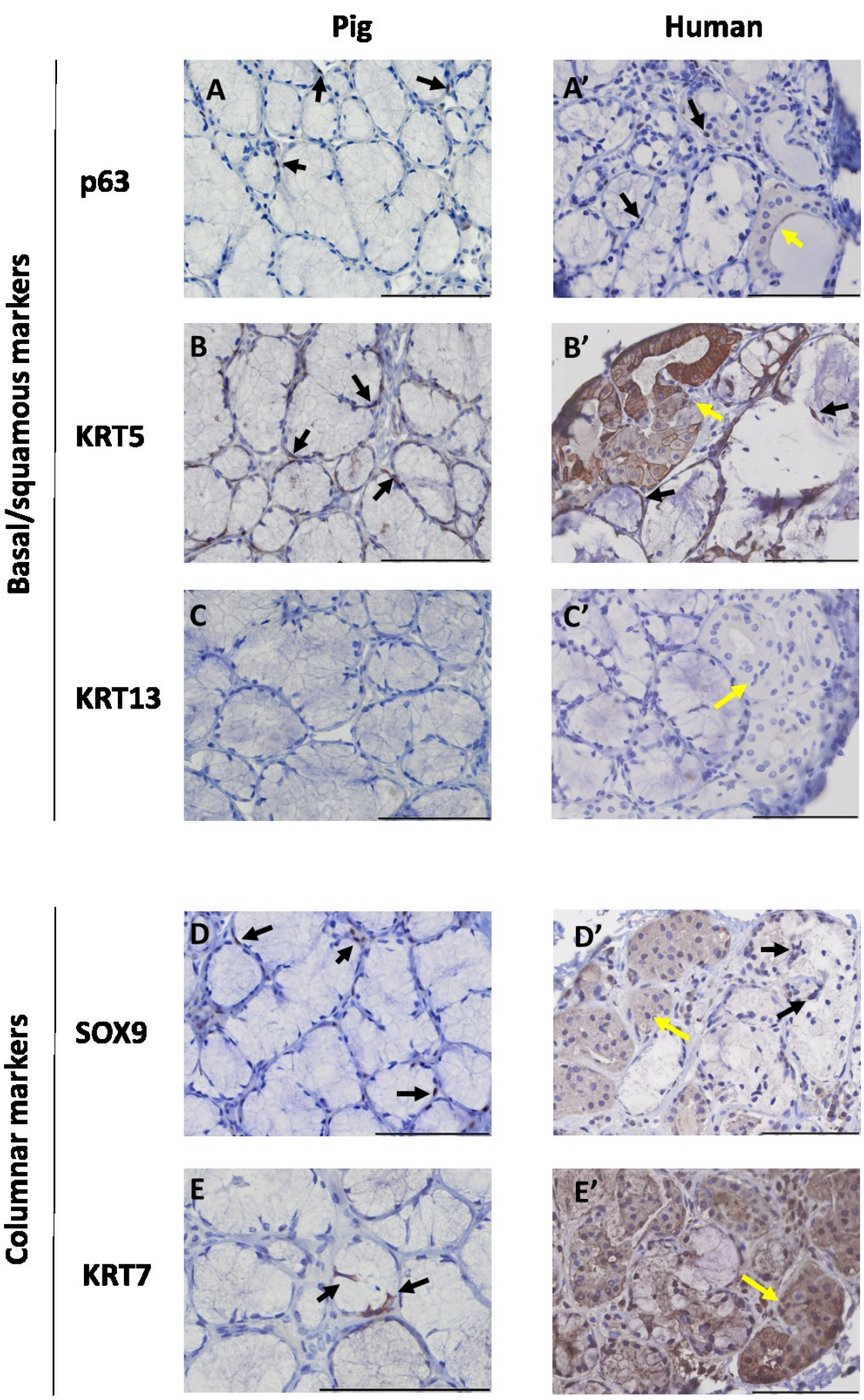

Pig
Human
Basal/squamous markers
p63
KRT5
KRT13
Columnar markers
SOX9
KRT7
A
A'
B
B'
C
C'
D
D'
E
E'

Figure 2.9: Immunohistologic comparison of known basal/squamous and columnar markers in dissected SMGs from pig and human oesophagi. Representative images of p63 (A-A', n=3) black arrows indicate myoepithelial cells positive for p63, KRT5 (B-B', n=3) black arrows indicate myoepithelial cells positive for KRT5, KRT13 (C-C', n=3), SOX9 (D-D', n=3), and KRT7 (E-E', n=3) staining of pig and human SMGs, respectively. Yellow arrows indicate oncocytes. Scalebars=100 µm.

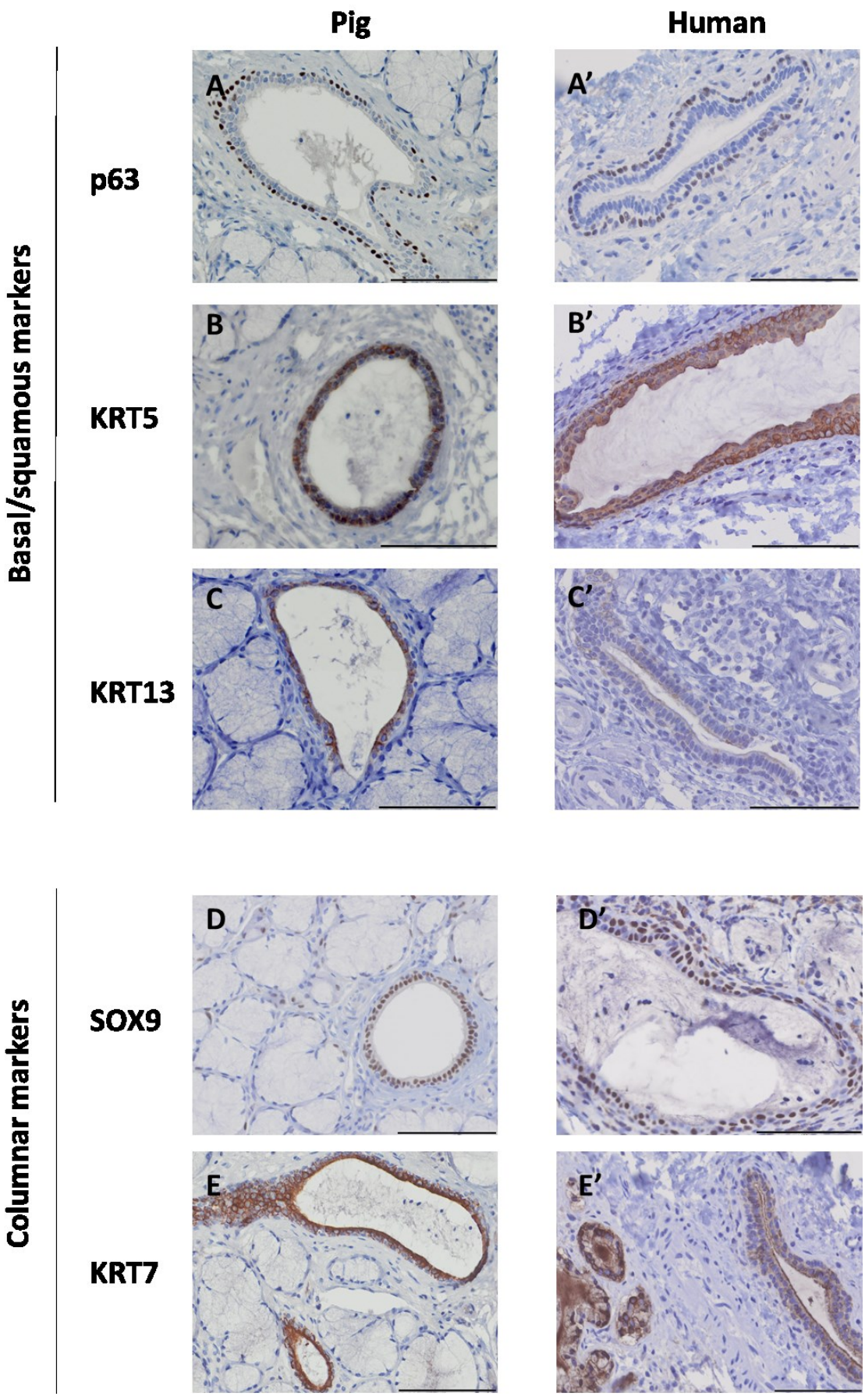

Pig
Human
Basal/squamous markers
p63
KRT5
KRT13
A
A'
B
B'
C
C'
Columnar markers
SOX9
KRT7
D
D'
E
E'

Figure 2.10: Immunohistologic comparison of known basal/squamous and columnar markers in dissected submucosal ducts from pig and human oesophagi. Representative images of p63 (A-A', n=3), KRT5 (B-B', n=3), KRT13 (C-C', n=3), SOX9 (D-D', n=3), and KRT7 (E-E', n=3) staining of pig and human SMGs, respectively. Scalebars=100 μm.

2.3.7 Known Barrett's markers are not present in human and pig normal SMGs

Human SMGs, including oncocytes, are negative for Barrett's markers MUC5AC, MUC2 and CDX2 (Figure 2.11A'-C'). Similarly, pig SMGs are also negative for the same markers (Figure 2.11A-C). Positive controls were included to ascertain staining procedure was successful.

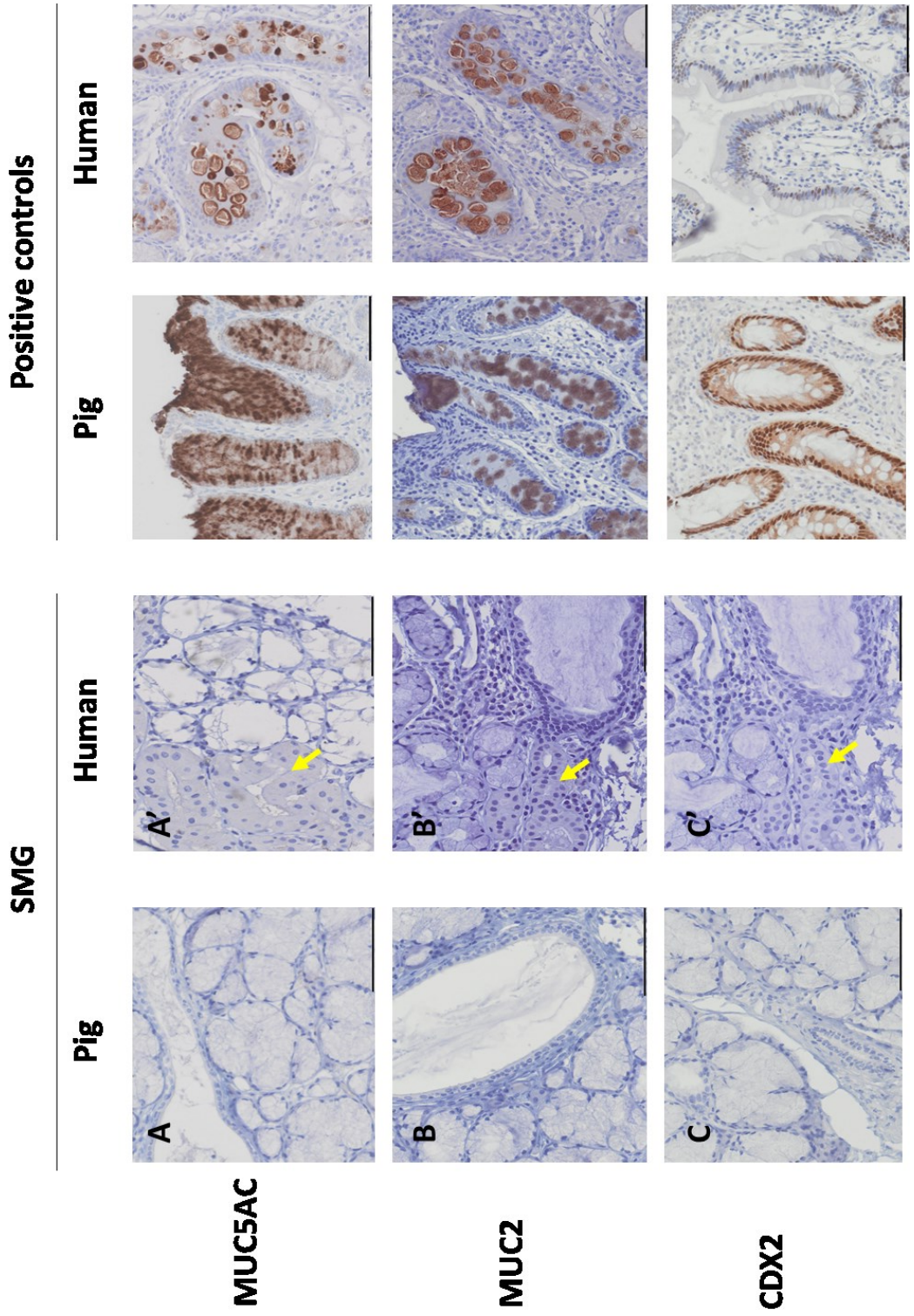
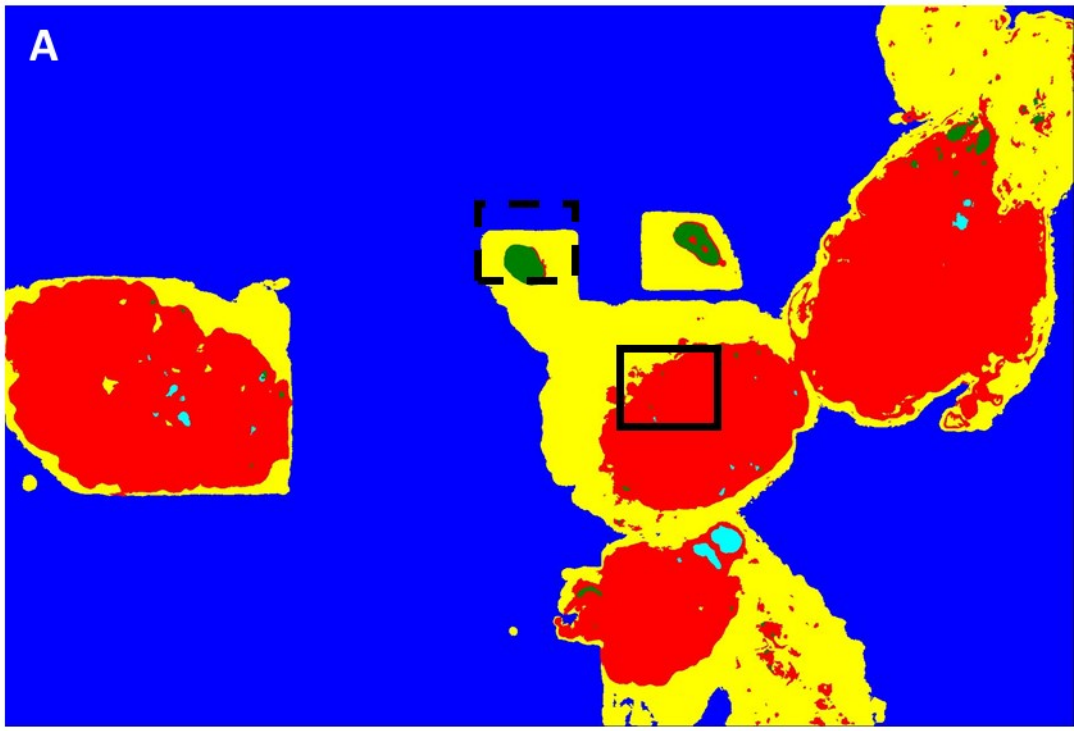
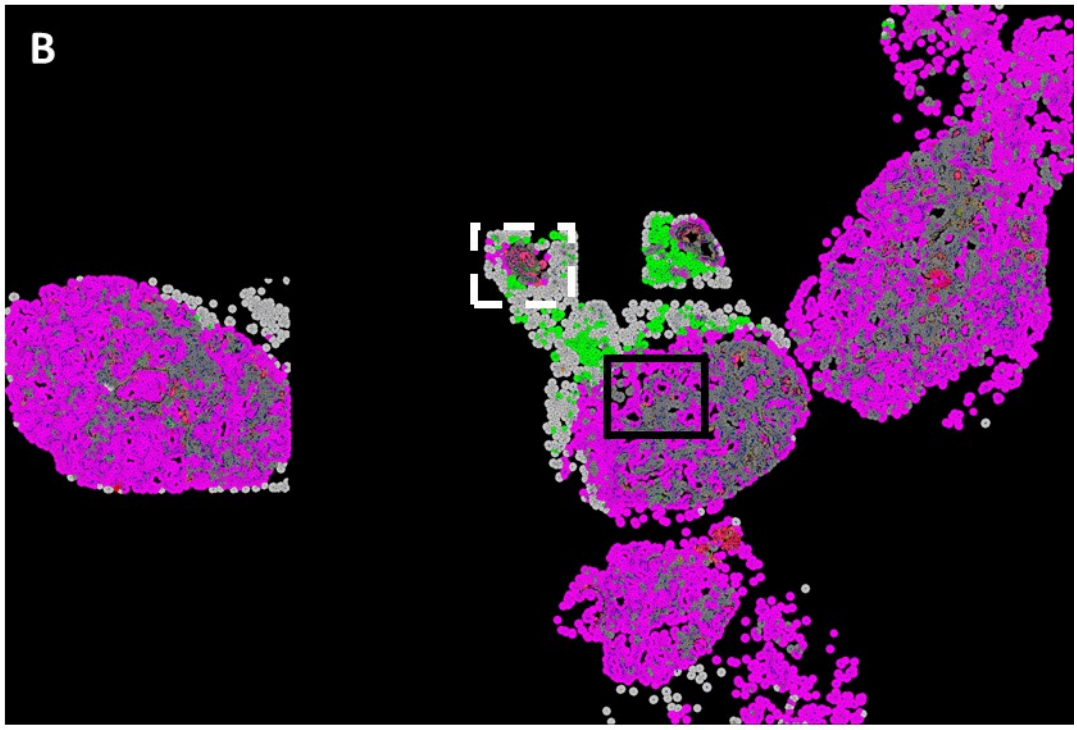


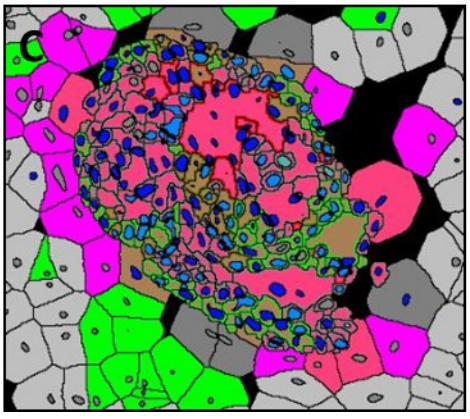
Figure 2.11: Immunohistologic comparison of known Barrett's markers in dissected SMGs from pig and human oesophagi. Representative images of MUC5AC (A-A', pig n=2, human n=3), MUC2 (B-B', pig n=1, human n=3) and CDX2 (C-C', pig n=1, human n=3) staining of pig and human SMGs, respectively. Yellow arrows indicate oncocytes. Positive controls MUC5AC: pig stomach (n=3) and human Barrett's (n=2); MUC2: pig colon (n=2) and human Barrett's (n=2); CDX2: pig colon (n=1) and human intestine (n=1). Scalebars=100 µm.

2.3.8 Multiplex immunofluorescence visualisation and quantification reproduces the observed phenotypes in IHC in pig and human duct cell subtypes

Multiplex immunofluorescence (m-IF) was used to further characterise the colocalization and intensity of multiple markers, including those described in the previous sections as well as some additional markers. M-IF was demonstrated to be a powerful tool for analysing marker distribution and signal strength at the single cell level[190,191]. Furthermore, m-IF allows for the segmentation of both tissue compartments (Figure 2.12A) and individual cells (Figure 2.12B, C and D) based on phenotypic characteristics defined by the user. This segmentation is performed using markers for which the cells are known to be positive or negative. Similarly, the phenotypic markers can also be subdivided into three distinct sub-cellular localisations (nuclear, cytoplasmic or membrane) based on antibody specificity and antigen localisation. Finally, the measured signal intensity for each marker is normalised to the exposure time, which allows for comparison of expression levels between samples. To achieve these goals, two antibody panels were optimised and used to stain pig and human SMGs.


A


B


C

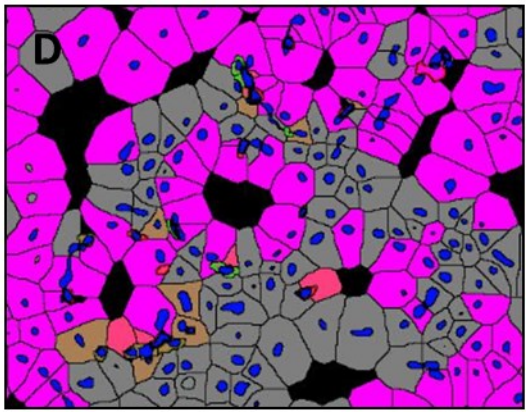

D

Figure 2.12: Representative images of tissue classification and segmentation outputs from HALO analysis software. Classification of human SMG imaged tissue (A, blue: background, red: acini, green: duct, yellow: mesenchyme, cyan: oncocytes). Cell segmentation of classified tissue (B) with insets showing areas of (C) duct and (D) acinar glands. Colours represent marker for which cells are positive; blue: Dapi, purple: KRT7, green: p75/NGFR, red: EPCAM, brown = KRT5, light grey: Stromal cells, dark grey: undefined/blurry cells imaging regions.

Figures 2.13 and 14 show m-IF using “panel one” performed on pig and human SMG sections, respectively. Panel one consists of the basal markers (KRT5, p63), columnar markers (KRT7, SOX9), epithelial marker (EPCAM), nerve growth factor receptor (NGFR/p75, a known progenitor marker[192,193]) and DAPI. In addition to the staining, Figure 2.15 shows the log transformed score (i.e. staining intensity) for each marker in the ductal compartments (basal and luminal duct cells) for both human and pig tissues.

As shown previously in Figure 2.10A and A’, p63 staining can be found specifically in the nucleus of basal duct cells, which was corroborated by m-IF (Figure 2.13 and 14) and staining intensity scores (Figure 2.15). Similarly, m-IF staining for KRT5 (Figure 2.13, 14) is consistent with the staining pattern observed using traditional IHC (Figure 2.10B and B’), with strong cytoplasmic expression in the basal layer duct cells and weaker expression in the luminal duct cells for both human and pig (Figure 2.15).

Likewise, m-IF staining and quantification validates the previous traditional IHC findings for the columnar markers SOX9 and KRT7 localisation in the nucleus and cytoplasm of luminal duct cells for both human and pig, respectively (Figure 2.13, 14 and 15).

EPCAM is known to be a membrane marker[194]. However, in pig SMGs cytoplasmic staining can also be observed, mainly in the gland acinar cells (Figures 2.17 and 18), in addition to its membrane localisation (Figure 2.13 and 14). This is also further quantified in Figure 2.15. Both cytoplasmic and membrane scores show a slightly higher score for luminal than for basal duct cells, and the width of the violin plots indicate a greater number of luminal duct cells to be positive for EPCAM in comparison to basal duct cells (Figure 2.15). Finally, p75 can be found to predominantly mark the membrane of basal but not luminal duct cells and some stromal cells (Figure 2.15).

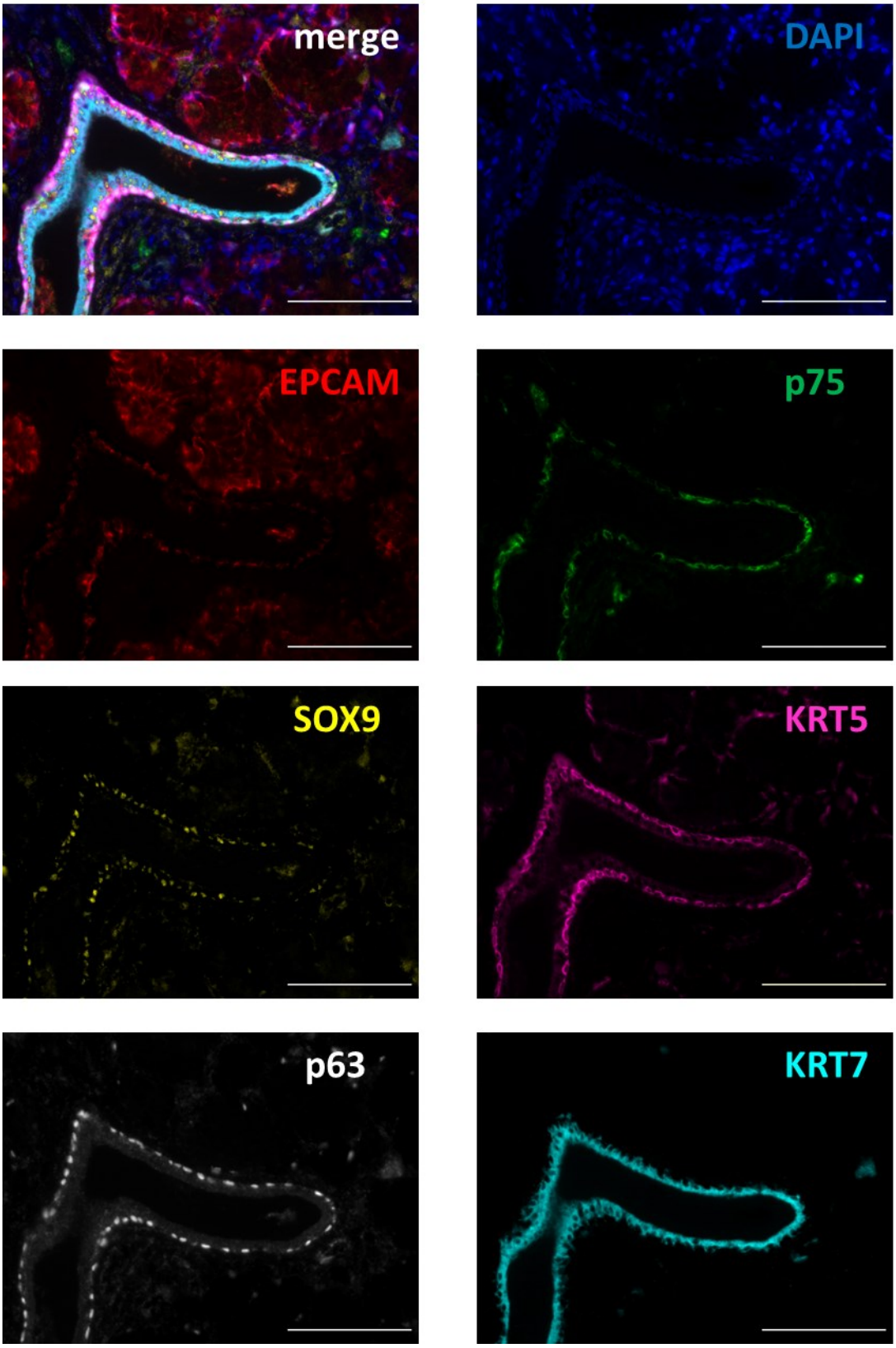


Figure 2.13: Representative multiplex-immunofluorescence (m-IF) images of pig SMG duct showing merged and individual channel images staining for panel one markers: DAPI, EPCAM, p75/NGFR, SOX9, KRT5, p63, KRT7. Scalebars=100 µm.

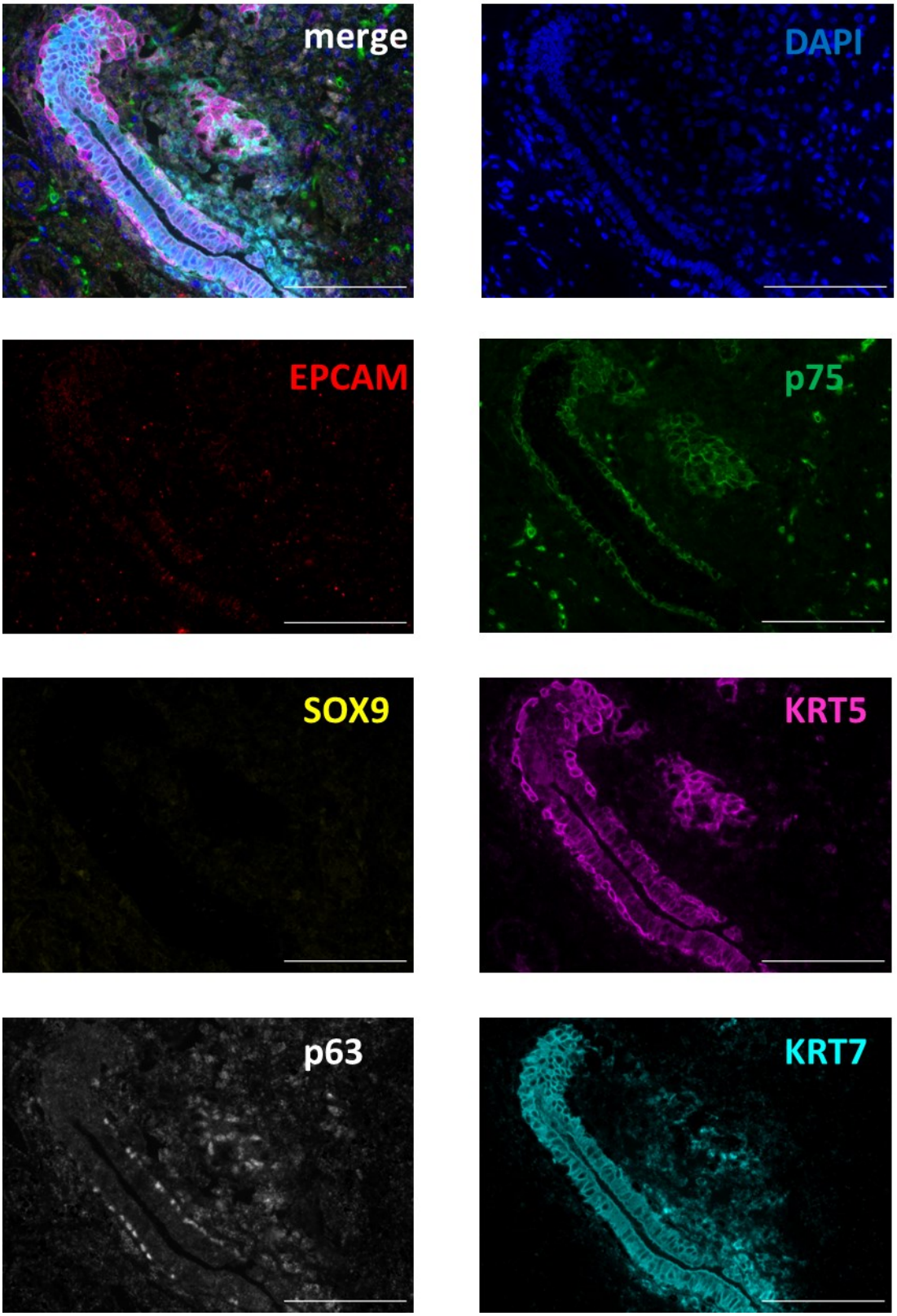


Figure 2.14: Representative m-IF images of human SMG duct showing merged and individual channel images staining for panel one markers: DAPI, EPCAM, p75/NGFR, SOX9, KRT5, p63, KRT7. Scalebars=100 µm.

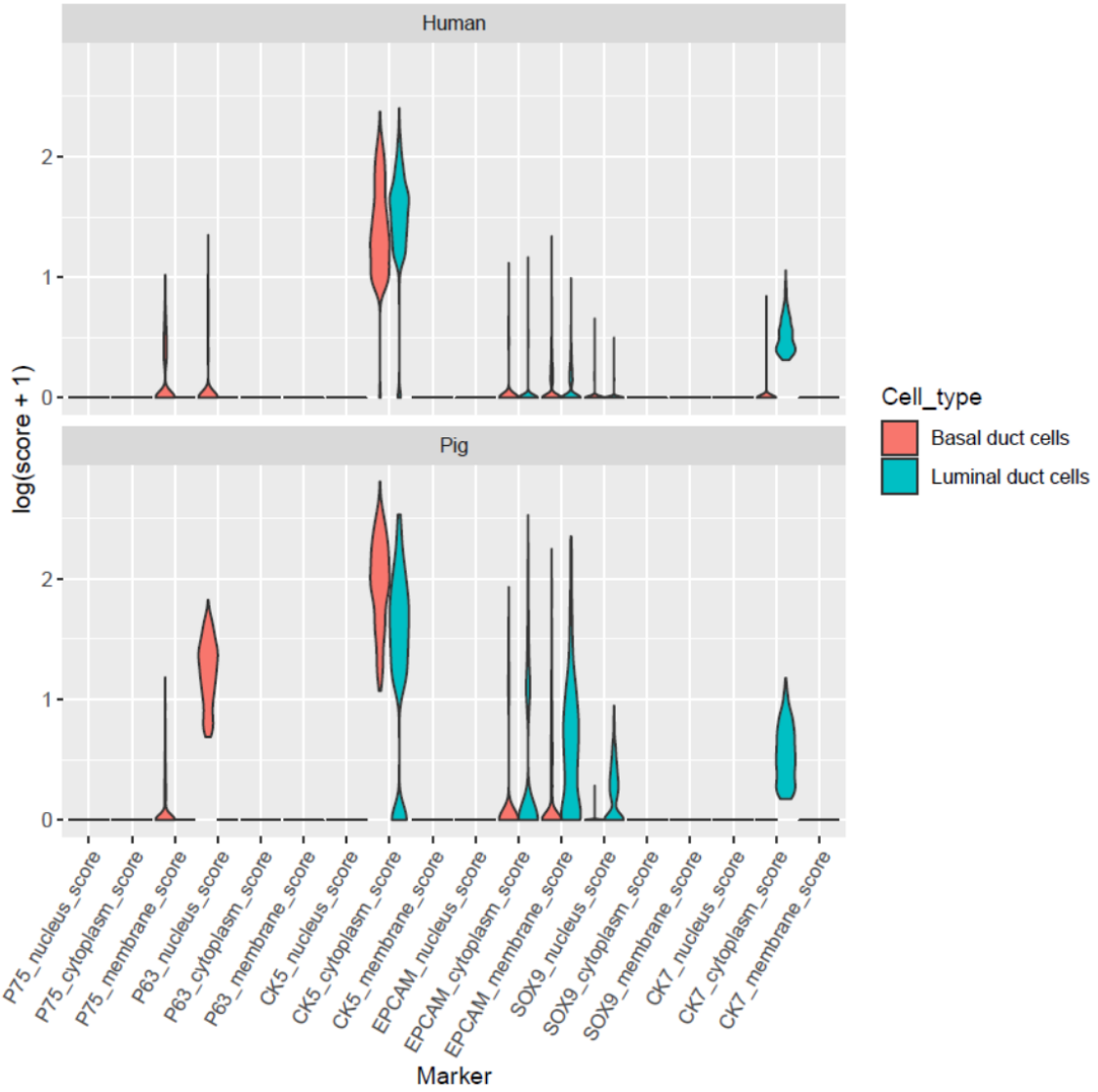


Figure 2.15: Violin plot comparing the phenotype of basal and luminal duct cells from pig and human SMG. X-axis indicates the staining marker and subcellular localisation while the y-axis shows the log(score+1) of the measured intensity normalised to exposure. Number of cells analysed: pig=1600 cells, human=2465 cells. Number of sections: pig n=1, human=2.

2.3.9 Principal component analysis shows the contribution of staining markers in ductal cell subtype specification

Dimensionality reduction using principal component analysis (PCA) allows for an easier visualisation of the differences in the contribution of the different markers to the different cell types. PCA of the m-IF data (Figure 2.16A and B) shows that the basal marker p63, KRT5 and progenitor marker p75 contribute to the distribution of basal duct cells for both pig and human as indicated by arrows/vectors. The degree of contribution however is different between the two species. KRT5 shows lower contribution for the distinction of human basal duct cells from luminal duct cells compared to pig. This is due to the prominent presence of KRT5 in luminal duct cells in human unlike in pig (Figure 2.15). Conversely, the columnar markers KRT7 and SOX9 for pig and KRT7 only for human contribute to the luminal duct cells distinction from basal duct cells. The marker EPCAM contributes to both cell types as both cell types are positive for the marker.

2.3.10 Correlation analysis can be used as surrogate for marker colocalization

Correlation analysis of marker intensity values was performed at the single cell level. Significant positive correlations were found between p63 and p75 as both these markers mark basal duct cells (Figure 2.16C and D). Similarly, p63 and KRT5 show strong correlation as basal duct cells have high KRT5 expression compared to luminal duct cells in pig but this correlation is weaker in human. SOX9 shows strong correlation with KRT7 as both markers specifically stain luminal duct cells in pig but no correlation is found in human. Membrane staining for EPCAM also shows high correlation with KRT7 in pig, which is consistent with the more prominent EPCAM staining in luminal vis-à-vis basal duct cells (Figure 2.16C and D).

Strong negative correlations can be observed between p63 and EPCAM/SOX9/KRT7 (Figure 2.16C and D) as the latter three markers are expressed in luminal duct cells unlike p63 in pig. Similarly, p75 and KRT5, which strongly mark basal duct cells, negatively correlate with KRT7 in pig. Likewise, human ducts show negative correlation between KRT7 and p75. However, human luminal ducts, which are strongly positive for KRT7, also show weak staining for KRT5. This is reflected in the low correlation between KRT5 and KRT7 in human.

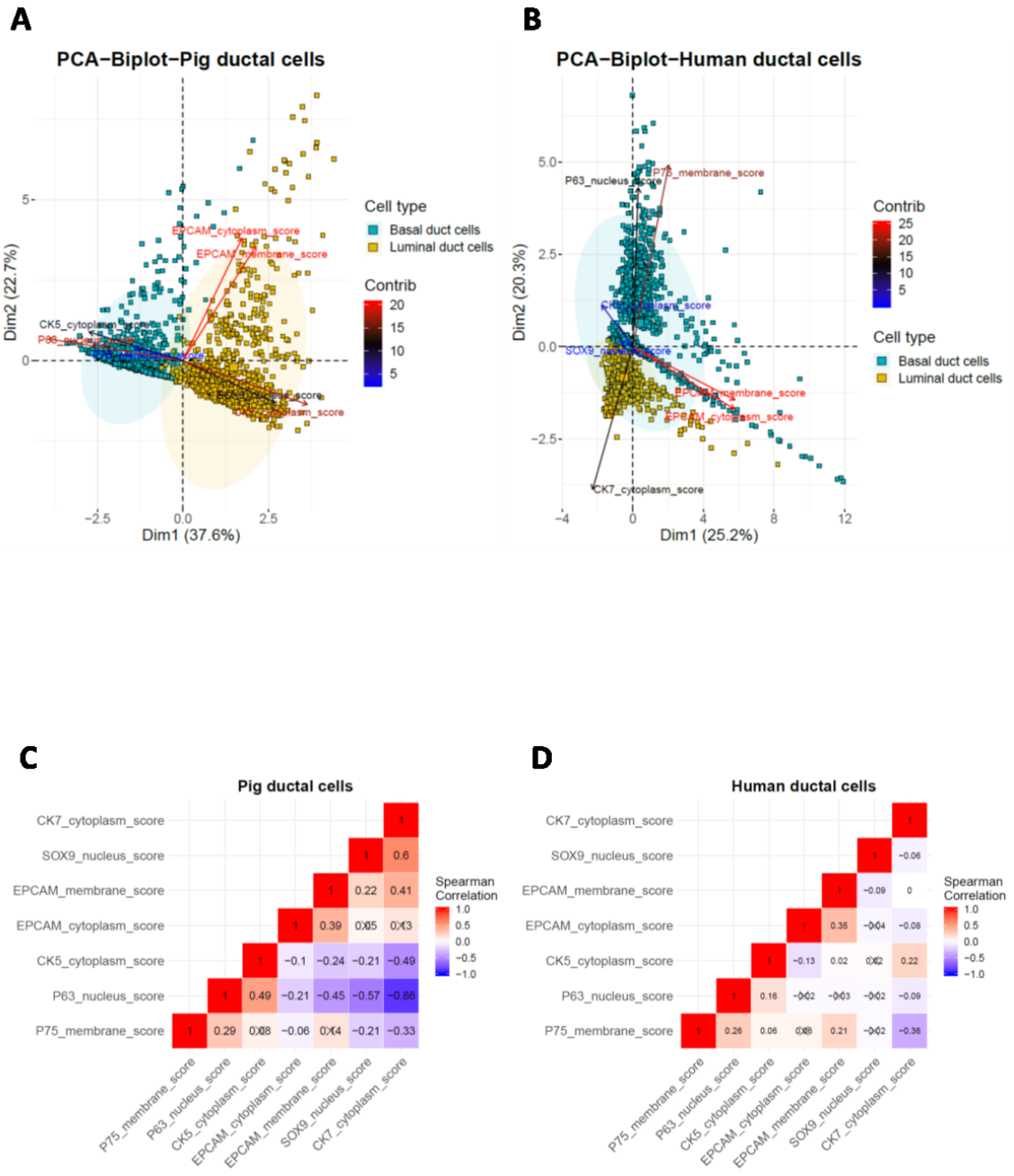


Figure 2.16: Contribution of 'panel one' staining markers for oesophageal submucosal duct cell subtype distinction. PCA biplot of the first and second principal components showing markers contributing to the distinction between basal and luminal duct cells for pig (A) and human (B). Arrows show the markers contributing to the points (cell) distribution while the colour of the arrows shows the degree to which these markers contribute. Half correlation matrix showing Spearman correlation coefficients used as a surrogate quantification of marker colocalization of the markers found in pig (C) and human (D) submucosal duct cells. Values marked with "X" did not reach statistical significance. Number of cells analysed: pig=1600 cells, human=2465 cells. Number of sections: pig n=1, human=2.

2.3.11 Principal component analysis and colocalization analysis of human and pig glands show little similarity

One of the biggest differences between pig and human is the presence of KRT7 in human gland cells in comparison to pig (Figures 2.17, 18, 19 and 20).

PCA biplot analysis shows the basal markers KRT5 and p63 to be strong contributors to the distinction between myoepithelial and gland's acinar cells in pig whereas in human only the progenitor marker p75 specifies the myoepithelial cells (Figures 2.20A and B). Interestingly, p75 does not contribute to either myoepithelial or gland's acinar cells distinction in pig. Confocal imaging of pig SMG frozen sections confirms p75 staining to be found on an EPCAM negative stromal cell (Figure 2.21). EPCAM is the only marker showing a strong contribution to the distinction of acinar cells in pig glands.

Oncocytes, which are not present in pig, show both columnar (KRT7) and basal (KRT5) markers contribute to their distinction from other cells in addition to EPCAM (Figure 2.18, 19 and 20B). For acinar cells, mainly EPCAM and KRT7 contribute to their distinction from other cell types in human.

Pig glands colocalization analysis shows mainly strong correlation between KRT5 and p63, which were shown be expressed by myoepithelial cells (Figure 2.20C). In human, the progenitor marker p75 and KRT5 show some degree of colocalization, indicating the myoepithelial cells in this tissue (Figure 2.20D). The highest correlation can be found between KRT5 and EPCAM indicating oncocytes.

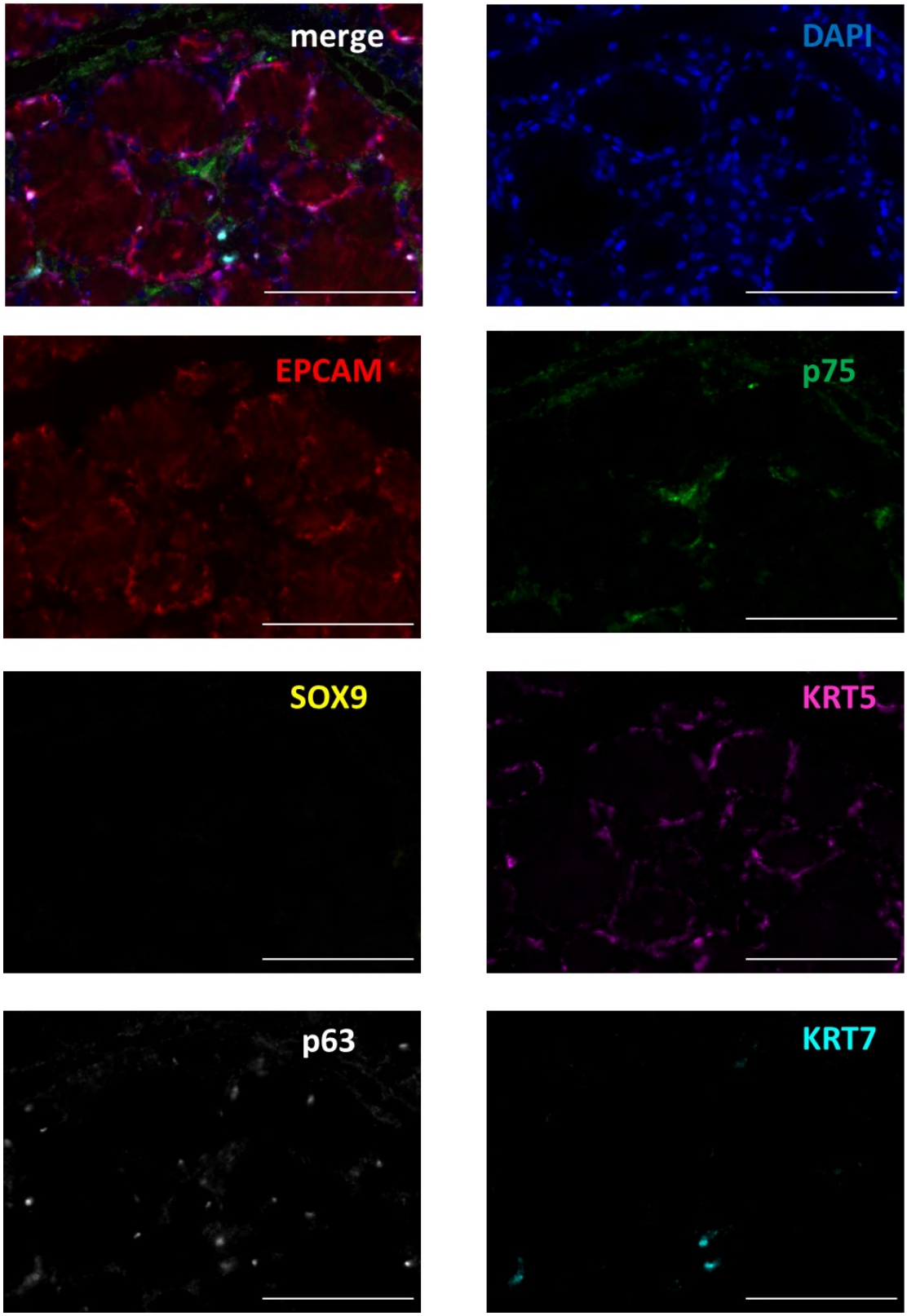


Figure 2.17: Representative m-IF images of pig submucosal gland acini showing merged and individual channel images staining for panel one markers: DAPI, EPCAM, p75/NGFR, SOX9, KRT5, p63, KRT7. Scalebars=100 µm.

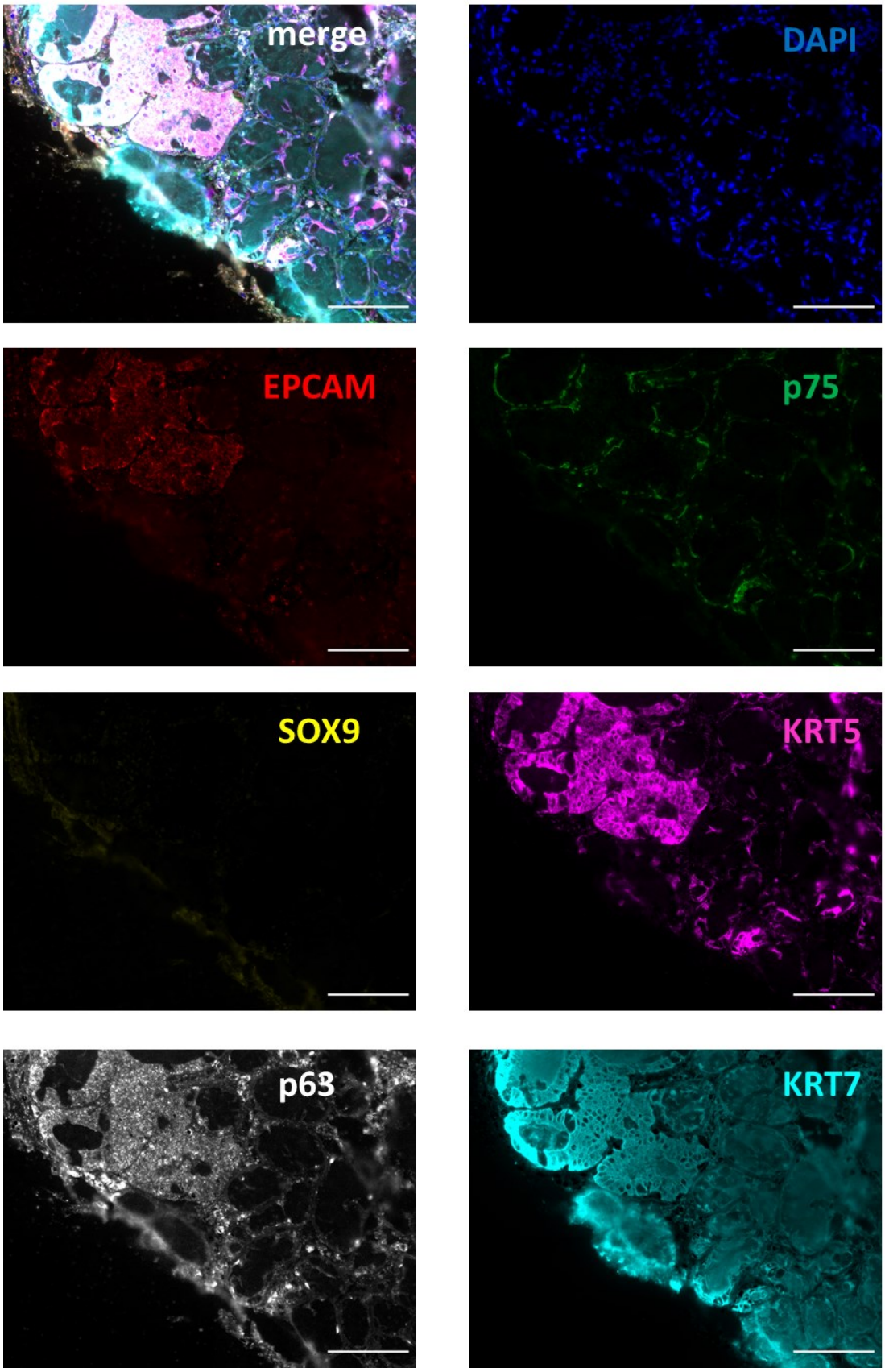


Figure 2.18: Representative m-IF images of human submucosal gland acini showing merged and individual channel images staining for panel one markers: DAPI, EPCAM, p75/NGFR, SOX9, KRT5, p63, KRT7. Scalebars=100 µm.

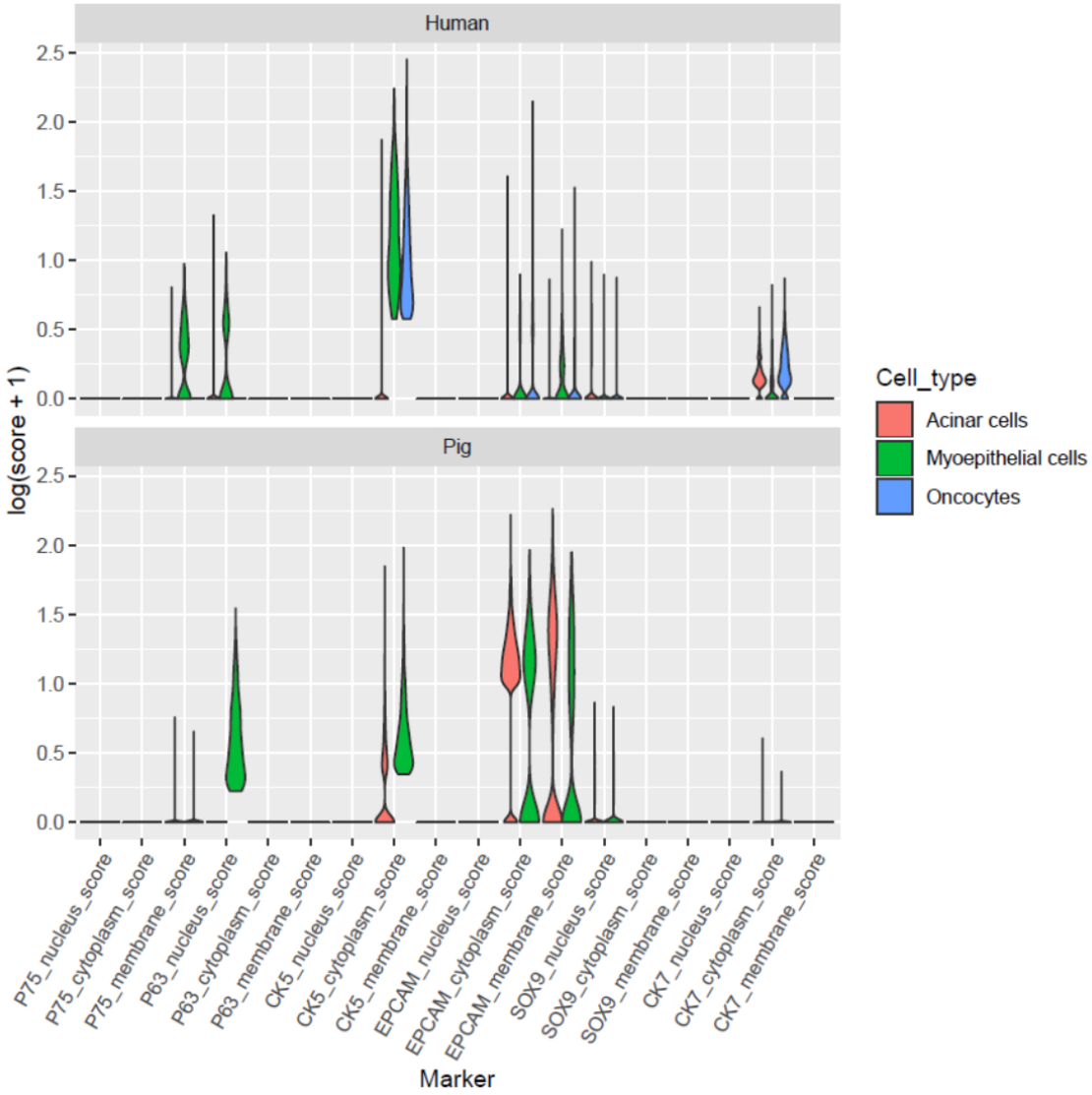


Figure 2.19: Violin plots comparing the phenotype of gland acinar and myoepithelial cells from pig and also oncocytic cells for human SMGs. X-axis indicates the staining marker and subcellular localisation while the y-axis shows the log(score+1) of the measured intensity normalised to exposure. Number of cells analysed: pig=5294 cells, human=10568 cells. Number of sections: pig n=1, human=2.

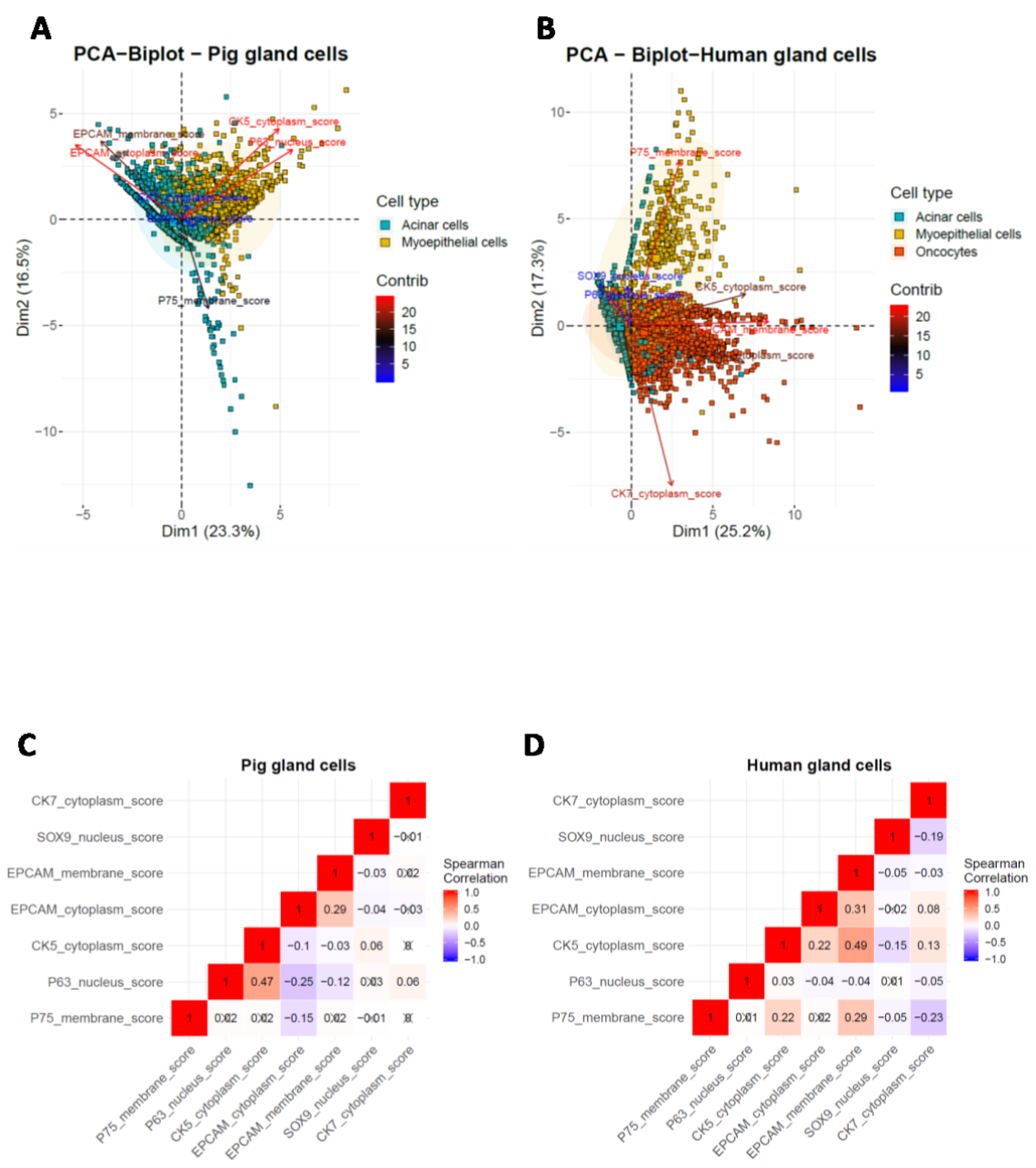


Figure 2.20: Contribution of 'panel one' staining markers for oesophageal SMG cell subtype distinction. PCA biplot of the first and second principal components plot showing the gland acinar and myoepithelial cells in pig (A) and human (B) oesophageal submucosal compartment of SMGs. Arrows show the markers contributing to the points (cell) distribution while the colour of the arrows shows the degree to which these markers contribute. Half correlation matrix showing Spearman correlation coefficients of the indicated markers pig (C) and human (D) as a surrogate quantification for marker colocalization. Values marked with "X" did not reach statistical significance. Number of cells analysed: pig=5294 cells, human=10568 cells. Number of sections: pig n=1, human=2.

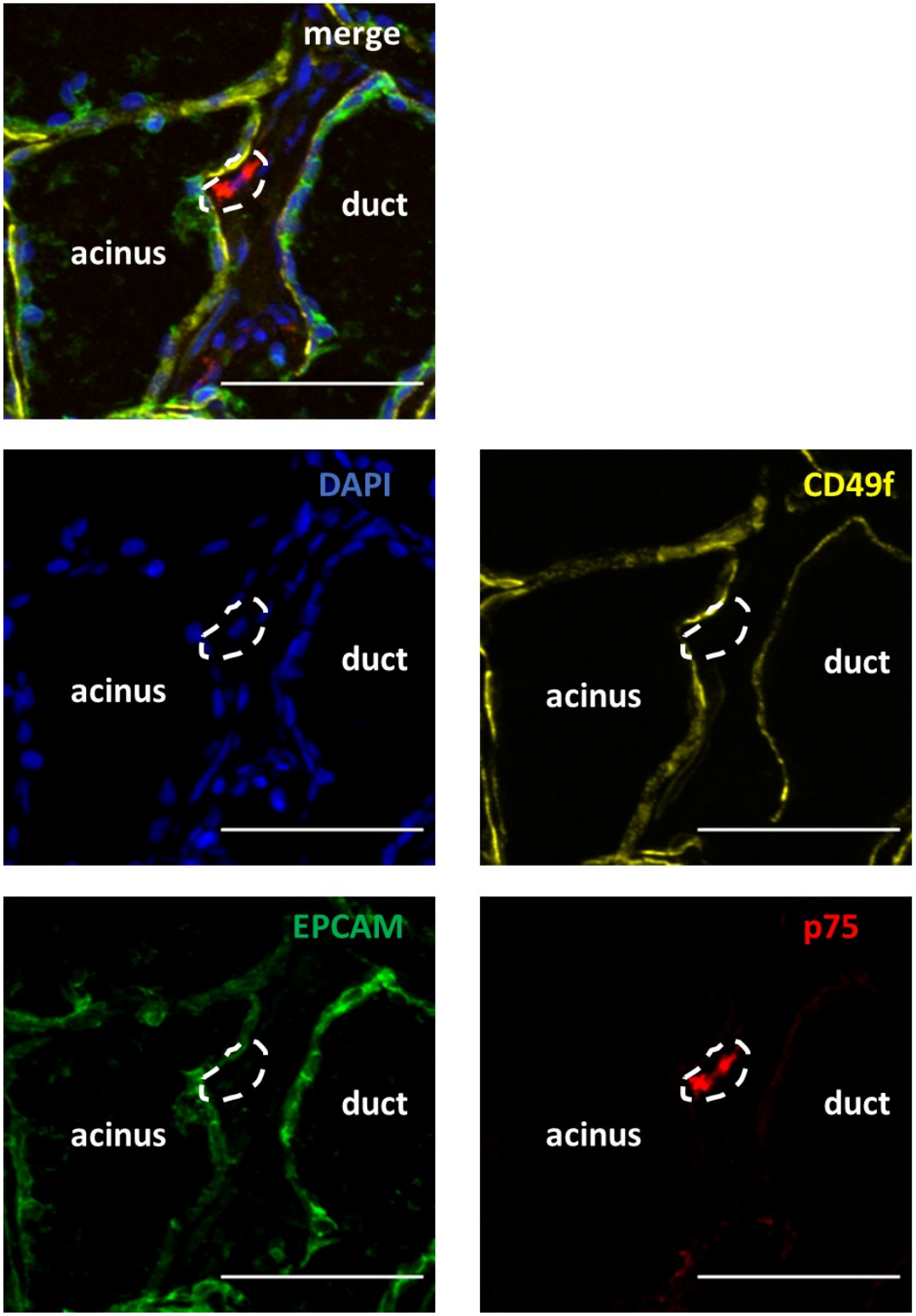


Figure 2.21: Confocal z-stack merged and individual channels images of pig SMGs frozen section stained with CD49f, EPCAM, p75 and counter stained with DAPI. Dotted area marks p75 staining in EPCAM and CD49f negative (stromal) cell adjacent to EPCAM and CD49f positive gland cells. Scalebars=100 µm (n=1).

2.3.12 Gland specific stains in combination with m-IF PCA and colocalization analysis indicate potential subtle differences between human and pig SMG secretions

In addition to KRT7, WGA and UEA staining showed big differences between human and pig SMGs (Figure 2.7). To investigate these differences further, another panel for m-IF was developed. Panel two consisted of WGA and UEA (both of which mark gland acini cells), ACTA2 (myoepithelial cells) and ITGA6/CD49f (a progenitor marker for different cell types including haematopoietic and mammary epithelial cells[195,196]).

Pig m-IF quantification and PCA biplot show ACTA2, and to a lower degree ITGA6, to be markers contributing to the distinction of myoepithelial cells, while the lectins WGA and UEA strongly mark gland's acinar cells (Figure 2.22,23 and 25A). Interestingly, the PCA arrows/vectors for UEA and WGA are pointing in different directions, indicating that these markers do not fully colocalise on the acinar cells (Figure 2.25A). This is further supported by the correlation matrix (Figure 2.25C), which shows a lower correlation (0.27) compared to human (0.7, Figure 2.25D). Conversely, the human PCA biplot (Figure 2.25B) shows UEA and WGA arrows/vectors to be almost parallel indicating strong correlation, which is demonstrated in colocalization analysis and m-IF staining (Figure 2.23, 24 and 25D). ACTA2 shows stronger correlation with ITGA6/CD49f in pig than in human, most likely due to human acinar cells having more prominent ITGA6 staining compared to pig (Figure 2.22 and 23).

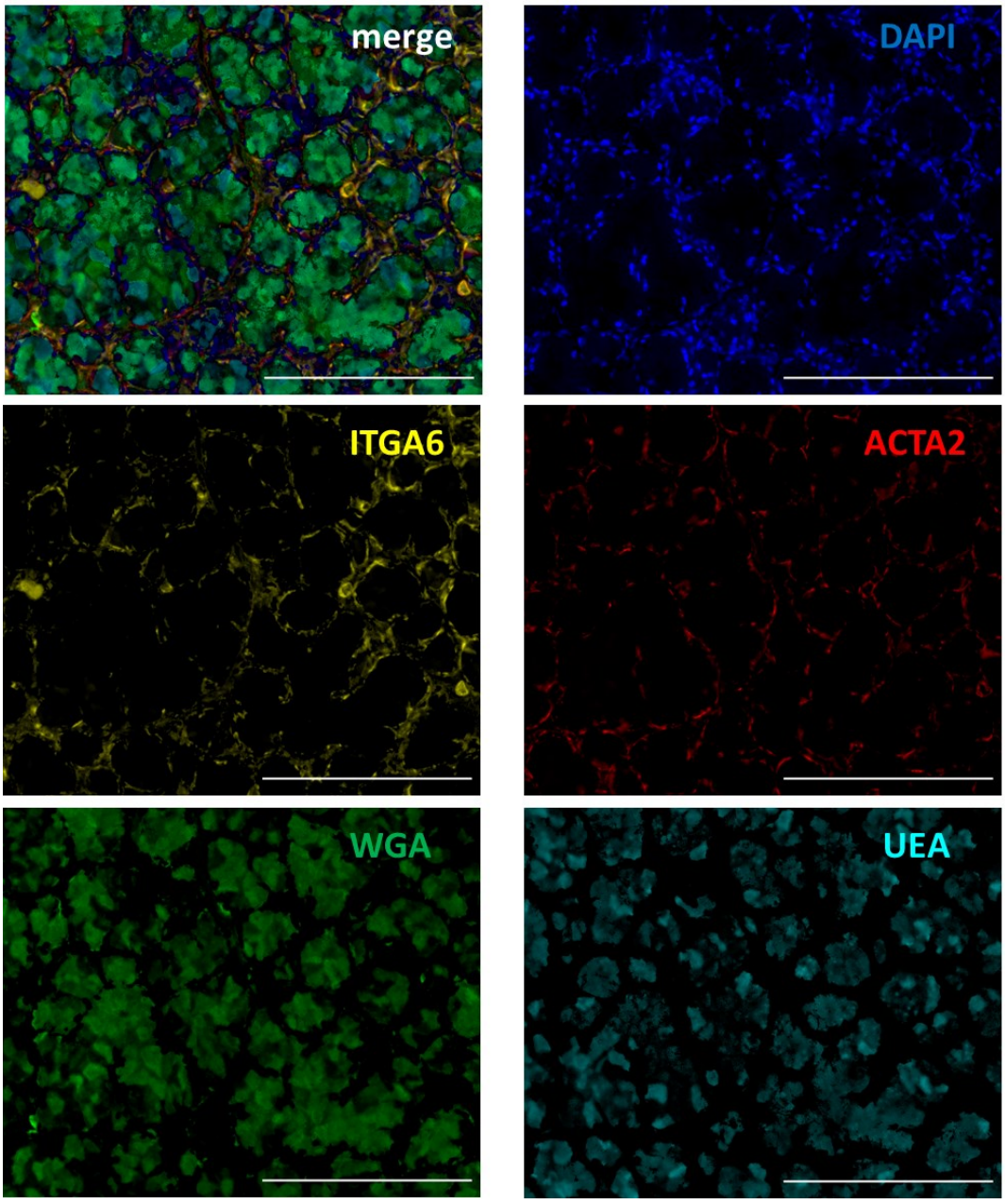


Figure 2.22: Representative m-IF images of pig submucosal gland acini showing merged and individual channel images staining for panel two markers: DAPI, ITGA6/CD49f, ACTA2, WGA and UEA. Scalebars=100 µm.

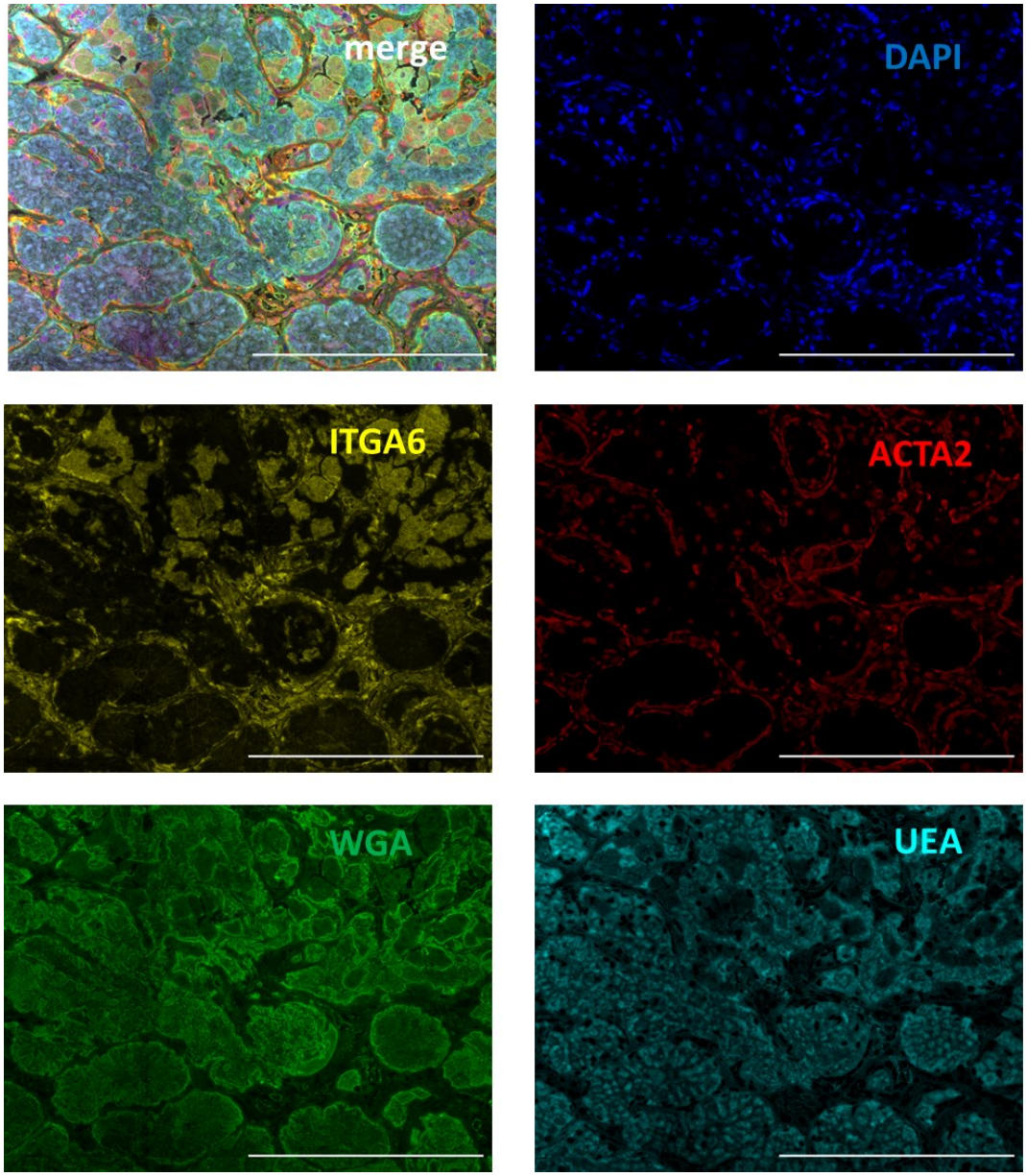


Figure 2.23: Representative m-IF images of human submucosal gland acini showing merged and individual channel images staining for panel two markers: DAPI, ITGA6/CD49f, ACTA2, WGA and UEA. Scalebars=100 µm.

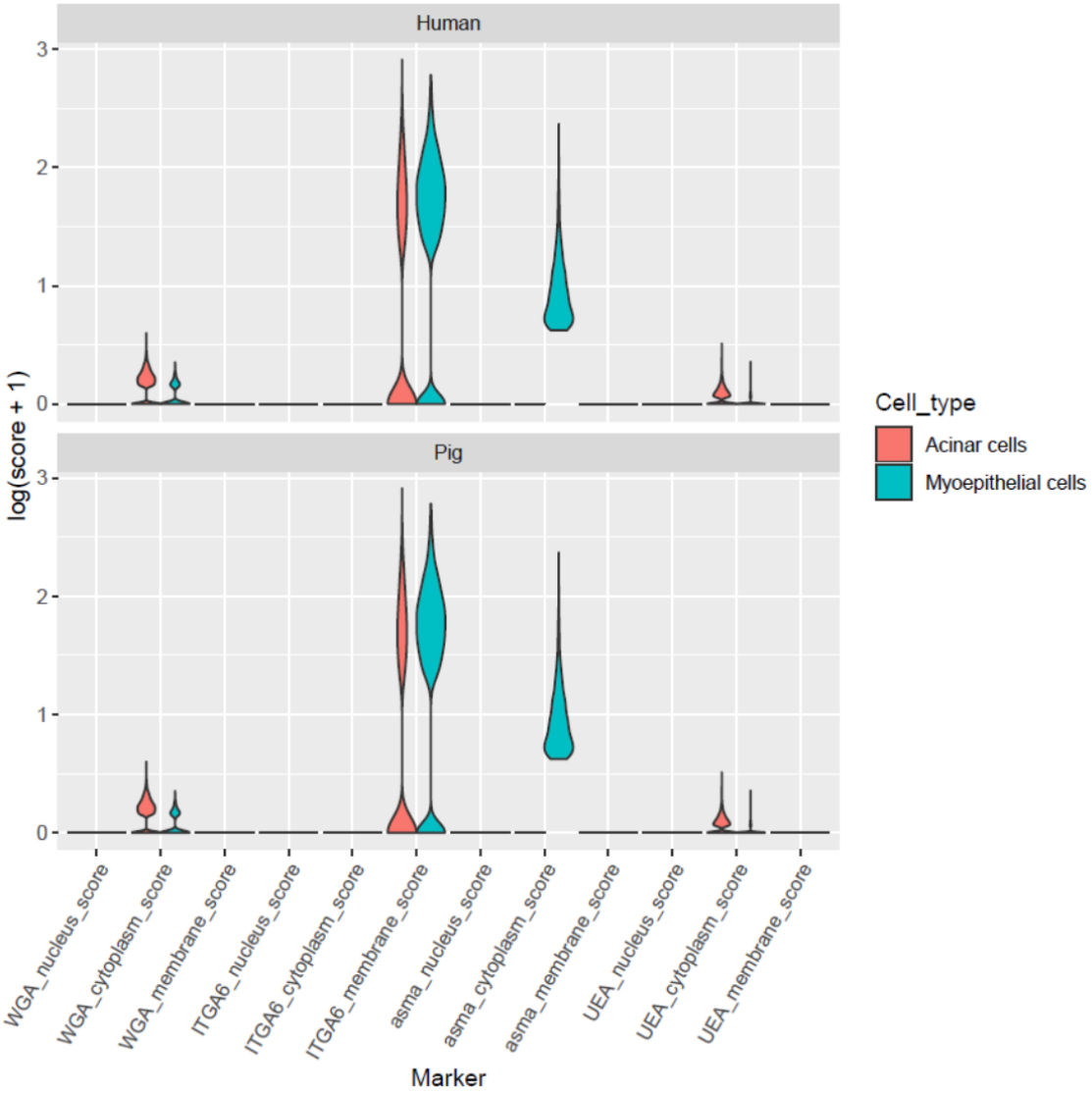


Figure 2.24: Violin plots comparing the phenotype of acinar and myoepithelial cells from pig and human SMGs. X-axis indicates the staining marker and subcellular localisation while the y-axis shows the log(score+1) of the measured intensity normalised to exposure. Number of cells analysed: pig=19761 cells, human=10429 cells. Number of sections: pig n=1, human=1.

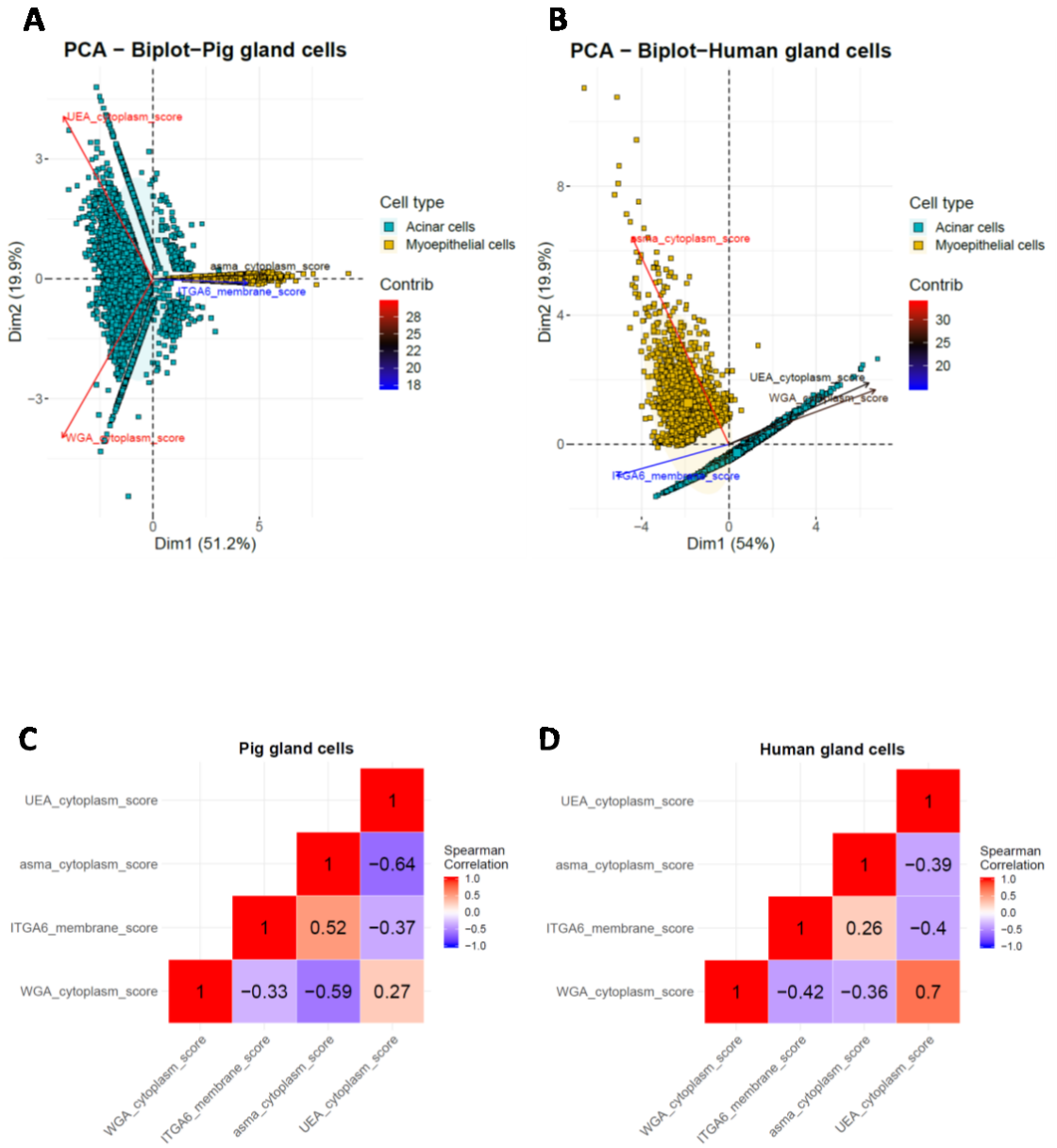


Figure 2.25: Contribution of 'panel two' staining markers for cell subtype distinction of gland compartment of SMGs. PCA biplot of the first and second principal components plot show the gland and myoepithelial cells in pig (A) and human (B) oesophageal submucosal gland compartment of SMGs. Arrows show the markers contributing to the points (cell) distribution while the colour of the arrows shows the degree to which these markers contribute. Half correlation matrix showing Spearman correlation coefficients of the indicated markers found in pig (C) and human (D) as a surrogate quantification of marker colocalization. Values marked with "X" did not reach statistical significance. Number of cells analysed: pig=19761 cells, human=10429 cells. Number of sections: pig n=1, human=1.

2.4 Discussion

In order to test the hypothesis that the oesophageal SMGs contain the cell of origin for BO, acquisition of sufficient good quality SMG tissue for experiments is imperative. As acquisition of human oesophageal resections from patients is not always feasible, finding an alternative source of tissue would enable further studies. Unfortunately, standard animal models such as mice and rats do not possess SMGs in their oesophagus. There are only a few animals, which are known to possess SMGs such as dogs, cats, primates and (guinea) pigs[162]. For this work, pig tissue was chosen as it is easy to acquire in abundant amounts. Here, pig SMGs have been histologically characterised and compared to human to determine if it could be used as a substitute for studying Barrett's.

Although no full-length human oesophagi were available, a clear difference in pig and human SMG distribution could be observed at the distal end of their respective oesophagi (section 2.3.1). This difference could be evolutionarily determined due to the mastication differences between species and the frequency of reflux. Cats, similar to pigs, are known to have SMGs only in the proximal third of their oesophagus, whereas dogs, which can suffer from frequent reflux as humans do, are known to have SMGs in the distal part of their oesophagus[197]. The current data (section 2.3.1) suggests a decreasing gradient in the distribution of SMGs from distal to proximal direction in the distal third of the human oesophagus. Whether SMGs are also present in the upper third of the human oesophagus could not be investigated due to the lack of proximal oesophageal resection specimens from patients.

The high variability in SMG numbers found between patient tissues (Figure 2.2B) was also observed by others[6,142]. Work by Goetsch E., reviewed by Long *et al*.[6], where the distribution of SMGs was studied in two full-length human oesophagi, showed large differences in the number and distribution of the SMGs between individuals along the full length oesophagus. Similarly, Lorinc *et al*. studied the distribution and number of SMGs in normal and metaplastic oesophagus segments from seven EAC patient resections. Likewise, they observed high inter-patient variability[142]. However, similar to the CASCADE tissues acquired in this work, they did not have access to full length oesophagi, only the proximal part of the stomach and a portion of the intrathoracic part of the oesophagus. Still, these data support the high interpatient variability in SMG numbers observed in the oesophagi of CASCADE patients, and that the findings are not due to dissecting artefacts.

Overall, the global structure of human and pig SMGs are highly similar, supporting the idea that they fulfil the same role in both species. SMGs in both species are formed of lobules with mucin producing glands connected to ducts that carry the secretions to the lumen of the oesophagus. However, finer inspection of the tissues shows a number of differences, such as the presence of oncocytes in human but not in pig SMGs. The prevalence of oncocytes in tissues such as the salivary gland[183] and lung SMGs is known to increase with age[198]. This could explain why oncocytes were not found in the pig tissues, which were acquired from young (6 months old) pigs.

Oncocytes do not show mucin production as shown in Figure 2.5 and by others[183]. Gonzalez *et al*. have shown that oncocytes express cystic fibrosis transmembrane conductance regulator (CFTR) and sodium-potassium (Na^+/K^+) ATPase, which indicates a role for oncocytes in the production of bicarbonate to neutralise stomach acid[183]. Diffuse ACTA2 staining in oncocytes indicates that these cells might arise from myoepithelial cells (Figure 2.7). Gonzalez *et al*. came to a similar conclusion by showing certain myoepithelial cells expressing KRT7 before differentiating into spindle cells, which are thought to give rise to oncocytes in human[183]. These spindle cells, however, were not shown to express CFTR or Na^+/K^+ ATPase, only KRT7. In pigs, rare cells were observed showing a similar shape and positivity for KRT7 (Figure 2.9) to spindle cells described by Gonzalez *et al*. However, as no oncocytes were observed in these young pigs, it is not clear whether a similar conclusion could be drawn for pigs.

Lectin staining in pig SMGs shows considerable differences compared to human. Interestingly, correlation analysis (section 2.3.12) showed little correlation (ie. colocalization) for WGA and UEA staining in pig compared to human, potentially indicating specialised or restricted types of mucins being produced by different gland cells in pig.

In contrast, PCA analysis of ductal cells showed few differences between human and pig, aside from SOX9 expression in human luminal duct cells, which was observed by IHC but also showed some variability among the samples stained. PCA analysis of the glandular compartment of SMGs on the other hand showed significant differences between human and pig. One important distinction is the presence of p75 in human but not pig myoepithelial cells. Moreover, p75 staining is not localised on epithelial cells in the glands (excluding basal duct cells). Confocal images of frozen pig SMGs confirm strong p75 staining in EPCAM negative stromal cells adjacent to gland cells. Similarly, p75 positive stromal cells were not identified in human gland tissues. The implications of these differences are at the moment unknown. Altogether, these results show large similarities between human and pig submucosal gland duct cells but a number of differences in the gland acini cells. These differences potentially reflect adaptation to the differences in diet, mastication, propensity for reflux and regurgitation between the two species.

The basal duct cells of the SMGs express the basal markers p63 and KRT5 and progenitor markers p75 and CD49f[199]. These markers would qualify the basal layer duct cells as the resident progenitor cells of the submucosal ducts in both human and pig. Similarly, the gland compartment also shows the myoepithelial cells to have a progenitor-like staining profile in human and less so in the pig. The myoepithelial cells in both species express both basal markers KRT5 and p63 and the progenitor marker CD49f yet, as mentioned previously, only in human do they also express p75. This suggests that the myoepithelial cells may be the resident progenitor cells of the gland compartment of the SMGs. A progenitor role for myoepithelial cells was observed in the lung submucosal glands through lineage tracing[200,201]. Interestingly, the lung submucosal gland myoepithelial

cells from pig were also shown to be positive for the basal/squamous markers KRT5 and p63 and progenitor markers CD49f and p75 similar to human SMG myoepithelial cells[201].

In conclusion, the histologic and colocalization analyses suggest that both ductal and glandular compartments of the SMGs contain their own progenitor cell in both human and pig. The markers used to stain the likely ductal progenitor cells (basal duct cells) show identical profiles between human and pig. The likely gland compartment progenitor cells (myoepithelial cells) are largely similar between the two species where they share two out of three progenitor markers (except p75 for pig). This large degree of similarity in regards to the presence of potential progenitor cells with similar morphologies, histologic characteristics and marker expression, supports the idea of using pig SMGs as substitute for human SMGs for further studies of SMG cells as a possible source of BO.

Chapter 3

Single cell RNA sequencing of porcine SMGs supports histologic analyses and alludes to the presence of distinct submucosal duct and gland progenitor cells

3.1 Introduction

In the previous chapter, a number of markers and staining methods were used to characterise and compare human and pig SMGs. These comparisons shed light on the similarities between human and pig submucosal ducts but also showed important distinctions in submucosal gland acini.

To further characterise the SMGs in a more high-throughput manner, single cell RNA sequencing was utilised to attempt to dissect the different cell populations in the SMGs *in silico* and compare the sequencing data to the histologic analyses performed.

A high-throughput method for single cell RNA sequencing involves the use of microfluidic microdroplets to sequester single cells (up to 10,000) in aqueous droplets (one cell per droplet) in an oil phase. Each droplet also contains a bead with barcoded oligonucleotides to capture RNA from each co-captured single cell. This allows for each transcript sequenced to be assigned to its cell of origin producing a transcriptome library for each captured cell, identifying transcribed genes, and their relative expression level on a single cell basis.

As there are a limited number of pig specific or cross-species reactive antibodies, it has not been possible to test some of the key markers known to be present in human SMGs, such as TFF3[202], using IHC. Such issues could potentially be resolved by looking at the transcriptome of pig SMGs instead. Furthermore, another important goal was to identify other markers that could be used to sort and purify cell populations of interest for *in vitro* studies.

3.2 Material and methods

3.2.1 Pig tissue digestion

Pig SMGs were manually dissected minced into small pieces (< 2 mm on any side) using a number 22 scalpel. Minced tissue was incubated in digestion media consisting of 2 mg/ml Collagenase A (10103578001, Roche, Basel, Switzerland), 0.5 mg/ml hyaluronidase (H3506, Sigma-Aldrich), 0.5 mg/ml DNAse I (11284932001, Roche), 1x antibiotic/antimycotic (15240062, Thermo Fischer Scientific), and 6.25 mM MgCl2 in Hank's Balanced salt solution (H6648, HBSS, Sigma-Aldrich).

Digestion was performed on a shaker at 200 rpm for 2 hrs at 37˚C with trituration every 30 min. After the final incubation, the digest was strained through a 70 µm strainer and washed with 25-30 ml of PBS, centrifuged at 1000 g for 15 min before collecting into a 2 ml Eppendorf tube.

Dispase II (4942078001, Roche) was added to the cell pellet to a final concentration of 5 mg/ml in 2 ml of PBS then incubated for 10 min at 37˚C in a water bath with periodic shaking. The digest was diluted in 8 ml of PBS/5%FBS (PF5) and centrifuged at 500 g for 5 min. The cell pellet was resuspended in 2 ml of 0.25% trypsin-EDTA (25200056, Thermo

Fischer Scientific) in a 2 ml Eppendorf tube, incubated for 5 min in a water bath, washed in 8 ml of PF5 with centrifugation at 300 g for 5 min and resuspended in 1 ml PF5.

3.2.2 Pig cells FACS staining

Pig SMG cells were stained with a pig specific EPCAM primary antibody (1:100, LS-C374109, LSBio). Incubations were performed on ice for 30 min in the dark followed by a single wash with 3 ml PF5 and centrifugation at 300 g for 5 min. Cell pellets were resuspended in PF5, containing F(ab')2 Dylight 405 (1:250, JI111476047, Jackson ImmunoResearch, West Grove, PA) secondary antibody incubation and Sytox Green (1:30000, S7020, Thermo Fisher Scientific) staining were performed simultaneously for 15 min on ice in the dark. Cells were washed once with 3 ml PF5, centrifuged at 300 g for 5 min and resuspended in PF5 for sorting.

3.2.3 Single cell RNA-sequencing pre-processing and filtering

Single cell RNA sequencing (sc-RNA seq) data generation was performed, with the help of Luciano Martelotto of the Single Cell Innovation Lab, University of Melbourne, using Single cell 3' v2 chemistry (10X Genomics, Pleasanton, CA) and processing of sequencing data was performed using Cell Ranger (Version 3.0.2). Sample data was demultiplexed to generate FASTQ files containing sequencing reads. Reads were aligned to Ensembl SScrofa11.1 reference genome allowing for mapping of 93.3% of reads. Of the mapped reads 54.7% were confidently mapped to annotated transcriptome (annotated genes). Empty droplets were filtered using cell barcodes yielding 3650 sequenced cells with a median of 1288 identified genes sequenced per cell (Appendix Figure 1).

Further processing and analysis of sc-RNA seq data was performed with IPython (v7.8.0) using the scanpy[203] package (v1.4.4.post1). Due to the limited annotation of the pig genome, only genes with official symbols were kept for further analyses. Subsequently, the number of genes detected per cell (n_genes), transcripts count per cell (n_counts), percentage of mitochondrial (percent_mito) and heat shock protein (percent_HSP) genes were determined (Appendix Figure 2). Transcripts count per cell and percentage mitochondrial genes were regressed out to prevent any effects on downstream clustering and data was scaled to unit variance as per standard pre-processing protocols[203]. For further analyses low quality cells and outliers were filtered out. Cells with 500 to 2200 genes detected, containing less than 5% mitochondrial gene transcripts and having read counts between 1500 to 9000 were kept. In addition, genes transcripts detected in ≤3 cells (0.1%) were also removed leaving 2442 cells with 9894 genes. Potential doublets were inferred using the Scrublet[204] packages, however the inferred doublets were not removed from the dataset as they did not affect downstream analysis (Appendix Figure 3).

3.2.4 Dimensionality reduction, clustering and marker genes identification

Principal component analysis (PCA) was performed to find the main axes of variation and data denoising. For clustering, Scanpy computes a neighbourhood graph using the

denoised PCA data matrix. Two-dimensional embedding was performed using Uniform Manifold Approximation and Projection (UMAP)[205]. Clustering of cells into sub-clusters in UMAP was performed using the Leiden algorithm[206] at 0.1 resolution for all cells and 0.2 resolution for ducts only cluster.

Cluster marker genes were determined by ranking the highly differentially expressed genes using the Wilcoxon rank-sum (Mann-Whitney-U) test[207].

3.2.5 Trajectory inference and pseudo-temporal gene dynamics

For differentiation trajectory inference, clusters inferred as ductal and gland cells were analysed separately with basal duct cells and myoepithelial cells being set as root cells to start differentiation trajectories. For duct cells analysis slingshot v1.0.0[208] was used and for acinar cells monocle2 2.10.1[209] was used in addition to slingshot. Top ductal gene expression describing the differentiation across pseudotime was determined using VGAM package (v1.1.1). To perform these analyses, scanpy data was converted to 'singlecellexperiment' using anndata2ri package (v0.1.dev69+635da2f).

3.2.6 Gene ontology analysis

Gene ontology analyses of each cluster was performed using GProfiler (v1.0.0) with human genome annotation as the reference[210].

3.3 Results

3.3.1 Sorting strategy for EPCAM positive submucosal duct and gland cells

Pig SMGs were characterised using sc-RNA seq. To focus solely on epithelial cells, dissected SMGs from five pig oesophagi were pooled, digested to generate single cells and stained with pig specific EPCAM antibody (which stains SMG acini and ductal epithelial cells, Figure 3.1) for sorting. The viability stain, Sytox green, was also used to enrich for good quality, viable single cells. Figure 3.2 shows a diagram of the gating and sorting strategy established.

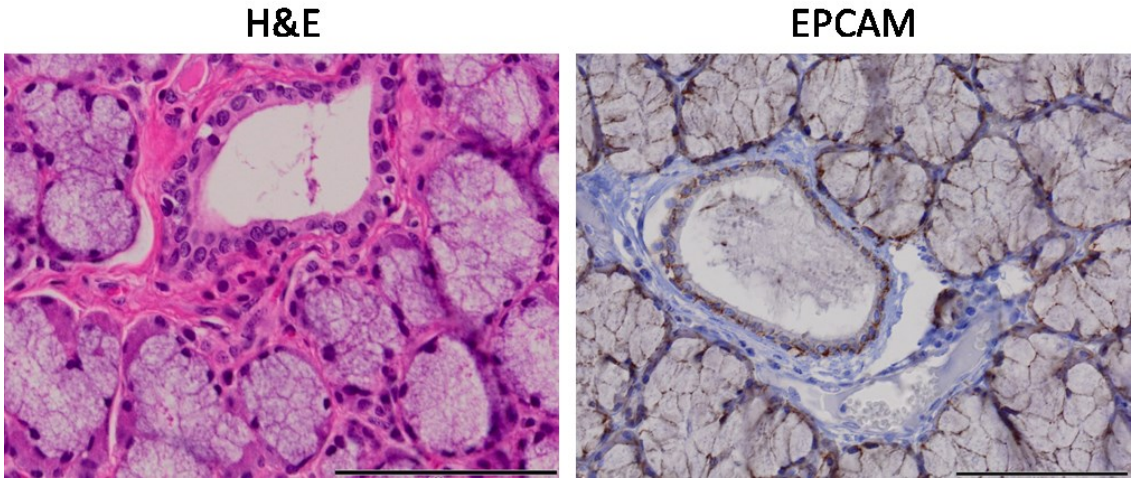


Figure 3.1: Representative images of H&E (left panel) and immunostaining for EPCAM (right panel) of pig SMGs showing EPCAM specificity in epithelial cells and absence in stromal cells (n=3). Scalebars=100 µm

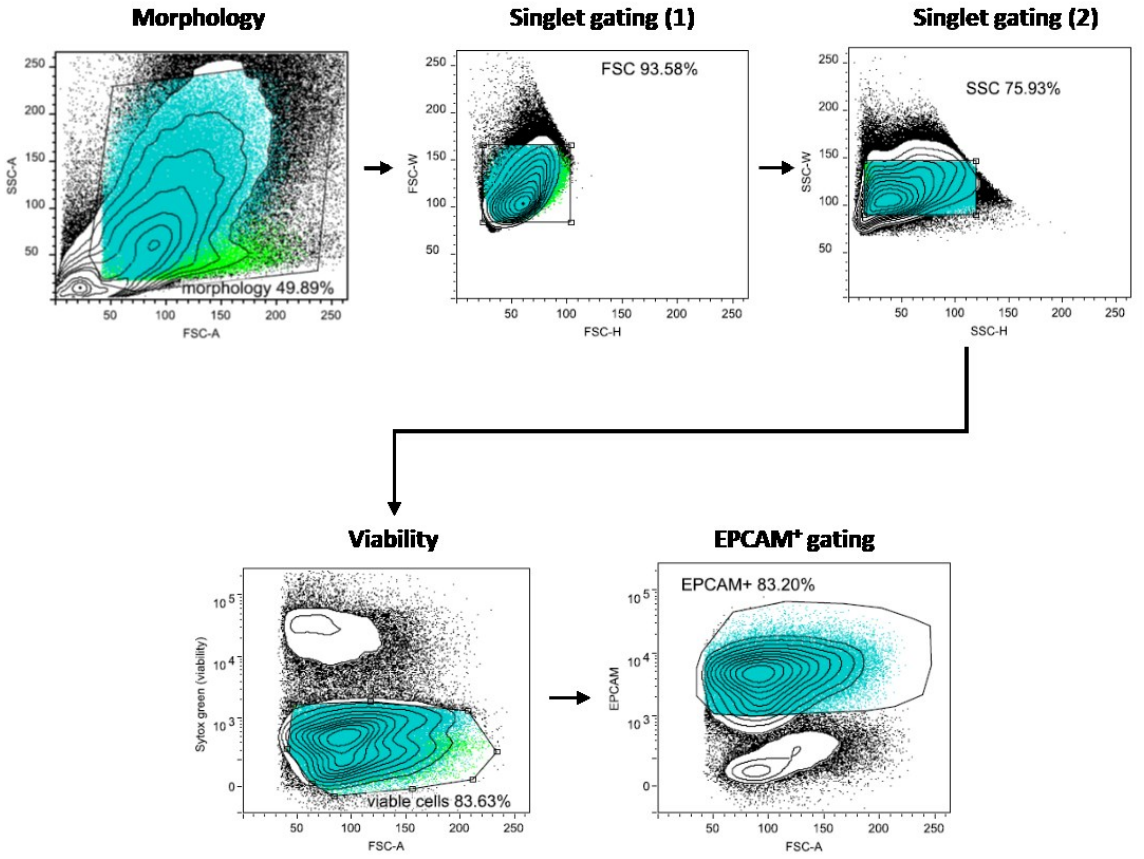


Figure 3.2: Gating strategy for sorting viable epithelial cells from digested pig SMGs for sc-RNA sequencing. Morphology panel shows gating of bulk cells based on morphology (eg. excluding debris), Singlet gating (1) panel shows FSC gating for selecting single cells, Singlet gating (2) panel shows FSC gating for selecting single cells, Viability panel shows gating for viable cells (ie. negative for Sytox green) and EPCAM$^+$ gating panel shows gating for EPCAM positive cells. Blue dots represent the backgated EPCAM$^+$ cells and green dots represent viable backgated cells.

3.3.2 Pig submucosal gland cells can be divided into four distinct clusters

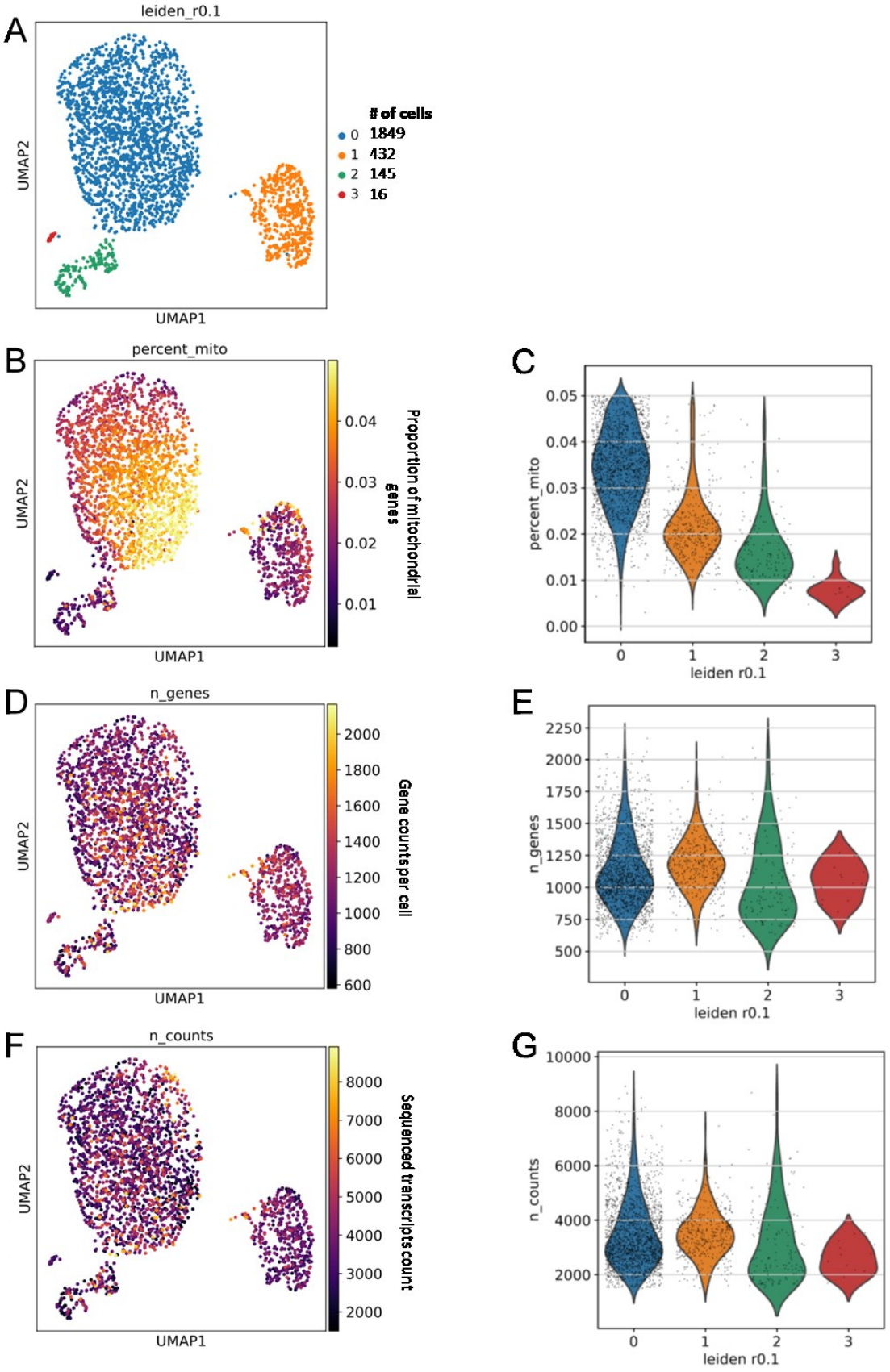
Clustering of the pre-processed 2442 pig SMG cells can be divided, using the Leiden algorithm, into four clusters containing different numbers of cells (Figure 3.3A). The two largest clusters (cluster 0 and 1) contain 1849 and 432 sequenced cells, respectively, while the smallest cluster (cluster 3) contains 16 cells. Cluster 2 seems to be represented by two small sub-clusters that combined together contains 145 cells (Figure 3.3A).

The proportion of mitochondrial genes is generally used to filter out low quality cells as these show an upregulation of these genes[211]. Figures 3.6B and C show the proportions of mitochondrial genes detected per cell in each of the clusters. Cluster 0 contains cells with the highest proportion of mitochondrial genes with the majority containing a proportion ≥0.025 after filtering. Conversely, cluster 3 contains cells with the lowest proportion of mitochondrial genes, while the majority of cells in clusters 1 and 2 have an intermediate proportion of mitochondrial genes. The implications of these differences are not clear.

Comparison of the number of detected genes per cell, visualised using UMAP and violin-jitter plots (Figures 3.6D and E), largely showed no differences, indicating that this is not the reason for the differences in mitochondrial gene proportions between clusters. Cluster 3 shows a smaller dispersion of the number of detected genes compared to the other clusters, likely due to the small number of cells in this group. Similarly, no large differences were observed in the number of sequenced reads (n_counts) per cell (Figures 3.6F and G). Most cells of each cluster have similar numbers of detected sequenced reads count (between 2000 and 4000) with a skew towards bigger counts for the bigger clusters (clusters 0, 1 and 2).

Finally, as EPCAM was used as marker for enriching for epithelial cells of the submucosal gland, this can be confirmed by assessing the level of EPCAM mRNA expression in the different clusters. EPCAM expression is relatively equal in most clusters except cluster 3, which shows clearly lower levels of expression (Figure 3.3H and I).

As cluster 2 seems to be composed of two groups of cells, the clustering resolution for this cluster was increased from 0.1 to 0.2 to differentiate the two groups, designated hereafter as clusters 2,0 and 2,1 (Figure 3.3J).



(continued on next page)

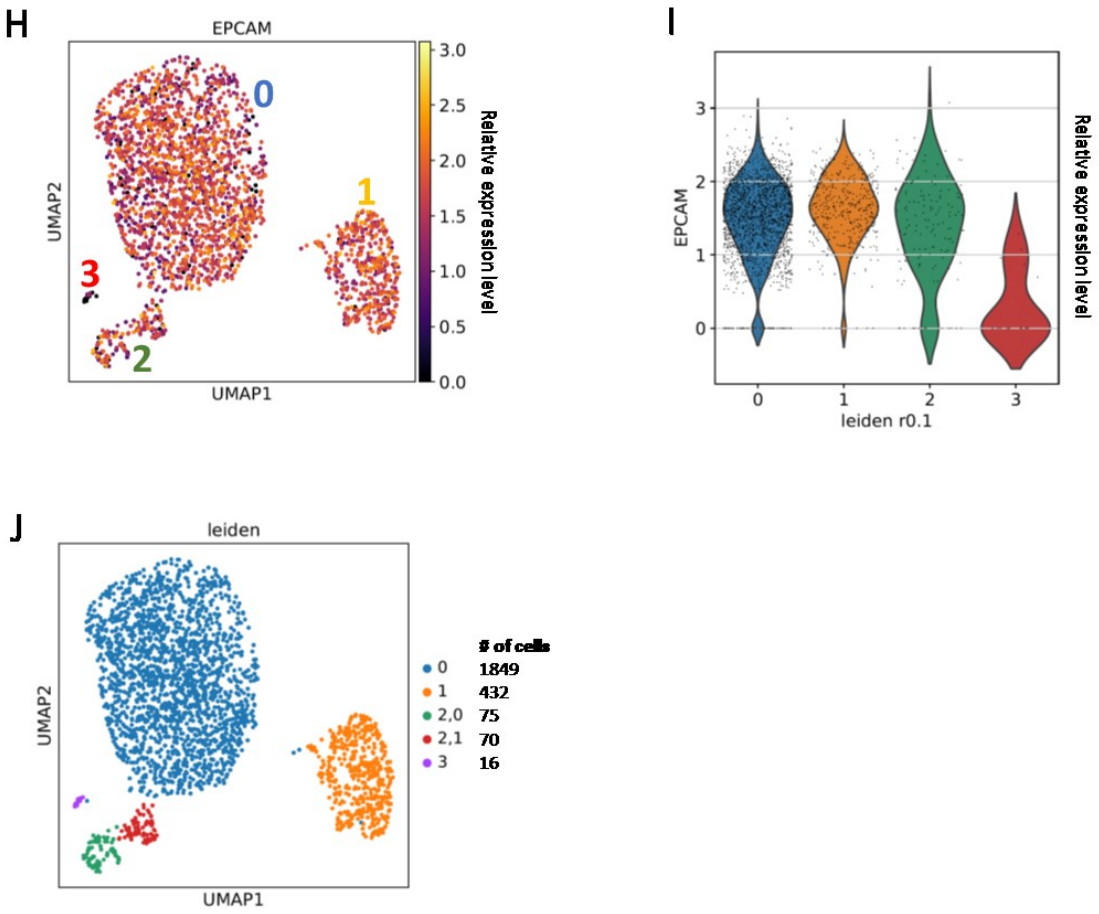


Figure 3.3: UMAP dimensionality reduction and clustering of pig SMGs single cell RNA sequencing data. UMAP projection of the Leiden algorithm-based identification of distinct clusters of pig SMG cells based on similarity of gene expression profile with the number of cells per cluster indicated (A). UMAP projection of the proportion of mitochondrial genes per cell (B). Violin-jitterplot of the Leiden algorithm defined clusters showing the proportion and distribution of mitochondrial genes per cell (C). UMAP projection of the number of detected genes per cell (D). Violin-jitterplot of the Leiden algorithm clusters showing the number and distribution of detected genes per cell (E). UMAP projection of the number of unique molecular identifier (UMI) counts per cell (F). Violin-jitterplot of the Leiden algorithm clusters showing the number and distribution of sequenced reads per cell (G). UMAP projection of the relative EPCAM expression level per cell (H). Violin-jitterplot of the Leiden clusters showing the relative EPCAM expression level per cell (I). UMAP projection of the adjusted Leiden algorithm-based identification of distinct clusters for cluster 2 (J).

3.3.3 Gene overlap analysis and feature plots reveal the identities of myoepithelial, secretory and ductal cell clusters

To get an idea of which cells within the submucosal glands are represented in each cluster, the top 100 differentially expressed genes defining each cluster (Appendix Table 1) were determined. These genes were compared to the Panglao database (Panglao DB)[212] and proportions of overlap with the cells in the database were determined. Panglao DB contains information on 4644 marker genes of 178 cell types from 29 tissues for both human and mouse. Figure 3.4 shows a heatmap of the top 15 cell types with the highest proportion of overlapping genes to the five pig SMG clusters.

Cluster 0 shows little overlap to the majority of cell types in the database. The highest similarities are with myofibroblasts and airway goblet cells suggesting that cells in cluster 0 have a secretory phenotype. Interestingly, the myofibroblast overlap does not seem to coincide with the data showing these cells to be positive for EPCAM as shown in Figure 3.3 and the cell sorting. Further evidence for this phenotype can be seen by the high expression of microsomal glutathione S-transferase 3 (MGST3), Solute Carrier Family 31 Member 2 (SLC31A2) and Aquaporin3 (AQP3) in cells of cluster 0 relative to the other clusters (Figure 3.5A-C and A'-C'). MGST3 is known to play a role in the production of prostaglandin E, which has been shown to be part of the secretions produced and released by oesophageal SMGs[6]. SLC31A2 and AQP3 are both known to be involved in transport and secretion of inorganic material and water[213,214], respectively, further supporting the secretory phenotype. SLC31A2 is known to have a specifically high expression in salivary glands whereas AQP3 is more ubiquitous according to the Human Protein Atlas database[215].

Cluster 1 shows the highest overlap with goblet cells thereby also indicating a secretory phenotype. This is supported by the high expression of secreted proteins including Trefoil factors 2 and 3 (TFF2, TFF3) and lysozyme (LYZ) in cells of cluster 1 (Figure 3.6A-C and A'-C').

Cluster 2,0 shows highest overlap with ductal cells and myofibroblasts potentially indicating this cluster represents cells with a ductal phenotype. As shown in Chapter 2, basal duct cells of the SMGs strongly express KRT5. Consistent with these findings, UMAP feature plot and violin-jitter plot show cluster 2,0 expressing high levels of KRT5 (Figure 3.7A and A'). In addition, in Chapter 2, NGFR/p75 was shown to be colocalised with KRT5 in basal duct cells. This can also be observed for a proportion of the cluster 2,0 KRT5 expressing cells (Figure 3.7C and C'). Similarly, cluster 2,1 shows the highest overlap with ductal cells as well (Figure 3.4). In addition, high overlap can also be found with luminal duct cells and mammary epithelial cells suggesting that these cells represent the luminal duct cells of the SMGs. In support of this observation, cells in Cluster 2,1 show high expression of KRT7 and no expression of KRT5 (Figure 3.7B and B'), consistent with the immunostaining pattern for luminal duct cells of the pig SMGs shown in Chapter 2.

Finally, cluster 3 shows high overlap with a number of different cell types. High overlap can be found with juxtaglomerular cells, myofibroblasts, myoepithelial cells and a number of smooth muscle cell types, suggesting that these cells may have a contractile function in the SMGs. In Chapter 2, ACTA2 staining was used to mark myoepithelial cells in the SMGs. Similarly, ACTA2 is highly expressed in cells in cluster 3 (Figure 3.8A and A'). Further, genes highly expressed in muscle cells, Myosin regulatory light polypeptide 9 (MYL9, Figure 3.8B and B') and Myosin light chain kinase (MYLK, Figure 3.8C and C'), were also found to be highly expressed in cluster 3. In addition, the basal marker KRT5 was also shown to be expressed in SMG myoepithelial cells in Chapter 2. This was also observed in cluster 3 (Figure 3.7A and A').

Using the current data, the individual clusters were renamed to the cell types that they putatively represent (Figure 3.8D). Cluster 0 was renamed to Acinar cells as these cells are likely derived from the gland compartment and show a secretory phenotype, but no clear type of secretory phenotype could be inferred from the current data. Cluster 1, also most likely derived from the gland compartment, was renamed to Mucinous cells as this cluster shows expression of trefoil factors, which are found in mucin producing cells. Clusters 2,0 and 2,1 were renamed to Basal duct cells and Luminal duct cells, respectively. Finally, cluster 3 was renamed to Myoepithelial cells.

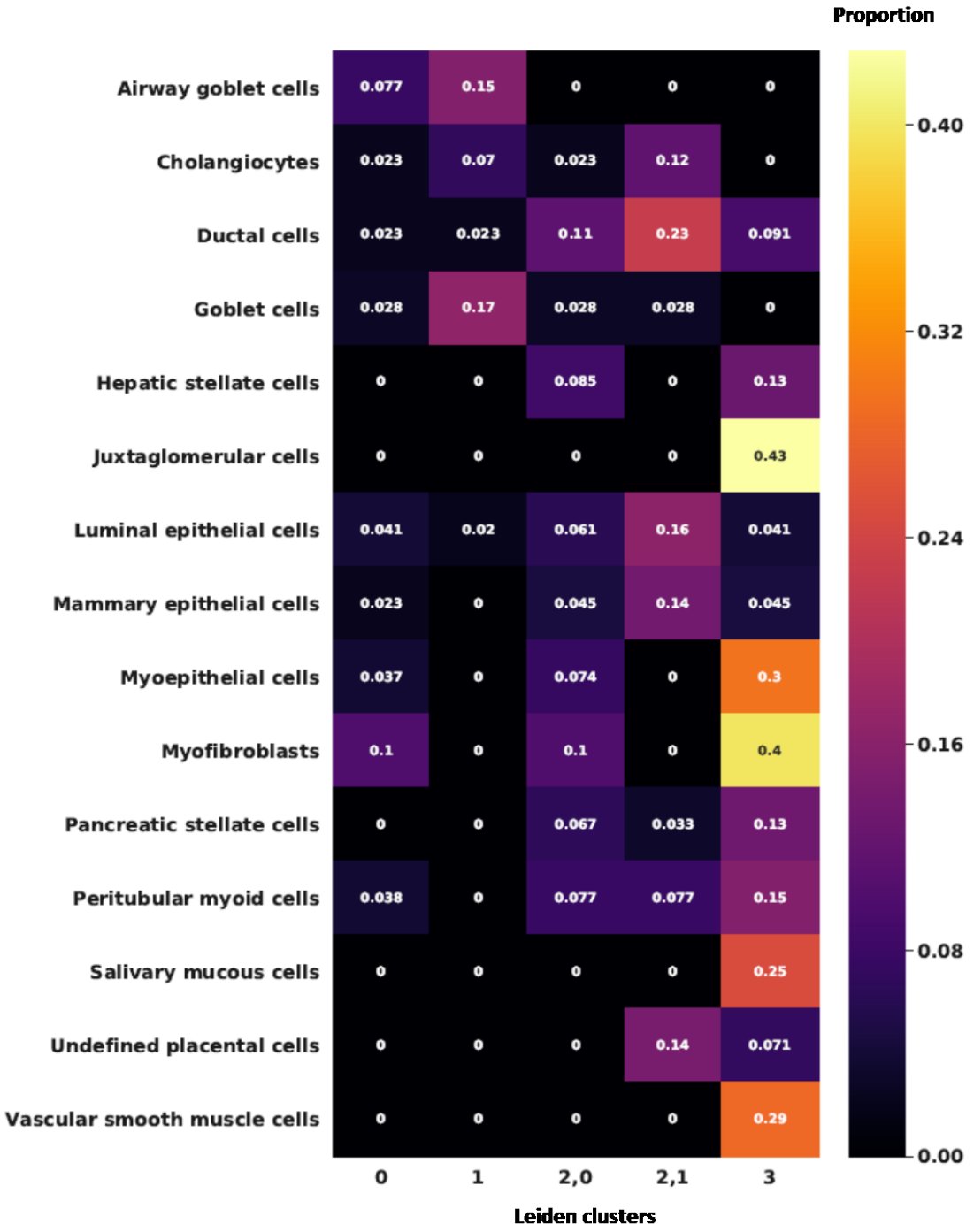


Figure 3.4: Gene expression profiles of individual clusters determined by the adjusted Leiden algorithm were compared to Panglao database containing information on 4644 marker genes of 178 cell types from 29 tissues for both human and mouse. Heatmap shows the top 15 cell types of the Panglao database showing the highest overlap in marker genes with the Leiden algorithm defined clusters of sequenced pig SMG cells.

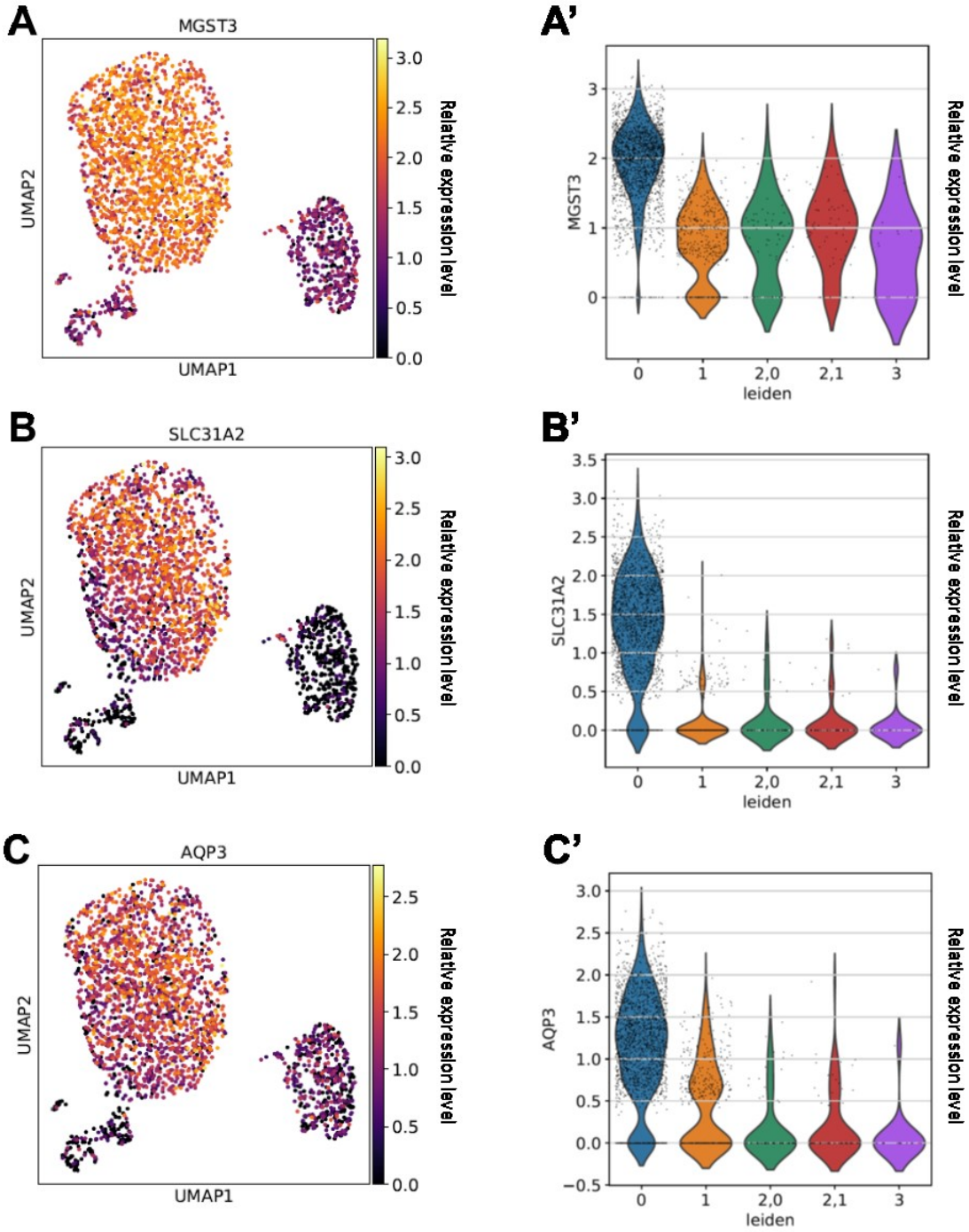


Figure 3.5: UMAP feature and violin-jitter plots showing relative expression levels and distribution of selected highly expressed genes in SMG cluster 0 MGST3 (A-A'), SLC31A2 (B-B') and AQP3 (C-C') compared to remaining clusters.

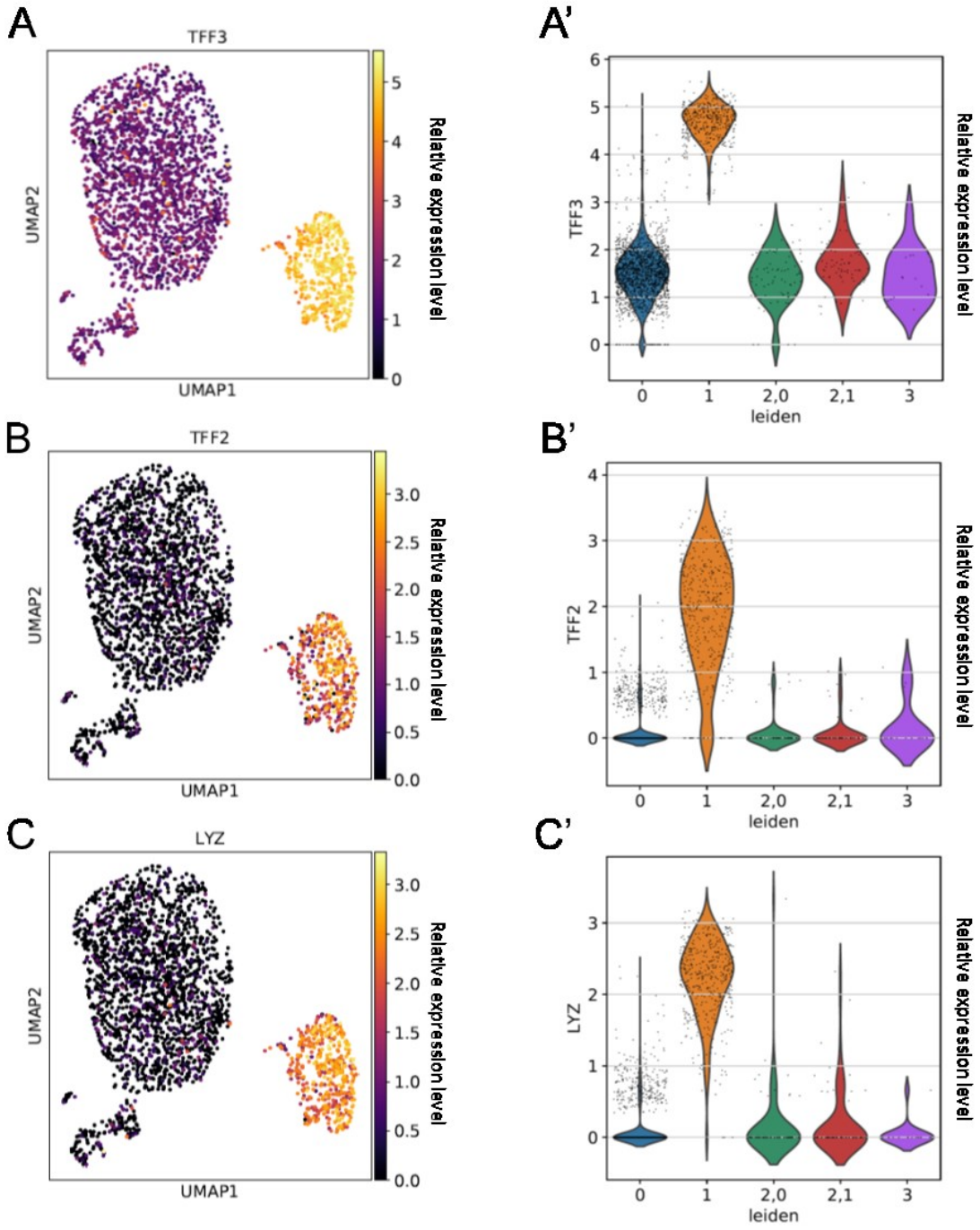


Figure 3.6: UMAP feature and violin-jitter plots showing relative expression levels and distribution of selected highly expressed genes in SMG cluster 1 TFF3 (A-A'), TFF2 (B-B') and LYZ (C-C') compared to remaining clusters.

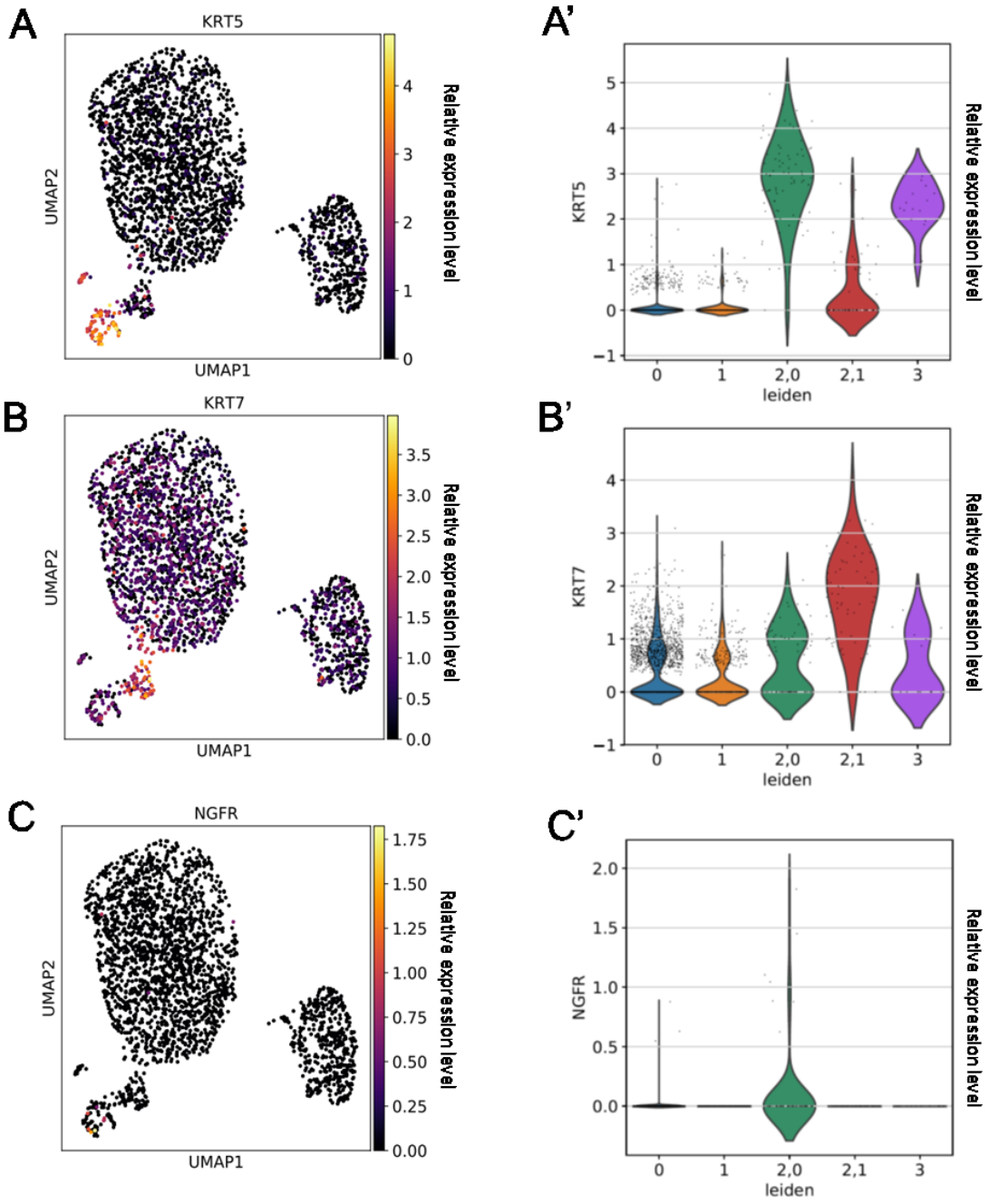


Figure 3.7: UMAP feature and violin-jitter plots showing relative expression levels and distribution of selected highly expressed genes in SMG clusters 2,0 and 2,1 KRT5 (A-A'), KRT7 (B-B') and p75/NGFR (C-C') compared to remaining clusters.

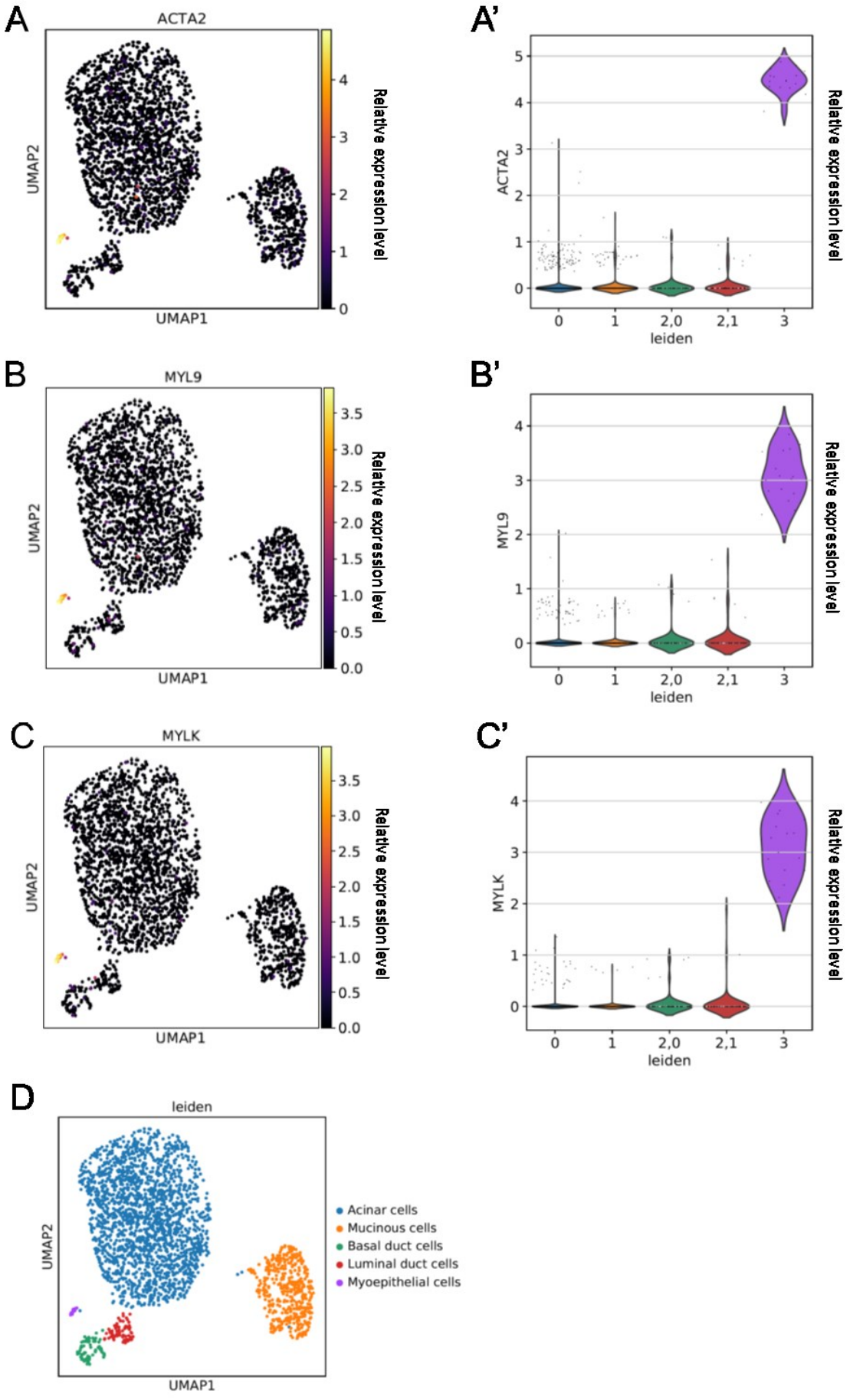

A
ACTA2
UMAP2
UMAP1
Relative expression level
A'
ACTA2
leiden
Relative expression level
B
MYL9
UMAP2
UMAP1
Relative expression level
B'
MYL9
leiden
Relative expression level
C
MYLK
UMAP2
UMAP1
Relative expression level
C'
MYLK
leiden
Relative expression level
D
leiden
UMAP2
UMAP1
Acinar cells
Mucinous cells
Basal duct cells
Luminal duct cells
Myoepithelial cells

Figure 3.8: UMAP feature and violin-jitter plots showing relative expression levels and distribution of selected highly expressed genes in SMG cluster 3 ACTA2 (A-A'), MYL9 (B-B') and MYLK (C-C') compared to remaining clusters. UMAP projection of the adjusted Leiden algorithm with renamed clusters based on identified cell types (D).

3.3.4 Differentiation trajectory inference of SMG cells does not separate between ductal and gland cells

In order to infer the differentiation trajectory of the SMG cells, two separate analyses were performed with different starting populations using the slingshot R package. As both myoepithelial and basal duct cells showed progenitor marker expression, these two populations were tested as starting populations. Figures 3.12A and C shows the connectivity between the individual clusters, thereby indicating which cells of each cluster are predicted to differentiate into another cluster. Figures 3.12A and 3.12C show identical cluster connectivity regardless of whether myoepithelial or basal duct cells were selected as the starting population, respectively. When the myoepithelial cells were selected as the starting population, then the cluster derived from it is inferred to be the basal duct cells cluster followed by luminal duct, acinar and mucinous cells, as shown in the pseudotime trajectory plot (Figure 3.9B). If the basal duct cells were selected as starting cluster, then the algorithm infers the presence of two lineages. One lineage where the basal duct cells differentiate into luminal duct, acinar and mucinous cells (Figure 3.9D) and one lineage where the basal duct cells differentiate into the myoepithelial cells (Figure 3.9E).

To confirm that these results are not biased to the algorithm used for inference, the same analysis was performed using the monocle2 R package. Contrary to Slingshot analysis, monocle2 analysis does not infer a clear differentiation trajectory between myoepithelial and basal duct cells. Instead, monocle2 analysis clusters the two clusters together (Figure 3.10A), indicating a strong similarity between the two populations. For the remaining clusters, monocle2 analysis shows largely a similar result as the Slingshot analysis (Figure 3.10A and B). Another difference the monocle2 analysis shows is a branching (indicated by '1' black dot) within the acinar cells cluster. This branching indicates more complexity within the acinar cells' population, where a portion of the cells is inferred to differentiate into mucinous cells and the rest remains acinar (Figure 3.10B and C). Interestingly, both methods suggest that the luminal duct cells can differentiate into the acinar cells. Furthermore, differentially expressed genes analysis (Appendix Table 1) shows luminal duct cells expressing the (cancer) stem cell marker aldehyde dehydrogenase 1A3 (ALDH1A3) , which was shown to require SOX9 for its expression[216,217], another marker highly expressed in luminal duct cells as shown in Chapter 2.

Based on the current available data analysing all the cell types together, it is not possible to conclusively determine the lineage hierarchy of the different cells within the pig SMGs. Therefore, the ductal and gland compartment were next analysed separately based on the hypothesis that two distinct progenitors could exist in the individual compartments.

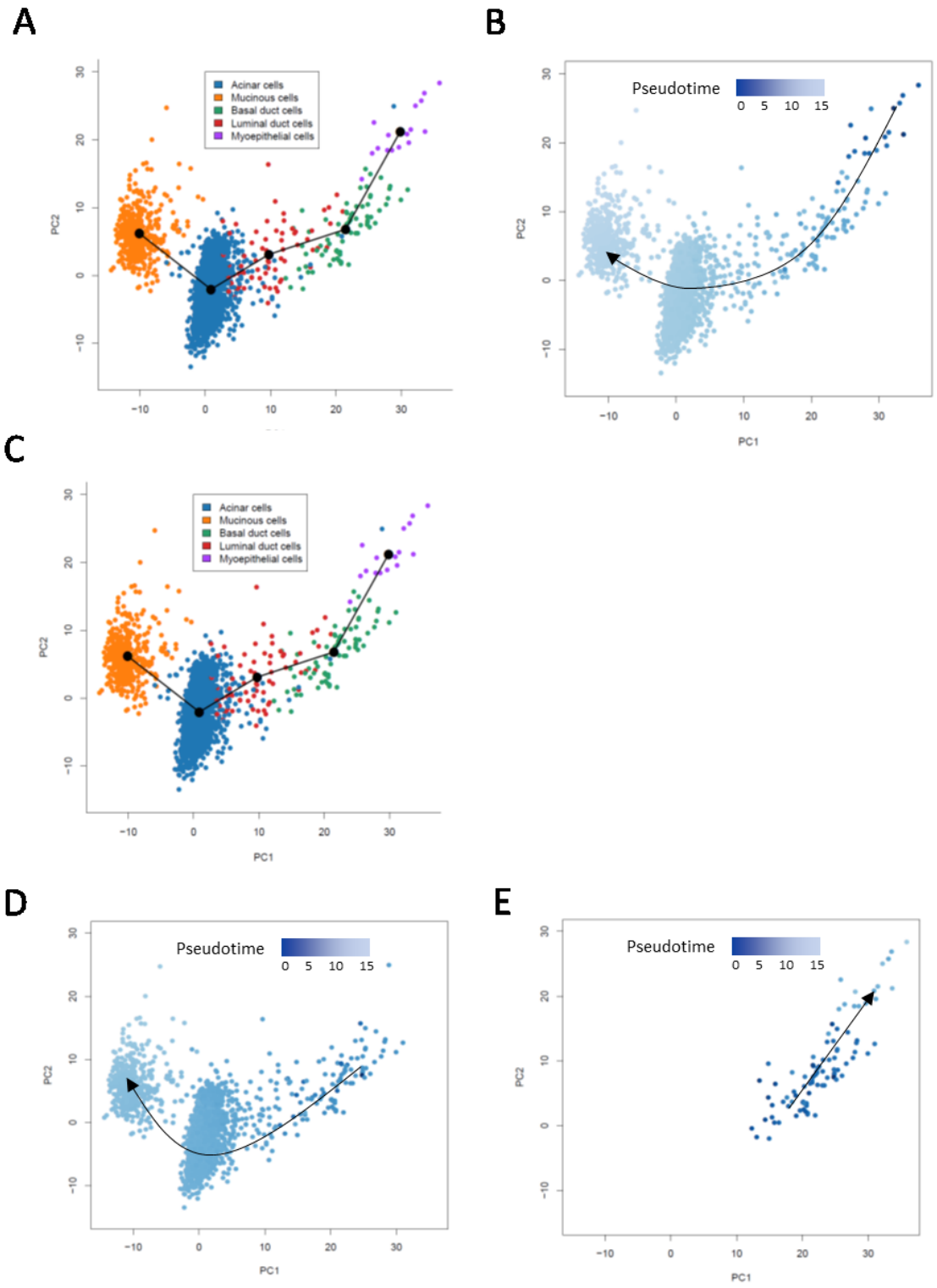
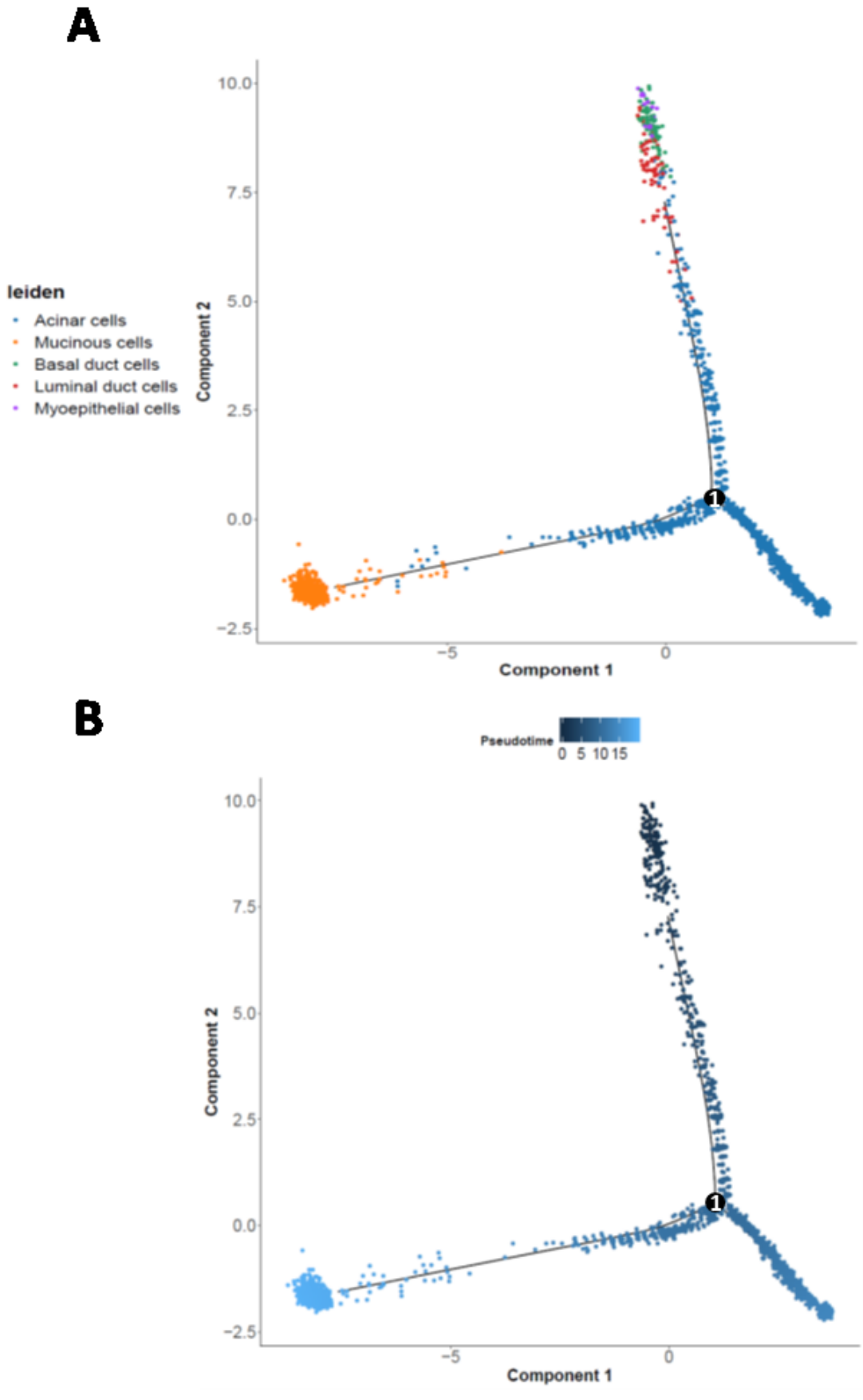


Figure 3.9: Diffusion pseudotime and Trajectory analyses of sc-RNA seq pig SMG cells using the Slingshot algorithm to determine how the different clusters are connected to each other. Based on the inter-cluster connectivity the differentiation trajectory of the clusters containing progenitor cells (myoepithelial or basal duct cells) were inferred based on changes in gene expression in pseudotime. Slingshot PCA plot of SMG cells inferred connectivity with myoepithelial cells (A) or basal duct cells (C) as starting population. Slingshot pseudotime diffusion plot of SMG cells with myoepithelial cells as starting cluster (B) and Basal duct cells as starting cluster (D and E). Arrows in B, D and E indicate the direction of differentiation.



(continued on next page)

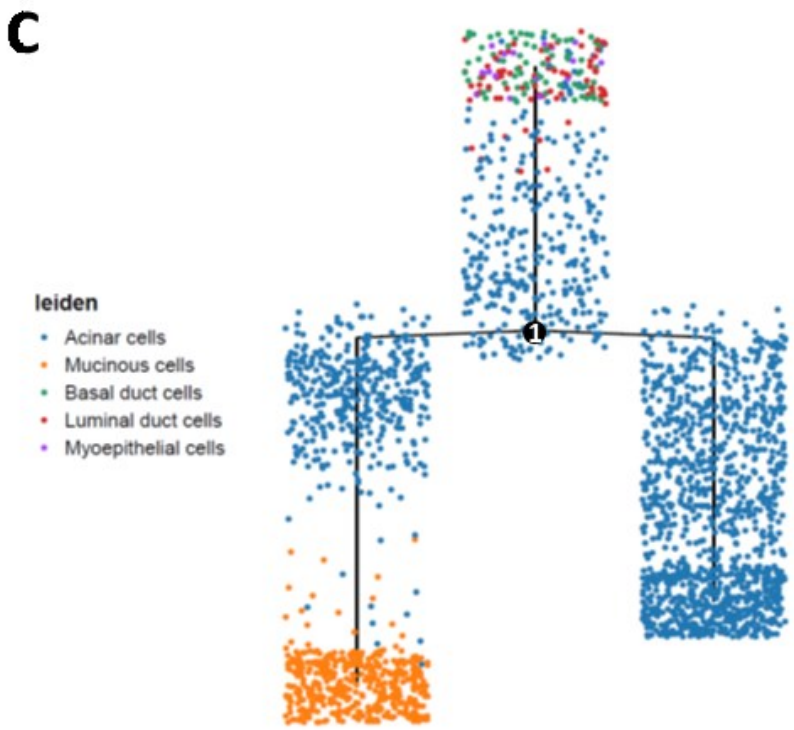


Figure 3.10: Diffusion pseudotime and Trajectory analyses of sc-RNA seq pig SMG cells using the Monocle2 algorithm was used to determine how the different clusters are connected to each other. Based on the inter-cluster connectivity the differentiation trajectory of the clusters containing progenitor cells (myoepithelial or basal duct cells) were inferred based on changes in gene expression in pseudotime. Monocle2 PCA plot of SMG cells inferred connectivity (A) and Monocle2 pseudotime diffusion plot (B) of SMG cells, which are similar when either myoepithelial or basal duct cells are used as the starting cluster. Monocle minimum spanning tree plot on SMG cells (C). Black point "1" indicates lineage branching in the acinar cell cluster leading to cells remaining acinar and other cells differentiating towards mucinous cells.

3.3.5 Gene expression change dynamics support the role of basal duct cells as resident progenitor cells for submucosal ducts

As shown in the previous chapter, basal duct cells show high expression of basal markers (KRT5 and p63) and progenitor markers (p75/NGFR and CD49f). These findings suggest that these cells may be the resident progenitor of the submucosal gland duct, and due to their location *in situ*, they most-likely give rise to luminal duct cells. Indeed, the inferred differentiation trajectory from the pseudotime analyses support this notion (Figure 3.9 and Figure 3.10). This is further supported in the heatmap showing the top 50 most changed genes along the inferred pseudotime trajectory (Figure 3.11). The heatmap shows genes highly expressed in basal duct cells (such as KRT5, TACSTD2; blue arrows) to progressively decrease along the differentiation trajectory towards luminal duct cells. Conversely, genes highly expressed in luminal duct cells (KRT7, ALDH1A3; orange arrows) are low in basal duct cells and increase along the same trajectory.

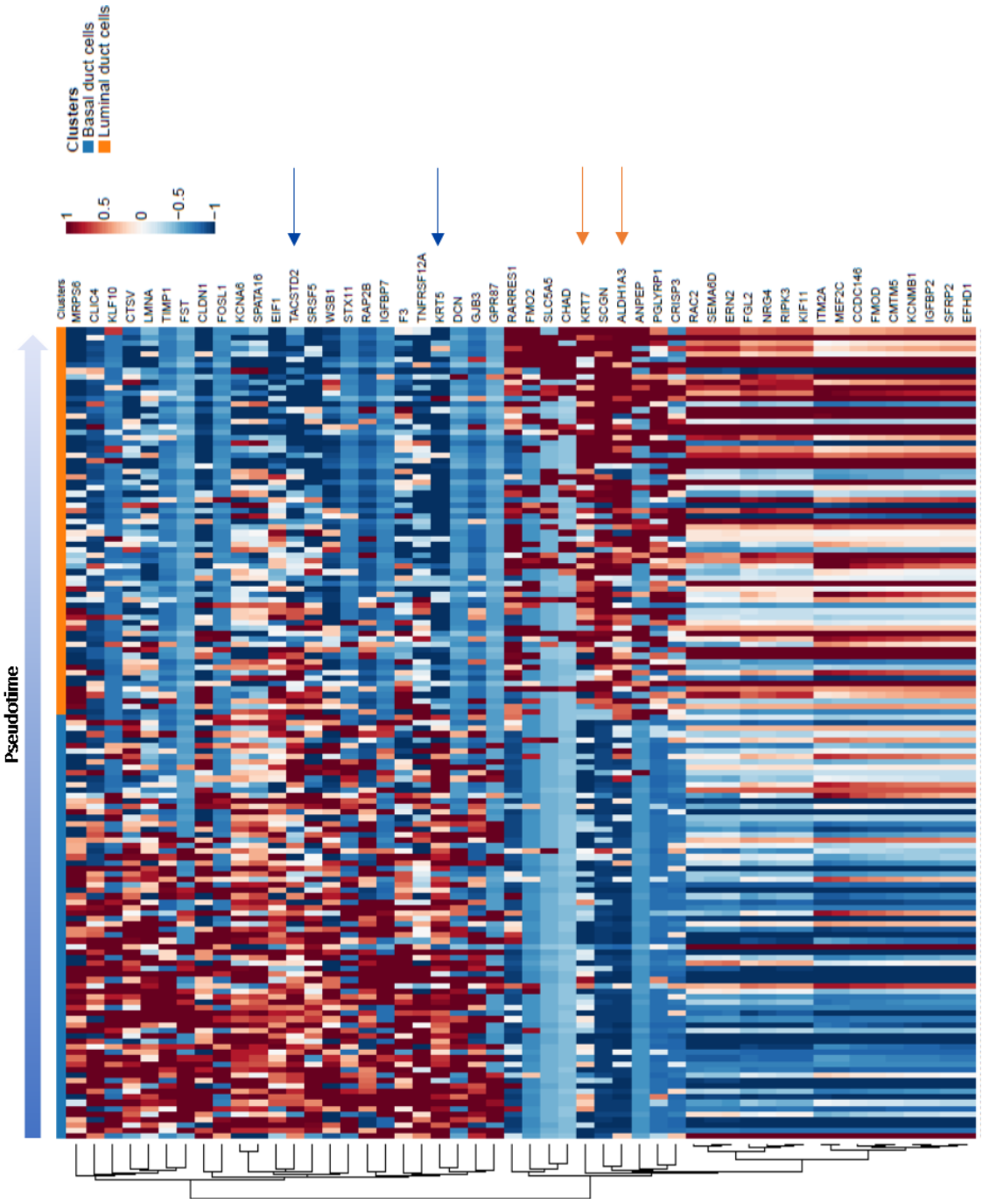

Clusters
Basal duct cells
Luminal duct cells
1
0.5
0
−0.5
−1
Pseudotime
Clusters
MRPS6
CLIC4
KLF10
CTSV
LMNA
TIMP1
FST
CLDN1
FOSL1
KCNA6
SPATA16
EIF1
TACSTD2
SRSF5
WSB1
STX11
RAP2B
IGFBP7
F3
TNFRSF12A
KRT5
DCN
GJB3
GPR87
RARRES1
FMO2
SLC5A5
CHAD
KRT7
SCGN
ALDH1A3
ANPEP
PGLYRP1
CRISP3
RAC2
SEMA6D
ERN2
FGL2
NRG4
RIPK3
KIF11
ITM2A
MEF2C
CCDC146
FMOD
CMTM5
KCNMB1
IGFBP2
SFRP2
EFHD1

Figure 3.11: Temporally changing gene expression along the inferred differentiation trajectory of pig SMG basal and luminal duct cells. Heatmap showing gene expression of the top 50 significantly changing genes along the differentiation trajectory of pig submucosal duct cells across pseudotime. Blue arrows indicate example genes highly expressed in the pig SMGs basal duct cells and orange arrows indicate genes highly expressed in the luminal duct cells.

3.3.6 Different trajectory inference methods suggest different paths of differentiation for gland acini cells

Similar to basal layer duct cells, myoepithelial cells showed expression of basal markers (KRT5 and p63) with one important distinction being that in humans but not in pig they also expressed the progenitor marker p75/NGFR as shown in Chapter 2 (Figures 2.19 and 2.21). These findings suggest that the myoepithelial cells could be the resident progenitor cells of the gland acini.

Connectivity and pseudotime differentiation trajectory of gland compartment cells using Slingshot with the myoepithelial cells as the origin, suggests a direct connectivity along a differentiation trajectory from myoepithelial to acinar to mucinous cells (Figure 3.12A and B). This inference suggests that acinar cells are a transitioning state between the myoepithelial cells and the mucinous cell type. Interestingly, these findings do not agree with the same analysis performed using monocle2 (Figure 3.12C and D). Monocle2 analysis suggests multiple branchings along the differentiation trajectory of myoepithelial cells where these cells can give rise to either mucinous or acinar cells directly (branching “4”). Additionally, in this analysis a number of cells with an acinar cell phenotype (blue dots) can be observed to cluster between the myoepithelial, acinar and mucinous cell clusters as indicated by black arrows in Figure 3.12C. This suggests that these cells are highly similar to acinar cells but may function as transitory cells towards the dichotomous lineages (acinar or mucinous). Furthermore, acinar cells can also be seen to branch into two individual branches (indicated by node “1”). From this branching, two other nodes (”2” and “3”) can be seen indicating more complexity within this population.

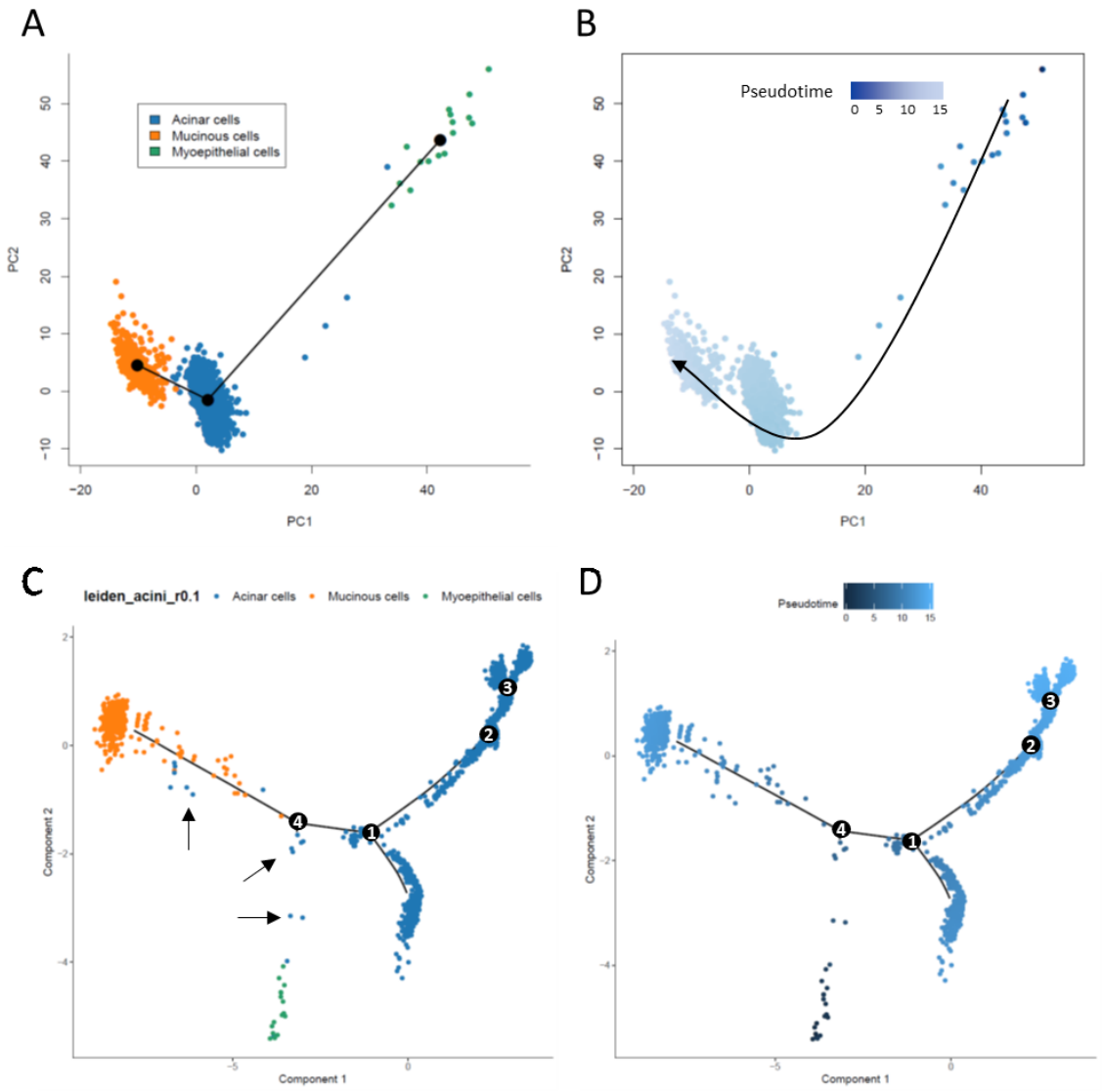

A
Acinar cells
Mucinous cells
Myoepithelial cells
PC1
PC2
B
Pseudotime
0 5 10 15
PC1
PC2
C
leiden_acini_r0.1
Acinar cells
Mucinous cells
Myoepithelial cells
Component 1
Component 2
D
Pseudotime
Component 1
Component 2

Figure 3.12: Trajectory inference and pseudotime analyses of submucosal gland cells only (ie ductal cells excluded). Slingshot PCA plot of gland cells inferred differentiation trajectory (A). Slingshot pseudotime diffusion plot of gland cells (B). Monocle PCA plot of gland cells inferred differentiation trajectories (C). Arrows indicate potential transitory cells of acinar origin. Monocle pseudotime diffusion plot of acinar cells (D). Numbered black dots in panels C and D represent lineage branching in the acinar cell cluster indicating more complexity within the cluster.

3.4 Discussion

In the previous chapter, a number of markers were used to characterise the different cell types found in pig and human SMGs using anti-human antibodies that showed cross-reactivity with pig. However, the number of markers that can be tested in this manner is limited. This could be remedied using single cell RNA-sequencing, which provides a more high-throughput way of characterising large number of cells and potentially identify cell specific markers.

In Chapter 2, myoepithelial, acinar, basal and luminal duct cells were histologically identified based on morphology and staining patterns in pig SMGs. Those findings were further confirmed with the sc-RNA seq results of $EPCAM^+$ sorted cells in this chapter. The sc-RNA seq results demonstrate the presence of myoepithelial, two ductal cell subtypes and two phenotypically secretory cell subtypes. The sc-RNA seq results confirm the presence of highly expressed markers of muscle tissues (ACTA2, MYL9) in myoepithelial cells cluster, basal marker (KRT5) in both myoepithelial and basal duct cells clusters and columnar marker (KRT7) in luminal duct cells cluster. The interesting findings pertain more to the acinar cells, which could be separated into the two largest clusters confirming some of the disparity found between human and pig SMGs in Chapter 2. The second largest cluster shows a clear mucinous phenotype with the expression of TFF3, a marker found in both normal human SMGs and BO. Furthermore, expression of the gastric marker TFF2, which is also found in BO[218] can also be found in this cluster. Additionally, expression of lysozyme indicates this cluster to have antibacterial function. This function fits with findings in human SMG gland cells, which were also shown to be positive for lysozymes[183].

The largest cluster, also showing a secretory phenotype (AQP3, SCGB1D1), seems to have a role in potentially promoting tissue protection and homeostasis. This is indicated through expression of genes such as MGST1 and MGST3, which are important for prostaglandin E_2 (PGE_2) production, and Bradykinin (KNG1). PGE_2 was shown to be secreted in human SMGs and to have an important role in bicarbonate (HCO_3^-) secretion to neutralize gastric acid[6,219,220]. KNG1 is known to have an antimicrobial activity[221]. Overall, these data support the role of pig SMG acinar cells in protecting the oesophageal epithelium from ingested material and clearing reflux components such as acid and bile[6].

Immunohistochemistry and multiplex-IF staining in the previous chapter showed distinct staining profiles for human and pig SMG acinar cells. UEA and WGA lectin staining showed lower correlation in pig compared to human SMGs potentially indicating two different acinar cells to be present in pig SMGs. Interestingly, gene ontology analysis (Appendix Figure 4 and Appendix Table 2) of the mucinous cells cluster shows enrichment of the GDP-L-fucose biosynthetic process. L-fucose is a hexose deoxy sugar residue found on cell membrane glycans and glycolipids[222] of multiple cells, which can be stained using the lectin UEA. These findings indicate that the subset of cells positive for UEA are most likely the mucinous cells within the pig SMGs.

In a recent paper by Owen *et al*.[223], sc-RNA seq was performed on human BO, normal oesophageal, gastric and duodenal tissues. In their work, Owen *et al*. showed that the population of oesophageal cells could be subdivided into two clusters (Appendix Figure 5A, black and red arrows). One cluster showed high expression of squamous markers indicating that the sequenced cells were derived from the squamous epithelium (Appendix Figure 5C and D). However, the second normal oesophageal tissue derived cluster showed expression of TFF3 (Appendix Figure 5B), which led the authors to the conclusion that they sequenced normal SMG cells[202]. Interestingly, in their clustering analysis, they showed that the TFF3$^+$ normal oesophageal derived cells clustered together with the BO sequenced cells(Appendix Figure 5A and B)[223], indicating high transcriptional similarity between BO and human SMGs. These data seem to further support the potential SMG origin of BO. However, upon further analysis of their published sc-RNA seq data, the TFF3$^+$ normal oesophageal cells the authors labelled as SMG cells can be found to also be MUC5AC$^+$ (Appendix Figure 6A and B). This contradicts the data shown in Chapter 2 (Figure2.11A and A') where both normal human and pig SMGs are negative for this BO marker. Consequently, the overlap of the oesophageal TFF3$^+$ positive and BO cells in the clustering analysis of Owen *et al*., coupled with their positivity for a bona fide Barrett's marker suggests that the authors may actually have analysed BO cells cross-contaminating the normal oesophageal cells, and not normal human SMGs. Furthermore, comparison of the pig SMG data to the BO data published by Owen *et al*., clustering analysis showed their BO cells to cluster in between the pig acinar and mucinous cells (Appendix Figure 7). This clustering result indicates a shared similarity between BO and the secretory cells in pig SMGs but not a complete overlap as the data of Owen *et al*. suggests for human SMGs. The implication of these findings is not clear at the moment as it is not known how comparable normal human and pig SMGs are at the transcriptional level. Attempts to do this for this thesis were unsuccessful due to insufficient human sample.

Clustering analysis of normal human SMGs and BO could be expected to show a similar result to clustering of pig SMGs with BO, with human secretory acinar cells clustering close, but not fully overlapping with BO cells. A key difference to keep in mind is that human SMGs may not show two different secretory cell types as pig SMGs show. As shown in Chapter 2 (Figures 2.25), human but not pig gland acinar cells show complete overlap for the lectin markers (WGA and UEA), suggestive of a more homogenous acinar population in human SMGs.

To determine the differentiation trajectories of stem cells at the single cell level, differentiation trajectory inference methods, tracking pseudo-temporal changes in gene expression between individual single cells, were developed and used in developmental studies[208]. As both ductal and gland acini tissues contain cells with bona fide stem/progenitor markers, it could be possible to assess their potential using this method. Not surprisingly, basal duct cells could easily be shown to have connectivity to luminal duct cells, and have a diffuse expression profile of marker genes, which decreases towards the luminal lineage after cell ranking in pseudotime. Interestingly, trajectory

inference suggests the luminal duct cells to be a precursor for the acinar cells in the gland compartment. Other organs such as the prostate and breast were shown to have luminal progenitor cells[217,224-226]. Specifically, in breast luminal progenitors were also shown to express SOX9 and ALDH1A3[217]. These cells were also shown to be able to differentiate into mature ductal and alveolar cells[226]. Consequently, the inferred differentiation trajectory data could suggest a similar role for the luminal duct cells whereby (a fraction of) these cells have progenitor capacity capable of giving rise to mature luminal duct cells and acinar cells. Still, the current histologic data did not show differences in staining profile within the luminal duct cells population nor did the transcriptomic analyses of the sc-RNA seq data.

In the gland compartment, the myoepithelial cells show expression of basal/progenitor cell markers suggesting these cells to be resident progenitor cells. Furthermore, the myoepithelial cells in the submucosal glands of a morphologically and functionally similar organ such as the trachea, were shown to contribute to airway epithelium repair following severe naphthalene damage[227]. A clear connectivity between myoepithelial cells and acinar or mucinous cells could not be conclusively demonstrated using this pseudotime analysis as evident by the gaps between the cell clusters in the diffusion and trajectory inference analyses. This is potentially due to the low number of myoepithelial cells leftover for analysis after filtering. Furthermore, as SMGs are neither a highly regenerative nor a proliferative tissue in homeostasis, it is conceivable that the pool of differentiating myoepithelial cells is too small to be able to show a clear connectivity to the acinar cells. These two issues could be remedied by sequencing more cells. Historically, stem cells are known to be a very small population within their resident tissue. The myoepithelial cells in the glands certainly fit this criterium as they represent 0.7% of total SMG cells according to the sequencing results.

The disparity between the trajectory inferences methods used for gland acini cells further reflects the need for a higher number of cells to be sequenced and/or higher sequencing resolution. For differentiation trajectory inference, the Slingshot algorithm was shown to perform better than monocle2[208]. Still, for the current dataset, it is not possible to determine which method is more accurate at modelling the differentiation state as no prior biological knowledge is available for comparison.

An unavoidable and definite concern of performing single cell RNA sequencing of solid tissues, such as the oesophageal SMGs, is the need for tissue digestion to generate single cells. A recent study[228] showed the effects of sample preparation on the quality of the RNA used for RNA sequencing. Tissue digestion led to technical and biological variation, and induction of stress related gene expression[228]. For the study of pig SMGs, the tissue digestion protocol was not adjusted to accommodate for the sensitivity of RNA to potential damage from prolonged digestion at 37˚C. O'Flanagan *et al*.[228] demonstrated in their work that the proto-oncogenes FOS and JUN were higher expressed in collagenase treated tissues at 37˚C when compared to cold active protease. In the current work, JUN mRNA was not detected but FOS mRNA was detected (not shown) and was found in all clusters. Whether this expression is due to the digestion process or

is naturally occurring within the SMGs is not clear from the current available data. Furthermore, O'Flanagan *et al*. found the nuclear factor kappa-light-chain-enhancer of activated B cells (NFκB) pathway to be highly activated in the collagenase at 37˚C compared to cold active protease digested tissue[228]. Gene ontology analysis in Appendix Table 2 did not show an enrichment of this pathway in any cluster. It is also possible that a number of stress related genes, activated from the digestion process, are not properly annotated in the pig genome and would therefore not be identified during sequencing. To conclusively determine the effect of the digestion process on the pig SMGs RNA, a similar approach should be taken as performed by O'Flanagan *et al*. while keeping all other variables such as digestion time and all manual procedures unchanged. Another potential issue to consider is the extent to which the pig genome is sequenced. In the current work, about half (54.7%) of sequenced reads were confidently mapped to the pig genome (Figure 3.1B). In order to minimize the effects the digestion process could have on the single cell analysis, strict selection of sequenced cells for downstream analyses was performed. The selection was based on the percentage of mitochondrial genes detected, and the exclusion of outliers based on the number of reads sequenced and genes identified as markers for cell integrity. Heat shock proteins were also considered as these would be induced by the 37˚C digestion process. However, the latter were not used for the selection as they are also known to play a role in chaperoning cell secretions and therefore might be vital for cell type characterisation[229].

In conclusion, sc-RNA seq showed further similarities between pig and human SMGs in terms of marker expression (TFF3) and surprisingly also BO (TFF2). Additionally, important markers for the role of SMGs in the protection of the oesophageal epithelium (acid clearance and antimicrobial) could be identified. These data further support the idea of using pig SMGs as a substitute for human SMGs in the study of BO. Chapter 2 showed basal duct cells expressing basal (KRT5 and p63) and progenitor markers (CD49f and p75). The latter could be used to purify this population of cells for *in vitro* studies. Due to the low number of identified sequenced genes, it was not possible to identify additional cell surface markers to use for further purification of the gland compartment cells. This could be remedied by using lower throughput/high resolution sequencing techniques, which could yield higher numbers of identified genes.

Chapter 4

3D culture and characterisation of porcine SMG progenitor cells

4.1 Introduction

The potential role of the submucosal glands in the development of Barrett's has not been appreciably explored due to the lack of suitable animal or *in vitro* models. In order to explore their possible role without the use of animal models, an *in vitro* culture system capable of recapitulating the *in vivo* niche is necessary.

Recent advances in understanding the composition and dynamics of diverse stem cell niches have allowed for the generation of organoids from diverse tissue types (primarily of epithelial origin)[168]. Such advances in the development of organoid culture systems have paved the way to model, isolate, and better characterise the role of multiple signalling pathways in tissue development and disease. The greatest advantage of using organoid culture systems, over two-dimensional (2D) culture systems, is their potential to reproduce (to some degree) the 3D structure, cellular distribution and polarity *in vitro*. This allows for more accurate modelling of the *in vivo* environment and function of the tissue of origin[230]. Consequently, this makes organoids a prime tool to use for studying the aetiology of different diseases.

Previous chapters demonstrated the presence of cells in the SMGs expressing stem/progenitor markers. Specifically, the basal duct cells of the ductal compartment and the myoepithelial cells of the gland compartment could be the resident progenitor cells of the pig SMGs. Whether only one or both cells are progenitors could not be conclusively determined from the trajectory inference data in Chapter 3. However, these questions could potentially be answered using 3D *in vitro* culture systems.

The goal of this chapter is to use the markers identified in the previous chapters to sort putative progenitor cells present in the submucosal glands and functionally characterise their progenitor potential. To achieve this, pig SMG cells were cultured in 3D in different culture conditions and the derived spheroids were characterised. Furthermore, an attempt at assessing their potential to undergo Barrett's (like) metaplastic changes was made, using ectopic expression of the intestinal transcription factor CDX2.

4.2 Material and methods

4.2.1 Pig tissue digestion

As described in section 3.2.1.

4.2.2 Pig cells FACS staining

As described in section 3.2.2. Primary antibodies used can be found in Table 4.1.

Table 4.1: Antibodies used for FACS of pig SMGs

Marker	Primary antibody / dilution	Catalogue number	Manufacturer	Secondary antibody / dilution
EPCAM (pig)	EPCAM Antibody (aa25-265) / 1:100	LS-C374109	LSBio	F(ab) Dylight 405 (711-476-152, immunojackson) / 1:250
CD49f-AF647	Alexa Fluor® 647 anti-human/mouse CD49f Antibody / 1:200	313609	Biolegend	NA
CD271-PE	CD271 (LNGFR)-PE, human / 1:200	130-112-790	Miltenyi Biotec	NA

NA: not applicable

4.2.3 Pig spheroids culture and human SMG culture

For bulk pig SMG spheroid culture, 48 well plates (CLS3548, Corning) were preheated in an incubator at 37˚C at least 10 min before cell seeding. Bulk cells were resuspended in cold growth factor reduced, phenol red free Matrigel™ (356231, Corning) in 20 µl per well. After seeding the cells, the plate was placed in the incubator for 15 min to allow the Matrigel™ to solidify, after which 250 µl mTEC or intestinal organoid culture media (Table 4.2 and 4.3) was added.

Sorted gland ($EPCAM^+/CD49f^+/p75^-$) or duct ($EPCAM^+/CD49f^+/p75^+$) were seeded at 5000 gland cells or 1000 duct cells in 20 µl per well, respectively. After seeding the cells, plate was placed in the incubator for 15 min to allow the Matrigel™ to solidify, after which 250 µl SMG culture media (Table 4.4) was added.

Passaging of gland derived spheroids was performed by resuspending spheroids in 1 ml of ice cold Advanced DMEM with no additives and triturated using a gauge 26 needle ten times. Broken spheroids were spun down at 400 g for 5 min. The spheroid pellet was resuspended in Matrigel™ and re-seeded in 48 well plate. Spheroids were split at a ratio of 1:4-1:8 every 7-12 days.

Passaging of dense ductal spheroids was performed by resuspending spheroids in TrypLE (12563011, Gibco) and incubating for 5 min at 37˚C after which spheroids were triturated with a P200 pipette. Cells were washed once with PF5, spun down at 400 g for 5 min and resuspended in Matrigel™ before seeding into a 48 well plate. Spheroids were split at a ratio 1:4-1:8 every 7-9 days. At seeding and passaging, SMG culture media (Table 4.4) was supplemented with 10 µM ROCK Inhibitor Y-27632 (Y0503, Sigma-Aldrich) which was omitted during subsequent media changes. Media was changed every other day (3 to 4 times a week).

For *ex vivo* culture of human SMGs, dissected SMGs were minced into small fragments (~2 mm) with surgical scissors, resuspended in Matrigel, seeded and cultured as described previously for pig SMG spheroids. Human SMG passaging was performed as

described for pig dense ductal spheroids in the previous paragraph. SMG culture media for human SMGs (Table 4.4) was changed every other day.

Table 4.2: mTEC culture media

Component	**Catalogue #**	**Manufacturer**	**Final concentration**
DMEM/F12	NA	*Made in-house (media kitchen)	NA
L-Glutamine	25030081	Gibco	2 mM
Penicillin/Streptomycin (100X)	15140122	Gibco	1 X
Foetal Bovine Serum (FBS)	10099	Gibco	5% (v/v)
EGF	100-15	Peprotech, Rocky Hill, NJ	50 ng/ml
Insulin	I0516	Sigma-Aldrich	0.1 µg/ml
bFGF	100-18B	Peprotech	25 ng/ml
Insulin Transferrin Selenium (ITS) (100X)	I3146	Sigma-Aldrich	1 X
Bovine Pituitary Extract	13028014	Thermo Fischer Scientific	140 µg/ml
Cholera toxin	C8052	Sigma-Aldrich	100 ng/ml
Retinoic acid	R2625	Sigma-Aldrich	10 nM

NA: not applicable

Table 4.3: intestinal organoid culture media

Component	Catalogue #	Manufacturer	Final concentration
Advanced DMEM/F12	12634010	Gibco	NA
HEPES	15630080	Gibco	10 mM
L-Glutamine	25030081	Gibco	2 mM
Penicillin/Streptomycin (100X)	15140122	Gibco	1 X
Nicotinamide	72340	Sigma-Aldrich	10 mM
B27 (50 X, serum replacement)	A3582801	Gibco	1 X
EGF	100-15	Peprotech	50 ng/ml
Noggin	120-10C	Peprotech	100 ng/ml
A8301 (TGF beta inhibitor)	SML0788	Sigma-Aldrich	1 µM
SB202190 (p38MAPK inhibitor)	S7067	Sigma-Aldrich	3 µM
Wnt conditioned media*	NA	NA	50% (v/v)
Rspondin conditioned media*	NA	NA	20% (v/v)

NA: not applicable

* wnt and Rspondin conditioned media were produced as described by Broutier *et al*.[231]

Table 4.4: Pig SMG spheroid and human SMG mince culture media

Component	Catalogue #	Manufacturer	Final concentration
Advanced DMEM/F12	12634010	Gibco	NA
HEPES	15630080	Gibco	10 mM
L-Glutamine	25030081	Gibco	2 mM
Penicillin/Streptomycin (100X)	15140122	Gibco	1 X
Nicotinamide	72340	Sigma-Aldrich	10 mM
B27 (50 X, serum replacement)	A3582801	Gibco	1 X
EGF	100-15	Peprotech	50 ng/ml
Noggin	120-10C	Peprotech	100 ng/ml
A8301 (TGF beta inhibitor)	SML0788	Sigma-Aldrich	1 µM
SB202190 (p38MAPK inhibitor)	S7067	Sigma-Aldrich	3 µM
bFGF	100-18B	Peprotech	50 ng/ml
Insulin	I0516	Sigma-Aldrich	10 µg/ml
ITS (100X)	I3146	Sigma-Aldrich	1 X
PGE_2	HY-101952	MedChemExpress, Monmouth Junction, NJ	10 nM
Wnt conditioned media*	NA	NA	50% (v/v)
Rspondin conditioned media*	NA	NA	20% (v/v)

NA: not applicable

* wnt and Rspondin conditioned media were produced as described by Broutier *et al*.[231]

4.2.5 BAR-T, CP-A and CP-B cells culture

Non dysplastic human Barrett's oesophagus BAR-T (developed by Rhonda Souza[232]), CP-A (ATCC® CRL-4027™; Manassas; VA) and high-grade dysplasia CP-B cell lines (ATCC® CRL-3216™) were seeded in T75 flasks (C7231, Greiner) and cultured in supplemented MCDB153 media (Table 4.5). Cells were passaged upon reaching confluency using 0.25% trypsin-EDTA (Thermo Fischer Scientific), washed with supplemented MCDB153 media, spun down at 400 g for 5 min and seeded into a new T75 flask.

Table 4.5: Supplemented MCDB153 culture media

Component	Catalogue #	Manufacturer	Final concentration
MCDB 153	12634010	Gibco	NA
L-Glutamine	25030081	Gibco	2 mM
Penicillin/Streptomycin (100X)	15140122	Gibco	1 X
Foetal Bovine Serum	10099	Gibco	5% (v/v)
EGF	100-15	Peprotech	50 ng/ml
ITS (100X)	I3146	Sigma-Aldrich	1 X
Bovine Pituitary Extract	13028014	Thermo Fischer Scientific	140 µg/ml
Adenine	A2786	Sigma-Aldrich	20 µg/ml
Hydrocortisone	H0888	Sigma-Aldrich	400 ng/ml

NA: not applicable

4.2.6 Immunohistochemistry

As described in sections 2.2.2 and 2.2.3.

4.2.7 Virus production

HEK293T cells (ATCC® CRL-3216™) were seeded at 4 to 5x10^6 cells in a T75 flask in DMEM with 10% FBS and 1X Penicillin/Streptomycin. The following day a transfection mix was prepared by mixing packaging vectors (pEQ-PAM3-E and pRDF-RD114) with retroviral expression vector (MSCV-hCDX2-GFP or MSCV-hHNF4A-DsRed) at a ratio of 1:1:1 to generate a total of 13.5 µg DNA mix for transfection in 340 µl DMEM (no supplements). The transfection mix was incubated at RT for 5 min. Polyethyleneimine (PEI, 23966, Polysciences, Warrington, PA) was prepared by diluting 67.5 µg in 200 µl DMEM (no supplement) for 5 min at RT. Finally, the two mixes were mixed together and incubated for 15 min at RT before adding to HEK293T cells for overnight incubation.

Media was refreshed the following day (day 2). Virus containing supernatant was collected once on day 3, twice on day 4 and once day 5. At collection, media was filtered through a 0.45 µM filter (CLS431225, Sigma-Aldrich) to remove any dead cells and stored at 4˚C until after the final collection on day 5. Finally, the viral supernatant was concentrated using Amicon® Ultra 15 ml Filters (C7715, Merck) followed by aliquoting in cryovials and stored at -80˚C.

4.2.8 Viral transduction

Pig SMG spheroids were incubated for 5-10 min in TrypLE at 37˚C followed by trituration. Cells were washed once with PF5, spun for 4 min at 500 g and resuspended in 100 µl Advanced DMEM/F12 with no supplements in a 96 well plate. To the cell suspension, 100 µl of viral supernatant supplemented with 8 µg/ml polybrene (S2667, Sigma-Aldrich) was added and mixed by gentle pipetting. The 96 well plate was sealed with parafilm

and centrifuged at 600 g at 32˚C for 1 hour. After spinning, the plate was incubated at 37˚C for 3 to 4 hours before collecting the cells, pelleting at 500 g for 4 min and resuspending in Matrigel™ for seeding in 48 well plates as described in section 4.2.3.

4.3 Results

4.3.1 Dissected pig SMG cells cultured in different media give rise to spheroids with different morphologies

To interrogate the potential presence of SMG progenitor cells, dissected pig SMGs were digested to generate a bulk pool of cells, which were resuspended in Matrigel and seeded into 48 wells plate. Media developed for the culture of lung (mTEC) and intestinal organoids were used as these are well established for culturing organoids (Figure 4.1).

Culture of pig SMG cells in mTEC media gave rise to a mixture of cystic and dense spheroids (Figure 4.2 Full well Brightfield, A and A'). The dense spheroids were lost upon passaging indicating that the mTEC media does not support their long-term growth and maintenance. Conversely, cystic spheroids were maintained after passaging (up to five passages were tested). Histologically, cystic spheroids show a columnar phenotype with single layer and multi-layered regions. The cystic spheroids are positive for Alcian blue indicating mucin production (Figure 4.2C). Dense spheroids show more basophilic basal layer cells and eosinophilic, flattened, and elongated apical cells and were negative for Alcian blue (Figure 4.2B', C').

Both cystic and dense spheroids show expression of the columnar marker SOX9 (Figure 4.2D, D'; black arrows). However, only cystic spheroids show weak KRT7 staining in the apical layer cells (Figure 4.2E, E'; black arrows). The squamous and basal marker KRT5 was found mainly in the basal layer cells of the cystic spheroids (Figure 4.2F). In the dense spheroids, KRT5 staining is strongest in the basal layer and seems weaker in the apical layers (Figure 4.2F'). The other squamous/basal marker, p63, is mostly located in the basal layers of the cystic spheroids, similar to KRT5 whereas in dense spheroids it can also be found in some of the supra-basal layers (Figure 4.2G, G'; black arrows). Finally, the squamous markers KRT13 is present in the apical layer whereas the basal layer is negative (Figure 4.2H, H').

As shown in chapter 2, the lectins UEA and WGA stain the mucin producing cells in the SMGs. UEA and WGA staining can be found at the basal and apical layers of the cystic mucin producing spheroids and is ubiquitously distributed (Figure 4.2I and J). Interestingly, UEA staining can also be found in the lumen of the dense spheroids (Figure 4.2I'). WGA staining is ubiquitous throughout the dense spheroids (Figure 4.2J').

Culture of bulk pig SMG cells in intestinal organoid media resulted in spheroids with only a dense morphology (Figure 4.3 Full well BF and A). H&E histologic characterisation shows the apical layer cells to be strongly eosinophilic and have an elongated squamoid morphology (Figure4.3B). This is similar to the dense spheroids grown in mTEC media. The basal/squamous marker KRT5 is present throughout the dense spheroids and in

some regions seems stronger in the basal layer compared to apical layer. P63 staining is found mainly in the basal layer cells and occasionally in the suprabasal cells (Figure 4.3D; black arrows). As with the dense and cystic spheroids in the mTEC media, KRT13 is absent in the basal layer and more prominent in the apical layers (Figure 4.3E). Finally, the columnar marker KRT7 is absent throughout the spheroids whereas SOX9 is prominent (Figures4.3F and G, black arrows).

These data suggest that the intestinal organoid media supports the growth of cells, which give rise to dense spheroids only. These dense spheroids show the same staining pattern for basal/squamous and columnar markers as the dense spheroids, that are progressively lost in the mTEC media suggesting that they may arise from the same cell. The data also suggest that the mTEC media contains factors not present in the intestinal organoid media, which are necessary for the culture of cystic spheroids.

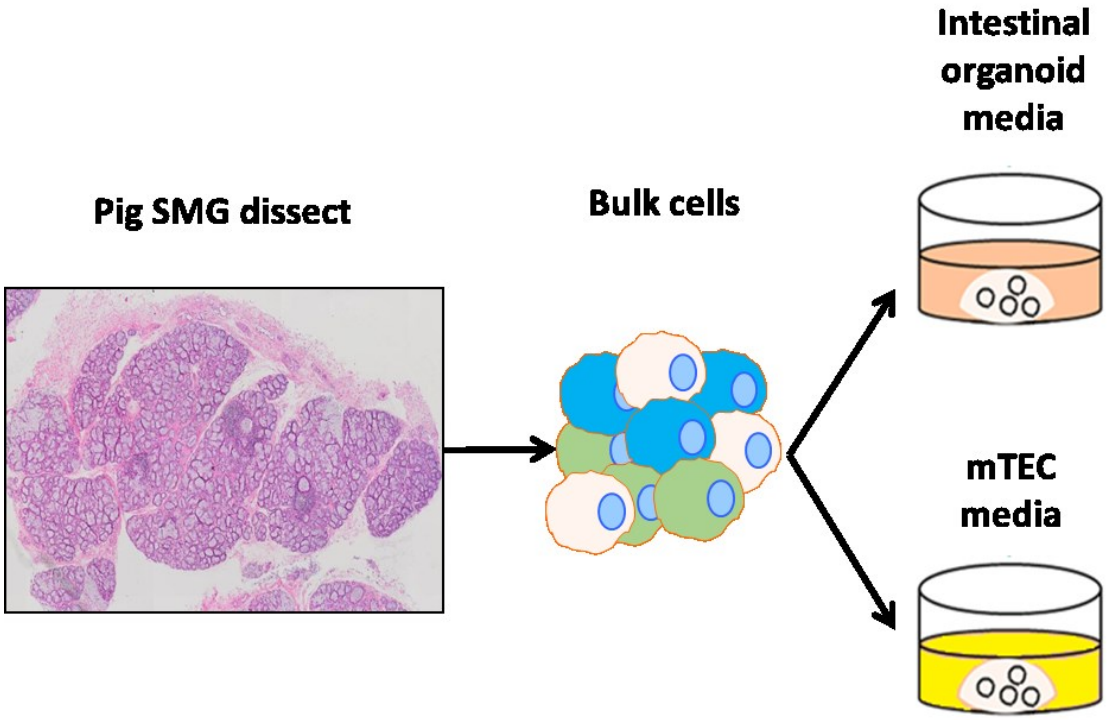


Figure 4.1: Schematic of the initial approach used to culture pig SMG cells. Dissected pig SMGs were digested and acquired cells were seeded in Matrigel and given either intestinal organoid or mTEC culture media for culturing.

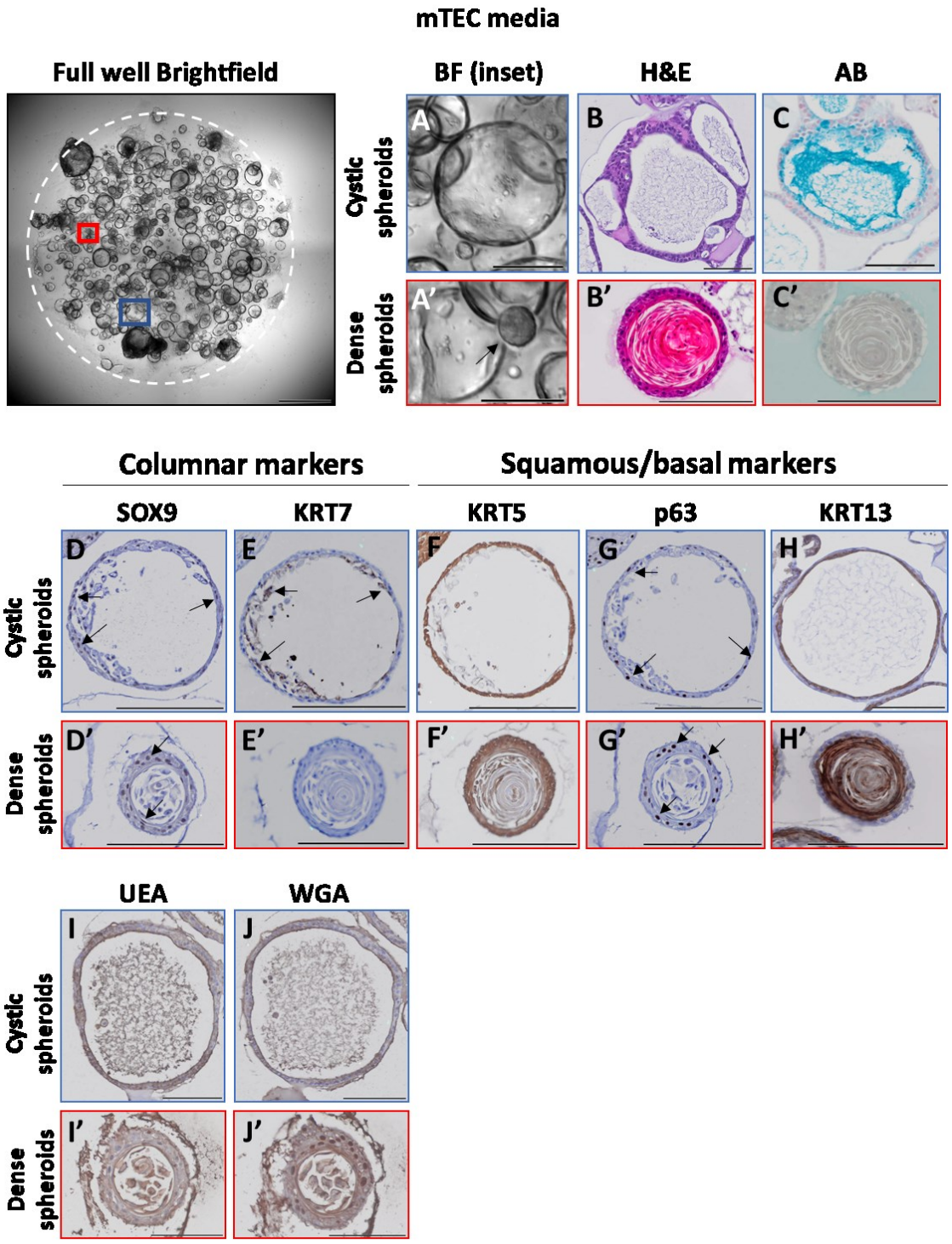


Figure 4.2: Bulk cells acquired from dissected and digested pig SMGs were seeded in Matrigel and cultured in mTEC media for two weeks. Full well Brightfield panel shows a representative image of the resulting spheroids. Red square in Full well panel shows the inset of dense spheroids and blue square shows inset for cystic spheroids (scalebar=1 mm). BF insets of cystic and dense pig SMGs spheroids grown in mTEC media (A-A'). Representative images of H&E (B-B', n=5), AB (C-C', n=3), SOX9 (D-D', n=2), KRT7 (E-E', n=3), KRT5 (F-F', n=3), p63 (G-G', n=3), KRT13 (H-H', n=3), UEA (I-I', n=2) and WGA (J-J', n=2) staining of spheroids grown in mTEC media. Scalebars=100 µm. Arrows indicate cells or nuclei positive for marker shown in the respective panel.

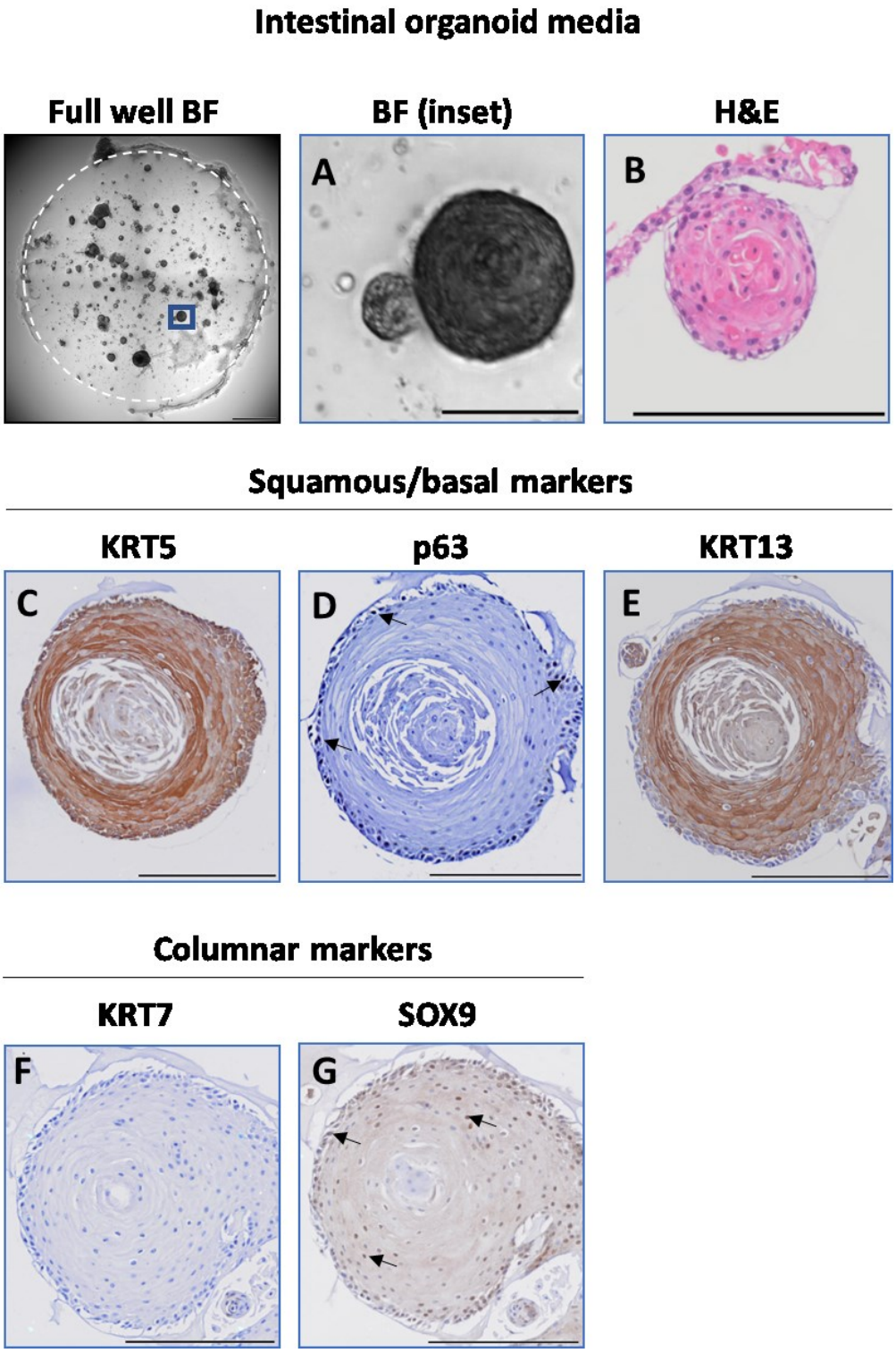


Figure 4.3: Bulk cells acquired from dissected and digested pig SMGs were seeded in Matrigel and cultured in intestinal organoid media for two weeks. Full well Brightfield panel shows a representative image of resulting spheroids. Blue square in Full well panel shows the inset of dense spheroids (scalebar=1 mm), BF inset of dense pig SMGs spheroids grown in intestinal organoid media (A). Representative images of H&E (B, n=3), KRT5 (C, n=2), p63 (D, n=2), KRT13 (E, n=2), KRT7 (F, n=2) and SOX9 (G, n=2) staining of spheroids grown in intestinal organoid media. Scalebars=200 µm. Arrows indicate nuclei positive for marker.

4.3.2 CD49f^{+}/p75^{-} cells generate cystic and dense spheroids

As shown in Chapter 2, the cell surface markers CD49f and p75 marked basal layer ductal cells, which were also positive for the basal markers p63 and KRT5. Similarly, in the gland compartment of the SMGs CD49f marked myoepithelial cells in pig (in human SMGs myoepithelial cells are also p75 positive). Correspondingly, these two populations can be identified by flow cytometric analysis of pig SMGs in addition to a third population of cells (Figure 4.4). Figure 4.4F shows a CD49f and p75 double positive population indicating the basal duct cells population, a CD49f single positive p75 negative population indicating gland cells and a third population of EPCAM^{+} but p75 and CD49f double negative cells (population R7). The CD49f single positive population comprises the majority of the EPCAM^{+} cells, consistent with gland cells being more abundant than duct cells, while the other two populations are near equal in proportions. This potentially indicates the double negative population to be luminal duct cells population as they are negative for CD49f and p75 and should be proportionately equal to the basal layer duct cells.

After sorting the EPCAM+/CD49f+/p75- population of cells, which comprises the submucosal gland compartment cells, culture in the intestinal organoid media as it is completely defined was attempted. Culture in this media showed a low number of very slow growing cystic spheroids after 30 days (not shown). In an attempt at remedying this slow growth, a new completely defined SMG culture media was developed to support the growth of this population (Table 4.4). The SMG media was developed based on the intestinal organoid media supplemented with defined factors of the mTEC media. This media still supports the growth of the sorted ductal cells population even after adding new factors.

Pig SMG gland cells (EPCAM^{+}/CD49f^{+}/p75^{-}) seeded in Matrigel and cultured for up to nine passages in SMG media, grow into two morphologically different spheroids (Figure 4.5A). The majority of spheroids show a cystic morphology (blue outlined inset) versus dense looking spheroids (red outlined inset).

Flow cytometry re-analysis of cultured spheroids after culturing for 2-2.5 weeks (P0) show largely two population of cells (Figure 4.5B). The two populations of cells comprise a CD49f^{+}/p75^{-} population similar to the seeded cells, and a CD49f^{+}/p75^{+} double positive population similar to the ductal population. Interestingly, after subsequent passaging and culture, the proportion of double positive cells diminishes whereas the proportion of CD49f^{+}/p75^{-} increases (P1 and P3, Figure 4.5B), indicating an enrichment of a progenitor cell population present in CD49f^{+}/p75^{-} sorted cells or a loss of a CD49f^{+}/p75^{+} progenitor pool.

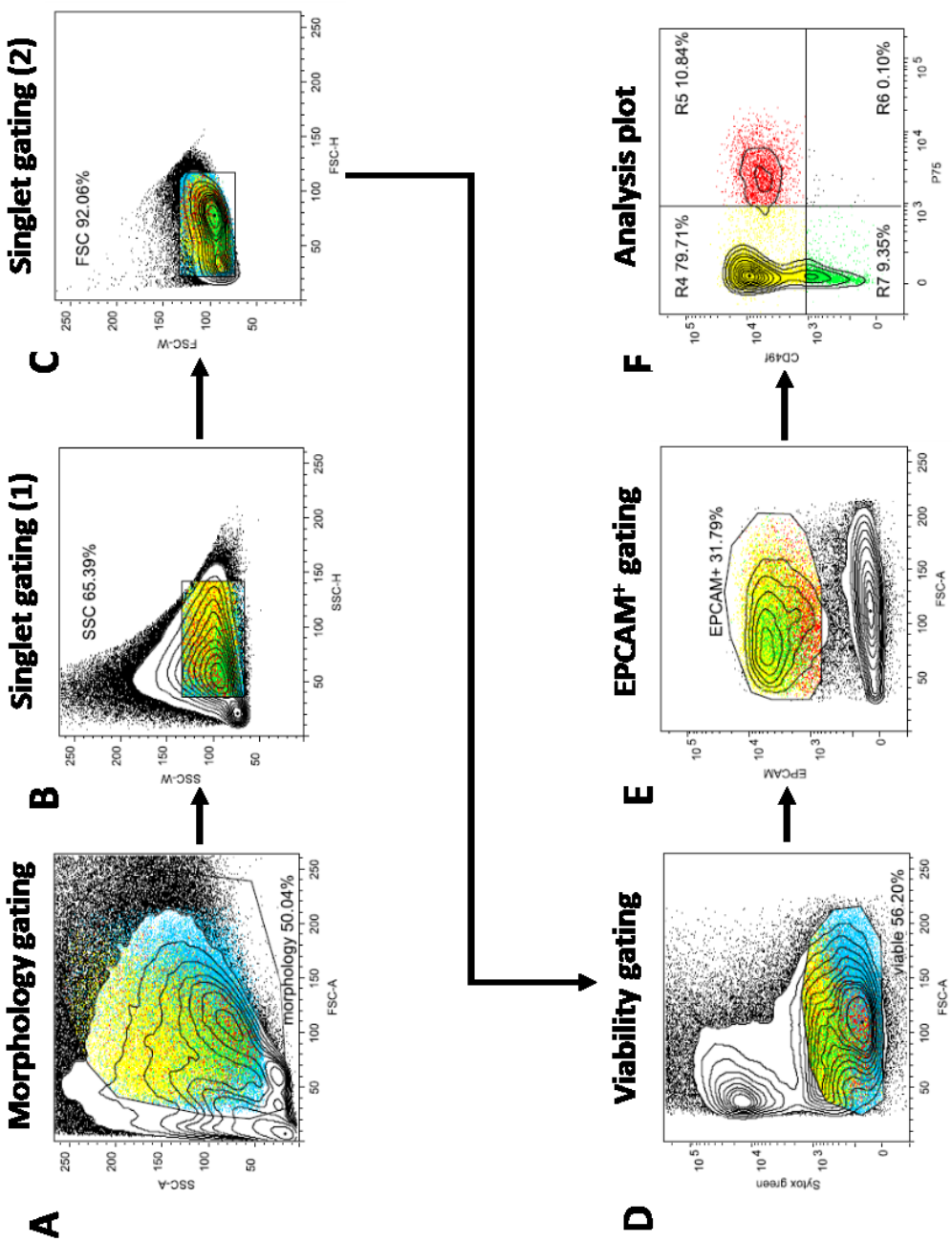

A
Morphology gating
morphology 50.04%
SSC-A
FSC-A
B
Singlet gating (1)
SSC 65.39%
SSC-W
SSC-H
C
Singlet gating (2)
FSC 92.06%
FSC-W
FSC-H
D
Viability gating
viable 56.20%
Sytox green
FSC-A
E
EPCAM+ gating
EPCAM+ 31.79%
EPCAM
FSC-A
F
Analysis plot
R5 10.84%
R4 79.71%
R7 9.35%
R6 0.10%
CD49f
P75

Figure 4.4: Gating strategy for FACS sorting $EPCAM^{+}CD49f^{+}p75^{-}$ gland cells and $EPCAM^{+}CD49f^{+}p75^{+}$ ductal cells from pig SMGs. Morphology panel (A) shows gating of bulk cells based on morphology (ie. excluding debris), Singlet gating 1 (B) shows FSC gating for selecting single cells, Singlet gating 2 (C) shows FSC gating for selecting single cells, Viability panel (D) shows gating for viable cells (negative for Sytox green) and $EPCAM^{+}$ gating panel (E) shows gating for EPCAM positive cells. Analysis plot (F) of $p75^{-}/CD49f^{+}$ (gland cells), $p75^{+}/CD49f^{+}$ (ductal cells) and $p75^{-}/CD49f^{-}$ cells. Red dots represent the $p75^{+}/CD49f^{+}$ (ductal cells) backgated population, yellow dots represent $p75^{-}/CD49f^{+}$ (gland cells) backgated population, green dots represent $EPCAM^{+}/p75^{-}/CD49f^{-}$ (undefined cells) backgated population and blue dots represent the viable (sytox green negative) backgated population.

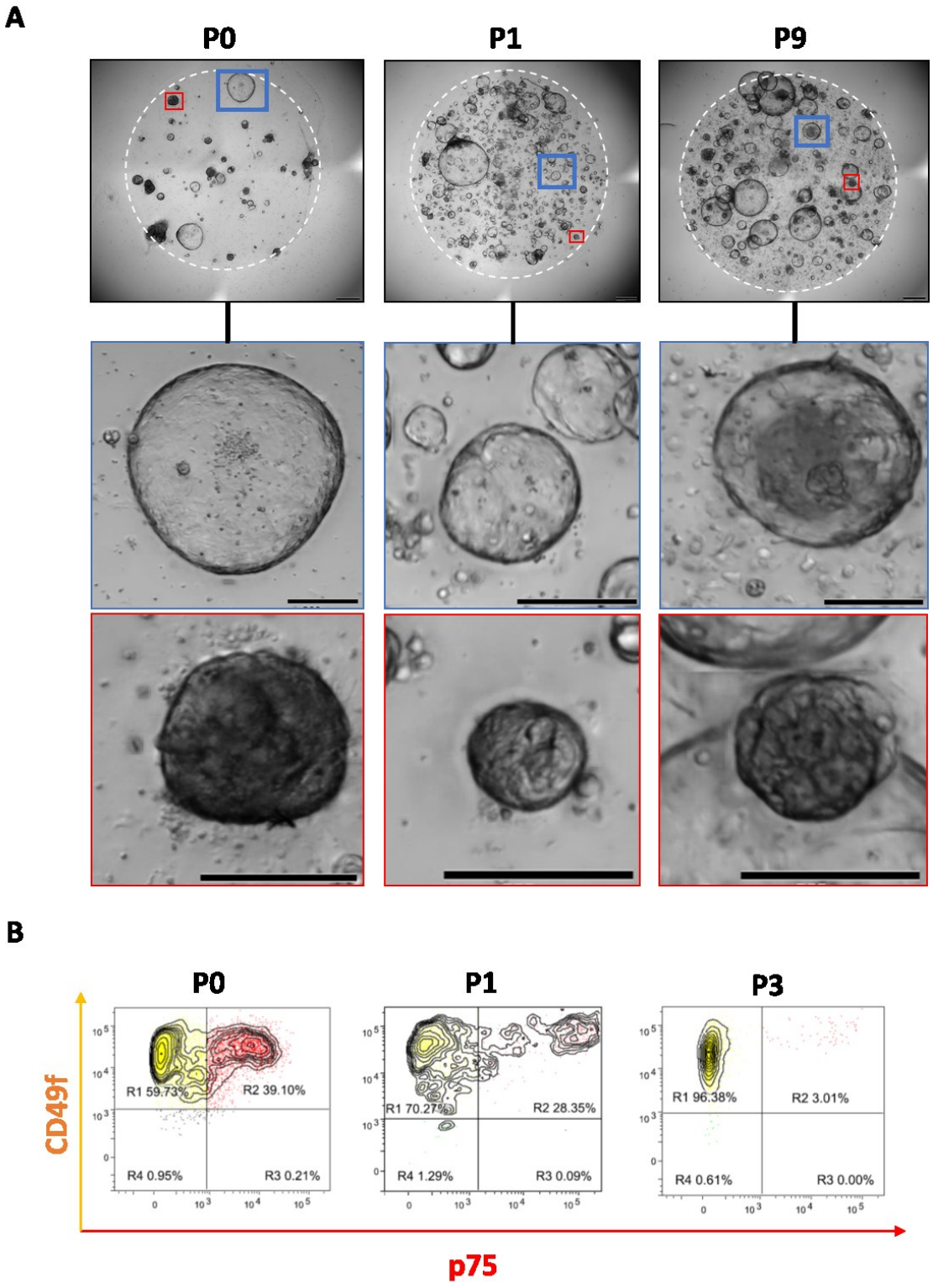


Figure 4.5: Sorted EPCAM^{+}CD49f^{+}p75^{-} cells acquired from digested pig SMGs were seeded in Matrigel and cultured in SMG spheroid media for one to two weeks. Representative BF images and insets of cystic spheroids of passages 0 (n=3), 1 (n=3) and 9 (n=1) derived from EPCAM^{+}CD49f^{+}p75^{-} cells sorted from SMGs (A). Cystic and dense spheroids are indicated by blue and red squares respectively. Scalebar=500 µm (Inset=200 µm). Flow cytometry reanalysis of cells from EPCAM^{+}CD49f^{+}p75^{-} cultured spheroids at passages 0, 1 and 3 for the markers CD49f and p75 (n=1) (B). Red dots represent the p75^{+}/CD49f^{+} backgated population, yellow dots represent p75^{-}/CD49f^{+} backgated population.

4.3.3 CD49f^{+}/p75^{-} derived dense and cystic spheroids co-express squamous and columnar markers and mucin found in BO

Cystic spheroids grown from the CD49f^{+}/p75^{-} sorted population show a mostly single layer columnar morphology with basally located nuclei and apically located mucin (Figure 4.6A and B, H&E and AB panel). Some cystic spheroids contained glandular structures protruding apically into their lumen (Figure 4.6A, H&E panel, black arrows). On the other hand, dense spheroids show a more squamoid morphology in their apical cells while the basally located cells show little cytoplasm and have a more pyramid like shape (Figure 4.6A').

Interestingly, p75 staining can be seen predominantly in the outer basal cells of the dense spheroids (Figure 4.6C-C'), indicating that the CD49f^{+}/p75^{+} cells in the flow cytometry re-analysis are mostly derived from these spheroids. Cystic spheroids are mostly negative for p75.

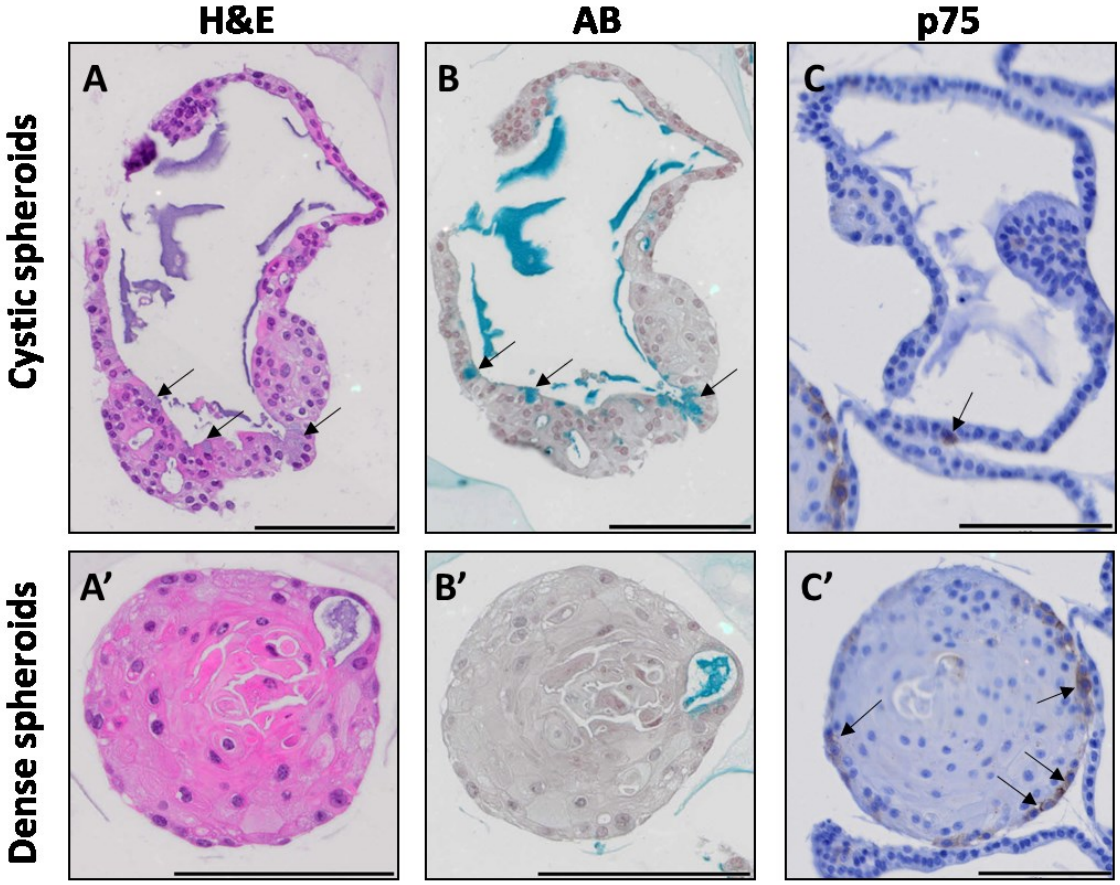


Figure 4.6: Representative images of histologic characterisation of passage 4 CD49f^{+}/p75^{-} derived cystic and dense spheroids stained for H&E (A-A', n=2), AB (B-B', n=2) and p75 (C-C', n=1). Scalebars=100 µm. Arrows indicate cells positive for marker shown in the respective panel.

Staining with basal/squamous markers KRT5, p63 and the squamous marker KRT13 shows both cystic (Figure 4.7A-C) and dense (Figure 4.7A'-C') spheroids to be positive for these markers. KRT5 staining in cystic spheroids is variable in intensity but always located at the basal layer, while in the dense spheroids' intensity is highest in the outer basally located cells and decreases towards the inner apical layers. P63 staining is found in a subset of cell nuclei in the cystic spheroids (Figure 4.7B, black arrows) and in a subset of basally located cells in the dense spheroids. Finally, the squamous differentiation marker KRT13, as with KRT5, shows ubiquitous staining in the single layered area of the cystic spheroids but weaker staining in the more multi-layered glandular areas. In the dense spheroids, KRT13 is negative in the basal layers cells while the apical cells are positive.

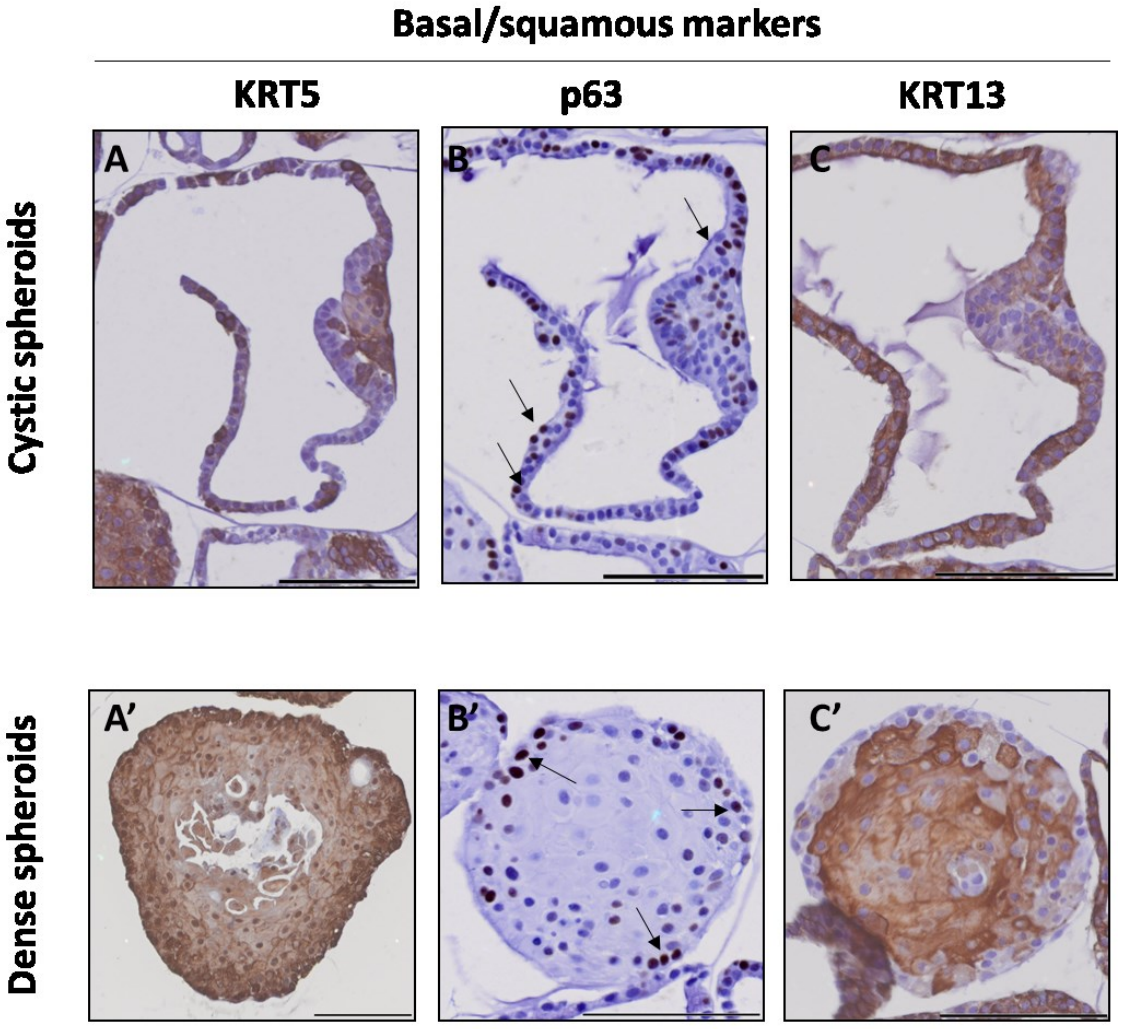


Figure 4.7: Representative images of histologic characterisation of passage 4 CD49f^{+}/p75^{-} derived cystic and dense spheroids stained for KRT5 (A-A', n=2), p63 (B-B', n=1) and KRT13 (C-C', n=2). Scalebars=100 µm. Arrows indicate cells positive for marker shown in the respective panel.

Similar to KRT13, single layered regions of cystic spheroids show ubiquitous staining for the columnar marker KRT7 while the glandular regions show very weak staining mostly in the apical region (Figure 4.8A). Dense spheroids are negative for KRT7 (Figure 4.8A'). The columnar marker SOX9 can be found in a subset of cells in cystic spheroids (Figure 4.8B, black arrows) while in the dense spheroids, SOX9 is located more in the apical cells and is not present in basal layer cells (Figure 4.8B').

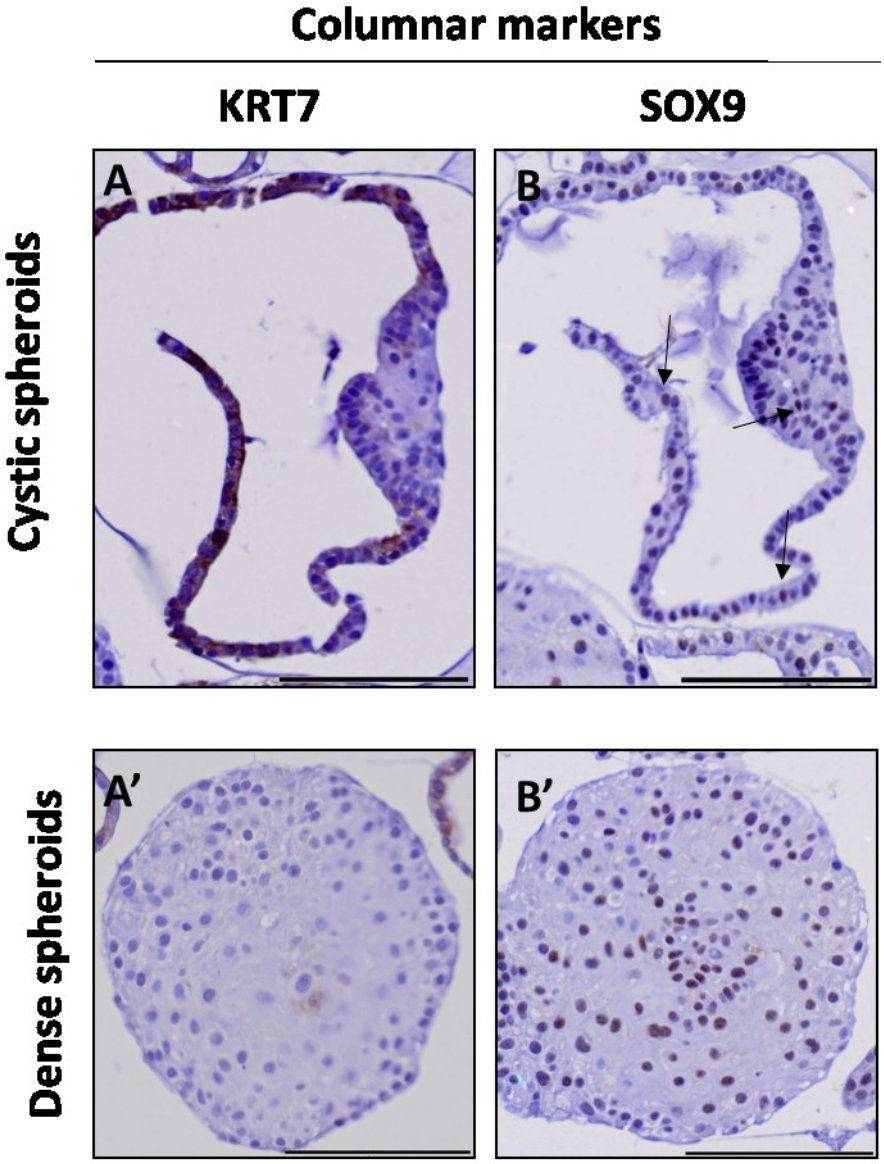


Figure 4.8: Representative images of histologic characterisation of passage 4 CD49f^{+}/p75^{-} derived cystic and dense spheroids stained for KRT7 (A-A', n=2) and SOX9 (B-B', n=2). Scalebars=100 µm. Arrows indicate cells positive for marker shown in the respective panel.

In the cystic spheroids, WGA lectin staining is present in the membrane of the basal cells and the lumen of apical cells (Figure 4.9A), while in the dense spheroids it is weakly present in the membrane of the luminal cells (Figure 4.9A'). UEA staining is more prominent in the membrane of the dense spheroids compared to WGA (Figure 4.9B', black arrows) while the cystic spheroids do not show much difference (Figure 4.9B). Similarly, the mucin marker MUC5AC can also be found in the lumen of both cystic and dense spheroids (Figure 4.9C and C'), indicating that the gland sorted ($CD49f^+/p75^-$) cells can have a squamous cell fate while concomitantly show expression of a marker expressed in BO and gastric epithelium[233].

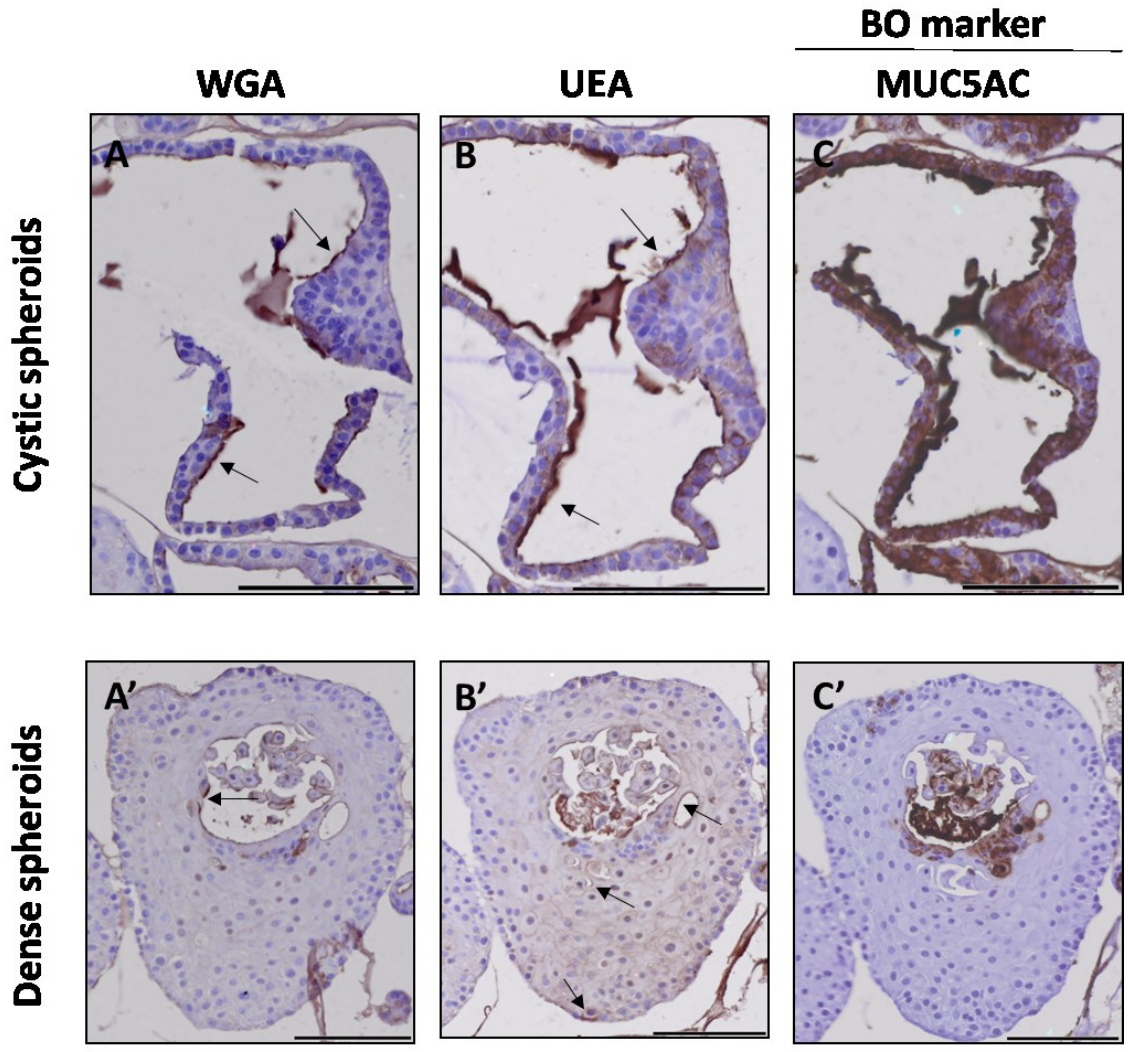


Figure 4.9: Representative images of histologic characterisation of passage 4 $CD49f^+/p75^-$ derived cystic and dense spheroids stained for WGA (A-A', n=2), UEA (B-B', n=2) and MUC5AC (C-C', n=2). Scalebars=100 µm. Arrows indicate cells positive for marker shown in the respective panel.

4.3.4 CD49f$^+$/p75$^+$ double positive sorted cells give rise to dense spheroids with squamous morphology

Basal ductal (CD49f$^+$/p75$^+$ double positive) cells sorted from pig SMGs grow into dense spheroids (Figure 4.10A), which is maintained with subsequent passages. Flow cytometric re-analyses of the cells derived from the dense spheroids show double positivity for the sorting markers used to purify their cell of origin (p75 and CD49f). However, the double positive population is decreasing over a few passages. This is potentially indicating a reduced pool of double positive cells after each passage (Figure 4.10B). Concomitantly, an increased proportion of CD49f single positive and double negative cells (R4 population) can be observed most likely indicating luminal differentiated cells.

H&E staining of spheroids derived from CD49f$^+$/p75$^+$ double positive cells shows basal layer cells to have a basophilic stain while the apical cells have a squamoid morphology and eosinophilic stain (Figure 4.11A). The basal marker p63 is localised in the basal and suprabasal layer cells and mostly absent in the apical cells. KRT5 staining is strongest in the basal layer cells and decreases towards the apical layer indicating differentiation. Conversely, the squamous differentiation marker KRT13 is absent in basal layer cells and found prominently in the apical squamoid cells.

Columnar marker KRT7 is absent in the dense spheroids (Figure 4.11B). On the other hand, the columnar marker SOX9 can be found in a few basal and supra-basal cells of the spheroids (SOX9 panel, arrows). Finally, unlike the dense spheroids derived from CD49f$^+$/p75$^-$ cells (Figure 4.9C'), these dense spheroids do not express the BO marker MUC5AC. This indicates that these cells have more restricted cell fate (only squamous) compared to the spheroids derived from CD49f$^+$/p75$^-$ cells.

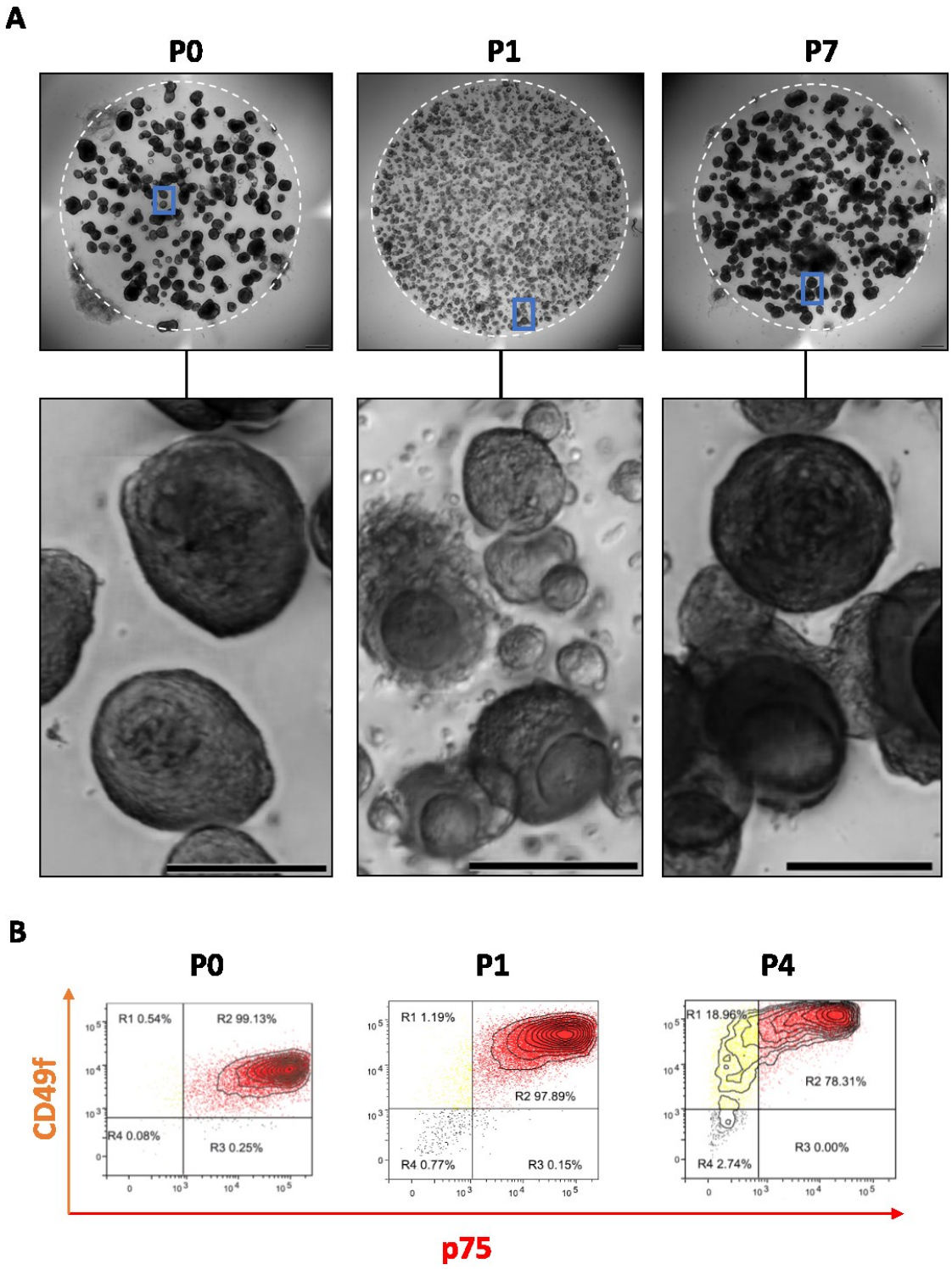


Figure 4.10: Sorted EPCAM^{+}CD49f^{+}p75^{+} cells acquired from digested pig SMGs were seeded in matrigel and cultured in SMG spheroid media for one to two weeks. Representative BF images and insets of dense spheroids of passages 0 (n=3), 1 (n=3) and 7 (n=1) derived from EPCAM^{+}CD49f^{+}p75^{+} cells sorted from SMGs (A). Scalebars=500 µm (Inset=200 µm). Flow cytometry reanalysis of cells from EPCAM^{+}CD49f^{+}p75^{+} derived spheroids cultured for one week in spheroid culture media at passages 0, 1 and 4 (n=1) (B). Red dots represent the p75^{+}/CD49f^{+} backgated population, yellow dots represent p75^{-}/CD49f^{+} backgated population.

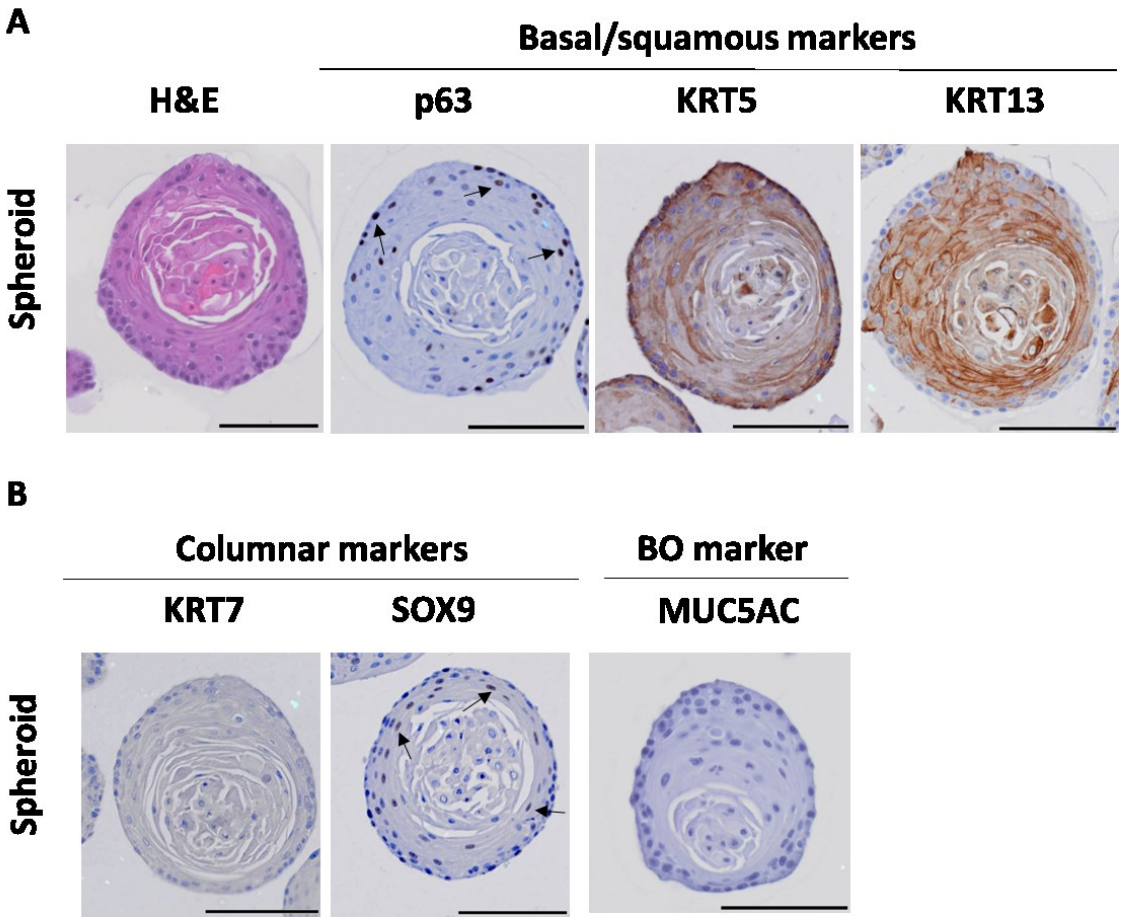


Figure 4.11: Representative images of histologic characterisation of passage 3 spheroids derived from EPCAM^{+}CD49f^{+}p75^{+} cells stained for H&E, p63, KRT5 and KRT13 (A) and KRT7, SOX9 and MUC5AC (B). Scalebars=100 µm. Arrows indicate positive cells for marker shown in panel (n=2).

4.3.5 Flow cytometry staining profile of non-dysplastic Barrett's cells resembles pig SMG gland cells compartment

In comparison to the pig SMG cells, non-dysplastic Barrett's cell lines BAR-T and CP-A and high-grade dysplasia cell line CP-B were analysed with the same markers (CD49f and p75) (Figure 4.12). Interestingly, BAR-T and CP-A show mainly positivity for CD49f only, similar to pig gland cells, whereas CP-B shows a more diffuse staining profile with majority of the cells being positive only for CD49f and the remaining cells positive for both CD49f and p75 (Figures 2.13 and 2.21), similar to the ductal cells in pig and both human ductal and myoepithelial cells.

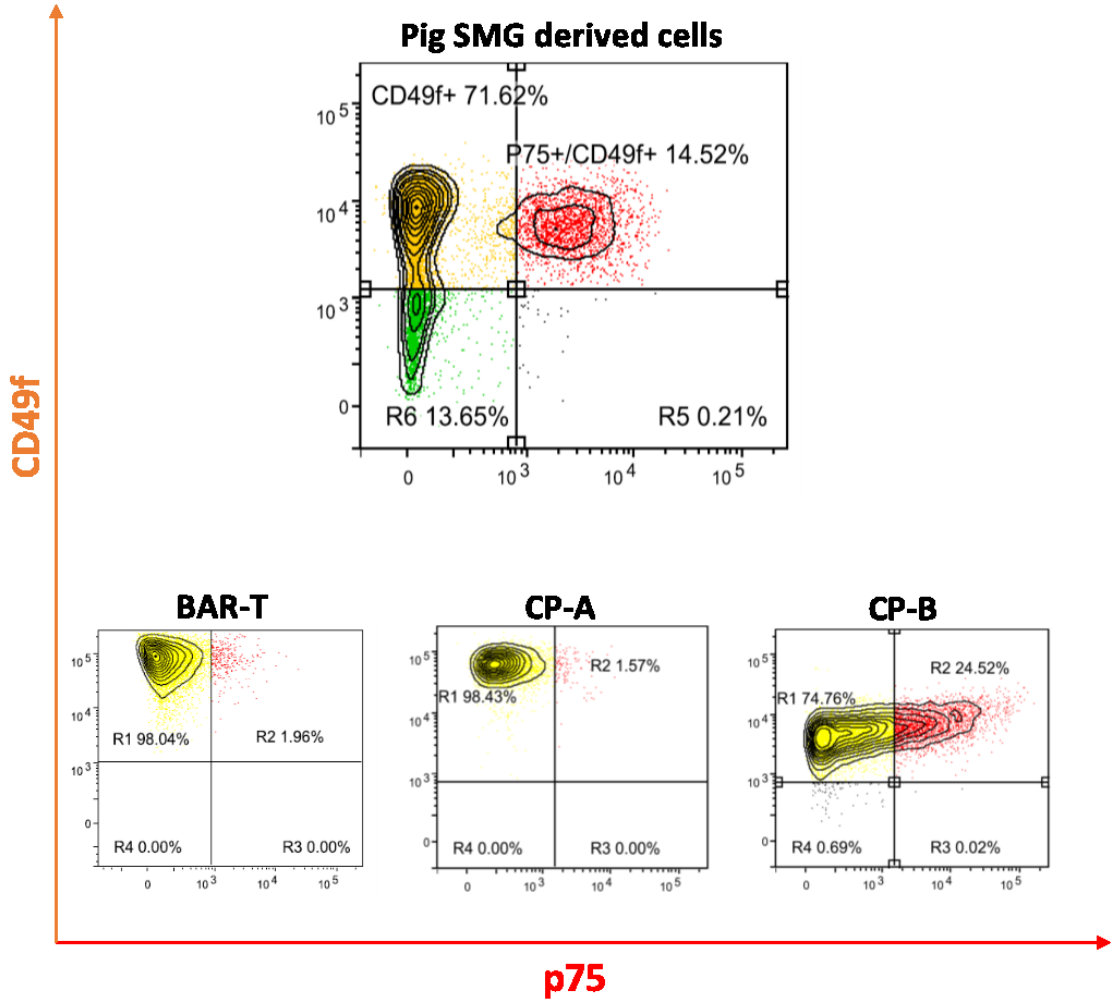


Figure 4.12: Flow cytometry analyses comparing EPCAM^{+} sorted pig SMG cells stained for p75 and CD49f to BAR-T (n=2), CP-A (n=1) and CP-B (n=1) cell lines stained for p75 and CD49f. Red dots represent the p75^{+}/CD49f^{+} backgated population, yellow dots represent p75^{-}/CD49f^{+} backgated population.

4.3.6 Viral overexpression of hCDX2 does not induce Barrett's-like differentiation in ductal cell derived squamous spheroids

Since BO is characterised by the replacement of the normal squamous epithelium by an intestinal-like epithelium[32], it could be possible to induce such a metaplastic change *in vitro* using the current 3D culture systems. Caudal-related homeobox transcription factor 2 (CDX2) has been shown to be essential for the development and patterning of intestinal tissue in mice[234]. Not surprising, CDX2 is expressed in BO with an intestinal phenotype[93,118,235]. Furthermore, culture of oesophageal keratinocytes at low pH or with bile acids have been shown to induce CDX2 expression in those cells[118,236,237]. Therefore, human CDX2 was cloned into a retroviral vector to allow overexpression of CDX2 with a Green Fluorescent Protein (GFP) reporter (Figure 4.13A).

Transduction of both ductal ($CD49f^+/p75^+$) and gland ($CD49f^+/p75^-$) cells was performed, however only ductal cells showed successful reporter and transgene integration. Figure 4.13B shows Brightfield (BF), GFP and merged channel of $CD49f^+/p75^+$ derived dense spheroids transduced with the MSCV-hCDX2-GFP vector. Immunohistochemistry staining shows a representative spheroid positive for both GFP and CDX2 indicating successful transduction and ectopic expression. However, H&E staining shows no evidence of metaplastic transformation. Basal and apical cells of the spheroid show similar morphology to empty vector control spheroids (Figure 4.14A-A'). Furthermore, staining for the intestinal marker MUC2 is also negative confirming no intestinalisation (Figure 4.14 D-D' and E).

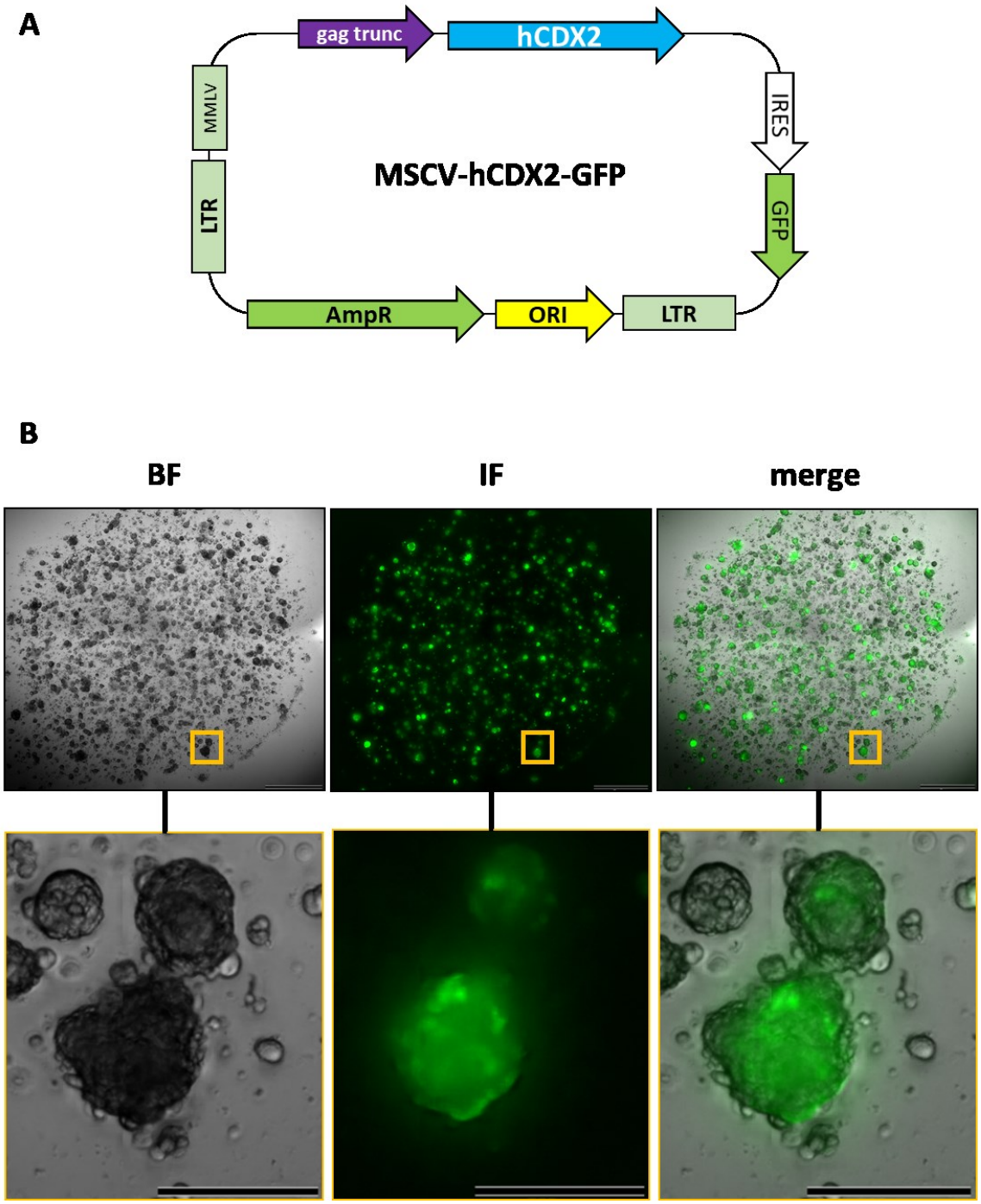


Figure 4.13: Schematic of MSCV retroviral construct for human CDX2 (A). gag trunc.: Group antigens truncated, IRES: Internal Ribosome Entry site, GFP: green fluorescent protein, ORI: Origin of Replication, AmpR: Ampicilin Resistance, LTR: long terminal repeat, MMLV: Moloney murine leukemia virus. Representative BF images and insets of spheroids derived from EPCAM+CD49f+p75+ cells transduced with MSCV-CDX2-GFP retroviral construct (n=2) (B). Scalebars=1 mm (Inset=200 µm)

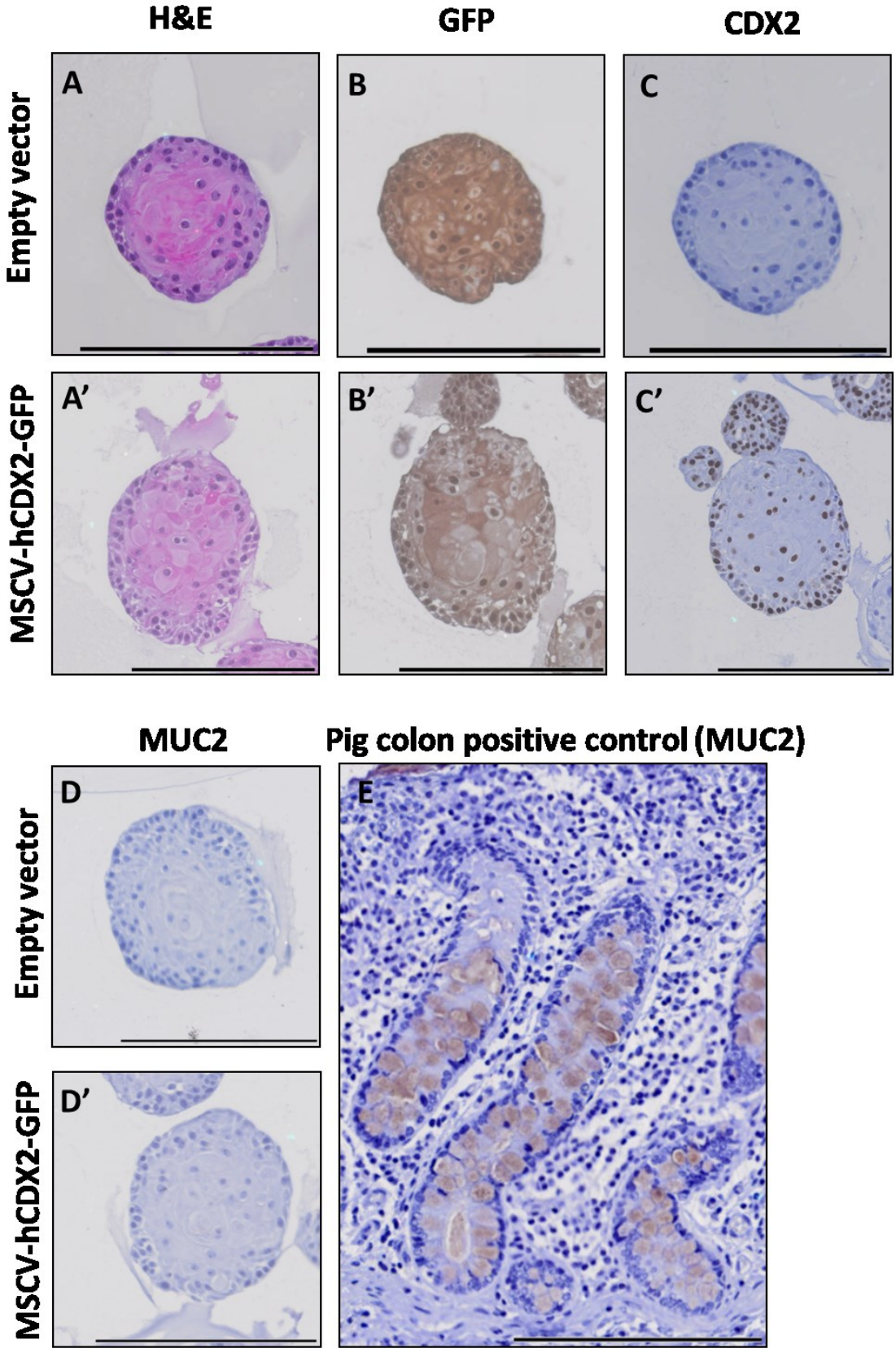


Figure 4.14: Representative images of histologic characterisation of spheroids derived from EPCAM^{+}CD49f+p75+ cells transduced with MSCV-CDX2-GFP retroviral construct stained for H&E(A-A'), GFP(B-B'), CDX2 (C-C') and MUC2 (D-D') (n=2). Pig colon tissue was used as a positive control for MUC2 (E). Scalebars=200 µm

4.3.7 Retinoic acid prevents squamous differentiation of CD49f$^+$/p75$^+$ ductal cells and induces expression of pig luminal duct marker KRT7

In section 4.3.4, it was shown that CD49f$^+$/p75$^+$ ductal cells grow into spheroids with a squamous morphology instead of the expected ductal morphology. This suggests that the culture system promotes the ductal cells to differentiate towards the squamous cell fate.

Retinoic acid (RA) was shown to play an important role in inducing glandular differentiation in the oesophagus[124], and inhibit squamous-cell differentiation in human head and neck squamous cell carcinoma[238]. This makes it an interesting candidate to try and suppress the squamous differentiation of CD49f$^+$/p75$^+$ sorted ductal cells.

Figure 4.15A shows BF images of spheroids grown from CD49f$^+$/p75$^+$ sorted ductal cells cultured for two weeks with 0, 0.1, 1 or 10 µM RA. At 0.1 µM RA, the spheroids have a dense morphology similar to untreated (DMSO) controls. However, this changes at 1 and 10 µM RA where a subset of the spheroids has a cystic morphology. H&E staining confirms the dense phenotype of the control and 0.1 µM RA condition cultured spheroids (Figure 4.15B). However, the apical cells in the dense spheroids at 0.1 µM do not show the same squamoid morphology as the control condition, indicating that RA at this concentration is already inhibiting squamous cell fate commitment. At 1 and 10 µM RA both cystic and multilayer spheroids can be distinguished (Figure 4.15B). Interestingly, the spheroids cultured in 1 and 10 µM can also be shown to produce low amounts of mucin as determine by AB staining (Figure 4.16C and D, black arrows).

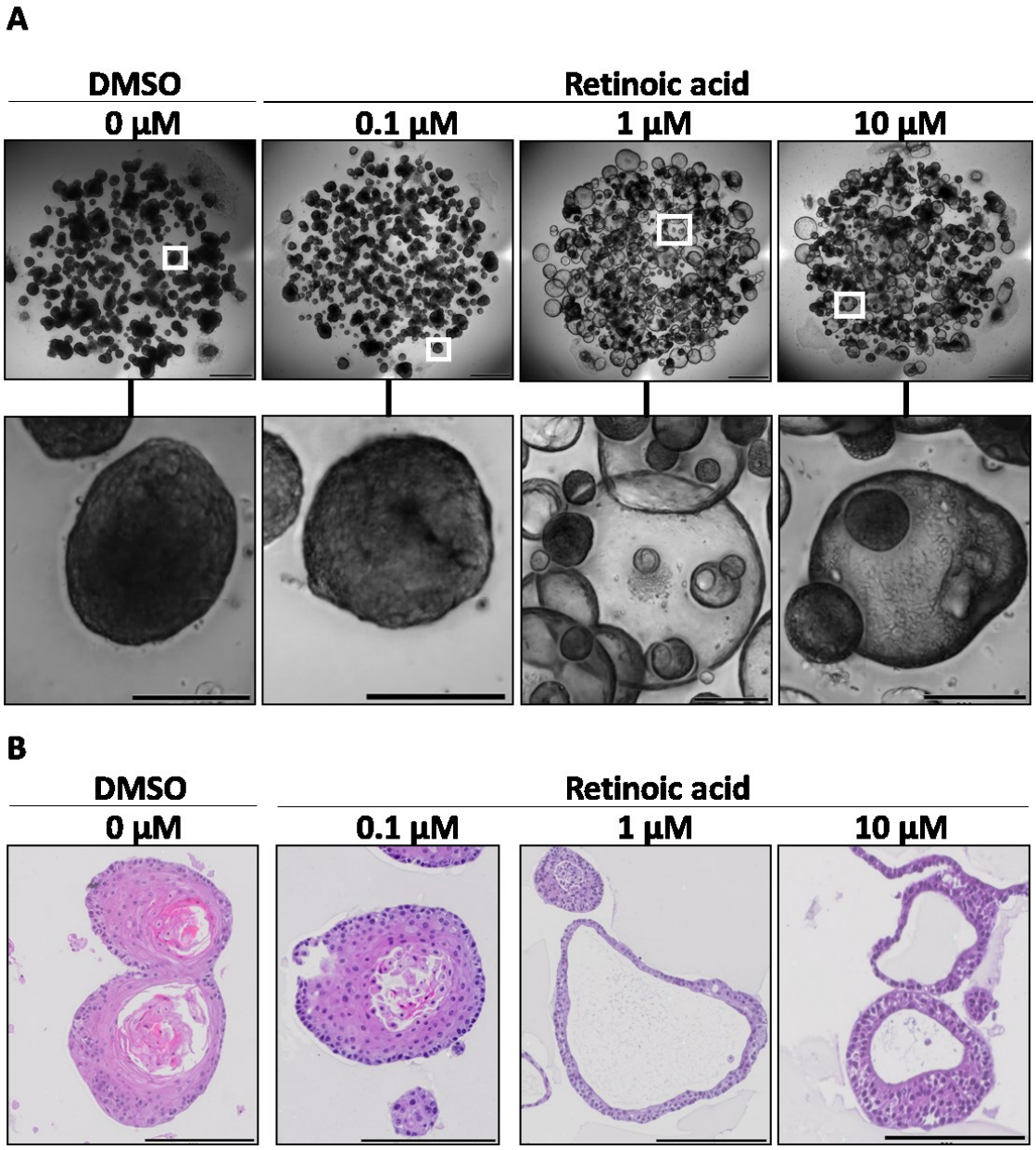


Figure 4.15: Representative BF images and insets of spheroids derived from EPCAM^{+}CD49f^{+}p75^{+} cells cultured in 0, 0.1, 1 or 10 µM Retinoic acid (A, n=2). Scalebar=1 mm (inset=200 µm). Representative images of H&E staining of dense spheroids cultured in 0 µM (n=3), 0.1 µM (n=2), 1 µM (n=2) and 10 µM RA (n=1) (B). Scalebars=200 µm.

KRT13 staining is present in the supra-basal and apical cells of the control and 0.1 µM RA treated spheroids and clearly absent in the basal layer. KRT13 can still be observed in the luminal cells of the cystic spheroids treated with 1 µM RA and seems to also extend to the basal layer in spheroids treated with 10 µM RA.

As shown in Figure 4.11, dense spheroids are negative for KRT7. Similarly, control or 0.1 µM RA culture dense spheroids are KRT7 negative. Following treatment with 1 µM RA, KRT7 can be found in the luminal cells of the cystic spheroids (Figure 4.16G, black arrows). This is a phenotype similar to the pig SMG luminal duct shown in Chapter 2 (Figure 2.10E). Following treatment with 10 µM RA, KRT7 is prominent in the supra-basal and luminal cells but can also be found in some of the basal cells (Figure 4.16H).

Finally, MUC5AC staining can be seen in the AB positive spheroids after treatment with 1 and 10 µM RA (Figure 4.16G'and H', black arrows), similar to the cystic spheroids derived from $CD49f^{+}/p75^{-}$ sorted cells (Figure 4.9C). This suggests that RA prevents ductal cells from undergoing squamous differentiation and might even be able to push them towards a glandular cell fate similar to the $CD49f^{+}/p75^{-}$ derived spheroids. This is also consistent with flow cytometric analysis of cells from RA cultured spheroids stained for the markers CD49f and p75 (Figure 4.17). The fraction of $CD49f^{+}/p75^{+}$ cells decreases at 1 and 10 µM while the proportion of $CD49f^{+}/p75^{-}$ cells increases. Interestingly, prolonging the culture period of dense spheroids to two weeks instead of one week (Figure 4.10B, P1 panel), increased the proportion of p75 negative (population R1) and double negative (population R4) in these cultures. This is most likely reflecting the proportion of apical cells in the spheroids, which are probably differentiating, due to no longer being exposed to the growth factors in the media. The double negative population proportion is much lower in the 1 and 10 µM RA conditions. This is most likely reflecting the fact that these spheroids have a less multi-layered morphology, where the apical cells could potentially still be exposed to growth factors in the media, and thereby not fully differentiate.

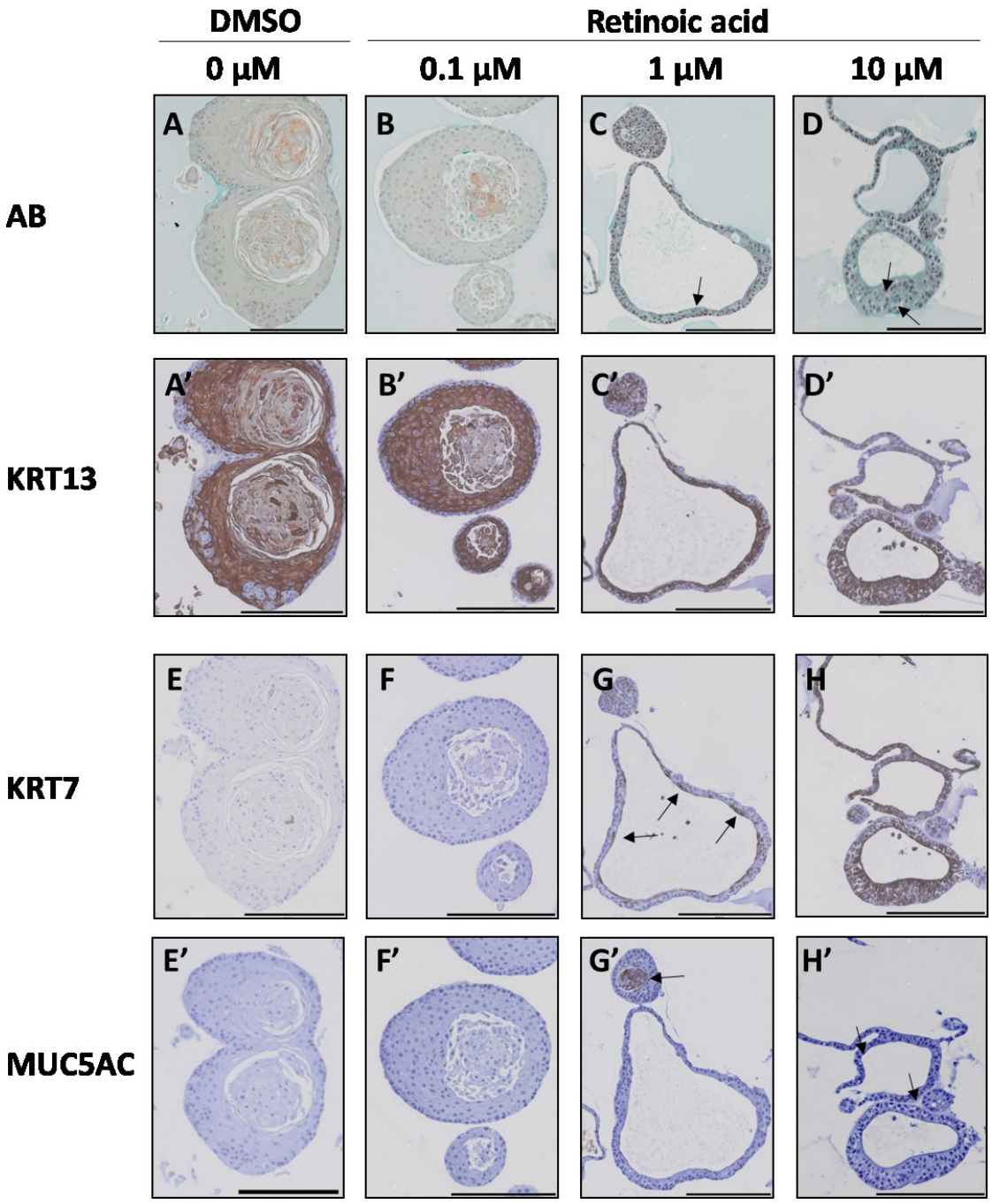


Figure 4.16: Representative images of histologic characterisation of spheroids derived from EPCAM^{+}CD49f^{+}p75^{+} cells cultured in 0, 0.1, 1 or 10 µM Retinoic acid stained for AB (A-D), KRT13 (A'-D'), KRT7 (E-H) and MUC5AC (E'-H') (n=1). Scalebars=200 µm. Arrows indicate cells positive for marker shown in the respective panel.

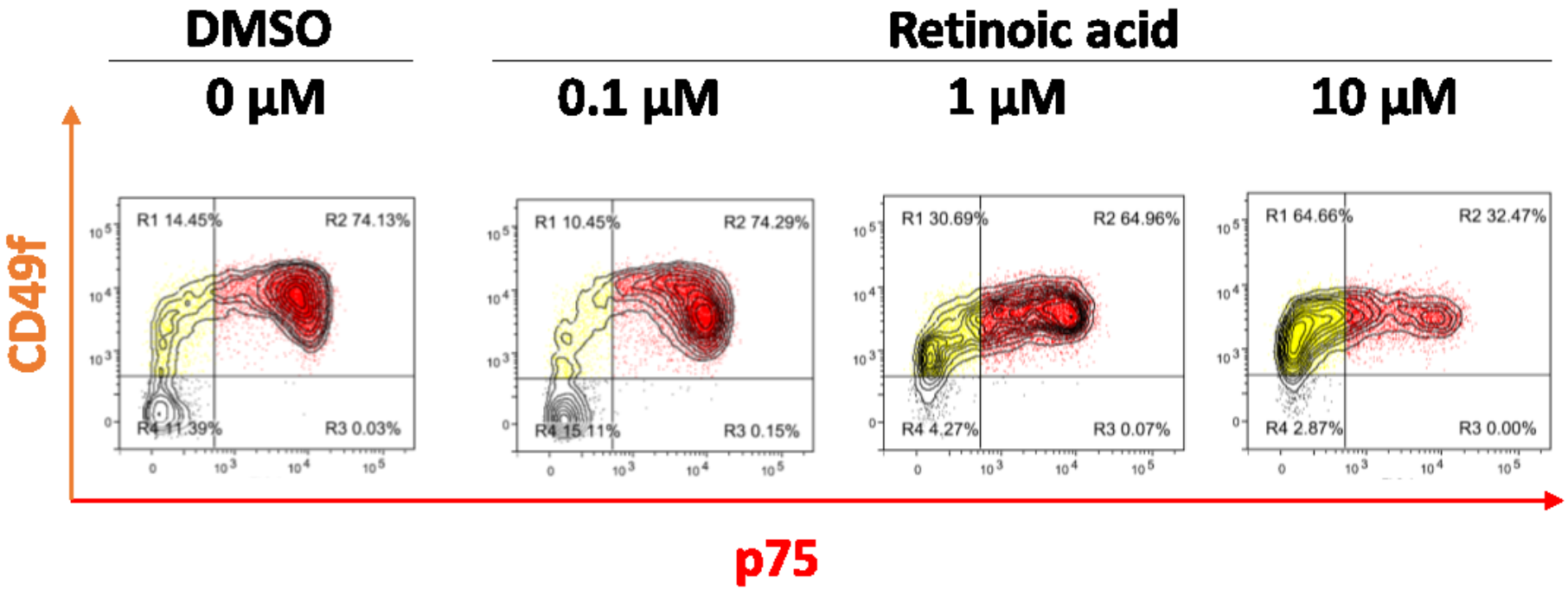

DMSO
Retinoic acid
0 µM
0.1 µM
1 µM
10 µM
CD49f
p75
R1 14.45%
R2 74.13%
R4 11.39%
R3 0.03%
R1 10.45%
R2 74.29%
R4 15.11%
R3 0.15%
R1 30.69%
R2 64.96%
R4 4.27%
R3 0.07%
R1 64.66%
R2 32.47%
R4 2.87%
R3 0.00%

Figure 4.17: Flow cytometry reanalysis of cells from EPCAM+CD49f+p75+ sorted cells derived spheroids cultured for two weeks in SMG spheroid culture media supplemented with 0 µM (DMSO), 0.1 µM, 1 µM and 10 µM retinoic acid (n=1). Red dots represent the $p75^{+}/CD49f^{+}$ backgated population, yellow dots represent $p75^{-}/CD49f^{+}$ backgated population.

4.3.8 Growth of organoids from human SMGs is adversely affected by pre-operation treatment

Multiple attempts at growing organoids from human SMG cells were performed but with limited success. Some success was possible with surgical resection specimens acquired from patients whom did not undergo chemoradiotherapy prior to surgery (18PM0248 and 19PM0024), as summarised in Table 4.6. Successful organoid growth up to passage 2 was possible when the tissue was minced into small fragments but not enzymatically digested, as opposed to performing digestion (17PM0273), before embedding in Matrigel and culturing as with pig SMGs (Figure 4.18). Unfortunately, insufficient material was available for histology.

Table 4.6: Summary of information regarding human specimens used for organoid culture attempts and outcome

Patient-ID	gender	digested	Radio-therapy	Chemo-therapy	growth at P0	growth at P1	growth at P2
17PM0209	**M**	yes	yes	yes	no	no	no
17PM0239	**M**	yes	yes	yes	no	no	no
17PM0273	**M**	yes	no	no	no	no	no
17PM0297	**M**	yes	yes	yes	yes	no	no
18PM0069	**F**	yes	no	yes	yes	yes	no
18PM0248	**M**	no	no	no	yes	yes	no
19PM0024	**M**	no	no	no	yes	no	no
CA0093	**M**	no	no	yes	no	no	no
CA0096	**F**	no	yes	yes	no	no	no
CA0097	**F**	no	no	yes	no	no	no
CA0099	**M**	no	no	Yes	no	no	no

Unsuccessful
Successful

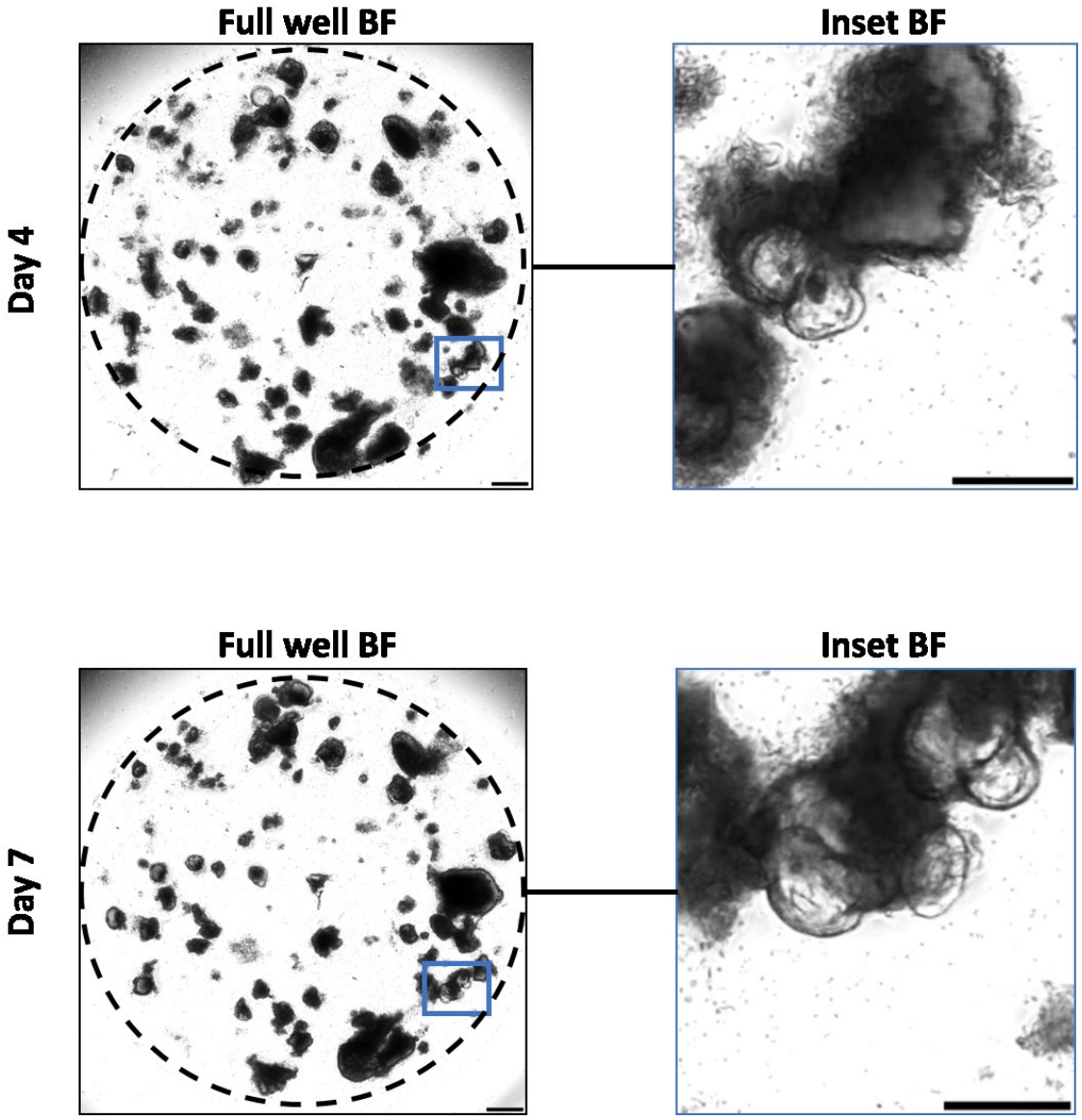


Figure 4.18: Human SMGs were dissected from surgical resection minced into small fragments, seeded in Matrigel and cultured in SMG spheroid media. Brightfield images and insets of human SMGs tissue fragments (patient-ID: 19PM0024) cultured for 4 and 7 days. Scalebars full well=500 µm, Scalebars insets=200 µm.

4.4 Discussion

The ability to culture single sorted cells over long periods of time over multiple passages, coupled to their expression of progenitor markers, supports the hypothesis of the SMGs containing cells with progenitor capacity expected of tissue resident adult stem cells.

The distinct morphologies acquired from the culture of pig SMGs in this work fits with the work performed by Furstenberg *et al.*[239] also using unsorted pig SMGs. In their work, Furstenberg *et al*. were also able to grow dense and cystic spheroids. One notable difference is that their cystic spheroids did not show p63 expression whereas in the current work they do. This reason for this disparity is unclear. One possible reason could be the culture media they used, which has a different composition compared to the media used in this work.

The presence of two morphologically distinct spheroids in the bulk cultures indicates the presence of more than one progenitor within the pig SMGs. This supports the hypothesis of the gland and duct compartments possessing their own pool of progenitor cells. To address this hypothesis, sorting and purifying the individual cells and culturing them in the same culture system could provide a definitive answer.

In the current work, gland cells could be partly purified using CD49f positivity and negativity for p75 as shown in Chapter 2. Two morphologically distinct types of spheroids (cystic and dense) could be grown from the $CD49f^{+}/p75^{-}$ glandular compartment sorted cells of the SMGs. Culture of this population of cells in the defined intestinal organoid media was slow (up to 30 days) and required further optimisation. As cystic spheroids grew rapidly (~1.5 week) in mTEC media, defined components of that media were added to the intestinal organoid media to improve on the growth of cystic spheroids.

CD49f labels the majority of gland cells (Figures 2.22 and 2.23), it is not clear which cell(s) within the glandular compartment give(s) rise to the different spheroids. Both histologic and sc-RNA seq analyses showed the myoepithelial cells to express basal/squamous markers (p63 and KRT5). These findings support the notion that the myoepithelial cells could be the resident progenitor cell of the submucosal gland compartment. Histologic analyses of both cystic and dense spheroids from the $CD49f^{+}/p75^{-}$ population both express the same basal/squamous markers found in the myoepithelial cells. This is weakly supported by the trajectory inference and diffusion pseudotime analyses of gland compartment cells shown in previous chapter (Figures 3.15). Using the myoepithelial cells as starting progenitor population, the data weakly suggests a direct link between myoepithelial and acinar cells of the gland compartment. However, as mentioned in previous chapter, the current data lacks enough datapoints to show a clear connectivity between the myoepithelial cluster and the acinar or mucinous clusters.

Cystic spheroids derived from the $CD49f^{+}/p75^{-}$ population show expression of the BO marker (MUC5AC). This phenotype is possibly induced by the culture system itself as MUC5AC expression was not found in the pig SMGs (Figure 2.11) and could therefore be

a coincidental artefact. Nonetheless, this demonstrates that the cells at least do have the potential to produce mucin found in BO, potentially indicating the submucosal gland cells as being a source of Barrett's. Interestingly, the presence of MUC5AC in the lumen of squamous looking dense spheroids from the same population ($CD49f^+/p75^-$) is reminiscent of the multilayer epithelium phenotype observed in BO[137] further supporting the $CD49f^+/p75^-$ population of cells as a potential origin of Barrett's. Furthermore, the finding that non-dysplastic Barrett's cell lines BAR-T and CP-A display a $CD49f^+/p75^-$ phenotype similar to the submucosal gland compartment cells lends further support to this hypothesis.

Squamous islands within Barrett's lesions were shown histologically to be connected to SMG ducts[142]. This fits with the phenotype of the squamous spheroids cultured from sorted $CD49^+/p75^+$ ductal SMG cells. Interestingly, the dense spheroids derived from the ductal ($CD49^+/p75^+$) cells resemble histologically the dense spheroids from the gland compartment ($CD49^+/p75^-$) cells. This may be due to sorting contamination from doublet cells. This however seems unlikely due to the strict FACS single cell gating strategy. Furthermore, the lack of MUC5AC staining in the lumen of the $CD49f^+/p75^+$ cultured dense spheroids (Figure 4.11), supports the idea that the dense spheroids derived from $CD49f^+/p75^-$ gland cells (Figure 4.9) originate from a different cell and are not be a consequence of sorting contamination. Both sorted cell types were treated and cultured in identical conditions indicating a difference in potential of spheroid initiating cells.

An interesting premise, which seems to be supported by the differentiation trajectory analyses in chapter 3, is that the myoepithelial cells in the gland compartment, could be a progenitor for the basal duct cells in the ductal compartment. As shown in Figure 3.9A and B, the myoepithelial cells and basal duct cells show a clear connectivity, and pseudotime diffusion analysis suggest that the myoepithelial cells could give rise to the basal duct cells. These data could suggest that the dense spheroids, from the gland compartment cells ($CD49^+/p75^-$), are derived from myoepithelial cells, and are less restricted in their differentiation capacity compared to the ductal ($CD49^+/p75^+$) derived spheroids. Interestingly, treatment with RA of ductal ($CD49^+/p75^+$) derived spheroids also induces expression of MUC5AC, thereby potentially reversing this restriction.

The rationale for using RA to prevent squamous differentiation, aside from published literature[124,238], is further supported by gene expression analysis performed in Chapter 3 showing luminal duct cells expressing high levels (6.5 fold) of Retinoic acid receptor responder protein 1 (RARRES1), a RA receptor-responsive gene. This indicates that RA plays a role in the homeostasis of luminal duct cells. Coupling these findings, with the morphologic changes upon RA treatment, and the expression of both squamous (KRT13) and columnar (KRT7) markers similar to the luminal duct cells, supports the notion that RA is an important factor in maintaining the ductal phenotype of ductal cells *in vitro*.

Overexpression of hCDX2 did not induce any morphologic changes reminiscent of BO. This could be due to the commitment of sorted $CD49f^+/p75^+$ cells towards a squamous lineage. A similar experiment (with a similar outcome) has previously been performed

in vivo through the overexpression of CDX2 in murine oesophageal squamous epithelium using a Cytokeratin 14 driven promoter[159]. The overexpression of CDX2, although leading to reduction in expression of Desmocollin-3, which affected the epithelial barrier at the ultrastructural level[159], did not change the histologic phenotype of the murine oesophageal epithelium. These findings could indicate that once the progenitor cells are committed to the squamous cell fate, it might not be possible to force them to undergo intestinal metaplasia. Jiang *et al.*[131] showed that a population of basal oesophageal cells expressing both basal/squamous (p63 and CK5) and the columnar marker CK7, which they propose to be the origin of BO, were able to grow into spheroids where CDX2 overexpression induced intestinal metaplasia. However, their intestinal metaplasia data was not shown through the expression of intestinal mucin (MUC2) but through the expression of Anterior Gradient 2 (AGR2), which is also expressed in non-intestinal tissues such as the trachea[215] and in the mucinous cells of pig SMGs (Appendix Table 1). Treatment of ductal ($CD49^+/p75^+$) derived spheroids with RA to prevent squamous commitment may allow these cells to undergo BO metaplastic change upon CDX2 overexpression.

$CD49f^+/p75^-$ derived cystic spheroids show similar marker positivity ($p63^+/CK5^+/CK7^+$) to the cells Jiang *et al*. transduced with CDX2 to induce 'intestinal metaplasia'; however, transduction of the cystic spheroids cells was not successful using the current protocol.

Human SMG culture using the pig SMG culture protocol was not successful, most likely due to multiple factors. A factor contributing to the success of the culture seems to be the kind of pre-operation treatment received by the patients. Minor success could be obtained from patients whom did not receive any chemotherapy or radiotherapy to the chest area prior to surgery. Furthermore, SMGs acquired from CASCADE patients donating their bodies post-mortem did not result in successful cultures. The tissues acquired from these patients were processed for culture rapidly (within 12hrs) from dissection. In contrast pig SMGs could be frozen for months at -80˚C and still yield spheroids, indicating that the lack of success from human tissue is likely due to difficulties in maintaining the viability of cells from the CASCADE tissues.

In addition to the source of tissues (surgical resection vs CASCADE), the pre-processing of the tissues prior to culture potentially also affects the success rate as tissue digestion can also have negative effects. Finally, it is also conceivable that human cells require additional factors or different culture conditions than pig cells for their long-term maintenance *in vitro*.

Chapter 5

General discussion and future directions

The stable, yet phenotypically diverse, phenotype of BO supports the notion of a multipotent progenitor cell as the origin of the metaplastic change[32]. Consequently, the different hypothesized cells of origin of Barrett's, led to multiple models being developed and tested (reviewed in Chapter 1). Among the different origins, the oesophageal SMGs have not been extensively studied in the context of BO development due to their absence in standard animal models (rodents).
To remedy this issue, finding alternative animal models would be the best option. Among the few animals known to possess SMGs, only dogs are also known to develop BO. However, dogs are difficult to acquire for experimentation, difficult to handle in a lab, and also take very long to develop BO[127,163]. Pigs represent an attractive alternative to dogs as they also possess SMGs, and their tissue is more readily available in abundant amounts from abattoirs. However, so far pigs have not been reported to develop BO. This raises a number of questions regarding the similarity of their oesophageal structure and their suitability as a replacement model.

In Chapter 2, the distribution, abundance and morphology of pig and human SMGs were compared. Furthermore, the expression of diverse markers was assessed. Characterisation of SMG distribution in human in this work suggests a decreasing gradient from distal to proximal region. Due to the high inter-patient variability found in SMG numbers, and the relatively low number of oesophagi dissected (n=6), a clear conclusion could not be drawn about the prevalence and distribution of SMGs in the distal part of the oesophagus. Still, the presence of SMGs at the distal end of the oesophagus fits with the expected pattern of BO prevalence, which also starts at the distal end of the oesophagus and extends proximally. Conversely, pig SMGs are mainly present at the proximal end of the oesophagus and are largely absent in the distal part. If the SMGs are the source of BO, this could explain the lack of reports on the prevalence of BO in the pig oesophagus as reflux may rarely reach the proximal end in a chronic fashion to induce metaplastic changes. The prevalence of SMGs at the proximal end of the human oesophagus could not be assessed in the current work.
It is important to recognize the potential implications of the large variation in the number of SMGs at the distal end of human oesophagi. The SMGs are known to produce and secrete factors that play a role in oesophageal protection and homeostasis (Chapter 1). Consequently, having a large number of SMGs could mean that GORD patients are better protected against reflux and reflux components, and therefore are less likely to develop BO. However, if BO does truly emerge from the SMGs then having more SMGs might increase the chance of GORD patients to develop metaplastic changes in SMG progenitor cells.

The histologic characterisation of human and pig SMGs showed largely similar morphologies (Chapter2). The glands and ducts seem to contain the same types of cells except for the oncocytes, which were only seen in human SMGs. As mentioned previously (Chapter 2), oncocyte numbers increase with age. This could explain why none was found in pig SMGs, which came from relatively young (6 months) pigs. The role of oncocytes in the SMGs is not clear. As oncocytes do not produce mucin, but do

express proteins related to secretion of bicarbonate[183], this suggests they may play a role in acid suppression. Abdulnour-Nakhoul *et al*.[240] measured bicarbonate secretion by pig SMGs and found it to be similar to human SMGs[241] thereby indicating that they have functional similarities.

Another disparity between human and pig SMGs is the distribution of the lectin staining on the mucinous cells as shown in Chapter 2. The low degree of co-localisation of lectins staining, in pig compared to human mucinous cells, suggests the pig mucinous cells are more specialised in the type of mucin they produce. Aside from these differences, human and pig SMGs show a remarkable similarity in marker expression. An interesting point to consider is the difference in the expression of the progenitor marker p75, which is present in human SMG myoepithelial cells but absent in pig. The staining pattern of human myoepithelial cells is identical to basal layer duct cells for the tested progenitor markers. This could mean that one cell is a progenitor for the other. This seems to be confirmed by the pig sc-RNA seq data, where diffusion analysis shows connectivity between pig myoepithelial and basal duct cells (Chapter 3, Figure 3.9), even though pig myoepithelial cells do not show expression of p75 (Chapter 2, Figure 2.21). Also, this is the case regardless of which cell cluster (myoepithelial or basal duct) was used as starting cluster.

The presence of multiple progenitors could be confirmed using the developed 3D culture systems (Chapter 4, Figures 4.2 and 4.3). The initial culture of bulk cells from digested pig SMGs in mTEC media shows spheroids of two distinct morphologies. The purification and subsequent culture of basal duct cells confirmed that these cells have the potential of growing into spheroids with a macroscopically dense morphology. Interestingly, in the current culture system, these cells seem predisposed to undergo differentiation towards a squamous cell fate (Chapter 4, Figure 4.11). This fits with the observed phenotype of squamous islands being connected to submucosal gland ducts in BO patients[142]. Furthermore, the work performed by Jeong *et al*.[178] and DeWard *et al*.[16], using oesophageal squamous epithelium basal layer cells from human and mouse, gave rise to spheroids with a similar morphology. These findings, in addition to the diffusion gene expression (Chapter 3, Figure 3.11), provide additional evidence that the basal duct cells are bipotent and can give rise to both columnar luminal duct cells as well as squamous epithelium cells. Interestingly, this could mean that basal duct cells might function as a pool of reserve progenitor cells for the squamous epithelium. A similar progenitor distribution was demonstrated in the skin, where a subpopulation of Lgr6 positive progenitor cells in the isthmus of the hair follicle was shown to be able to regenerate all cell lineages of the skin[242].

The gland compartment cells ($CD49f^{+}/p75^{-}$) could at this stage not be purified due to the lack of appropriate cell surface markers. Culture of sorted gland compartment cells, led to the growth of two morphologically distinct spheroids. The presence of two morphologically distinct spheroids suggests the presence of more than one multipotent progenitor cell within this population, which does not correspond with the histological nor transcriptomic analyses. The dense spheroids derived from this population of cells

resemble the dense spheroids from the basal duct cells purified population. Interestingly, these spheroids also produce BO associated mucin (MUC5AC) in their apical layer, a phenotype strongly resembling the multi-layered epithelium observed in BO. This difference suggests the presence of another bipotent cell within the gland compartment. The cystic spheroids also express squamous markers suggesting a bipotent progenitor cell of origin. Further purification and characterisation of this population is necessary to determine which cells give rise to which subtype of spheroids.

SMGs do not express the BO marker MUC5AC *in situ* (Chapter 2, Figure 2.11), but this seems to be the case after *in vitro* culture (Chapter 4, Figure 4.9). A potential explanation for this could the fact that the 3D culture system is unable to model a homeostatic environment. This means that the SMG cells once extracted from the natural environment and put in culture, respond to their new environment as they would during tissue damage/repair. This could also be concluded of BO where the SMGs could be damaged by reflux components. The other potential reason for this phenotype change *in vitro* could be an artefact from a (number of) growth factor(s) in the culture media inducing the expression of MUC5AC. A similar phenotype change was observed in a recent publication with cultured liver organoids as the liver does not normally express MUC5AC *in situ*[243]. The likely explanation for the expression of this marker is the activation of the Wnt signalling pathway through the addition of Wnt and Rspondin-1 conditioned media to the SMG culture media (Table 4.4). Wnt pathway activation has previously been shown to induce beta-catenin accumulation in the nucleus[244], which in turn induces expression of MUC5AC[245]. This may seem contradictory to the results from the cystic spheroids cultured in mTEC media, which contains no added Wnt and Rspondin-1, unlike the SMG culture media. However, the mTEC media contains non-defined components (BPE and FBS) which may induce a similar effect and therefore no clear conclusion can be drawn from the differences in culture conditions.

BO is characterised by expression of intestinal mucin (MUC2), which is controlled by the expression of the intestinal transcription factor CDX2[246]. Induction of intestinalisation in spheroids derived from the gland compartment ($CD49f^+/p75^-$) was unsuccessful and requires further optimisation. Currently, it is not possible to determine if any cells within the $CD49f^+/p75^-$ population have the potential to undergo intestinal metaplasia.
Transduction of the basal duct cells with CDX2 was successful but did not show any morphologic or histologic changes indicating metaplasia. A potential reason for that is that the commitment of these cells to the squamous lineage precludes them from undergoing intestinalisation. As this would not be the case *in vivo*, preventing the ductal cells from undergoing squamous differentiation *in vitro* through addition of RA, may enable intestinal metaplasia. Interestingly, sc-RNA seq analysis showed Retinoic acid receptor responder protein 1 (RARRES1) to be one of the top 100 genes expressed in the luminal duct cell population only (Appendix Table 1). This indicates that RA may be important for the homeostasis of the ducts and should be part of the culture media for these cells.

Another point to consider is that BO could arise from multiple cells but only from one cell type can it progress to OAC. As mentioned previously, BO is stable and rarely progresses to OAC through a sequence of non-dysplastic, low grade to high grade dysplasia. When staining non-dysplastic and high-grade dysplasia cell lines for the progenitor markers CD49f and p75, the high-grade dysplasia cell line CP-B shows a similar profile as the basal duct cells. Conversely, the non-dysplastic BO cells show only positivity for CD49f similar to the gland compartment cells. Having multiple potential cellular origins could explain the diversity in Barrett's phenotypes, and also explain the difference between progressors and non-progressors to OAC.

The current data are insufficient to show a strong connection between SMGs and BO. The pig SMGs clearly demonstrate the presence of at least two progenitor cells. One progenitor cell is located in the basal layer of the submucosal ducts and shows potential for oesophageal repair and is most likely the source of squamous islands in BO. The potential of these cells to undergo Barrett's metaplasia could not be tested without the cells undergoing squamous differentiation. The use of RA to prevent the squamous differentiation process could potentially resolve this issue, and subsequent overexpression of CDX2 in the basal duct cells in the presence of RA might induce a Barrett's-like phenotype in these cells. Interestingly, RA levels were shown to be elevated in the oesophagus of BO patients compared to healthy controls[124]. This might indicate that in BO patients the repair of the squamous epithelium might be hampered by abnormally high concentrations of RA, thereby preventing the basal layer duct cells and most likely the basal layer of the squamous epithelium from undergoing squamous differentiation.

The myoepithelial cells in the gland compartment are likely the second progenitor cells of the SMGs. The potential of this progenitor is not clear from the current data as no marker was found that could be used to specifically isolate these cells by flow cytometry. Enriching for the gland cell compartment by sorting for $CD49f^{+}/p75^{-}$ cells and re-sequencing, might enable for the detection of additional markers to identify and sort individual cell populations. Similar to the basal duct cells, overexpression of CDX2 in the purified cells could show if the cells from the SMGs have the potential to undergo Barrett's like metaplasia. This could also be extended to other factors such as acid and bile acids to test the effects these components have on these SMG derived cells. Treatment with these factors could potentially induce phenotype and gene expression changes similar to BO.

In conclusion, this work demonstrates the similarities and differences between human and pig SMGs. The characterisation of these tissues allowed for the identification of cells with progenitor phenotypes. Single cell RNA sequencing of the epithelial cells of the pig SMGs showed functional similarities documented for human SMGs. The progenitor markers identified, digestion and cell sorting strategies developed allowed for the purification of the progenitor cells in the submucosal ducts and enrichment of the gland compartment progenitor(s). Finally, the developed organoid culture systems allowed for long-term growth and expansion of these enriched cell population and their

characterisation *in vitro.* These culture systems can be used to test the potential of SMG cells in giving rise to BO.

References

1. Orlando RC. Review article: Oesophageal mucosal resistance. *Aliment Pharmacol Ther*. 1998;12(3):191-197.

2. Powell DW. Barrier function of epithelia. *Am J Physiol*. 1981;241(4):G275-88.

3. Hopwood D, Logan KR, Bouchier IA. The electron microscopy of normal human oesophageal epithelium. *Virchows Arch B Cell Pathol*. 1978;26(4):345-358.

4. Homberg M, Magin TM. Beyond expectations: Novel insights into epidermal keratin function and regulation. *Int Rev Cell Mol Biol*. 2014;311:265-306.

5. Zhang X, Patil D, Odze RD, et al. The microscopic anatomy of the esophagus including the individual layers, specialized tissues, and unique components and their responses to injury. *Ann N Y Acad Sci*. 2018;1434(1):304-318.

6. Long JD, Orlando RC. Esophageal submucosal glands: Structure and function. *Am J Gastroenterol*. 1999;94(10):2818-2824.

7. Mittal RK, Hong SJ, Bhargava V. Longitudinal muscle dysfunction in achalasia esophagus and its relevance. *J Neurogastroenterol Motil*. 2013;19(2):126-136.

8. Gray SW, Skandalakis JE. *Embryology for surgeons; the embryological basis for the treatment of congenital defects.* Philadelphia: Saunders; 1972:918.

9. Que J, Okubo T, Goldenring JR, et al. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development*. 2007;134(13):2521-2531.

10. Domyan ET, Ferretti E, Throckmorton K, Mishina Y, Nicolis SK, Sun X. Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. *Development*. 2011;138(5):971-981.

11. Sakai N, Suenaga T, Tanaka K. Electron microscopic study on the esophageal mucosa in human fetuses. *Auris Nasus Larynx*. 1989;16(3):177-183.

12. Yu WY, Slack JM, Tosh D. Conversion of columnar to stratified squamous epithelium in the developing mouse oesophagus. *Dev Biol*. 2005;284(1):157-170.

13. Eguchi G, Kodama R. Transdifferentiation. *Curr Opin Cell Biol*. 1993;5(6):1023-1028.

14. Rodriguez P, Da Silva S, Oxburgh L, Wang F, Hogan BL, Que J. BMP signaling in the development of the mouse esophagus and forestomach. *Development*. 2010;137(24):4171-4176.

15. Jiang M, Ku WY, Zhou Z, et al. BMP-driven NRF2 activation in esophageal basal cell differentiation and eosinophilic esophagitis. *J Clin Invest*. 2015;125(4):1557-1568.

16. DeWard AD, Cramer J, Lagasse E. Cellular heterogeneity in the mouse esophagus implicates the presence of a nonquiescent epithelial stem cell population. *Cell Rep*. 2014;9(2):701-711.

17. Daniely Y, Liao G, Dixon D, et al. Critical role of p63 in the development of a normal esophageal and tracheobronchial epithelium. *Am J Physiol Cell Physiol*. 2004;287(1):C171-81.

18. Koster MI, Kim S, Mills AA, DeMayo FJ, Roop DR. P63 is the molecular switch for initiation of an epithelial stratification program. *Genes Dev*. 2004;18(2):126-131.

19. Romano RA, Smalley K, Magraw C, et al. DeltaNp63 knockout mice reveal its indispensable role as a master regulator of epithelial development and differentiation. *Development*. 2012;139(4):772-782.

20. Seery JP, Watt FM. Asymmetric stem-cell divisions define the architecture of human oesophageal epithelium. *Curr Biol*. 2000;10(22):1447-1450.

21. Barbera M, di Pietro M, Walker E, et al. The human squamous oesophagus has widespread capacity for clonal expansion from cells at diverse stages of differentiation. *Gut*. 2015;64(1):11-19.

22. Kalabis J, Oyama K, Okawa T, et al. A subpopulation of mouse esophageal basal cells has properties of stem cells with the capacity for self-renewal and lineage specification. *J Clin Invest*. 2008;118(12):3860-3869.

23. Croagh D, Phillips WA, Redvers R, Thomas RJ, Kaur P. Identification of candidate murine esophageal stem cells using a combination of cell kinetic studies and cell surface markers. *Stem Cells*. 2007;25(2):313-318.

24. Giroux V, Lento AA, Islam M, et al. Long-lived keratin 15+ esophageal progenitor cells contribute to homeostasis and regeneration. *J Clin Invest*. 2017;127(6):2378-2391.

25. Doupe DP, Alcolea MP, Roshan A, et al. A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science*. 2012;337(6098):1091-1093.

26. Clarke MF, Dick JE, Dirks PB, et al. Cancer stem cells--perspectives on current status and future directions: AACR workshop on cancer stem cells. *Cancer Res*. 2006;66(19):9339-9344.

27. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut*. 2015;64(3):381-387.

28. Siewert JR, Ott K. Are squamous and adenocarcinomas of the esophagus the same disease? *Semin Radiat Oncol*. 2007;17(1):38-44.

29. Stewart B. World cancer report . lyon, france: International agency for research on cancer. . 2003.

30. Smyth EC, Lagergren J, Fitzgerald RC, et al. Oesophageal cancer. *Nat Rev Dis Primers*. 2017;3:17048.

31. McColl KEL. What is causing the rising incidence of esophageal adenocarcinoma in the west and will it also happen in the east? *J Gastroenterol*. 2019;54(8):669-673.

32. Phillips WA, Lord RV, Nancarrow DJ, Watson DI, Whiteman DC. Barrett's esophagus. *J Gastroenterol Hepatol*. 2011;26(4):639-648.

33. Blot WJ, Devesa SS, Kneller RW, Fraumeni JF,Jr. Rising incidence of adenocarcinoma of the esophagus and gastric cardia. *JAMA*. 1991;265(10):1287-1289.

34. Pohl H, Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J Natl Cancer Inst*. 2005;97(2):142-146.

35. Pohl H, Sirovich B, Welch HG. Esophageal adenocarcinoma incidence: Are we reaching the peak? *Cancer Epidemiol Biomarkers Prev*. 2010;19(6):1468-1470.

36. Kong CY, Kroep S, Curtius K, et al. Exploring the recent trend in esophageal adenocarcinoma incidence and mortality using comparative simulation modeling. *Cancer Epidemiol Biomarkers Prev*. 2014;23(6):997-1006.

37. Xie SH, Mattsson F, Lagergren J. Incidence trends in oesophageal cancer by histological type: An updated analysis in sweden. *Cancer Epidemiol*. 2017;47:114-117.

38. Corley DA, Kubo A, Levin TR, et al. Race, ethnicity, sex and temporal differences in barrett's oesophagus diagnosis: A large community-based study, 1994-2006. *Gut*. 2009;58(2):182-188.

39. Johansson J, Hakansson HO, Mellblom L, et al. Risk factors for barrett's oesophagus: A population-based approach. *Scand J Gastroenterol*. 2007;42(2):148-156.

40. Gerson LB, Triadafilopoulos G. Screening for esophageal adenocarcinoma: An evidence-based approach. *Am J Med*. 2002;113(6):499-505.

41. Brown LM, Swanson CA, Gridley G, et al. Adenocarcinoma of the esophagus: Role of obesity and diet. *J Natl Cancer Inst*. 1995;87(2):104-109.

42. Vaughan TL, Davis S, Kristal A, Thomas DB. Obesity, alcohol, and tobacco as risk factors for cancers of the esophagus and gastric cardia: Adenocarcinoma versus squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev*. 1995;4(2):85-92.

43. Smith KJ, O'Brien SM, Green AC, Webb PM, Whiteman DC, Study of Digestive Health. Current and past smoking significantly increase risk for barrett's esophagus. *Clin Gastroenterol Hepatol*. 2009;7(8):840-848.

44. Yousef F, Cardwell C, Cantwell MM, Galway K, Johnston BT, Murray L. The incidence of esophageal cancer and high-grade dysplasia in barrett's esophagus: A systematic review and meta-analysis. *Am J Epidemiol*. 2008;168(3):237-249.

45. Cook MB, Corley DA, Murray LJ, et al. Gastroesophageal reflux in relation to adenocarcinomas of the esophagus: A pooled analysis from the barrett's and esophageal adenocarcinoma consortium (BEACON). *PLoS One*. 2014;9(7):e103508.

46. Gotley DC, Morgan AP, Ball D, Owen RW, Cooper MJ. Composition of gastro-oesophageal refluxate. *Gut*. 1991;32(10):1093-1099.

47. Asaoka D, Miwa H, Hirai S, et al. Altered localization and expression of tight-junction proteins in a rat model with chronic acid reflux esophagitis. *J Gastroenterol*. 2005;40(8):781-790.

48. Jovov B, Que J, Tobey NA, Djukic Z, Hogan BL, Orlando RC. Role of E-cadherin in the pathogenesis of gastroesophageal reflux disease. *Am J Gastroenterol*. 2011;106(6):1039-1047.

49. Dekel R, Morse C, Fass R. The role of proton pump inhibitors in gastro-oesophageal reflux disease. *Drugs*. 2004;64(3):277-295.

50. Naya MJ, Pereboom D, Ortego J, Alda JO, Lanas A. Superoxide anions produced by inflammatory cells play an important part in the pathogenesis of acid and pepsin induced oesophagitis in rabbits. *Gut*. 1997;40(2):175-181.

51. Souza RF, Huo X, Mittal V, et al. Gastroesophageal reflux might cause esophagitis through a cytokine-mediated mechanism rather than caustic acid injury. *Gastroenterology*. 2009;137(5):1776-1784.

52. Dvorakova K, Payne CM, Ramsey L, et al. Apoptosis resistance in barrett's esophagus: Ex vivo bioassay of live stressed tissues. *Am J Gastroenterol*. 2005;100(2):424-431.

53. Poetsch AR. The genomics of oxidative DNA damage, repair, and resulting mutagenesis. *Comput Struct Biotechnol J*. 2020;18:207-219.

54. Whiteman DC, Sadeghi S, Pandeya N, et al. Combined effects of obesity, acid reflux and smoking on the risk of adenocarcinomas of the oesophagus. *Gut*. 2008;57(2):173-180.

55. Kendall BJ, Macdonald GA, Hayward NK, et al. Leptin and the risk of barrett's oesophagus. *Gut*. 2008;57(4):448-454.

56. Kamangar F, Chow WH, Abnet CC, Dawsey SM. Environmental causes of esophageal cancer. *Gastroenterol Clin North Am*. 2009;38(1):27-57, vii.

57. Clark GW, Smyrk TC, Mirvish SS, et al. Effect of gastroduodenal juice and dietary fat on the development of barrett's esophagus and esophageal neoplasia: An experimental rat model. *Ann Surg Oncol*. 1994;1(3):252-261.

58. Kubo A, Block G, Quesenberry CP,Jr, Buffler P, Corley DA. Effects of dietary fiber, fats, and meat intakes on the risk of barrett's esophagus. *Nutr Cancer*. 2009;61(5):607-616.

59. Kubo A, Corley DA. Meta-analysis of antioxidant intake and the risk of esophageal and gastric cardia adenocarcinoma. *Am J Gastroenterol*. 2007;102(10):2323-30; quiz 2331.

60. Jakszyn P, Gonzalez CA. Nitrosamine and related food intake and gastric and oesophageal cancer risk: A systematic review of the epidemiological evidence. *World J Gastroenterol*. 2006;12(27):4296-4303.

61. Gonzalez CA, Jakszyn P, Pera G, et al. Meat intake and risk of stomach and esophageal adenocarcinoma within the european prospective investigation into cancer and nutrition (EPIC). *J Natl Cancer Inst*. 2006;98(5):345-354.

62. Freedman ND, Abnet CC, Leitzmann MF, et al. A prospective study of tobacco, alcohol, and the risk of esophageal and gastric cancer subtypes. *Am J Epidemiol*. 2007;165(12):1424-1433.

63. Thrift AP, Nagle CM, Fahey PP, Smithers BM, Watson DI, Whiteman DC. Predictors of survival among patients diagnosed with adenocarcinoma of the esophagus and gastroesophageal junction. *Cancer Causes Control*. 2012;23(4):555-564.

64. Pandeya N, Williams G, Green AC, Webb PM, Whiteman DC, Australian Cancer Study. Alcohol consumption and the risks of adenocarcinoma and squamous cell carcinoma of the esophagus. *Gastroenterology*. 2009;136(4):1215-24, e1-2.

65. Mohammed I, Nightingale P, Trudgill NJ. Risk factors for gastro-oesophageal reflux disease symptoms: A community study. *Aliment Pharmacol Ther*. 2005;21(7):821-827.

66. Locke GR,3rd, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ,3rd. Risk factors associated with symptoms of gastroesophageal reflux. *Am J Med*. 1999;106(6):642-649.

67. Cook MB, Kamangar F, Whiteman DC, et al. Cigarette smoking and adenocarcinomas of the esophagus and esophagogastric junction: A pooled analysis from the international BEACON consortium. *J Natl Cancer Inst*. 2010;102(17):1344-1353.

68. Stanciu C, Bennett JR. Smoking and gastro-oesophageal reflux. *Br Med J*. 1972;3(5830):793-795.

69. Chattopadhyay DK, Greaney MG, Irvin TT. Effect of cigarette smoking on the lower oesophageal sphincter. *Gut*. 1977;18(10):833-835.

70. Trudgill NJ, Smith LF, Kershaw J, Riley SA. Impact of smoking cessation on salivary function in healthy volunteers. *Scand J Gastroenterol*. 1998;33(6):568-571.

71. Yamaoka Y, Kita M, Kodama T, Sawai N, Kashima K, Imanishi J. Induction of various cytokines and development of severe mucosal inflammation by cagA gene positive helicobacter pylori strains. *Gut*. 1997;41(4):442-451.

72. Schutze K, Hentschel E, Dragosics B, Hirschl AM. Helicobacter pylori reinfection with identical organisms: Transmission by the patients' spouses. *Gut*. 1995;36(6):831-833.

73. Labenz J, Blum AL, Bayerdorffer E, Meining A, Stolte M, Borsch G. Curing helicobacter pylori infection in patients with duodenal ulcer may provoke reflux esophagitis. *Gastroenterology*. 1997;112(5):1442-1447.

74. Vicari JJ, Peek RM, Falk GW, et al. The seroprevalence of cagA-positive helicobacter pylori strains in the spectrum of gastroesophageal reflux disease. *Gastroenterology*. 1998;115(1):50-57.

75. Eross B, Farkas N, Vincze A, et al. Helicobacter pylori infection reduces the risk of barrett's esophagus: A meta-analysis and systematic review. *Helicobacter*. 2018;23(4):e12504.

76. Ye W, Held M, Lagergren J, et al. Helicobacter pylori infection and gastric atrophy: Risk of adenocarcinoma and squamous-cell carcinoma of the esophagus and adenocarcinoma of the gastric cardia. *J Natl Cancer Inst*. 2004;96(5):388-396.

77. Aghayeva S, Mara KC, Katzka DA. The impact of helicobacter pylori on the presence of barrett's esophagus in azerbaijan, a high-prevalence area of infection. *Dis Esophagus*. 2019.

78. BARRETT NR. Chronic peptic ulcer of the oesophagus and 'oesophagitis'. *Br J Surg*. 1950;38(150):175-182.

79. HAYWARD J. The lower end of the oesophagus. *Thorax*. 1961;16:36-41.

80. Chandrasoma P, DeMeester T. GERD : Reflux to esophageal adenocarcinoma. In: Academic Press/Elsevier, c2006; 2006.

81. Yantiss RK. Diagnostic challenges in the pathologic evaluation of barrett esophagus. *Arch Pathol Lab Med*. 2010;134(11):1589-1600.

82. McDonald SA, Graham TA, Lavery DL, Wright NA, Jansen M. The barrett's gland in phenotype space. *Cell Mol Gastroenterol Hepatol*. 2014;1(1):41-54.

83. Naini BV, Souza RF, Odze RD. Barrett's esophagus: A comprehensive and contemporary review for pathologists. *Am J Surg Pathol*. 2016;40(5):e45-66.

84. Paull A, Trier JS, Dalton MD, Camp RC, Loeb P, Goyal RK. The histologic spectrum of barrett's esophagus. *N Engl J Med*. 1976;295(9):476-480.

85. Haggitt RC, Tryzelaar J, Ellis FH, Colcher H. Adenocarcinoma complicating columnar epithelium-lined (barrett's) esophagus. *Am J Clin Pathol*. 1978;70(1):1-5.

86. Skinner DB, Walther BC, Riddell RH, Schmidt H, Iascone C, DeMeester TR. Barrett's esophagus. comparison of benign and malignant cases. *Ann Surg*. 1983;198(4):554-565.

87. Oh DS, Demeester SR. Pathophysiology and treatment of barrett's esophagus. *World J Gastroenterol*. 2010;16(30):3762-3772.

88. Oberg S, Johansson J, Wenner J, Walther B. Metaplastic columnar mucosa in the cervical esophagus after esophagectomy. *Ann Surg*. 2002;235(3):338-345.

89. DeMeester SR, DeMeester TR. Columnar mucosa and intestinal metaplasia of the esophagus: Fifty years of controversy. *Ann Surg*. 2000;231(3):303-321.

90. Chandrasoma PT, Der R, Ma Y, Dalton P, Taira M. Histology of the gastroesophageal junction: An autopsy study. *Am J Surg Pathol*. 2000;24(3):402-409.

91. Suh E, Traber PG. An intestine-specific homeobox gene regulates proliferation and differentiation. *Mol Cell Biol*. 1996;16(2):619-625.

92. Groisman GM, Amar M, Meir A. Expression of the intestinal marker Cdx2 in the columnar-lined esophagus with and without intestinal (barrett's) metaplasia. *Mod Pathol*. 2004;17(10):1282-1288.

93. Vallbohmer D, DeMeester SR, Peters JH, et al. Cdx-2 expression in squamous and metaplastic columnar epithelia of the esophagus. *Dis Esophagus*. 2006;19(4):260-266.

94. Yamanaka Y, Shiotani A, Fujimura Y, et al. Expression of sonic hedgehog (SHH) and CDX2 in the columnar epithelium of the lower oesophagus. *Dig Liver Dis*. 2011;43(1):54-59.

95. Jensen P, Krogsgaard MR, Christiansen J, Braendstrup O, Johansen A, Olsen J. Observer variability in the assessment of type and dysplasia of colorectal adenomas, analyzed using kappa statistics. *Dis Colon Rectum*. 1995;38(2):195-198.

96. Schlemper RJ, Riddell RH, Kato Y, et al. The vienna classification of gastrointestinal epithelial neoplasia. *Gut*. 2000;47(2):251-255.

97. Dixon MF. Gastrointestinal epithelial neoplasia: Vienna revisited. *Gut*. 2002;51(1):130-131.

98. Woodman A, Jankowski J, Shepherd N. The metaplasia-dysplasia-carcinoma sequence of barrett's esophagus. In: Tilanus H.W. ASEA, ed. *Barrett's esophagus.* 2001: Springer, Dordrecht; 2001:167-180. 10.1007/978-94-017-0829-6_13.

99. Spechler SJ. Barrett's esophagus. *Semin Gastrointest Dis*. 1996;7(2):51-60.

100. Qumseya BJ, Bukannan A, Gendy S, et al. Systematic review and meta-analysis of prevalence and risk factors for barrett's esophagus. *Gastrointest Endosc*. 2019;90(5):707-717.e1.

101. Yamasaki T, Sakiani S, Maradey-Romero C, et al. Barrett's esophagus patients are becoming younger: Analysis of a large united states dataset. *Esophagus*. 2020.

102. Shaheen NJ, Falk GW, Iyer PG, Gerson LB, American College of Gastroenterology. ACG clinical guideline: Diagnosis and management of barrett's esophagus. *Am J Gastroenterol*. 2016;111(1):30-50; quiz 51.

103. ASGE Standards of Practice Committee, Evans JA, Early DS, et al. The role of endoscopy in barrett's esophagus and other premalignant conditions of the esophagus. *Gastrointest Endosc*. 2012;76(6):1087-1094.

104. Fitzgerald RC, di Pietro M, Ragunath K, et al. British society of gastroenterology guidelines on the diagnosis and management of barrett's oesophagus. *Gut*. 2014;63(1):7-42.

105. Old O, Moayyedi P, Love S, et al. Barrett's oesophagus surveillance versus endoscopy at need study (BOSS): Protocol and analysis plan for a multicentre randomized controlled trial. *J Med Screen*. 2015;22(3):158-164.

106. Bhat SK, McManus DT, Coleman HG, et al. Oesophageal adenocarcinoma and prior diagnosis of barrett's oesophagus: A population-based study. *Gut*. 2015;64(1):20-25.

107. Verbeek RE, Leenders M, Ten Kate FJ, et al. Surveillance of barrett's esophagus and mortality from esophageal adenocarcinoma: A population-based cohort study. *Am J Gastroenterol*. 2014;109(8):1215-1222.

108. Macdonald CE, Wicks AC, Playford RJ. Final results from 10 year cohort of patients undergoing surveillance for barrett's oesophagus: Observational study. *BMJ*. 2000;321(7271):1252-1255.

109. Corley DA, Mehtani K, Quesenberry C, Zhao W, de Boer J, Weiss NS. Impact of endoscopic surveillance on mortality from barrett's esophagus-associated esophageal adenocarcinomas. *Gastroenterology*. 2013;145(2):312-9.e1.

110. Paterson AL, Gehrung M, Fitzgerald RC, O'Donovan M. Role of TFF3 as an adjunct in the diagnosis of barrett's esophagus using a minimally invasive esophageal sampling device-the cytosponge(TM). *Diagn Cytopathol*. 2019.

111. Lao-Sirieix P, Boussioutas A, Kadri SR, et al. Non-endoscopic screening biomarkers for barrett's oesophagus: From microarray analysis to the clinic. *Gut*. 2009;58(11):1451-1459.

112. Heberle CR, Omidvari AH, Ali A, et al. Cost effectiveness of screening patients with gastroesophageal reflux disease for barrett's esophagus with a minimally invasive cell sampling device. *Clin Gastroenterol Hepatol*. 2017;15(9):1397-1404.e7.

113. Chandrasoma PT, Lokuhetty DM, Demeester TR, et al. Definition of histopathologic changes in gastroesophageal reflux disease. *Am J Surg Pathol*. 2000;24(3):344-351.

114. Going JJ, Fletcher-Monaghan AJ, Neilson L, et al. Zoning of mucosal phenotype, dysplasia, and telomerase activity measured by telomerase repeat assay protocol in barrett's esophagus. *Neoplasia*. 2004;6(1):85-92.

115. van den Brink GR, Hardwick JC, Tytgat GN, et al. Sonic hedgehog regulates gastric gland morphogenesis in man and mouse. *Gastroenterology*. 2001;121(2):317-328.

116. Wang DH, Clemons NJ, Miyashita T, et al. Aberrant epithelial-mesenchymal hedgehog signaling characterizes barrett's metaplasia. *Gastroenterology*. 2010;138(5):1810-1822.

117. Clemons NJ, Wang DH, Croagh D, et al. Sox9 drives columnar differentiation of esophageal squamous epithelium: A possible role in the pathogenesis of barrett's esophagus. *Am J Physiol Gastrointest Liver Physiol*. 2012;303(12):G1335-46.

118. Kazumori H, Ishihara S, Rumi MA, Kadowaki Y, Kinoshita Y. Bile acids directly augment caudal related homeobox gene Cdx2 expression in oesophageal keratinocytes in barrett's epithelium. *Gut*. 2006;55(1):16-25.

119. Liu T, Zhang X, So CK, et al. Regulation of Cdx2 expression by promoter methylation, and effects of Cdx2 transfection on morphology and gene expression of human esophageal epithelial cells. *Carcinogenesis*. 2007;28(2):488-496.

120. Hu Y, Jones C, Gellersen O, Williams VA, Watson TJ, Peters JH. Pathogenesis of barrett esophagus: Deoxycholic acid up-regulates goblet-specific gene MUC2 in concert with CDX2 in human esophageal cells. *Arch Surg*. 2007;142(6):540-4; discussion 544-5.

121. Huo X, Zhang HY, Zhang XI, et al. Acid and bile salt-induced CDX2 expression differs in esophageal squamous cells from patients with and without barrett's esophagus. *Gastroenterology*. 2010;139(1):194-203.e1.

122. O'Riordan JM, Abdel-latif MM, Ravi N, et al. Proinflammatory cytokine and nuclear factor kappa-B expression along the inflammation-metaplasia-dysplasia-adenocarcinoma sequence in the esophagus. *Am J Gastroenterol*. 2005;100(6):1257-1264.

123. Abdel-Latif MM, O'Riordan J, Windle HJ, et al. NF-kappaB activation in esophageal adenocarcinoma: Relationship to barrett's metaplasia, survival, and response to neoadjuvant chemoradiotherapy. *Ann Surg*. 2004;239(4):491-500.

124. Chang CL, Lao-Sirieix P, Save V, De La Cueva Mendez G, Laskey R, Fitzgerald RC. Retinoic acid-induced glandular differentiation of the oesophagus. *Gut*. 2007;56(7):906-917.

125. Sasaki Y, Iwai N, Tsuda T, Kimura O. Sonic hedgehog and bone morphogenetic protein 4 expressions in the hindgut region of murine embryos with anorectal malformations. *J Pediatr Surg*. 2004;39(2):170-3; discussion 170-3.

126. Yip YL, Tsao SW. Regulation of p63 expression in primary and immortalized nasopharyngeal epithelial cells. *Int J Oncol*. 2008;33(4):713-724.

127. Bremner CG, Lynch VP, Ellis FH,Jr. Barrett's esophagus: Congenital or acquired? an experimental study of esophageal mucosal regeneration in the dog. *Surgery*. 1970;68(1):209-216.

128. Quante M, Bhagat G, Abrams JA, et al. Bile acid and inflammation activate gastric cardia stem cells in a mouse model of barrett-like metaplasia. *Cancer Cell*. 2012;21(1):36-51.

129. Fitzgerald RC, Abdalla S, Onwuegbusi BA, et al. Inflammatory gradient in barrett's oesophagus: Implications for disease complications. *Gut*. 2002;51(3):316-322.

130. Wang X, Ouyang H, Yamamoto Y, et al. Residual embryonic cells as precursors of a barrett's-like metaplasia. *Cell*. 2011;145(7):1023-1035.

131. Jiang M, Li H, Zhang Y, et al. Transitional basal cells at the squamous-columnar junction generate barrett's oesophagus. *Nature*. 2017;550(7677):529-533.

132. Sarosi G, Brown G, Jaiswal K, et al. Bone marrow progenitor cells contribute to esophageal regeneration and metaplasia in a rat model of barrett's esophagus. *Dis Esophagus*. 2008;21(1):43-50.

133. Wang DH. The esophageal squamous epithelial cell-still a reasonable candidate for the barrett's esophagus cell of origin? *Cell Mol Gastroenterol Hepatol*. 2017;4(1):157-160.

134. Shields HM, Zwas F, Antonioli DA, Doos WG, Kim S, Spechler SJ. Detection by scanning electron microscopy of a distinctive esophageal surface cell at the junction of squamous and barrett's epithelium. *Dig Dis Sci*. 1993;38(1):97-108.

135. Evans JA, McDonald SA. The complex, clonal, and controversial nature of barrett's esophagus. *Adv Exp Med Biol*. 2016;908:27-40.

136. Nicholson AM, Graham TA, Simpson A, et al. Barrett's metaplasia glands are clonal, contain multiple stem cells and share a common squamous progenitor. *Gut*. 2012;61(10):1380-1389.

137. Shields HM, Rosenberg SJ, Zwas FR, Ransil BJ, Lembo AJ, Odze R. Prospective evaluation of multilayered epithelium in barrett's esophagus. *Am J Gastroenterol*. 2001;96(12):3268-3273.

138. Boch JA, Shields HM, Antonioli DA, Zwas F, Sawhney RA, Trier JS. Distribution of cytokeratin markers in barrett's specialized columnar epithelium. *Gastroenterology*. 1997;112(3):760-765.

139. Chen X, Qin R, Liu B, et al. Multilayered epithelium in a rat model and human barrett's esophagus: Similar expression patterns of transcription factors and differentiation markers. *BMC Gastroenterol*. 2008;8:1-230X-8-1.

140. Gillen P, Keeling P, Byrne PJ, West AB, Hennessy TP. Experimental columnar metaplasia in the canine oesophagus. *Br J Surg*. 1988;75(2):113-115.

141. Coad RA, Woodman AC, Warner PJ, Barr H, Wright NA, Shepherd NA. On the histogenesis of barrett's oesophagus and its associated squamous islands: A three-dimensional study of their morphological relationship with native oesophageal gland ducts. *J Pathol*. 2005;206(4):388-394.

142. Lorinc E, Oberg S. Submucosal glands in the columnar-lined oesophagus: Evidence of an association with metaplasia and neosquamous epithelium. *Histopathology*. 2012;61(1):53-58.

143. Leedham SJ, Preston SL, McDonald SA, et al. Individual crypt genetic heterogeneity and the origin of metaplastic glandular epithelium in human barrett's oesophagus. *Gut*. 2008;57(8):1041-1048.

144. Bartel MJ, Srivastava A, Gordon S, Rothstein RI, Pohl H. Subsquamous intestinal metaplasia is common in treatment-naive barrett's esophagus. *Gastrointest Endosc*. 2018;87(1):67-74.

145. Garman KS, Orlando RC, Chen X. Review: Experimental models for barrett's esophagus and esophageal adenocarcinoma. *Am J Physiol Gastrointest Liver Physiol*. 2012;302(11):G1231-43.

146. Chen KH, Mukaisho K, Sugihara H, Araki Y, Yamamoto G, Hattori T. High animal-fat intake changes the bile-acid composition of bile juice and enhances the development of barrett's esophagus and esophageal adenocarcinoma in a rat duodenal-contents reflux model. *Cancer Sci*. 2007;98(11):1683-1688.

147. Kumagai H, Mukaisho K, Sugihara H, Miwa K, Yamamoto G, Hattori T. Thioproline inhibits development of esophageal adenocarcinoma induced by gastroduodenal reflux in rats. *Carcinogenesis*. 2004;25(5):723-727.

148. Terasaki M, Totsuka Y, Nishimura K, et al. Detection of endogenous DNA adducts, O-carboxymethyl-2'-deoxyguanosine and 3-ethanesulfonic acid-2'-deoxycytidine, in the rat stomach after duodenal reflux. *Cancer Sci*. 2008;99(9):1741-1746.

149. Pera M, Cardesa A, Bombi JA, Ernst H, Pera C, Mohr U. Influence of esophagojejunostomy on the induction of adenocarcinoma of the distal esophagus in sprague-dawley rats by subcutaneous injection of 2,6-dimethylnitrosomorpholine. *Cancer Res*. 1989;49(23):6803-6808.

150. Attwood SE, Smyrk TC, DeMeester TR, Mirvish SS, Stein HJ, Hinder RA. Duodenoesophageal reflux and the development of esophageal adenocarcinoma in rats. *Surgery*. 1992;111(5):503-510.

151. Shirvani VN, Ouatu-Lascar R, Kaur BS, Omary MB, Triadafilopoulos G. Cyclooxygenase 2 expression in barrett's esophagus and adenocarcinoma: Ex vivo induction by bile salts and acid exposure. *Gastroenterology*. 2000;118(3):487-496.

152. Souza RF, Shewmake K, Beer DG, Cryer B, Spechler SJ. Selective inhibition of cyclooxygenase-2 suppresses growth and induces apoptosis in human esophageal adenocarcinoma cells. *Cancer Res*. 2000;60(20):5767-5772.

153. Buttar NS, Wang KK, Leontovich O, et al. Chemoprevention of esophageal adenocarcinoma by COX-2 inhibitors in an animal model of barrett's esophagus. *Gastroenterology*. 2002;122(4):1101-1112.

154. Fein M, Peters JH, Baril N, et al. Loss of function of Trp53, but not apc, leads to the development of esophageal adenocarcinoma in mice with jejunoesophageal reflux. *J Surg Res*. 1999;83(1):48-55.

155. Pham TH, Genta RM, Spechler SJ, Souza RF, Wang DH. Development and characterization of a surgical mouse model of reflux esophagitis and barrett's esophagus. *J Gastrointest Surg*. 2014;18(2):234-40; discussion 240-1.

156. Fong LY, Ishii H, Nguyen VT, et al. P53 deficiency accelerates induction and progression of esophageal and forestomach tumors in zinc-deficient mice. *Cancer Res*. 2003;63(1):186-195.

157. Sui G, Bonde P, Dhara S, et al. Epidermal growth factor receptor and hedgehog signaling pathways are active in esophageal cancer cells from rat reflux model. *J Surg Res*. 2006;134(1):1-9.

158. Hao J, Liu B, Yang CS, Chen X. Gastroesophageal reflux leads to esophageal cancer in a surgical model with mice. *BMC Gastroenterol*. 2009;9:59-230X-9-59.

159. Kong J, Crissey MA, Funakoshi S, Kreindler JL, Lynch JP. Ectopic Cdx2 expression in murine esophagus models an intermediate stage in the emergence of barrett's esophagus. *PLoS One*. 2011;6(4):e18280.

160. Gibson CJ, Parry NM, Jakowski RM, Cooper J. Adenomatous polyp with intestinal metaplasia of the esophagus (barrett esophagus) in a dog. *Vet Pathol*. 2010;47(1):116-119.

161. Chambers JK, Saito T, Fukushima K, et al. Adenocarcinoma of barrett's esophagus in a dog. *J Toxicol Pathol*. 2017;30(3):239-243.

162. Kapoor H, Lohani KR, Lee TH, Agrawal DK, Mittal SK. Animal models of barrett's esophagus and esophageal adenocarcinoma-past, present, and future. *Clin Transl Sci*. 2015;8(6):841-847.

163. Narbona-Arnau B, Argente-Navarro P, Lloris-Carsi JM, Calvo-Bermudez MA, Cejalvo-Lapena D. Experimental model of barrett esophagus in dogs. *J Chir (Paris)*. 1994;131(5):261-265.

164. Li H, Walsh TN, O'Dowd G, Gillen P, Byrne PJ, Hennessy TP. Mechanisms of columnar metaplasia and squamous regeneration in experimental barrett's esophagus. *Surgery*. 1994;115(2):176-181.

165. Abdulnour-Nakhoul S, Nakhoul NL, Wheeler SA, et al. Characterization of esophageal submucosal glands in pig tissue and cultures. *Dig Dis Sci*. 2007;52(11):3054-3065.

166. Kruger L, Gonzalez LM, Pridgen TA, et al. Ductular and proliferative response of esophageal submucosal glands in a porcine model of esophageal injury and repair. *Am J Physiol Gastrointest Liver Physiol*. 2017;313(3):G180-G191.

167. Lancaster MA, Knoblich JA. Organogenesis in a dish: Modeling development and disease using organoid technologies. *Science*. 2014;345(6194):1247125.

168. Fatehullah A, Tan SH, Barker N. Organoids as an in vitro model of human development and disease. *Nat Cell Biol*. 2016;18(3):246-254.

169. Sato T, Vries RG, Snippert HJ, et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*. 2009;459(7244):262-265. doi: 10.1038/nature07935 [doi].

170. Leung C, Tan SH, Barker N. Recent advances in Lgr5(+) stem cell research. *Trends Cell Biol*. 2018;28(5):380-391.

171. Barker N, Huch M, Kujala P, et al. Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. *Cell Stem Cell*. 2010;6(1):25-36. doi: 10.1016/j.stem.2009.11.013 [doi].

172. Stange DE, Koo BK, Huch M, et al. Differentiated troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. *Cell*. 2013;155(2):357-368. doi: 10.1016/j.cell.2013.09.008 [doi].

173. Bartfeld S, Bayram T, van de Wetering M, et al. In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. *Gastroenterology*. 2015;148(1):126-136.e6.

174. Huch M, Dorrell C, Boj SF, et al. In vitro expansion of single Lgr5+ liver stem cells induced by wnt-driven regeneration. *Nature*. 2013;494(7436):247-250. doi: 10.1038/nature11826 [doi].

175. Huch M, Gehart H, van Boxtel R, et al. Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell*. 2015;160(1-2):299-312. doi: 10.1016/j.cell.2014.11.050 [doi].

176. Nanduri LS, Baanstra M, Faber H, et al. Purification and ex vivo expansion of fully functional salivary gland stem cells. *Stem Cell Reports*. 2014;3(6):957-964.

177. Maimets M, Rocchi C, Bron R, et al. Long-term in vitro expansion of salivary gland stem cells driven by wnt signals. *Stem Cell Reports*. 2016;6(1):150-162.

178. Jeong Y, Rhee H, Martin S, et al. Identification and genetic manipulation of human and mouse oesophageal stem cells. *Gut*. 2016;65(7):1077-1086.

179. Drost J, van Jaarsveld RH, Ponsioen B, et al. Sequential cancer mutations in cultured human intestinal stem cells. *Nature*. 2015;521(7550):43-47.

180. De Lisle RC, Borowitz D. The cystic fibrosis intestine. *Cold Spring Harb Perspect Med*. 2013;3(9):a009753.

181. Dekkers JF, Wiegerinck CL, de Jonge HR, et al. A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat Med*. 2013;19(7):939-945.

182. Schwank G, Koo BK, Sasselli V, et al. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell*. 2013;13(6):653-658.

183. Gonzalez G, Huang Q, Mashimo H. Characterization of oncocytes in deep esophageal glands. *Dis Esophagus*. 2016;29(6):670-680.

184. Lorinc E, Oberg S. Hyperplasia of the submucosal glands of the columnar-lined oesophagus. *Histopathology*. 2015;66(5):726-731.

185. Alsop K, Thorne H, Sandhu S, et al. A community-based model of rapid autopsy in end-stage cancer patients. *Nat Biotechnol*. 2016;34(10):1010-1014.

186. Reid PE, Iagallo M, Nehr S, Jankunis M, Morrow L, Trueman L. Mechanism of connective tissue techniques. I. the effect of dye concentration and staining time on anionic dye procedures. *Histochem J*. 1993;25(11):821-829.

187. Fu DA, Campbell-Thompson M. Periodic acid-schiff staining with diastase. *Methods Mol Biol*. 2017;1639:145-149.

188. Gudjonsson T, Adriance MC, Sternlicht MD, Petersen OW, Bissell MJ. Myoepithelial cells: Their origin and function in breast morphogenesis and neoplasia. *J Mammary Gland Biol Neoplasia*. 2005;10(3):261-272.

189. Shah AA, Mulla AF, Mayank M. Pathophysiology of myoepithelial cells in salivary glands. *J Oral Maxillofac Pathol*. 2016;20(3):480-490.

190. Kalra J, Baker J. Multiplex immunohistochemistry for mapping the tumor microenvironment. *Methods Mol Biol*. 2017;1554:237-251.

191. Dixon SJ, Lemberg KM, Lamprecht MR, et al. Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell*. 2012;149(5):1060-1072.

192. Bernabeu RO, Longo FM. The p75 neurotrophin receptor is expressed by adult mouse dentate progenitor cells and regulates neuronal and non-neuronal cell genesis. *BMC Neurosci*. 2010;11:136-2202-11-136.

193. Kameda Y. Expression of glial progenitor markers p75NTR and S100 protein in the developing mouse parathyroid gland. *Cell Tissue Res*. 2007;327(1):15-23.

194. Litvinov SV, Velders MP, Bakker HA, Fleuren GJ, Warnaar SO. Ep-CAM: A human epithelial antigen is a homophilic cell-cell adhesion molecule. *J Cell Biol*. 1994;125(2):437-446.

195. Ghebeh H, Sleiman GM, Manogaran PS, et al. Profiling of normal and malignant breast tissue show CD44high/CD24low phenotype as a predominant stem/progenitor marker when used in combination with ep-CAM/CD49f markers. *BMC Cancer*. 2013;13:289-2407-13-289.

196. Notta F, Doulatov S, Laurenti E, Poeppl A, Jurisica I, Dick JE. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science*. 2011;333(6039):218-221.

197. Vasile R, Flavia R, C R, Adrian FG, V. M. The presence and significance of the esophageal glands in the abdominal esophagus in dog. *Annals of The Romanian Society for Cell Biology*. 2016;20(2):11-14.

198. Matsuba K, Takizawa T, Thurlbeck WM. Oncocytes in human bronchial mucous glands. *Thorax*. 1972;27(2):181-184.

199. Rock JR, Onaitis MW, Rawlins EL, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc Natl Acad Sci U S A*. 2009;106(31):12771-12775.

200. Anderson PJ, Lynch TJ, Engelhardt JF. Multipotent myoepithelial progenitor cells are born early during airway submucosal gland development. *Am J Respir Cell Mol Biol*. 2017;56(6):716-726.

201. Tata A, Kobayashi Y, Chow RD, et al. Myoepithelial cells of submucosal glands can function as reserve stem cells to regenerate airways after injury. *Cell Stem Cell*. 2018;22(5):668-683.e6.

202. Kouznetsova I, Kalinski T, Peitz U, et al. Localization of TFF3 peptide in human esophageal submucosal glands and gastric cardia: Differentiation of two types of gastric pit cells along the rostro-caudal axis. *Cell Tissue Res*. 2007;328(2):365-374.

203. Wolf FA, Angerer P, Theis FJ. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1):15-017-1382-0.

204. Wolock SL, Lopez R, Klein AM. Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst*. 2019;8(4):281-291.e9.

205. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2018(37):38-44.

206. Traag VA, Waltman L, van Eck NJ. From louvain to leiden: Guaranteeing well-connected communities. *Sci Rep*. 2019;9(1):5233-019-41695-z.

207. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods*. 2018;15(4):255-261.

208. Street K, Risso D, Fletcher RB, et al. Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. 2018;19(1):477-018-4772-0.

209. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32(4):381-386.

210. Raudvere U, Kolberg L, Kuzmin I, et al. G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019;47(W1):W191-W198.

211. Ilicic T, Kim JK, Kolodziejczyk AA, et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol*. 2016;17:29-016-0888-1.

212. Franzen O, Gan LM, Bjorkegren JLM. PanglaoDB: A web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford)*. 2019;2019:10.1093/database/baz046.

213. van den Berghe PV, Folmer DE, Malingre HE, et al. Human copper transporter 2 is localized in late endosomes and lysosomes and facilitates cellular copper uptake. *Biochem J*. 2007;407(1):49-59.

214. Sasaki S, Ishibashi K, Marumo F. Aquaporin-2 and -3: Representatives of two subgroups of the aquaporin family colocalized in the kidney collecting duct. *Annu Rev Physiol*. 1998;60:199-220.

215. Uhlen M, Bjorling E, Agaton C, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics*. 2005;4(12):1920-1932.

216. Marcato P, Dean CA, Pan D, et al. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells*. 2011;29(1):32-45.

217. Domenici G, Aurrekoetxea-Rodriguez I, Simoes BM, et al. A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. *Oncogene*. 2019;38(17):3151-3169.

218. Van De Bovenkamp JH, Korteland-Van Male AM, Warson C, et al. Gastric-type mucin and TFF-peptide expression in barrett's oesophagus is disturbed during increased expression of MUC2. *Histopathology*. 2003;42(6):555-565.

219. Sellers ZM, Illek B, Figueira MF, et al. Impaired PGE2-stimulated cl- and HCO3- secretion contributes to cystic fibrosis airway disease. *PLoS One*. 2017;12(12):e0189894.

220. Sarosiek J, Murty VL, Nadziejko C, Slomiany A, Slomiany BL. Prostaglandin effect on the physical properties of gastric mucus glycoprotein and its susceptibility to pepsin. *Prostaglandins*. 1986;32(5):635-646.

221. Scharfstein J, Schmitz V, Svensjo E, Granato A, Monteiro AC. Kininogens coordinate adaptive immunity through the proteolytic release of bradykinin, an endogenous danger signal driving dendritic cell maturation. *Scand J Immunol*. 2007;66(2-3):128-136.

222. Becker DJ, Lowe JB. Fucose: Biosynthesis and biological function in mammals. *Glycobiology*. 2003;13(7):41R-53R.

223. Owen RP, White MJ, Severson DT, et al. Single cell RNA-seq reveals profound transcriptional similarity between barrett's oesophagus and oesophageal submucosal glands. *Nat Commun*. 2018;9(1):4261-018-06796-9.

224. Karthaus WR, Iaquinta PJ, Drost J, et al. Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell*. 2014;159(1):163-175.

225. Wang X, Kruithof-de Julio M, Economides KD, et al. A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature*. 2009;461(7263):495-500.

226. Yang X, Wang H, Jiao B. Mammary gland stem cells and their application in breast cancer. *Oncotarget*. 2017;8(6):10675-10691.

227. Lynch TJ, Anderson PJ, Rotti PG, et al. Submucosal gland myoepithelial cells are reserve stem cells that can regenerate mouse tracheal epithelium. *Cell Stem Cell*. 2018;22(5):653-667.e5.

228. O'Flanagan CH, Campbell KR, Zhang AW, et al. Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. *Genome Biol*. 2019;20(1):210-019-1830-0.

229. Whitley D, Goldberg SP, Jordan WD. Heat shock proteins: A review of the molecular chaperones. *J Vasc Surg*. 1999;29(4):748-751.

230. Dutta D, Heo I, Clevers H. Disease modeling in stem cell-derived 3D organoid systems. *Trends Mol Med*. 2017;23(5):393-410.

231. Broutier L, Andersson-Rolf A, Hindley CJ, et al. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. *Nat Protoc*. 2016;11(9):1724-1743.

232. Jaiswal KR, Morales CP, Feagins LA, et al. Characterization of telomerase-immortalized, non-neoplastic, human barrett's cell line (BAR-T). *Dis Esophagus*. 2007;20(3):256-264.

233. Arul GS, Moorghen M, Myerscough N, Alderson DA, Spicer RD, Corfield AP. Mucin gene expression in barrett's oesophagus: An in situ hybridisation and immunohistochemical study. *Gut*. 2000;47(6):753-761.

234. Verzi MP, Shin H, Ho LL, Liu XS, Shivdasani RA. Essential and redundant functions of caudal family proteins in activating adult intestinal genes. *Mol Cell Biol*. 2011;31(10):2026-2039.

235. Phillips RW, Frierson HF,Jr, Moskaluk CA. Cdx2 as a marker of epithelial intestinal differentiation in the esophagus. *Am J Surg Pathol*. 2003;27(11):1442-1447.

236. Marchetti M, Caliot E, Pringault E. Chronic acid exposure leads to activation of the cdx2 intestinal homeobox gene in a long-term culture of mouse esophageal keratinocytes. *J Cell Sci*. 2003;116(Pt 8):1429-1436.

237. Hu Y, Williams VA, Gellersen O, Jones C, Watson TJ, Peters JH. The pathogenesis of barrett's esophagus: Secondary bile acids upregulate intestinal differentiation factor CDX2 expression in esophageal cells. *J Gastrointest Surg*. 2007;11(7):827-834.

238. Jetten AM, Kim JS, Sacks PG, et al. Inhibition of growth and squamous-cell differentiation markers in cultured human head and neck squamous carcinoma cells by beta-all-trans retinoic acid. *Int J Cancer*. 1990;45(1):195-202.

239. von Furstenberg RJ, Li J, Stolarchuk C, et al. Porcine esophageal submucosal gland culture model shows capacity for proliferation and differentiation. *Cell Mol Gastroenterol Hepatol*. 2017;4(3):385-404.

240. Abdulnour-Nakhoul S, Nakhoul NL, Wheeler SA, Wang P, Swenson ER, Orlando RC. HCO3- secretion in the esophageal submucosal glands. *Am J Physiol Gastrointest Liver Physiol*. 2005;288(4):G736-44.

241. Brown CM, Snowdon CF, Slee B, Sandle LN, Rees WD. Measurement of bicarbonate output from the intact human oesophagus. *Gut*. 1993;34(7):872-880.

242. Snippert HJ, Haegebarth A, Kasper M, et al. Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science*. 2010;327(5971):1385-1389.

243. Aizarani N, Saviano A, Sagar, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature*. 2019;572(7768):199-204.

244. Valenta T, Hausmann G, Basler K. The many faces and functions of beta-catenin. *EMBO J*. 2012;31(12):2714-2736.

245. Raja SB, Murali MR, Devaraj H, Devaraj SN. Differential expression of gastric MUC5AC in colonic epithelial cells: TFF3-wired IL1 beta/akt crosstalk-induced mucosal immune response against shigella dysenteriae infection. *J Cell Sci*. 2012;125(Pt 3):703-713.

246. Moons LM, Bax DA, Kuipers EJ, et al. The homeodomain protein CDX2 is an early marker of barrett's oesophagus. *J Clin Pathol*. 2004;57(10):1063-1068.

Appendices

A

Estimated Number of Cells
3,650

Mean Reads per Cell	Median Genes per Cell
121,299	1,288

Sequencing

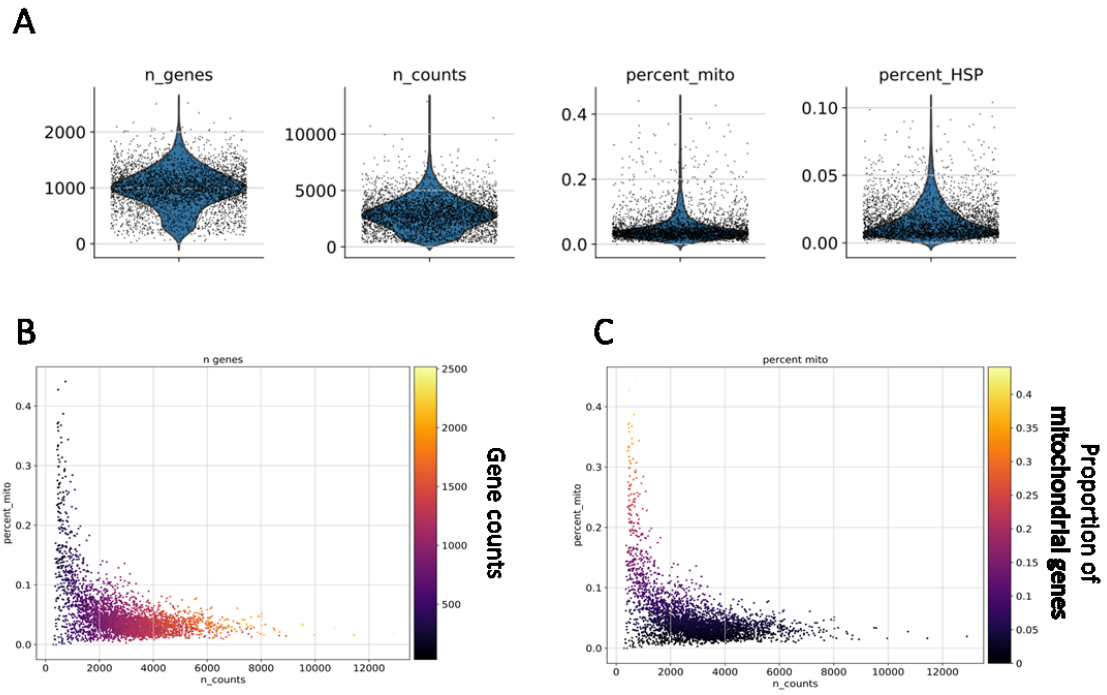
Number of Reads	442,743,026
Valid Barcodes	97.3%
Sequencing Saturation	90.5%
Q30 Bases in Barcode	95.8%
Q30 Bases in RNA Read	93.1%
Q30 Bases in UMI	92.8%

B

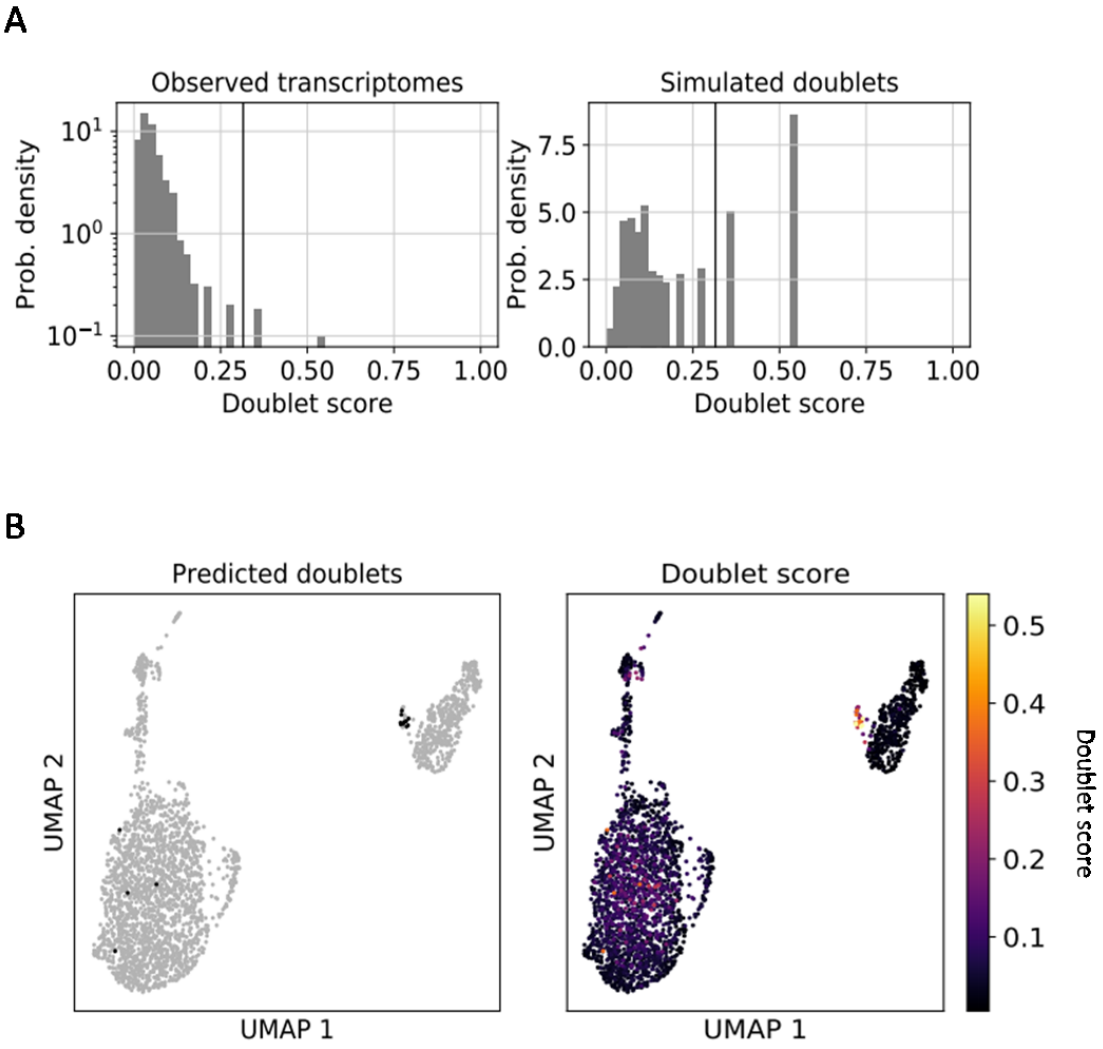
Mapping

Reads Mapped to Genome	93.3%
Reads Mapped Confidently to Genome	91.7%
Reads Mapped Confidently to Intergenic Regions	13.7%
Reads Mapped Confidently to Intronic Regions	18.1%
Reads Mapped Confidently to Exonic Regions	59.9%
Reads Mapped Confidently to Transcriptome	54.7%
Reads Mapped Antisense to Gene	1.0%

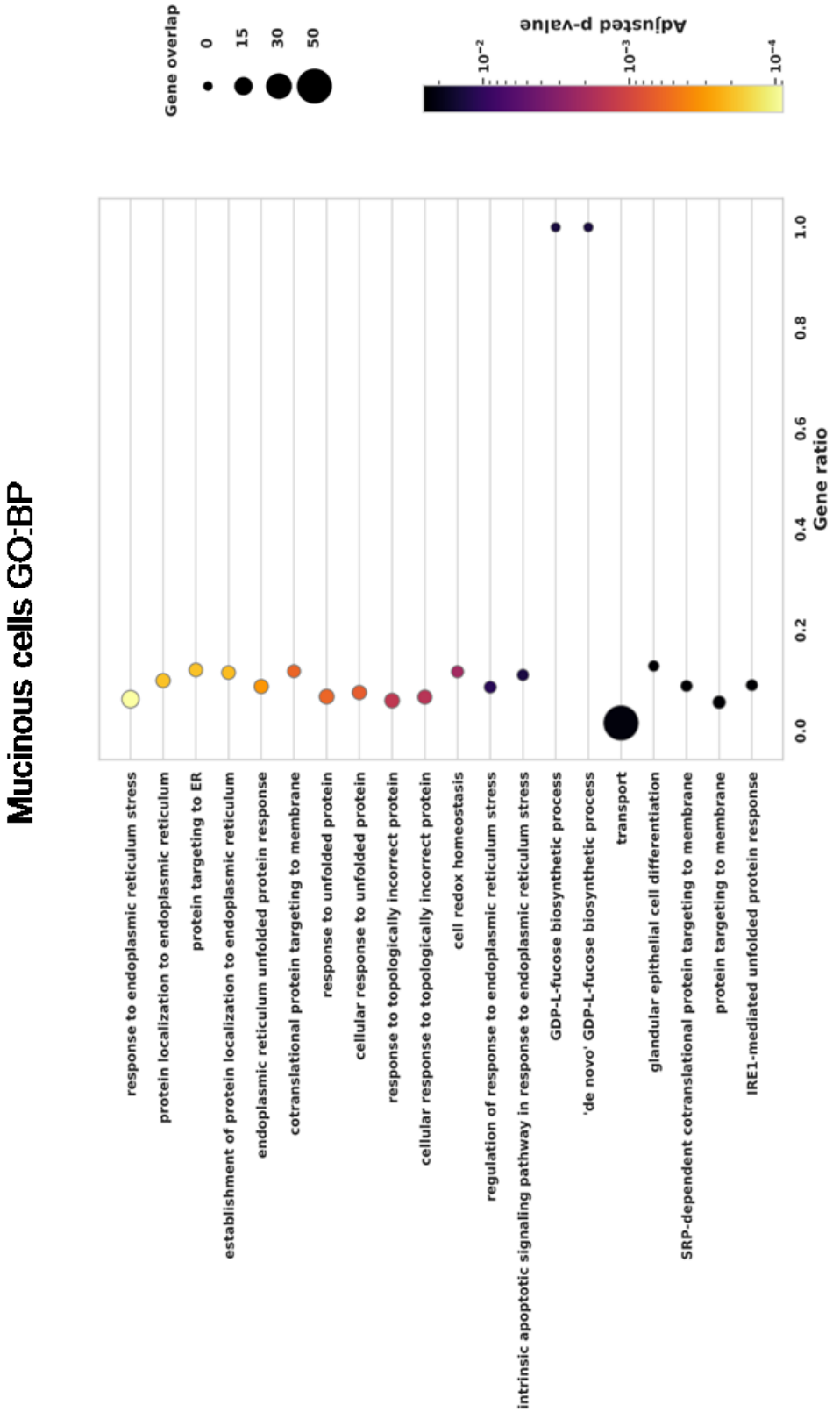
Appendix Figure 1: Cell ranger output of QC metrics of sorted EPCAM positive (EPCAM^{+}) pig SMG cells sequenced using the 10X Genomics Chromium Single Cell 3' v2 chemistry technology. Output summary of the number of sequenced cells, mean sequenced reads per cell and the median genes per cell in addition to other metrics (A). Proportion of genes mapped to Ensembl SScrofa11.1 reference genome (B).



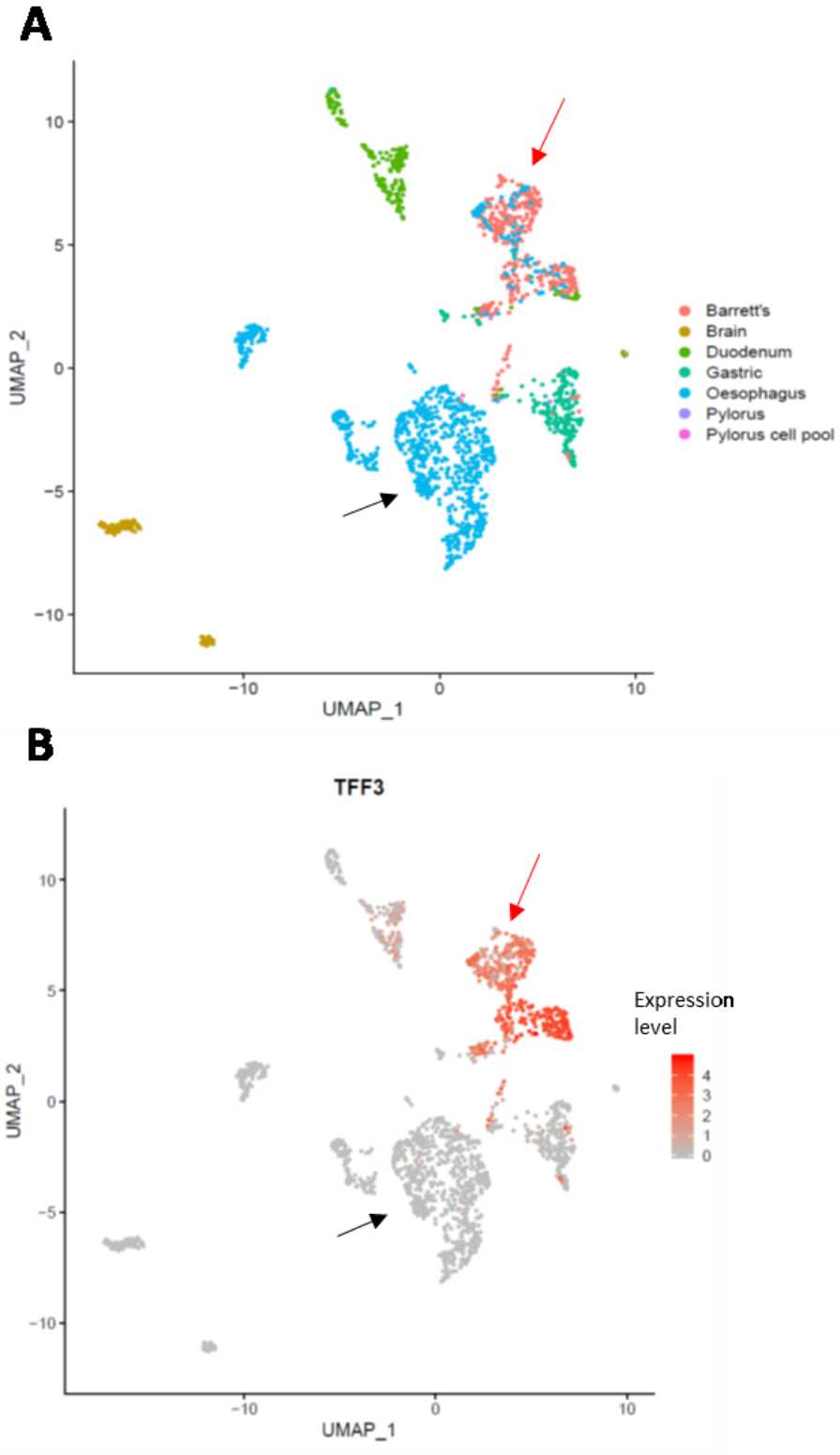
Appendix Figure 2: Preprocessing and analysis of sc-RNA seq data of EPCAM^{+} sorted pig SMG cells to remove outliers based on the proportion of mitochondrial and heat shock protein (HSP) genes, the number of read counts per cell. Violin plots show the number and distribution of genes detected per cell (n_genes), read counts per cell (n_counts), percentage of mitochondrial (percent_mito) and heat shock protein (percent_HSP) genes of pig SMG sequenced cells (A). Scatterplot of sequenced transcripts count per cell in the x-axis, proportion of mitochondrial genes in the y-axis and colour scale represents gene counts (B). Scatterplot of sequenced transcripts counts in the x-axis, proportion of mitochondrial genes in the y-axis and colour scale represents proportion of mitochondrial genes (C).



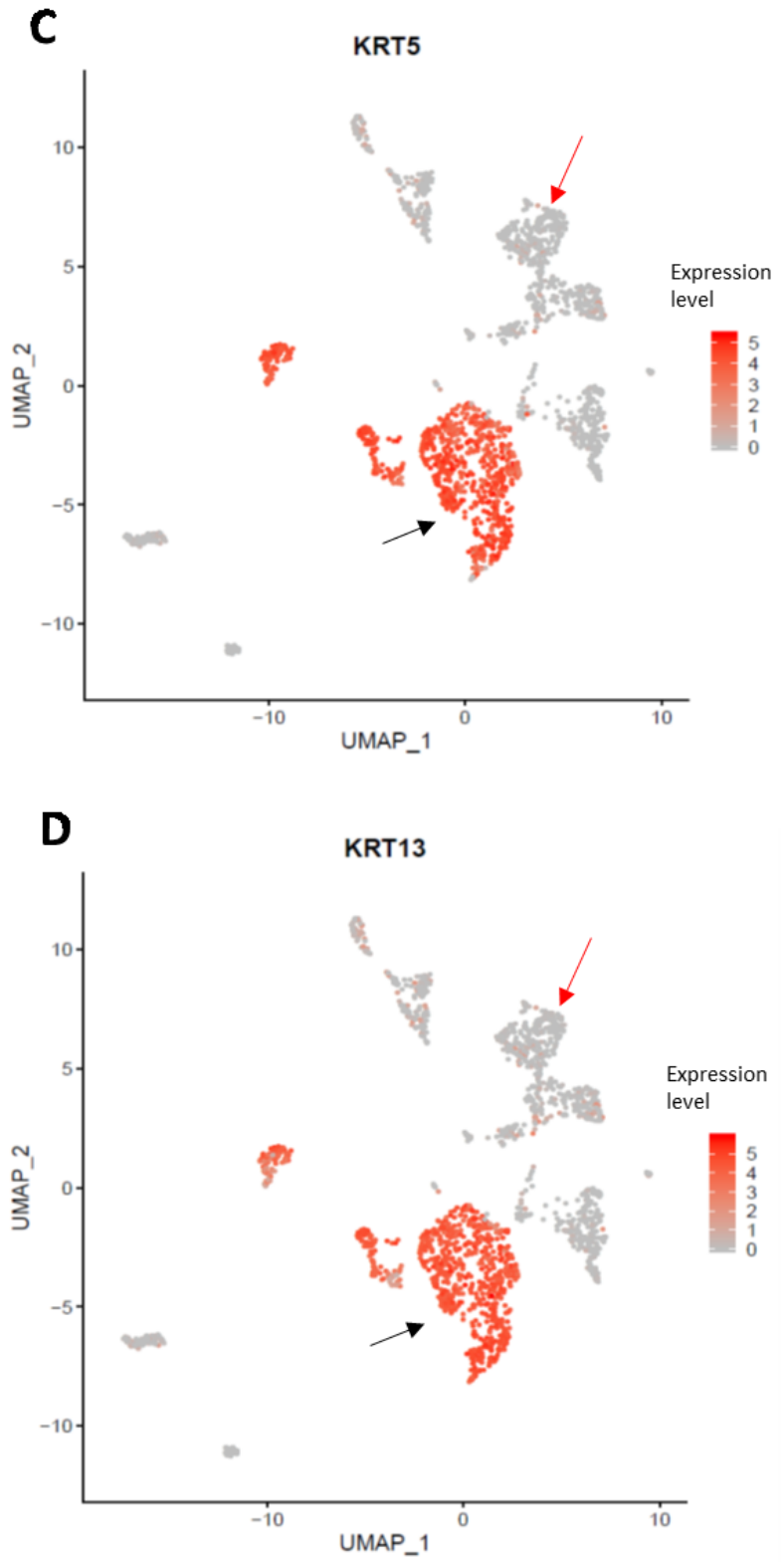
Appendix Figure 3: Pre-processing and analysis of sc-RNA seq data of EPCAM^{+} sorted pig SMG cells to identify potential doublets using the Scrublet package. Bar graph shows the probability density distribution of observed transcriptomes in pig sc-RNA seq data (left graph) and the simulated distribution of doublets (right graph) (A). UMAP scatter plot indicating doublet cells in black (left plot) and with the doublet score as colour scale (right plot) (B).



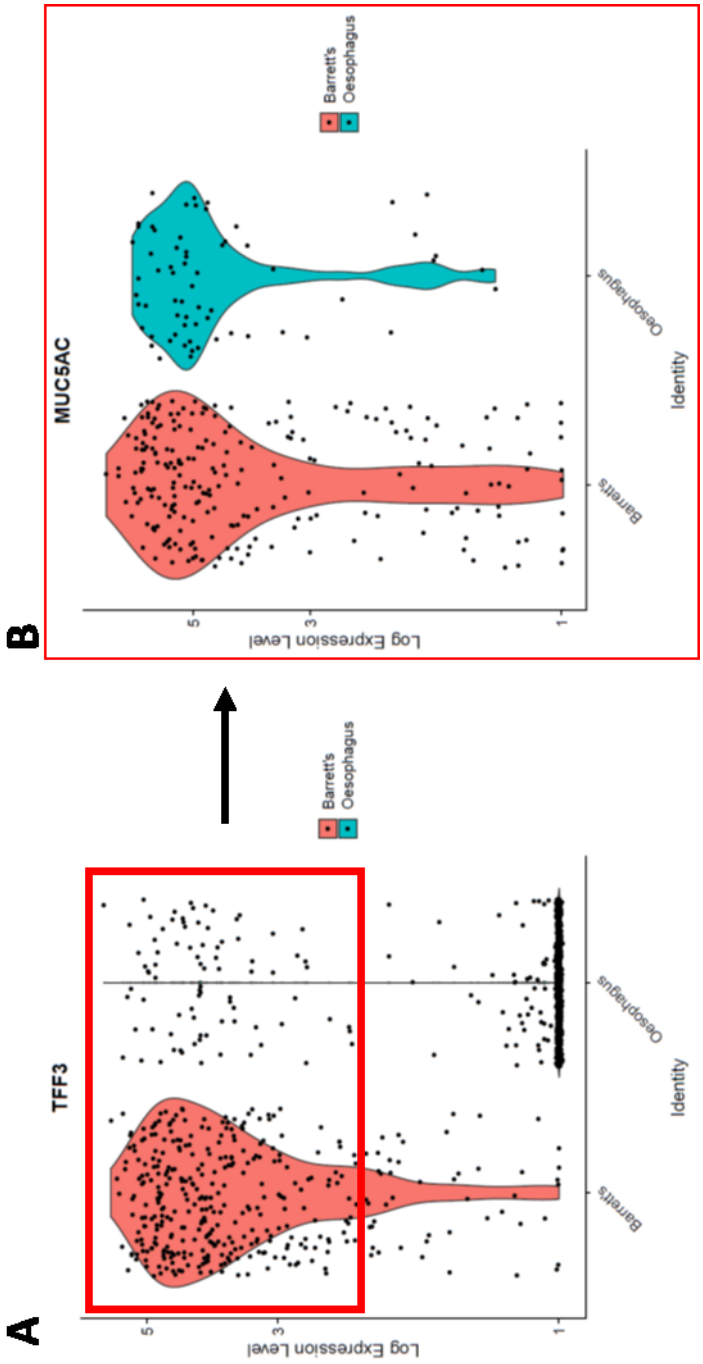
Appendix Figure 4: Gene ontology analysis of sc-RNA seq pig SMG mucinous cells cluster showing the top 20 biological process ontologies enriched in these cells. Enrichment of the GDP-L-fucose biosynthetic process indicates L-fucose residues to be present, which can be stained with the lectin UEA as shown in chapter 2.



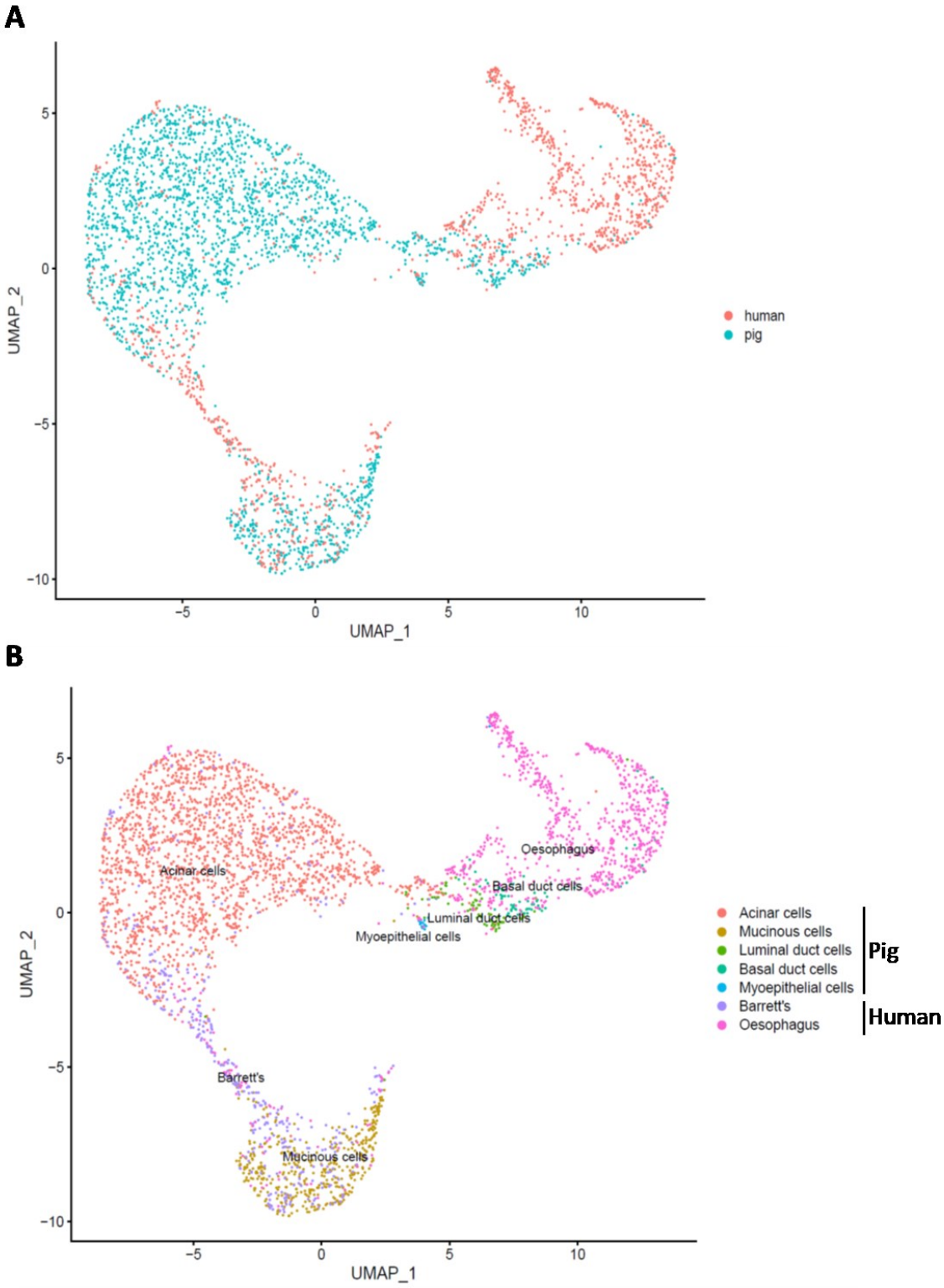
(Legend on next page)



Appendix Figure 5: UMAP projection and feature plots of the re-analysis of the single cell RNA sequenced tissues from four BO patients' data published by Owen *et al*.[223]. UMAP projection of the Leiden algorithm-based identification of distinct clusters of BO patients' sequenced tissues (A). UMAP featureplot showing the expression level of TFF3 in BO patients' sequenced tissues (B). UMAP featureplot showing the expression level of KRT5 in BO patients' sequenced tissues (C). UMAP featureplot showing the expression level of KRT13 in BO patients' sequenced tissues (D).



Appendix Figure 6: Violin-jitter plots of the re-analysis of the single cell RNA sequenced tissues from four BO patients' data published by Owen *et al*. [223]. Violin-jitter plots showing the expression levels of TFF3 for the normal oesophagus and Barrett's lesion derived cells (A). Red rectangle shows the subset of oesophagus cells positive for TFF3 thought to be human SMGs. The subset of cells was used to assess the expression of the Barrett's marker MUC5AC in Figure B. Violin-jitter plot showing the expression level of MUC5AC in the subset of oesophagus and Barrett's cells positive for TFF3 showing positivity for MUC5AC in both populations (B).



Appendix Figure 7: Integration of sc-RNA seq data from pig SMGs with the subset of oesophagus and Barrett's tissue dataset from Owen *et al*.[223]. UMAP projection of human (salmon) and pig (blue) integrated single cell data showing the majority of pig SMG cells not completely overlapping with human cells (either normal oesophagus or Barrett's derived) (A). UMAP projection of human and pig integrated single cell data showing the individual clusters for pig (Acinar, Mucinous, Luminal duct, basal duct and myoepithelial cells) and human (oesophagus and Barrett's cells). Data integration was performed in R using the Seurat package (v3.1.1).

Appendix Table 1: Top 100 differentially expressed genes in the pig SMG identified clusters

Acinar cells				
Acinar cells_names	**Acinar cells_scores**	**Acinar cells_pvals**	**Acinar cells_pvals_adj**	**Acinar cells_logfoldchanges**
PIGR	35.09556961	7.876E-270	7.7925E-266	2.748999596
TAC4	34.70037079	7.7777E-264	3.8476E-260	3.420089006
LDHB	33.19480515	1.2794E-241	2.5317E-238	2.057349205
COX7A1	31.96469688	3.3762E-224	4.7721E-221	2.603935003
MGST3	31.88702202	4.0408E-223	4.9974E-220	2.206213236
SLC31A2	31.74107742	4.2162E-221	4.635E-218	4.568911552
CHCHD10	29.8445282	1.0338E-195	1.0229E-192	2.742977619
COX3	29.43724823	1.8334E-190	1.6491E-187	0.880287886
TESC	29.26473045	2.9166E-188	2.4048E-185	2.458452225
ATP6	28.88924217	1.6298E-183	1.2404E-180	0.844631553
PABPC1	28.76816368	5.3689E-182	3.7943E-179	1.839092374
MGST1	28.70996475	2.8651E-181	1.8898E-178	1.838107467
CYTB	28.63814735	2.252E-180	1.3926E-177	1.078532815
NDRG2	27.84423447	1.265E-170	6.9531E-168	2.10512495
ETV1	27.74062538	2.2609E-169	1.1773E-166	2.48474741
COX1	27.57369614	2.3011E-167	1.1383E-164	0.99373728
TMEM254	26.50456047	8.5876E-155	3.1469E-152	2.303415537
SLC12A2	26.46875191	2.22E-154	7.8447E-152	1.74241817
PRDX5	26.34455872	5.9247E-153	2.0213E-150	1.235369921
RPL23	26.33639336	7.3487E-153	2.4236E-150	0.708165586
ATP1B1	25.54455566	6.3113E-144	1.8366E-141	1.888541102
CALM1	25.46376801	4.9695E-143	1.4048E-140	1.447326541
EPHX1	24.71188927	7.9688E-135	2.0748E-132	1.634375215
ND4	24.61062241	9.7225E-134	2.4665E-131	0.798799515
COX6C	24.56992722	2.649E-133	6.5524E-131	0.67387563
PRDX3	24.56765175	2.8016E-133	6.7608E-131	1.884429932
ATP2A3	24.33941269	7.5049E-131	1.7268E-128	2.222595453
COX2	24.05716133	7.0247E-128	1.448E-125	0.630399644
AQP3	23.45795441	1.0966E-121	2.2142E-119	2.083213806
VMO1	23.38757515	5.7181E-121	1.1315E-118	1.866246223
ELF5	23.37054062	8.5217E-121	1.6532E-118	3.464596748
CHP2	23.17014694	9.109E-119	1.7332E-116	3.572888374
COX5B	22.680439	6.9887E-114	1.172E-111	1.152809978
SLC26A4	22.23012161	1.7566E-109	2.7155E-107	4.043806553
PAM	22.22928619	1.7896E-109	2.7241E-107	3.09431839
KNG1	22.204319	3.1198E-109	4.6769E-107	2.680311203
ATP5MC2	21.97089386	5.4681E-107	8.0748E-105	1.053187013
CST3	21.92638588	1.4554E-106	2.0869E-104	1.888903379
CD99	21.82993698	1.2059E-105	1.7044E-103	1.82498312
RPS15	21.80580711	2.0437E-105	2.8479E-103	0.511272609
PHEROC	21.74499893	7.7033E-105	1.0586E-102	2.063805819
ST3GAL5	21.50742531	1.3267E-102	1.7501E-100	3.790534735

ETFB	21.46037102	3.6537E-102	4.7565E-100	1.838229299
GPT2	21.05677223	1.98228E-98	2.51445E-96	3.120462656
ATP5PO	20.96015549	1.51588E-97	1.8985E-95	1.198361039
RPL7	20.70875931	2.88794E-95	3.52757E-93	0.496008068
CYB5A	20.59223175	3.22181E-94	3.79484E-92	1.495782971
SCGB1D1	20.45287895	5.66316E-93	6.44037E-91	1.877626777
CDC26	19.811203	2.38319E-87	2.61992E-85	1.242602587
ND1	19.69537735	2.36228E-86	2.56839E-84	0.85457319
SLC25A3	19.33237267	2.86914E-83	2.98814E-81	1.015988231
SERINC2	19.20118713	3.61735E-82	3.6897E-80	2.767758369
MT3	19.05858803	5.57595E-81	5.62943E-79	2.081139565
F5	18.87817383	1.72449E-79	1.70621E-77	2.539270878
SEC11C	18.58778954	4.0347E-77	3.80184E-75	1.268070221
ZKSCAN1	18.51791191	1.48069E-76	1.38207E-74	2.271423817
GGH	18.40620232	1.17154E-75	1.07326E-73	1.650136352
MAP1LC3A	18.18870735	6.3418E-74	5.70416E-72	1.229333639
FAU	18.08697128	4.03672E-73	3.56601E-71	0.410433531
ATP1A1	17.40754509	7.23187E-68	6.06374E-66	1.291412234
NDUFS8	17.21438026	2.07154E-66	1.69387E-64	1.261193991
SUCLG1	17.08095169	2.05737E-65	1.62845E-63	1.19651103
PGAM1	17.02446175	5.40884E-65	4.21378E-63	1.804200649
MPZL1	16.94481277	2.10205E-64	1.61223E-62	1.750296116
NDUFA6	16.87399864	6.99011E-64	5.32001E-62	1.411848783
ND3	16.66182327	2.48335E-62	1.8336E-60	1.218705058
RPL27	16.62875938	4.31434E-62	3.13869E-60	0.584894598
ATP5MC3	16.4893074	4.37944E-61	3.16279E-59	1.182370782
RPL8	16.42070198	1.35985E-60	9.61025E-59	0.42689839
PLCB1	16.35785294	3.82375E-60	2.66424E-58	3.766625643
LGALS1	16.24888802	2.27454E-59	1.57373E-57	1.703380108
PHYH	16.17860985	7.1389E-59	4.90502E-57	1.169215322
ATP5F1A	16.15049744	1.12649E-58	7.58198E-57	0.988760769
CALD1	16.14283371	1.27548E-58	8.52675E-57	1.912445664
BCAP29	16.14069176	1.32052E-58	8.76862E-57	1.545891523
ANO1	16.08908844	3.04288E-58	2.00709E-56	2.343875885
UCHL3	15.83554935	1.76906E-56	1.13657E-54	1.529733181
RAMP1	15.79328251	3.46093E-56	2.20919E-54	2.128176689
CGNL1	15.76533794	5.38832E-56	3.41744E-54	2.316249609
DBI	15.36779594	2.6917E-53	1.62389E-51	1.253502369
SELENOP	15.35015965	3.53316E-53	2.11861E-51	1.188367963
PPP3CC	15.34774971	3.66685E-53	2.18553E-51	2.625086546
CLU	14.96128559	1.31462E-50	7.69638E-49	0.869400144
ST3GAL4	14.934412	1.96802E-50	1.14539E-48	3.049704552
ALDH9A1	14.85971642	6.01796E-50	3.48197E-48	1.648265481
WFDC2	14.75379658	2.90833E-49	1.67297E-47	0.384126782
NDUFA13	14.5943985	3.04891E-48	1.7437E-46	0.767293513
RPL7A	14.57371616	4.1281E-48	2.34732E-46	0.375507414

ATP5PB	14.50367165	1.14839E-47	6.45575E-46	0.946179986
RPL18	14.49942207	1.22175E-47	6.82936E-46	0.414580315
NIPSNAP2	14.10070896	3.75976E-45	1.99995E-43	1.608693123
TNFSF13	14.02089214	1.16145E-44	6.11245E-43	3.223759413
ND5	13.84894466	1.29096E-43	6.75805E-42	0.933117568
NPM1	13.82521725	1.7957E-43	9.35089E-42	0.741493881
COX6A1	13.61053276	3.46707E-42	1.78662E-40	0.656265378
LAPTM4B	13.59731388	4.15412E-42	2.12958E-40	0.973125637
ITM2B	13.56384754	6.56024E-42	3.34572E-40	0.54839319
SPHK1	13.53657341	9.5124E-42	4.82645E-40	2.502974272
RPL11	13.44263458	3.4011E-41	1.70815E-39	0.313049316
ETFA	13.37543488	8.41603E-41	4.20547E-39	1.546870112
Mucinous cells				
Mucinous cells_names	**Mucinous cells_scores**	**Mucinous cells_pvals**	**Mucinous cells_pvals_adj**	**Mucinous cells_logfoldchanges**
BPIFB2	32.63071442	1.5047E-233	1.2898E-229	5.806097984
CLCA1	32.60235977	3.7973E-233	1.2898E-229	6.383059978
TFF3	32.60145569	3.9108E-233	1.2898E-229	4.953651905
CGREF1	32.57693863	8.7019E-233	2.1524E-229	5.366503716
AGR2	32.45389175	4.7735E-231	9.4459E-228	4.824347019
LMAN1L	32.39793396	2.935E-230	4.8399E-227	6.045088291
LYZ	32.08764267	6.5571E-226	9.268E-223	6.112739086
DNAJC10	31.63817596	1.1029E-219	1.364E-216	4.52314043
TSPAN1	31.14775276	5.4401E-213	5.9804E-210	5.256000042
GMDS	31.06832886	6.4526E-212	6.3842E-209	3.060869217
SELENOM	30.03149033	3.8096E-198	3.4265E-195	3.717355013
FCGBP	29.97267342	2.2289E-197	1.8377E-194	5.021360397
CDC42EP5	29.89855194	2.055E-196	1.564E-193	3.901674747
TFF2	29.82198524	2.0269E-195	1.4324E-192	5.816668987
GFPT1	29.32799149	4.561E-189	3.0085E-186	3.514245272
CHIA	29.18113899	3.3652E-187	2.0809E-184	5.320664883
SPDEF	28.89454269	1.3981E-183	8.1371E-181	4.863546848
EIF4EBP3	28.85479355	4.4117E-183	2.2973E-180	3.741248131
SEC61B	28.59843254	7.0273E-180	3.4764E-177	1.86843729
S100A6	28.55804253	2.2319E-179	1.0038E-176	4.087219238
SELENOF	28.5130291	8.0765E-179	3.4743E-176	2.344937325
RABAC1	28.23218536	2.3555E-175	9.322E-173	1.935914636
TMSB10	28.04532051	4.5555E-173	1.7336E-170	2.363133669
FAM3D	28.02384758	8.3236E-173	3.0501E-170	5.690233707
NKX3-1	27.97443199	3.3262E-172	1.1348E-169	4.049115181
GOLM1	27.88440323	4.1244E-171	1.3602E-168	4.590656281
MARC2	26.83003998	1.4423E-158	4.3242E-156	3.232474804
TM9SF3	26.45529366	3.1713E-154	8.9647E-152	2.610259771
XBP1	26.37925529	2.3707E-153	6.5154E-151	1.819938183
CD63	26.16659164	6.3812E-151	1.6615E-148	1.430883169
TMED3	25.92767906	3.2469E-148	8.1419E-146	3.126304626
IRF3	25.85998917	1.8786E-147	4.4254E-145	4.095769882

ARF4	25.75796318	2.6249E-146	6.0397E-144	2.101312876
SSR4	25.70441246	1.0434E-145	2.3462E-143	1.713208437
S100A13	25.69053459	1.4912E-145	3.2787E-143	2.58526969
SSR3	25.314888	2.1902E-141	4.7108E-139	1.993435502
F3	25.12813568	2.4507E-139	5.159E-137	3.136665344
CREB3L1	24.92133904	4.3687E-137	9.005E-135	2.53363657
SEC61G	24.75256348	2.9094E-135	5.7572E-133	2.016571999
RPS12	24.72969818	5.1273E-135	9.947E-133	0.693690181
ST6GALNAC1	24.72221375	6.1714E-135	1.1742E-132	5.964443684
SSR2	24.66621971	2.4654E-134	4.6024E-132	1.239546895
SMIM14	24.44637299	5.4998E-132	1.0077E-129	2.838172913
S100A14	23.82624817	1.7856E-125	3.1548E-123	2.587273359
PRDX4	23.81131744	2.5497E-125	4.3495E-123	2.66458106
PYCR1	23.45409775	1.2006E-121	2.0133E-119	2.960612297
MARCKS	23.22537613	2.5236E-119	4.0932E-117	2.982030392
SFT2D1	23.07389832	8.4678E-118	1.3513E-115	1.563184857
ACSL5	22.82483292	2.5993E-115	4.0821E-113	4.443758488
ATOX1	22.48513603	5.8021E-112	8.8317E-110	1.160480618
KDELR2	22.29199219	4.4189E-110	6.6244E-108	2.51722455
S100A1	22.20568466	3.0264E-109	4.4692E-107	1.377551198
TSTA3	22.12919426	1.6551E-108	2.4081E-106	2.146559
CD164	22.01208878	2.2059E-107	3.1631E-105	1.747808933
TMED9	21.96064377	6.8521E-107	9.685E-105	2.005913973
PKIB	21.7525692	6.5318E-105	9.1022E-103	3.884446859
SEC62	21.59199142	2.1361E-103	2.9354E-101	1.264436841
PPIB	21.50369072	1.4378E-102	1.9488E-100	1.55399096
GALNT7	21.13793373	3.5634E-99	4.76437E-97	4.347745419
TCEA3	20.97310638	1.15474E-97	1.48376E-95	4.713586807
OSTC	20.77785492	6.86587E-96	8.49136E-94	2.117159367
GPX8	20.60836411	2.30906E-94	2.71974E-92	4.728268623
TMEM263	20.35079956	4.56767E-92	5.25494E-90	3.013186932
CREB3L4	20.28412437	1.776E-91	2.01974E-89	4.735647202
GALE	20.12809563	4.18783E-90	4.65555E-88	4.038036823
RPS29	20.01689339	3.92435E-89	4.31416E-87	0.468028635
NUPR1	19.99568367	6.00496E-89	6.52891E-87	3.19033289
SUPT4H1	19.8615799	8.75125E-88	9.3102E-86	2.173021555
QSOX1	19.70070076	2.12653E-86	2.21473E-84	2.48411417
SPCS1	19.68456841	2.92416E-86	3.01372E-84	1.209230781
TSPAN3	19.6055584	1.38624E-85	1.39953E-83	1.692321658
TMEM258	19.4097805	6.37993E-84	6.37606E-82	1.471977592
CRISP3	18.79864311	7.74738E-79	7.30025E-77	1.80486691
PRSS23	18.71079445	4.04302E-78	3.73847E-76	2.99834156
POLE4	18.57473564	5.1458E-77	4.67088E-75	1.329256535
OAZ1	18.33747864	4.15635E-75	3.67169E-73	0.723877847
MAGED1	18.22067261	3.53774E-74	3.04369E-72	3.013552904
FOXA1	18.13087082	1.81859E-73	1.53788E-71	3.348271847

SELENOS	18.10969734	2.67212E-73	2.2405E-71	2.052390814
CHID1	18.00902557	1.65515E-72	1.35339E-70	1.779970765
FAM114A1	17.90192413	1.13918E-71	9.08957E-70	2.860770702
ZNF423	17.53800964	7.34564E-69	5.72266E-67	5.721113205
FKBP11	17.43718719	4.30762E-68	3.32966E-66	1.28930068
YIPF1	17.41263008	6.6173E-68	5.07531E-66	2.260247946
FABP5	17.35339928	1.85907E-67	1.4149E-65	4.01434803
LMAN1	17.07507706	2.27525E-65	1.65525E-63	1.949146748
RRBP1	16.96530533	1.48328E-64	1.06345E-62	1.922451496
PAPSS1	16.71492195	1.02052E-62	7.06085E-61	1.980296731
PDIA4	16.53527832	2.04432E-61	1.40462E-59	1.821804523
RPLP2	16.06396103	4.56454E-58	3.03098E-56	0.498378396
ATF4	15.84050465	1.63504E-56	1.06428E-54	1.184660316
ERGIC2	15.80899048	2.69755E-56	1.74442E-54	1.363103509
MORF4L2	15.70779228	1.33756E-55	8.59342E-54	1.69182384
RHEB	15.44925022	7.63265E-54	4.84086E-52	1.302172661
ARL1	15.43131161	1.008E-53	6.35235E-52	1.463537693
CMAS	15.42631054	1.08922E-53	6.77783E-52	1.976357579
MYDGF	15.40201664	1.58644E-53	9.81013E-52	1.495921969
SLC39A7	15.25531483	1.51761E-52	9.32623E-51	2.375982285
MANF	15.14238358	8.50734E-52	5.19578E-50	1.605432034
SLC17A9	15.04833031	3.54056E-51	2.136E-49	4.835169315
Basal duct cells				
Basal duct cells_names	**Basal duct cells_scores**	**Basal duct cells_pvals**	**Basal duct cells_pvals_adj**	**Basal duct cells_logfoldchanges**
KRT5	14.24568653	4.76836E-46	4.71782E-42	7.585154533
TACSTD2	14.13390541	2.34755E-45	1.16133E-41	4.384650707
TUBA1B	13.98261833	1.99028E-44	6.56393E-41	3.557124376
ANXA1	13.58098888	5.19209E-42	1.02741E-38	5.539387703
F3	13.36549377	9.61945E-41	1.58625E-37	4.741092682
CTSV	13.33663368	1.41716E-40	1.75267E-37	3.518001318
EIF1	13.25030231	4.49387E-40	4.94026E-37	1.762192726
LMNA	13.08762169	3.87556E-39	3.83448E-36	4.215621471
S100A6	13.05268955	6.13428E-39	5.51751E-36	4.624591827
DDX5	13.02915287	8.35303E-39	6.88707E-36	2.379168987
S100A10	12.98149586	1.55808E-38	1.18582E-35	3.130486727
EMP1	12.62768841	1.48579E-36	8.64728E-34	5.947024345
EIF4A1	12.38383293	3.19702E-35	1.58157E-32	2.601768017
CFL1	12.27346516	1.25753E-34	5.65547E-32	2.064155579
TMSB10	12.23553944	2.00765E-34	8.63641E-32	2.351490736
S100A2	11.98511314	4.25261E-33	1.68301E-30	5.895956516
SRSF5	11.92589664	8.67454E-33	3.301E-30	3.402903795
ANXA2	11.88614082	1.39714E-32	5.11976E-30	2.163376808
PERP	11.7534008	6.78329E-32	2.39692E-29	2.550554514
CLDN1	11.67322445	1.74675E-31	5.76078E-29	6.759562492
H2AFZ	11.66482449	1.92801E-31	6.15346E-29	2.91798377
HSPA5	11.58148766	5.11508E-31	1.58152E-28	2.515587568

ATF4	11.56701565	6.05524E-31	1.81547E-28	2.10720396
S100A11	11.47694206	1.72266E-30	4.8171E-28	2.577698469
ARID5B	11.47544479	1.75274E-30	4.8171E-28	4.037361622
CXCL2	11.46446609	1.98982E-30	5.32089E-28	5.384566307
ACTG1	11.3531847	7.151E-30	1.86189E-27	1.37654078
TPM4	11.29188824	1.43912E-29	3.65094E-27	2.505033016
TUBB2B	10.97201443	5.20987E-28	1.09875E-25	3.136251926
TNFRSF12A	10.97184753	5.21947E-28	1.09875E-25	3.871252775
CYCS	10.81573677	2.89953E-27	5.85469E-25	2.383727551
BTG1	10.71809483	8.37099E-27	1.53375E-24	2.203956604
RAP2B	10.49619579	8.99316E-26	1.50811E-23	5.319732666
WSB1	10.46592236	1.23865E-25	1.99934E-23	3.853515625
HNRNPF	10.46484089	1.25287E-25	1.99934E-23	2.166184664
NFKBIA	10.40761948	2.28873E-25	3.53824E-23	3.052914381
BIRC3	10.35613728	3.92496E-25	5.97439E-23	3.271393538
EZR	10.34083366	4.60514E-25	6.90353E-23	2.581821203
EHD1	10.33052063	5.12816E-25	7.57284E-23	4.904401779
CCNL1	10.23371029	1.40049E-24	2.03772E-22	2.966477633
SERTAD1	10.2130003	1.73418E-24	2.41661E-22	2.955565929
TGIF1	10.20909119	1.80547E-24	2.48102E-22	3.432138205
TIMP1	10.00615597	1.4321E-23	1.94098E-21	6.809280396
MRPS6	9.905020714	3.9589E-23	5.29315E-21	3.263017178
TIMP3	9.787583351	1.27302E-22	1.61477E-20	3.288098812
GJB3	9.730861664	2.22699E-22	2.75423E-20	5.505795479
SDC4	9.415062904	4.72799E-21	5.63599E-19	2.614628315
SQSTM1	9.246560097	2.31834E-20	2.73068E-18	2.011641741
ARF6	9.222023964	2.91543E-20	3.39356E-18	1.893460512
ZFP36L1	9.01842308	1.90816E-19	2.07465E-17	1.870861411
JMJD1C	8.956045151	3.36532E-19	3.61918E-17	3.207073689
FTL	8.91213131	5.00608E-19	5.32583E-17	0.871973753
STX11	8.812826157	1.2203E-18	1.25767E-16	5.851186275
DUSP5	8.69239521	3.54878E-18	3.4764E-16	4.181968212
HSP90AB1	8.683828354	3.82666E-18	3.71186E-16	1.215799332
FUS	8.678754807	4.00125E-18	3.84353E-16	2.383859634
BZW2	8.67559433	4.11396E-18	3.9138E-16	3.067888021
ARF4	8.624444962	6.44044E-18	6.06874E-16	1.544389129
WFDC2	8.575457573	9.86939E-18	9.12596E-16	1.429088473
MANF	8.540775299	1.33324E-17	1.21019E-15	1.991502047
LITAF	8.530878067	1.45241E-17	1.29461E-15	2.860285044
IGFBP7	8.473157883	2.38829E-17	2.10311E-15	5.810620308
CLIC4	8.465006828	2.56138E-17	2.22301E-15	2.38319397
IER3	8.398553848	4.52012E-17	3.82239E-15	2.911284685
KLF4	8.278205872	1.25045E-16	1.04847E-14	2.623888969
BZW1	8.263900757	1.40988E-16	1.17221E-14	1.816059947
SRSF6	8.227472305	1.91206E-16	1.57649E-14	2.091897726
RND3	8.221234322	2.0142E-16	1.64698E-14	3.054683685

GPX1	8.139227867	3.97806E-16	3.19991E-14	2.224319696
DCN	8.122344971	4.57263E-16	3.64852E-14	6.363537788
FOSL1	8.082755089	6.33196E-16	4.98228E-14	6.071890831
CNN3	8.08250618	6.34493E-16	4.98228E-14	2.085437536
KRT8	8.004574776	1.1988E-15	9.19451E-14	0.987173438
CSTB	7.947021008	1.9105E-15	1.44294E-13	2.047523499
SFPQ	7.933630943	2.1283E-15	1.58327E-13	2.326987267
HNRNPM	7.903356552	2.71491E-15	2.00458E-13	1.587733865
DUSP14	7.899613857	2.79768E-15	2.03532E-13	3.630604982
ACTB	7.898449421	2.82394E-15	2.03942E-13	1.766815901
GSTM3	7.805797577	5.91264E-15	4.2391E-13	5.442749023
MCL1	7.804134369	5.99114E-15	4.26448E-13	1.863733888
TPM1	7.720631123	1.15755E-14	8.00896E-13	2.084782362
TXNRD1	7.683454037	1.54856E-14	1.03431E-12	3.631524324
TMEM165	7.672475338	1.68708E-14	1.1128E-12	2.816466808
MAT2A	7.593879223	3.10467E-14	2.00769E-12	2.120256662
CTNNB1	7.576413631	3.55238E-14	2.28229E-12	1.474372149
BAZ1A	7.529755116	5.08356E-14	3.22415E-12	3.95126915
SSB	7.342288494	2.09972E-13	1.30658E-11	1.563972712
TPI1	7.328648567	2.32485E-13	1.43763E-11	1.516256452
MDK	7.295796394	2.96898E-13	1.82454E-11	2.023135662
BHLHE40	7.267934322	3.65026E-13	2.22936E-11	3.480073452
FST	7.239905834	4.48996E-13	2.72538E-11	8.320314407
TOB1	7.233751297	4.69832E-13	2.81728E-11	2.432250261
TSC22D1	7.204807758	5.81257E-13	3.44579E-11	2.081151485
ACAT2	7.203310966	5.87678E-13	3.461E-11	2.743899345
ETV3	7.200649261	5.99265E-13	3.50836E-11	4.083187103
PLAUR	7.176197529	7.16771E-13	4.17161E-11	6.644611359
TM4SF1	7.171872616	7.39789E-13	4.25551E-11	3.378238678
UAP1	7.161559105	7.97645E-13	4.56179E-11	1.95170176
KLF10	7.133447647	9.78855E-13	5.53417E-11	2.686289787
COL17A1	7.018007278	2.25055E-12	1.25095E-10	5.216197491
Luminal duct cells				
Luminal duct cells_names	**Luminal duct cells_scores**	**Luminal duct cells_pvals**	**Luminal duct cells_pvals_adj**	**Luminal duct cells_logfoldchanges**
WFDC2	13.99200916	1.74414E-44	1.72565E-40	5.057377815
SCGN	13.68989563	1.16677E-42	5.77202E-39	7.756253719
ANXA2	12.88262463	5.63851E-38	1.58356E-34	2.596545935
B2M	12.72473049	4.30959E-37	8.52781E-34	2.435435295
ALDH1A3	12.55591488	3.68893E-36	6.08304E-33	5.884058475
ANXA1	12.06855965	1.54821E-33	1.91474E-30	4.462289333
SAT1	11.83687878	2.51645E-32	2.48978E-29	2.875611305
KRT7	11.71501827	1.06768E-31	9.60331E-29	3.331536293
TPM1	10.61715698	2.47999E-26	1.44335E-23	2.727565765
RARRES1	10.60322571	2.87877E-26	1.58236E-23	6.528450966
TMSB10	10.51619434	7.27529E-26	3.59908E-23	2.058733225
TACSTD2	10.42942142	1.81992E-25	8.57444E-23	2.806557417

PGLYRP1	10.26817417	9.80427E-25	4.40925E-22	4.004784107
CLU	9.901217461	4.1124E-23	1.56493E-20	1.565190315
CLDN4	9.574679375	1.02174E-21	3.61038E-19	2.355985165
TSTD1	9.545353889	1.35644E-21	4.62779E-19	2.412966967
LGALS3	9.09239769	9.68783E-20	2.73861E-17	3.40776515
TIMP3	9.026265144	1.77631E-19	4.62495E-17	3.208081484
S100A2	8.822447777	1.11984E-18	2.76994E-16	3.479946375
S100A6	8.775835991	1.69638E-18	4.09366E-16	2.595886946
S100A11	8.442763329	3.09923E-17	6.52422E-15	1.747721791
IFI6	8.29441452	1.09121E-16	2.15928E-14	3.205000401
ELF3	8.246943474	1.62497E-16	3.09182E-14	3.120151043
F3	8.187948227	2.65717E-16	4.86852E-14	2.519296885
TUBA1B	8.153635025	3.53146E-16	6.35278E-14	1.848203897
KLF4	8.104701996	5.28752E-16	9.34191E-14	2.607831001
ACSF2	8.057574272	7.78234E-16	1.35085E-13	2.333749056
PRDX6	8.041406631	8.88132E-16	1.48935E-13	1.854438066
PERP	7.92943573	2.20144E-15	3.51306E-13	1.424842477
S100A10	7.895466328	2.89232E-15	4.54231E-13	1.69799459
KRT8	7.868118763	3.60014E-15	5.47996E-13	0.942900717
ARID5B	7.774724007	7.56121E-15	1.10016E-12	3.091472149
EIF1	7.751160145	9.10566E-15	1.30567E-12	0.876094818
GADD45A	7.633599758	2.28288E-14	3.09408E-12	2.716470242
LMNA	7.520941257	5.43832E-14	7.17423E-12	2.12406373
PHLDA2	7.41550684	1.2116E-13	1.57732E-11	2.38927722
SOD1	7.369325638	1.71493E-13	2.17533E-11	1.012197614
ANPEP	7.264578819	3.74202E-13	4.46067E-11	3.889760017
FMO2	7.198101997	6.10565E-13	6.9436E-11	6.356227875
CLDN7	7.118208885	1.09339E-12	1.18879E-10	1.486944199
RND3	7.091807365	1.32372E-12	1.42357E-10	2.603367567
ISG12(A)	7.077359676	1.46927E-12	1.56311E-10	1.920326471
SERTAD1	7.058697701	1.6807E-12	1.75041E-10	2.338320971
DNAJA1	6.99806881	2.59515E-12	2.62004E-10	1.625909209
GSN	6.82882309	8.56142E-12	8.3868E-10	2.605179787
CRYAB	6.772665501	1.26431E-11	1.2028E-09	2.082057476
MDK	6.758820057	1.3912E-11	1.31091E-09	2.051629066
CTSV	6.725796223	1.74635E-11	1.6148E-09	1.505966306
HSPA8	6.707306385	1.98249E-11	1.81619E-09	1.264956951
EMP1	6.679184914	2.40275E-11	2.16116E-09	4.021345615
DUSP5	6.648225307	2.96648E-11	2.64417E-09	3.658400536
DYNLL1	6.64125967	3.11014E-11	2.74748E-09	1.308357835
KRT23	6.629477501	3.36877E-11	2.94961E-09	4.309949875
GPX1	6.614341736	3.73208E-11	3.156E-09	1.882938027
NFKBIA	6.610558033	3.82874E-11	3.2103E-09	1.829728603
DDX5	6.597486019	4.18188E-11	3.39144E-09	1.079487443
SRI	6.562312603	5.29796E-11	4.26163E-09	1.409671545
HSPA5	6.472185612	9.65953E-11	7.58503E-09	1.71857667

CD9	6.444838047	1.15724E-10	9.01553E-09	0.650788426
EZR	6.317473888	2.65873E-10	1.99284E-08	1.679646134
CNN3	6.298984051	2.99603E-10	2.22877E-08	1.863590837
BTG1	6.243858814	4.26905E-10	3.08307E-08	1.41571188
ZFAND5	6.239129066	4.40014E-10	3.15471E-08	1.704846382
ACTG1	6.197333336	5.74277E-10	4.00134E-08	0.722686946
CXCL2	6.114344597	9.69546E-10	6.70817E-08	3.165088415
SQSTM1	6.068421364	1.29174E-09	8.87531E-08	1.24309516
EIF4A1	6.056467533	1.39143E-09	9.49437E-08	1.351320624
MAFF	6.049157619	1.45605E-09	9.86725E-08	2.364295721
SPP1	5.994978428	2.03513E-09	1.36051E-07	4.953618526
GPRC5A	5.95842886	2.54675E-09	1.66871E-07	4.446273804
ISG15	5.909409523	3.43336E-09	2.22024E-07	2.758169889
EDN1	5.873032093	4.27896E-09	2.74909E-07	4.122108459
TUBB2B	5.798213005	6.70253E-09	4.19714E-07	2.150914192
SLC4A7	5.787033081	7.16405E-09	4.45793E-07	4.476169109
SLC5A5	5.74291563	9.30601E-09	5.7546E-07	5.540167809
FUS	5.691918373	1.2562E-08	7.57856E-07	1.758180618
EFEMP1	5.645908833	1.64311E-08	9.70095E-07	4.835658073
TPI1	5.556986332	2.74472E-08	1.56973E-06	1.145304203
CSTB	5.50736475	3.64245E-08	2.04763E-06	1.081536055
ACTB	5.498679161	3.82647E-08	2.12691E-06	1.089945436
SORBS2	5.421366215	5.91453E-08	3.21529E-06	1.671015739
PPP1R15A	5.400210381	6.65628E-08	3.55985E-06	1.42783463
TNFRSF12A	5.387224674	7.1554E-08	3.78586E-06	1.787687063
S100A1	5.354458809	8.58128E-08	4.51613E-06	0.812449872
MYL6	5.332357407	9.69459E-08	5.04833E-06	0.651112437
MIF	5.234060764	1.65826E-07	8.28627E-06	0.660315394
CEBPB	5.209207058	1.89649E-07	9.28907E-06	1.062621951
SRSF5	5.177473545	2.24911E-07	1.0855E-05	1.60416615
ATF4	5.170937538	2.32922E-07	1.11871E-05	1.03702867
HSPB1	5.167669773	2.37031E-07	1.13294E-05	0.798219264
IFITM3	5.152275562	2.57344E-07	1.21826E-05	0.752930641
CXCL8	5.096806526	3.45431E-07	1.61976E-05	5.854331493
MYOF	5.071093082	3.95538E-07	1.8373E-05	2.848552227
S100A14	4.984233856	6.22077E-07	2.84946E-05	1.155113697
CRISP3	4.970646381	6.67301E-07	3.04252E-05	1.67818296
HES1	4.907781124	9.21126E-07	4.06858E-05	1.206656933
DDX3X	4.867705822	1.12901E-06	4.94268E-05	1.375676036
BIRC3	4.844227791	1.27105E-06	5.51568E-05	1.728967786
CHAD	4.805442333	1.5441E-06	6.6423E-05	6.748852253
PLK2	4.743179321	2.1039E-06	8.88815E-05	2.076183081
Myoepithelial cells				
Myoepithelial cells_names	**Myoepithelial cells_scores**	**Myoepithelial cells_pvals**	**Myoepithelial cells_pvals_adj**	**Myoepithelial cells_logfoldchanges**
MYLK	6.904055595	5.05385E-12	5.11517E-09	10.77928066
ACTA2	6.904055595	5.05385E-12	5.11517E-09	11.02142429

MYL9	6.904055595	5.05385E-12	5.11517E-09	9.941904068
CBR1	6.904055595	5.05385E-12	5.11517E-09	11.00211811
TPM2	6.903699875	5.06653E-12	5.11517E-09	10.10844803
CNN1	6.901921272	5.13039E-12	5.11517E-09	10.99533176
MGP	6.901565552	5.14325E-12	5.11517E-09	10.09555626
PCP4	6.898008347	5.27367E-12	5.11517E-09	10.03317738
TPM1	6.893383503	5.44807E-12	5.11517E-09	6.121230602
MYL6	6.885913372	5.7418E-12	5.11517E-09	3.118827581
TAGLN	6.881644726	5.91655E-12	5.11517E-09	8.90015316
MYH11	6.874885559	6.20396E-12	5.11517E-09	5.641300678
KRT23	6.855320454	7.11529E-12	5.41528E-09	7.262481213
LGALS1	6.719608307	1.82214E-11	1.12676E-08	4.174121857
CSRP1	6.584252357	4.57181E-11	2.38071E-08	3.751354694
KRT5	6.556149483	5.5215E-11	2.73148E-08	5.799221516
PI15	6.465259552	1.01125E-10	4.16888E-08	10.26639748
SPARC	6.417591572	1.38447E-10	5.2993E-08	7.298383236
SSPN	6.416702271	1.39258E-10	5.2993E-08	7.311267853
DHRS3	6.380417347	1.76606E-10	6.2405E-08	2.284596682
CALD1	6.171780586	6.75253E-10	2.02453E-07	3.584749222
ACTG2	6.03642416	1.57567E-09	4.33046E-07	10.88177967
CXCL14	6.02077198	1.73587E-09	4.64182E-07	9.189081192
CRISPLD1	5.988044739	2.12379E-09	5.12507E-07	6.617483139
VIM	5.965989113	2.43156E-09	5.51173E-07	7.141620159
PDLIM7	5.961720467	2.49596E-09	5.51173E-07	6.135601997
DCN	5.961009026	2.50685E-09	5.51173E-07	7.169484138
TIMP3	5.891107559	3.83616E-09	8.25108E-07	3.545794249
F3	5.870830536	4.33617E-09	9.1281E-07	3.34959054
COL17A1	5.864249706	4.51168E-09	9.2997E-07	5.510717869
ID4	5.67428875	1.39266E-08	2.46054E-06	2.48239851
IGFBP2	5.602430344	2.11367E-08	3.48544E-06	10.3977232
POSTN	5.575039387	2.47474E-08	3.94921E-06	8.164180756
CFL1	5.556540966	2.75173E-08	4.32153E-06	1.719209433
MAP1B	5.522035122	3.35096E-08	5.10067E-06	6.317865849
FBXO32	5.503892899	3.71496E-08	5.56906E-06	3.655197144
PERP	5.4866395	4.09652E-08	6.04939E-06	2.140425444
ISG12(A)	5.398240089	6.72978E-08	9.79182E-06	3.480771303
SERPINH1	5.388279438	7.11353E-08	1.02002E-05	3.658302546
SCRG1	5.375117302	7.65327E-08	1.0665E-05	3.591746092
PDLIM3	5.286895752	1.24409E-07	1.66339E-05	3.1183846
CAV2	5.217350006	1.81501E-07	2.37422E-05	3.882938623
CAV1	5.163812637	2.4197E-07	2.91749E-05	8.687297821
MSRB3	5.161677837	2.44746E-07	2.91749E-05	8.423083305
GPC3	5.14780426	2.63553E-07	3.06776E-05	7.711488247
ACTB	5.118100643	3.08628E-07	3.46996E-05	2.123898745
PIK3R1	5.080215454	3.77007E-07	3.9682E-05	3.458825588
SH3BGR	5.071321964	3.95062E-07	4.11446E-05	5.603885174

C3	5.056381226	4.27287E-07	4.31385E-05	5.761633873
FTL	5.02792263	4.95822E-07	4.90566E-05	1.135875463
FOS	5.01885128	5.19813E-07	5.035E-05	2.552000761
SFRP2	4.745648861	2.07839E-06	0.000188657	12.77942657
PGM5	4.745648861	2.07839E-06	0.000188657	11.98099136
VIT	4.743336678	2.10226E-06	0.000189089	11.01522541
CLIC4	4.740135193	2.13576E-06	0.000190371	3.010185719
ACTN4	4.730885983	2.23542E-06	0.000197475	2.571720123
LMOD1	4.727506638	2.27294E-06	0.000199013	8.152926445
COL4A2	4.724127293	2.31106E-06	0.000200576	8.148460388
APP	4.60033226	4.21818E-06	0.000344914	1.459274888
S100A11	4.593217373	4.36464E-06	0.000353965	2.16161871
CAMK2G	4.581834316	4.60915E-06	0.000370756	5.283492565
IGFBP5	4.574897289	4.76454E-06	0.000380164	5.481784344
CDC42EP3	4.417129993	1.0002E-05	0.000755419	4.091356754
IL17B	4.402900696	1.06813E-05	0.000800612	3.872057676
PRNP	4.386714935	1.15075E-05	0.000849602	2.020035744
TUBA1B	4.344027042	1.39894E-05	0.001010303	1.744088054
DYNC1I2	4.331398487	1.48165E-05	0.001062281	2.000421286
YBX1	4.31076622	1.6269E-05	0.001148968	1.270438075
ITGB1	4.310410023	1.62952E-05	0.001148968	2.101074219
FGL2	4.309342861	1.6374E-05	0.001148968	9.77157402
TINAGL1	4.287465572	1.80723E-05	0.001259209	3.538012981
ARHGEF25	4.284797668	1.82906E-05	0.001265503	7.726000309
BTG1	4.277504921	1.89E-05	0.00128963	1.926113844
HMGN2	4.252781391	2.11131E-05	0.001430777	1.364615202
EGR1	4.235706329	2.27834E-05	0.001512881	2.387117147
TPM4	4.231793404	2.31835E-05	0.001529186	1.623806119
SERPING1	4.225746155	2.38151E-05	0.001560438	5.80975771
YWHAB	4.117781162	3.82538E-05	0.002380395	1.616153717
PLIN2	4.086654663	4.37638E-05	0.002689435	3.031196833
CAVIN1	4.077761173	4.54714E-05	0.002777125	4.616788864
LEPROT	3.981002092	6.86254E-05	0.00397064	2.884827852
ANXA4	3.956456423	7.60697E-05	0.004375779	2.136193752
ARPC5	3.933867455	8.35899E-05	0.004753094	1.948257804
PPP1R14A	3.910211325	9.22154E-05	0.005154687	3.699189186
GPX1	3.907898903	9.31022E-05	0.005175019	2.216738939
TGFB1I1	3.856317759	0.000115108	0.006223371	7.758759975
RBPMS	3.851159573	0.00011756	0.006287227	3.684630156
CPM	3.84955883	0.000118331	0.006294432	7.104443073
POMP	3.813452005	0.000137039	0.00721205	1.149559855
NUCKS1	3.811317682	0.000138228	0.007236127	1.120139956
GOLIM4	3.809538841	0.000139226	0.007248464	3.312335968
LPP	3.808293819	0.000139929	0.007248464	3.294735432
EMP3	3.783214808	0.000154816	0.007977848	5.461044312
CNN2	3.780368805	0.000156596	0.008027785	3.059648991

SERPINB5	3.764894485	0.000166619	0.008410879	5.437126637
PSAP	3.698194742	0.000217138	0.010741833	1.378947735
AMIGO2	3.651060104	0.00026116	0.012604477	4.568192482
SDC2	3.603569746	0.000313876	0.01478806	4.21712923
GADD45B	3.588273287	0.000332875	0.015462286	3.385512114
RALBP1	3.578312874	0.000345819	0.015914124	2.123115301

Appendix Table 2: Biological process enriched gene ontologies

Myoepithelial cells					
native	**p_value**	**term_size**	**intersection_size**	**recall**	**name**
GO:0048856	1.38E-09	3254	67	0.02059	anatomical structure development
GO:0032502	1.38E-09	3495	70	0.020029	developmental process
GO:0009888	1.38E-09	1162	39	0.033563	tissue development
GO:0003012	1.04E-08	235	18	0.076596	muscle system process
GO:0032501	3.68E-08	3804	70	0.018402	multicellular organismal process
GO:0006936	7.8E-08	174	15	0.086207	muscle contraction
GO:0061061	1.08E-07	397	21	0.052897	muscle structure development
GO:0048869	2.88E-07	2379	52	0.021858	cellular developmental process
GO:0009653	2.88E-07	1567	41	0.026165	anatomical structure morphogenesis
GO:0030154	4.22E-07	2258	50	0.022143	cell differentiation
GO:0003008	6.59E-07	802	28	0.034913	system process
GO:0030029	1.34E-06	474	21	0.044304	actin filament-based process
GO:0048731	1.34E-06	2677	54	0.020172	system development
GO:0097435	1.34E-06	426	20	0.046948	supramolecular fiber organization
GO:0048513	4.22E-06	1964	44	0.022403	animal organ development
GO:0010466	4.29E-06	118	11	0.09322	negative regulation of peptidase activity
GO:0007275	6.08E-06	2983	56	0.018773	multicellular organism development
GO:0030198	6.08E-06	220	14	0.063636	extracellular matrix organization
GO:0007155	6.12E-06	737	25	0.033921	cell adhesion
GO:0006928	6.18E-06	1160	32	0.027586	movement of cell or subcellular component
GO:0022610	6.33E-06	742	25	0.033693	biological adhesion
GO:0060537	1.26E-05	238	14	0.058824	muscle tissue development
GO:0052547	1.34E-05	240	14	0.058333	regulation of peptidase activity
GO:0030049	1.34E-05	13	5	0.384615	muscle filament sliding
GO:0033275	1.34E-05	13	5	0.384615	actin-myosin filament sliding
GO:0043062	1.63E-05	247	14	0.05668	extracellular structure organization
GO:0040011	1.63E-05	1029	29	0.028183	locomotion
GO:0030036	2.58E-05	434	18	0.041475	actin cytoskeleton organization
GO:0048870	5.74E-05	910	26	0.028571	cell motility
GO:0051239	5.74E-05	1804	39	0.021619	regulation of multicellular organismal process
GO:0051674	5.74E-05	910	26	0.028571	localization of cell
GO:0030239	0.000105	35	6	0.171429	myofibril assembly
GO:0007167	0.000115	648	21	0.032407	enzyme linked receptor protein signaling pathway
GO:0016477	0.000145	838	24	0.02864	cell migration
GO:0009611	0.000145	396	16	0.040404	response to wounding

GO:0010951	0.000145	112	9	0.080357	negative regulation of endopeptidase activity
GO:0045861	0.000233	190	11	0.057895	negative regulation of proteolysis
GO:0050793	0.000252	1547	34	0.021978	regulation of developmental process
GO:0055002	0.000285	93	8	0.086022	striated muscle cell development
GO:0051336	0.000285	753	22	0.029216	regulation of hydrolase activity
GO:0030855	0.000285	372	15	0.040323	epithelial cell differentiation
GO:0051346	0.0003	239	12	0.050209	negative regulation of hydrolase activity
GO:1901652	0.0003	328	14	0.042683	response to peptide
GO:0030048	0.0003	67	7	0.104478	actin filament-based movement
GO:0072359	0.000359	651	20	0.030722	circulatory system development
GO:0042060	0.000367	335	14	0.041791	wound healing
GO:0031032	0.000389	131	9	0.068702	actomyosin structure organization
GO:0060429	0.000422	720	21	0.029167	epithelium development
GO:0055001	0.000441	101	8	0.079208	muscle cell development
GO:0070252	0.000585	50	6	0.12	actin-mediated cell contraction
GO:0007010	0.000627	866	23	0.026559	cytoskeleton organization
GO:0052548	0.000706	221	11	0.049774	regulation of endopeptidase activity
GO:0048468	0.000875	1223	28	0.022895	cell development
GO:0014706	0.000875	227	11	0.048458	striated muscle tissue development
GO:0001568	0.000997	423	15	0.035461	blood vessel development
GO:0050896	0.001028	4834	69	0.014274	response to stimulus
GO:0007166	0.001201	1690	34	0.020118	cell surface receptor signaling pathway
GO:0040012	0.001499	609	18	0.029557	regulation of locomotion
GO:0070887	0.001538	1947	37	0.019004	cellular response to chemical stimulus
GO:0001944	0.001569	443	15	0.03386	vasculature development
GO:0051093	0.001601	556	17	0.030576	negative regulation of developmental process
GO:0040017	0.001603	341	13	0.038123	positive regulation of locomotion
GO:0072358	0.001747	449	15	0.033408	cardiovascular system development
GO:0032989	0.001846	684	19	0.027778	cellular component morphogenesis
GO:2000026	0.001846	1215	27	0.022222	regulation of multicellular organismal development
GO:0051146	0.001928	167	9	0.053892	striated muscle cell differentiation
GO:2000145	0.001928	569	17	0.029877	regulation of cell motility
GO:0006955	0.001928	1081	25	0.023127	immune response
GO:0043086	0.002005	458	15	0.032751	negative regulation of catalytic activity

GO:0010927	0.002101	66	6	0.090909	cellular component assembly involved in morphogenesis
GO:1901700	0.002145	1022	24	0.023483	response to oxygen-containing compound
GO:0048584	0.002195	1379	29	0.02103	positive regulation of response to stimulus
GO:0045596	0.002225	410	14	0.034146	negative regulation of cell differentiation
GO:0009719	0.002333	1030	24	0.023301	response to endogenous stimulus
GO:1902905	0.002391	135	8	0.059259	positive regulation of supramolecular fiber organization
GO:0010243	0.002684	651	18	0.02765	response to organonitrogen compound
GO:0030334	0.002684	532	16	0.030075	regulation of cell migration
GO:0098609	0.002684	419	14	0.033413	cell-cell adhesion
GO:0044092	0.002749	653	18	0.027565	negative regulation of molecular function
GO:0071310	0.002907	1635	32	0.019572	cellular response to organic substance
GO:0042493	0.003005	598	17	0.028428	response to drug
GO:0007519	0.003082	105	7	0.066667	skeletal muscle tissue development
GO:2000147	0.00335	325	12	0.036923	positive regulation of cell motility
GO:0048646	0.00353	669	18	0.026906	anatomical structure formation involved in morphogenesis
GO:0060538	0.003555	108	7	0.064815	skeletal muscle organ development
GO:0051270	0.003687	611	17	0.027823	regulation of cellular component movement
GO:0043434	0.00382	281	11	0.039146	response to peptide hormone
GO:0051272	0.003825	332	12	0.036145	positive regulation of cellular component movement
GO:0007517	0.003825	234	10	0.042735	muscle organ development
GO:0030162	0.004547	447	14	0.03132	regulation of proteolysis
GO:0044794	0.005087	11	3	0.272727	positive regulation by host of viral process
GO:1901698	0.005132	694	18	0.025937	response to nitrogen compound
GO:0048523	0.005327	2872	46	0.016017	negative regulation of cellular process
GO:0006897	0.005341	346	12	0.034682	endocytosis
GO:0042221	0.005487	2528	42	0.016614	response to chemical
GO:0007178	0.005578	201	9	0.044776	transmembrane receptor protein serine/threonine kinase signaling pathway
GO:0002376	0.005675	1631	31	0.019007	immune system process
GO:0048583	0.005942	2454	41	0.016707	regulation of response to stimulus
GO:0050790	0.005956	1408	28	0.019886	regulation of catalytic activity
GO:0006939	0.005956	54	5	0.092593	smooth muscle contraction

GO:0002576	0.006354	86	6	0.069767	platelet degranulation
GO:1901342	0.006354	206	9	0.043689	regulation of vasculature development
GO:0001525	0.006816	306	11	0.035948	angiogenesis
GO:0007015	0.007397	259	10	0.03861	actin filament organization
GO:0098657	0.007397	418	13	0.0311	import into cell
GO:0019065	0.007544	3	2	0.666667	receptor-mediated endocytosis of virus by host cell
GO:0030335	0.007544	315	11	0.034921	positive regulation of cell migration
GO:0070836	0.007544	3	2	0.666667	caveola assembly
GO:0048514	0.007544	366	12	0.032787	blood vessel morphogenesis
GO:0050720	0.007544	3	2	0.666667	interleukin-1 beta biosynthetic process
GO:0075509	0.007544	3	2	0.666667	endocytosis involved in viral entry into host cell
GO:0090131	0.007544	3	2	0.666667	mesenchyme migration
GO:0045362	0.007544	3	2	0.666667	positive regulation of interleukin-1 biosynthetic process
GO:0010033	0.007544	1998	35	0.017518	response to organic substance
GO:0048729	0.007544	423	13	0.030733	tissue morphogenesis
GO:0036146	0.007544	3	2	0.666667	cellular response to mycotoxin
GO:0050725	0.007544	3	2	0.666667	positive regulation of interleukin-1 beta biosynthetic process
GO:0050722	0.007544	3	2	0.666667	regulation of interleukin-1 beta biosynthetic process
GO:0042222	0.007544	3	2	0.666667	interleukin-1 biosynthetic process
GO:0008285	0.007544	425	13	0.030588	negative regulation of cell population proliferation
GO:0045360	0.007544	3	2	0.666667	regulation of interleukin-1 biosynthetic process
GO:0071495	0.007694	867	20	0.023068	cellular response to endogenous stimulus
GO:0042127	0.008017	939	21	0.022364	regulation of cell population proliferation
GO:0090257	0.008476	132	7	0.05303	regulation of muscle system process
GO:0065009	0.008481	1770	32	0.018079	regulation of molecular function
GO:0071363	0.008481	432	13	0.030093	cellular response to growth factor stimulus
GO:0009887	0.008605	614	16	0.026059	animal organ morphogenesis
GO:0042692	0.009018	223	9	0.040359	muscle cell differentiation
GO:0009967	0.009018	1022	22	0.021526	positive regulation of signal transduction
GO:0009725	0.009018	618	16	0.02589	response to hormone
GO:0009628	0.009018	814	19	0.023342	response to abiotic stimulus
GO:0045765	0.010671	182	8	0.043956	regulation of angiogenesis
GO:0048518	0.010895	3644	53	0.014544	positive regulation of biological process

GO:0010647	0.010927	1110	23	0.020721	positive regulation of cell communication
GO:0071417	0.011647	392	12	0.030612	cellular response to organonitrogen compound
GO:0070848	0.011847	451	13	0.028825	response to growth factor
GO:0023056	0.01199	1119	23	0.020554	positive regulation of signaling
GO:0044854	0.012937	4	2	0.5	plasma membrane raft assembly
GO:0009629	0.012937	4	2	0.5	response to gravity
GO:0120036	0.013116	913	20	0.021906	plasma membrane bounded cell projection organization
GO:1902903	0.013602	239	9	0.037657	regulation of supramolecular fiber organization
GO:0050817	0.01414	192	8	0.041667	coagulation
GO:0007169	0.014424	463	13	0.028078	transmembrane receptor protein tyrosine kinase signaling pathway
GO:0048519	0.015022	3230	48	0.014861	negative regulation of biological process
GO:0035239	0.015112	527	14	0.026565	tube morphogenesis
GO:0001775	0.015388	789	18	0.022814	cell activation
GO:0001503	0.016044	246	9	0.036585	ossification
GO:1901701	0.016248	726	17	0.023416	cellular response to oxygen-containing compound
GO:0014070	0.01639	596	15	0.025168	response to organic cyclic compound
GO:1901653	0.016656	248	9	0.03629	cellular response to peptide
GO:0030030	0.016668	936	20	0.021368	cell projection organization
GO:0022603	0.016824	664	16	0.024096	regulation of anatomical structure morphogenesis
GO:0071709	0.016824	19	3	0.157895	membrane assembly
GO:0032269	0.016987	665	16	0.02406	negative regulation of cellular protein metabolic process
GO:0010812	0.017021	43	4	0.093023	negative regulation of cell-substrate adhesion
GO:0045595	0.017392	1086	22	0.020258	regulation of cell differentiation
GO:0016043	0.018053	3926	55	0.014009	cellular component organization
GO:0044788	0.018609	20	3	0.15	modulation by host of viral process
GO:0044857	0.018609	5	2	0.4	plasma membrane raft organization
GO:0010046	0.018609	5	2	0.4	response to mycotoxin
GO:0045080	0.018609	5	2	0.4	positive regulation of chemokine biosynthetic process
GO:1901699	0.019363	424	12	0.028302	cellular response to nitrogen compound
GO:1904018	0.019734	116	6	0.051724	positive regulation of vasculature development
GO:0007165	0.019864	3192	47	0.014724	signal transduction
GO:0050808	0.019869	207	8	0.038647	synapse organization

GO:0009605	0.021131	1489	27	0.018133	response to external stimulus
GO:0000902	0.021846	620	15	0.024194	cell morphogenesis
GO:0050673	0.021908	262	9	0.034351	epithelial cell proliferation
GO:0007568	0.021941	211	8	0.037915	aging
GO:0045214	0.023193	22	3	0.136364	sarcomere organization
GO:0034762	0.023193	265	9	0.033962	regulation of transmembrane transport
GO:0060416	0.023193	22	3	0.136364	response to growth hormone
GO:0034109	0.024868	49	4	0.081633	homotypic cell-cell adhesion
GO:0051702	0.024868	49	4	0.081633	interaction with symbiont
GO:0060056	0.025356	6	2	0.333333	mammary gland involution
GO:1905475	0.025444	124	6	0.048387	regulation of protein localization to membrane
GO:0007219	0.025444	124	6	0.048387	Notch signaling pathway
GO:0044091	0.025444	23	3	0.130435	membrane biogenesis
GO:0071711	0.025444	23	3	0.130435	basement membrane organization
GO:0000904	0.025874	444	12	0.027027	cell morphogenesis involved in differentiation
GO:0034097	0.025974	703	16	0.02276	response to cytokine
GO:0008283	0.026373	1134	22	0.0194	cell population proliferation
GO:0051248	0.026459	705	16	0.022695	negative regulation of protein metabolic process
GO:0031589	0.027069	221	8	0.036199	cell-substrate adhesion
GO:0061448	0.027082	172	7	0.040698	connective tissue development
GO:0014909	0.027082	51	4	0.078431	smooth muscle cell migration
GO:0048009	0.027634	24	3	0.125	insulin-like growth factor receptor signaling pathway
GO:0048522	0.029076	3263	47	0.014404	positive regulation of cellular process
GO:0051128	0.02908	1614	28	0.017348	regulation of cellular component organization
GO:0090288	0.029622	88	5	0.056818	negative regulation of cellular response to growth factor stimulus
GO:0035295	0.029631	648	15	0.023148	tube development
GO:0090647	0.031488	7	2	0.285714	modulation of age-related behavioral decline
GO:0045073	0.031488	7	2	0.285714	regulation of chemokine biosynthetic process
GO:0050678	0.031488	228	8	0.035088	regulation of epithelial cell proliferation
GO:0072203	0.031488	7	2	0.285714	cell proliferation involved in metanephros development
GO:0014733	0.031488	7	2	0.285714	regulation of skeletal muscle adaptation
GO:0090092	0.031545	132	6	0.045455	regulation of transmembrane receptor protein serine/threonine kinase signaling pathway
GO:0030216	0.031545	90	5	0.055556	keratinocyte differentiation

GO:0051897	0.035368	93	5	0.053763	positive regulation of protein kinase B signaling
GO:0030155	0.035368	405	11	0.02716	regulation of cell adhesion
GO:0008219	0.035368	1324	24	0.018127	cell death
GO:0007159	0.035368	183	7	0.038251	leukocyte cell-cell adhesion
GO:0050818	0.035368	56	4	0.071429	regulation of coagulation
GO:0048662	0.035474	27	3	0.111111	negative regulation of smooth muscle cell proliferation
GO:0010828	0.035474	27	3	0.111111	positive regulation of glucose transmembrane transport
GO:0001937	0.035474	27	3	0.111111	negative regulation of endothelial cell proliferation
GO:0010035	0.036194	348	10	0.028736	response to inorganic substance
GO:0071840	0.036679	4072	55	0.013507	cellular component organization or biogenesis
GO:0006878	0.037675	8	2	0.25	cellular copper ion homeostasis
GO:0009612	0.037675	139	6	0.043165	response to mechanical stimulus
GO:0010941	0.037675	1030	20	0.019417	regulation of cell death
GO:1990535	0.037675	8	2	0.25	neuron projection maintenance
GO:0050755	0.037675	8	2	0.25	chemokine metabolic process
GO:0016322	0.037675	8	2	0.25	neuron remodeling
GO:0042033	0.037675	8	2	0.25	chemokine biosynthetic process
GO:0002063	0.037675	28	3	0.107143	chondrocyte development
GO:0051716	0.037675	4082	55	0.013474	cellular response to stimulus
GO:0007596	0.039403	189	7	0.037037	blood coagulation
GO:0023052	0.040641	3427	48	0.014006	signaling
GO:0043207	0.04085	750	16	0.021333	response to external biotic stimulus
GO:0050680	0.04085	98	5	0.05102	negative regulation of epithelial cell proliferation
GO:0051707	0.04085	750	16	0.021333	response to other organism
GO:0051271	0.04085	191	7	0.036649	negative regulation of cellular component movement
GO:0051591	0.040889	60	4	0.066667	response to cAMP
GO:0007599	0.041681	192	7	0.036458	hemostasis
GO:0008150	0.042005	9139	98	0.010723	biological_process
GO:0090130	0.042524	193	7	0.036269	tissue migration
GO:0045055	0.043293	486	12	0.024691	regulated exocytosis
GO:0008544	0.043379	194	7	0.036082	epidermis development
GO:0007154	0.044175	3450	48	0.013913	cell communication
GO:0014812	0.044175	62	4	0.064516	muscle cell migration
GO:1900025	0.044175	9	2	0.222222	negative regulation of substrate adhesion-dependent cell spreading
GO:1900272	0.044175	9	2	0.222222	negative regulation of long-term synaptic potentiation

GO:0002861	0.044175	9	2	0.222222	regulation of inflammatory response to antigenic stimulus
GO:0072109	0.044175	9	2	0.222222	glomerular mesangium development
GO:0019221	0.044175	426	11	0.025822	cytokine-mediated signaling pathway
GO:0032060	0.044175	9	2	0.222222	bleb assembly
GO:0008625	0.045759	63	4	0.063492	extrinsic apoptotic signaling pathway via death domain receptors
GO:0031295	0.045759	31	3	0.096774	T cell costimulation
GO:0006950	0.047244	2381	36	0.01512	response to stress
GO:0006979	0.047244	311	9	0.028939	response to oxidative stress
GO:0016192	0.047244	1293	23	0.017788	vesicle-mediated transport
GO:0030850	0.047244	32	3	0.09375	prostate gland development
GO:0031294	0.047244	32	3	0.09375	lymphocyte costimulation
GO:0009607	0.047244	768	16	0.020833	response to biotic stimulus
GO:0051495	0.047244	149	6	0.040268	positive regulation of cytoskeleton organization
GO:1903076	0.047244	64	4	0.0625	regulation of protein localization to plasma membrane
GO:0045766	0.047244	104	5	0.048077	positive regulation of angiogenesis
GO:0032970	0.047244	254	8	0.031496	regulation of actin filament-based process
GO:0045321	0.047244	698	15	0.02149	leukocyte activation
GO:0070268	0.047244	32	3	0.09375	cornification
GO:0032722	0.047244	32	3	0.09375	positive regulation of chemokine production
GO:0035690	0.047244	253	8	0.031621	cellular response to drug
Mucinous cells					
native	**p_value**	**term_size**	**intersection_size**	**recall**	**name**
GO:0072599	0.000212	69	8	0.115942	establishment of protein localization to endoplasmic reticulum
GO:0070972	0.000212	90	9	0.1	protein localization to endoplasmic reticulum
GO:0045047	0.000212	66	8	0.121212	protein targeting to ER
GO:0034976	0.000212	222	13	0.058559	response to endoplasmic reticulum stress
GO:0006613	0.000794	59	7	0.118644	cotranslational protein targeting to membrane
GO:0030968	0.002815	102	8	0.078431	endoplasmic reticulum unfolded protein response
GO:0006986	0.004872	147	9	0.061224	response to unfolded protein
GO:0034620	0.00617	118	8	0.067797	cellular response to unfolded protein
GO:0035966	0.00995	166	9	0.054217	response to topologically incorrect protein
GO:0035967	0.012275	134	8	0.059701	cellular response to topologically incorrect protein

GO:1905897	0.012275	69	6	0.086957	regulation of response to endoplasmic reticulum stress
GO:0070059	0.013385	45	5	0.111111	intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress
GO:0006810	0.013385	3026	49	0.016193	transport
GO:0042351	0.013917	2	2	1	'de novo' GDP-L-fucose biosynthetic process
GO:0042350	0.013917	2	2	1	GDP-L-fucose biosynthetic process
GO:0045454	0.01876	51	5	0.098039	cell redox homeostasis
GO:0051234	0.01876	3100	49	0.015806	establishment of localization
GO:0002067	0.025814	31	4	0.129032	glandular epithelial cell differentiation
GO:0006612	0.025814	123	7	0.056911	protein targeting to membrane
GO:0006614	0.025814	56	5	0.089286	SRP-dependent cotranslational protein targeting to membrane
GO:0060482	0.025923	3	2	0.666667	lobar bronchus development
GO:0060481	0.025923	3	2	0.666667	lobar bronchus epithelium development
GO:0046368	0.025923	3	2	0.666667	GDP-L-fucose metabolic process
GO:0060480	0.025923	3	2	0.666667	lung goblet cell differentiation
GO:1902236	0.040434	16	3	0.1875	negative regulation of endoplasmic reticulum stress-induced intrinsic apoptotic signaling pathway
GO:0009226	0.040434	16	3	0.1875	nucleotide-sugar biosynthetic process
GO:0006890	0.044146	68	5	0.073529	retrograde vesicle-mediated transport, Golgi to endoplasmic reticulum
GO:0031584	0.044146	4	2	0.5	activation of phospholipase D activity
GO:0006984	0.048235	40	4	0.1	ER-nucleus signaling pathway
Acinar cells					
native	**p_value**	**term_size**	**intersection_size**	**recall**	**name**
GO:0006119	4.57E-24	87	23	0.264368	oxidative phosphorylation
GO:1902600	1.34E-19	86	20	0.232558	proton transmembrane transport
GO:0046034	1.28E-18	194	25	0.128866	ATP metabolic process
GO:0006091	1.61E-16	356	29	0.081461	generation of precursor metabolites and energy
GO:0022904	5.9E-15	75	16	0.213333	respiratory electron transport chain
GO:0022900	4.16E-14	122	18	0.147541	electron transport chain
GO:0015672	4.25E-14	247	23	0.093117	monovalent inorganic cation transport
GO:0098660	4.25E-14	411	28	0.068127	inorganic ion transmembrane transport
GO:0042773	7.15E-14	59	14	0.237288	ATP synthesis coupled electron transport

GO:0042775	7.15E-14	59	14	0.237288	mitochondrial ATP synthesis coupled electron transport
GO:0098662	3.5E-13	377	26	0.068966	inorganic cation transmembrane transport
GO:0045333	7.48E-13	126	17	0.134921	cellular respiration
GO:0098655	6.15E-12	428	26	0.060748	cation transmembrane transport
GO:0006812	1.93E-11	578	29	0.050173	cation transport
GO:0015980	3.86E-11	188	18	0.095745	energy derivation by oxidation of organic compounds
GO:0034220	5.69E-11	561	28	0.049911	ion transmembrane transport
GO:0055114	4.36E-10	659	29	0.044006	oxidation-reduction process
GO:0042776	1.59E-09	13	7	0.538462	mitochondrial ATP synthesis coupled proton transport
GO:0015985	2.84E-09	14	7	0.5	energy coupled proton transport, down electrochemical gradient
GO:0015986	2.84E-09	14	7	0.5	ATP synthesis coupled proton transport
GO:0006123	7.03E-09	9	6	0.666667	mitochondrial electron transport, cytochrome c to oxygen
GO:0019646	7.03E-09	9	6	0.666667	aerobic electron transport chain
GO:0006811	1.29E-08	820	30	0.036585	ion transport
GO:0055085	1.98E-08	836	30	0.035885	transmembrane transport
GO:0006754	2.33E-08	29	8	0.275862	ATP biosynthetic process
GO:0009060	2.55E-08	60	10	0.166667	aerobic respiration
GO:0007005	2.85E-08	366	20	0.054645	mitochondrion organization
GO:0042407	4.37E-08	20	7	0.35	cristae formation
GO:0007007	4.81E-08	32	8	0.25	inner mitochondrial membrane organization
GO:0009142	8.77E-08	50	9	0.18	nucleoside triphosphate biosynthetic process
GO:0009206	1.59E-07	37	8	0.216216	purine ribonucleoside triphosphate biosynthetic process
GO:0000184	1.76E-07	74	10	0.135135	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
GO:0009145	1.87E-07	38	8	0.210526	purine nucleoside triphosphate biosynthetic process
GO:0006614	2.23E-07	56	9	0.160714	SRP-dependent cotranslational protein targeting to membrane
GO:0006810	2.42E-07	3026	58	0.019167	transport
GO:0009201	3.27E-07	41	8	0.195122	ribonucleoside triphosphate biosynthetic process
GO:0006613	3.31E-07	59	9	0.152542	cotranslational protein targeting to membrane
GO:0051234	6.02E-07	3100	58	0.01871	establishment of localization
GO:0045047	8.75E-07	66	9	0.136364	protein targeting to ER
GO:0070972	9.84E-07	90	10	0.111111	protein localization to endoplasmic reticulum

GO:0009205	1.07E-06	48	8	0.166667	purine ribonucleoside triphosphate metabolic process
GO:0072599	1.21E-06	69	9	0.130435	establishment of protein localization to endoplasmic reticulum
GO:0009141	1.53E-06	71	9	0.126761	nucleoside triphosphate metabolic process
GO:0009199	1.93E-06	52	8	0.153846	ribonucleoside triphosphate metabolic process
GO:0006413	2.45E-06	128	11	0.085938	translational initiation
GO:0009144	2.51E-06	54	8	0.148148	purine nucleoside triphosphate metabolic process
GO:0006120	5.33E-06	40	7	0.175	mitochondrial electron transport, NADH to ubiquinone
GO:1901566	5.93E-06	1163	31	0.026655	organonitrogen compound biosynthetic process
GO:0007006	1.32E-05	92	9	0.097826	mitochondrial membrane organization
GO:0032981	1.82E-05	48	7	0.145833	mitochondrial respiratory chain complex I assembly
GO:0010257	1.82E-05	48	7	0.145833	NADH dehydrogenase complex assembly
GO:0000956	4.73E-05	139	10	0.071942	nuclear-transcribed mRNA catabolic process
GO:0090150	4.83E-05	212	12	0.056604	establishment of protein localization to membrane
GO:0033108	9.16E-05	61	7	0.114754	mitochondrial respiratory chain complex assembly
GO:0019083	9.82E-05	118	9	0.076271	viral transcription
GO:0006612	0.000136	123	9	0.073171	protein targeting to membrane
GO:0015988	0.000151	4	3	0.75	energy coupled proton transmembrane transport, against electrochemical gradient
GO:0015990	0.000151	4	3	0.75	electron transport coupled proton transport
GO:0046688	0.000155	25	5	0.2	response to copper ion
GO:1990542	0.000173	68	7	0.102941	mitochondrial transmembrane transport
GO:0043603	0.000207	719	21	0.029207	cellular amide metabolic process
GO:0019080	0.000207	131	9	0.068702	viral gene expression
GO:0016310	0.000212	1532	33	0.02154	phosphorylation
GO:0051179	0.000266	3875	60	0.015484	localization
GO:0006402	0.000269	255	12	0.047059	mRNA catabolic process
GO:0072657	0.000334	405	15	0.037037	protein localization to membrane
GO:1903018	0.000351	30	5	0.166667	regulation of glycoprotein metabolic process
GO:0006979	0.000377	311	13	0.041801	response to oxidative stress
GO:0006839	0.000403	182	10	0.054945	mitochondrial transport
GO:0009152	0.000412	110	8	0.072727	purine ribonucleotide biosynthetic process

GO:0017144	0.000617	327	13	0.039755	drug metabolic process
GO:0006401	0.000646	281	12	0.042705	RNA catabolic process
GO:0006164	0.000694	119	8	0.067227	purine nucleotide biosynthetic process
GO:0009260	0.000694	119	8	0.067227	ribonucleotide biosynthetic process
GO:0046390	0.000871	123	8	0.065041	ribose phosphate biosynthetic process
GO:0072594	0.000915	393	14	0.035623	establishment of protein localization to organelle
GO:0072522	0.001069	127	8	0.062992	purine-containing compound biosynthetic process
GO:0010035	0.001079	348	13	0.037356	response to inorganic substance
GO:0006605	0.001198	302	12	0.039735	protein targeting
GO:0009636	0.001303	305	12	0.039344	response to toxic substance
GO:0006796	0.001371	2092	38	0.018164	phosphate-containing compound metabolic process
GO:0009150	0.001453	261	11	0.042146	purine ribonucleotide metabolic process
GO:0006793	0.001593	2108	38	0.018027	phosphorus metabolic process
GO:1901137	0.001668	419	14	0.033413	carbohydrate derivative biosynthetic process
GO:0009259	0.001969	271	11	0.04059	ribonucleotide metabolic process
GO:1901293	0.002244	184	9	0.048913	nucleoside phosphate biosynthetic process
GO:0019693	0.002342	277	11	0.039711	ribose phosphate metabolic process
GO:0006163	0.002391	278	11	0.039568	purine nucleotide metabolic process
GO:0097213	0.002528	2	2	1	regulation of lysosomal membrane permeability
GO:0097212	0.002528	2	2	1	lysosomal membrane organization
GO:0010559	0.00294	26	4	0.153846	regulation of glycoprotein biosynthetic process
GO:0044237	0.003488	6380	80	0.012539	cellular metabolic process
GO:0046907	0.003488	1247	26	0.02085	intracellular transport
GO:0072521	0.00402	297	11	0.037037	purine-containing compound metabolic process
GO:0046677	0.004139	202	9	0.044554	response to antibiotic
GO:0019439	0.004295	407	13	0.031941	aromatic compound catabolic process
GO:0051186	0.004462	302	11	0.036424	cofactor metabolic process
GO:1901135	0.004462	712	18	0.025281	carbohydrate derivative metabolic process
GO:0061024	0.00448	526	15	0.028517	membrane organization
GO:0051649	0.005289	1513	29	0.019167	establishment of localization in cell
GO:1901361	0.00556	420	13	0.030952	organic cyclic compound catabolic process
GO:0055075	0.005753	13	3	0.230769	potassium ion homeostasis

GO:0034655	0.005988	368	12	0.032609	nucleobase-containing compound catabolic process
GO:0006518	0.006457	547	15	0.027422	peptide metabolic process
GO:0044281	0.006457	1230	25	0.020325	small molecule metabolic process
GO:0042743	0.007368	34	4	0.117647	hydrogen peroxide metabolic process
GO:0098869	0.007774	62	5	0.080645	cellular oxidant detoxification
GO:0010038	0.007861	224	9	0.040179	response to metal ion
GO:0055078	0.008442	15	3	0.2	sodium ion homeostasis
GO:0009165	0.00881	181	8	0.044199	nucleotide biosynthetic process
GO:0046700	0.009638	391	12	0.030691	heterocycle catabolic process
GO:0044270	0.01023	394	12	0.030457	cellular nitrogen compound catabolic process
GO:1990748	0.011316	68	5	0.073529	cellular detoxification
GO:0042744	0.011488	17	3	0.176471	hydrogen peroxide catabolic process
GO:0006883	0.011488	4	2	0.5	cellular sodium ion homeostasis
GO:0030007	0.011488	4	2	0.5	cellular potassium ion homeostasis
GO:0015705	0.011488	4	2	0.5	iodide transport
GO:0045861	0.011488	190	8	0.042105	negative regulation of proteolysis
GO:0008152	0.01255	6706	81	0.012079	metabolic process
GO:0097237	0.014139	151	7	0.046358	cellular response to toxic substance
GO:0009117	0.014621	356	11	0.030899	nucleotide metabolic process
GO:0033365	0.014632	665	16	0.02406	protein localization to organelle
GO:1990573	0.015481	19	3	0.157895	potassium ion import across plasma membrane
GO:0010951	0.015527	112	6	0.053571	negative regulation of endopeptidase activity
GO:0006753	0.015527	360	11	0.030556	nucleoside phosphate metabolic process
GO:0044248	0.015851	1556	28	0.017995	cellular catabolic process
GO:1901700	0.015888	1022	21	0.020548	response to oxygen-containing compound
GO:0000056	0.017162	5	2	0.4	ribosomal small subunit export from nucleus
GO:0036376	0.017162	5	2	0.4	sodium ion export across plasma membrane
GO:0010248	0.017162	5	2	0.4	establishment or maintenance of transmembrane electrochemical gradient
GO:0098754	0.017251	77	5	0.064935	detoxification
GO:0043604	0.017685	551	14	0.025408	amide biosynthetic process
GO:0010466	0.019168	118	6	0.050847	negative regulation of peptidase activity

GO:0032414	0.023777	49	4	0.081633	positive regulation of ion transmembrane transporter activity
GO:0009056	0.024111	1768	30	0.016968	catabolic process
GO:0050821	0.030854	130	6	0.046154	protein stabilization
GO:1901575	0.030961	1472	26	0.017663	organic substance catabolic process
GO:0042273	0.030961	53	4	0.075472	ribosomal large subunit biogenesis
GO:0098661	0.032949	54	4	0.074074	inorganic anion transmembrane transport
GO:0010940	0.033035	7	2	0.285714	positive regulation of necrotic cell death
GO:0051641	0.033512	1977	32	0.016186	cellular localization
GO:0032411	0.034521	55	4	0.072727	positive regulation of transporter activity
GO:0042493	0.035458	598	14	0.023411	response to drug
GO:0055086	0.038191	410	11	0.026829	nucleobase-containing small molecule metabolic process
GO:0090559	0.040846	58	4	0.068966	regulation of membrane permeability
GO:0086013	0.040846	8	2	0.25	membrane repolarization during cardiac muscle cell action potential
GO:0033750	0.040846	8	2	0.25	ribosome localization
GO:0071428	0.040846	8	2	0.25	rRNA-containing ribonucleoprotein complex export from nucleus
GO:1904667	0.040846	8	2	0.25	negative regulation of ubiquitin protein ligase activity
GO:0000054	0.040846	8	2	0.25	ribosomal subunit export from nucleus
GO:0019725	0.041433	545	13	0.023853	cellular homeostasis
GO:0017001	0.043976	29	3	0.103448	antibiotic catabolic process
Basal duct cells					
native	**p_value**	**term_size**	**intersection_size**	**recall**	**name**
GO:0042221	7.2E-10	2528	60	0.023734	response to chemical
GO:0070887	7.2E-10	1947	52	0.026708	cellular response to chemical stimulus
GO:0010033	1E-07	1998	49	0.024525	response to organic substance
GO:0071310	2.08E-07	1635	43	0.0263	cellular response to organic substance
GO:0009611	2.08E-07	396	21	0.05303	response to wounding
GO:0032502	2.29E-06	3495	64	0.018312	developmental process
GO:0042060	2.51E-06	335	18	0.053731	wound healing
GO:0050896	2.58E-06	4834	77	0.015929	response to stimulus
GO:0040011	4.41E-06	1029	31	0.030126	locomotion
GO:0009888	4.88E-06	1162	33	0.028399	tissue development
GO:0034097	5.65E-06	703	25	0.035562	response to cytokine
GO:0030154	6.8E-06	2258	48	0.021258	cell differentiation
GO:0016477	9.01E-06	838	27	0.03222	cell migration

GO:0048869	1.02E-05	2379	49	0.020597	cellular developmental process
GO:0048870	1.02E-05	910	28	0.030769	cell motility
GO:0051674	1.02E-05	910	28	0.030769	localization of cell
GO:0071345	1.17E-05	639	23	0.035994	cellular response to cytokine stimulus
GO:0048856	1.17E-05	3254	59	0.018132	anatomical structure development
GO:0008544	1.18E-05	194	13	0.06701	epidermis development
GO:0012501	1.38E-05	1247	33	0.026464	programmed cell death
GO:0008219	1.47E-05	1324	34	0.02568	cell death
GO:0042981	1.47E-05	939	28	0.029819	regulation of apoptotic process
GO:0043588	1.47E-05	166	12	0.072289	skin development
GO:0043067	1.94E-05	956	28	0.029289	regulation of programmed cell death
GO:0030855	2.11E-05	372	17	0.045699	epithelial cell differentiation
GO:0032501	2.24E-05	3804	64	0.016824	multicellular organismal process
GO:0009719	2.25E-05	1030	29	0.028155	response to endogenous stimulus
GO:0006928	2.3E-05	1160	31	0.026724	movement of cell or subcellular component
GO:0022610	2.64E-05	742	24	0.032345	biological adhesion
GO:0006915	4.11E-05	1194	31	0.025963	apoptotic process
GO:0006950	4.51E-05	2381	47	0.01974	response to stress
GO:0051716	4.59E-05	4082	66	0.016169	cellular response to stimulus
GO:0009913	4.83E-05	125	10	0.08	epidermal cell differentiation
GO:1901700	5.5E-05	1022	28	0.027397	response to oxygen-containing compound
GO:0010941	6.26E-05	1030	28	0.027184	regulation of cell death
GO:0007155	7.58E-05	737	23	0.031208	cell adhesion
GO:1903034	0.000149	111	9	0.081081	regulation of response to wounding
GO:0060429	0.000187	720	22	0.030556	epithelium development
GO:0048523	0.000238	2872	51	0.017758	negative regulation of cellular process
GO:0030216	0.000269	90	8	0.088889	keratinocyte differentiation
GO:0007275	0.000295	2983	52	0.017432	multicellular organism development
GO:0042127	0.00036	939	25	0.026624	regulation of cell population proliferation
GO:0045446	0.000438	69	7	0.101449	endothelial cell differentiation
GO:0071363	0.000458	432	16	0.037037	cellular response to growth factor stimulus
GO:0030036	0.000475	434	16	0.036866	actin cytoskeleton organization
GO:0043066	0.000504	543	18	0.033149	negative regulation of apoptotic process
GO:0048646	0.000653	669	20	0.029895	anatomical structure formation involved in morphogenesis
GO:0060548	0.000653	611	19	0.031097	negative regulation of cell death
GO:0043069	0.000656	556	18	0.032374	negative regulation of programmed cell death

GO:0070848	0.000694	451	16	0.035477	response to growth factor
GO:0050793	0.000757	1547	33	0.021332	regulation of developmental process
GO:0003158	0.000824	78	7	0.089744	endothelium development
GO:0052548	0.000832	221	11	0.049774	regulation of endopeptidase activity
GO:0051094	0.000832	807	22	0.027261	positive regulation of developmental process
GO:0006979	0.000838	311	13	0.041801	response to oxidative stress
GO:0009653	0.000908	1567	33	0.021059	anatomical structure morphogenesis
GO:0051291	0.00094	54	6	0.111111	protein heterooligomerization
GO:0010951	0.000947	112	8	0.071429	negative regulation of endopeptidase activity
GO:0080090	0.000959	3494	56	0.016027	regulation of primary metabolic process
GO:0098609	0.001015	419	15	0.0358	cell-cell adhesion
GO:0048519	0.001015	3230	53	0.016409	negative regulation of biological process
GO:0048468	0.001022	1223	28	0.022895	cell development
GO:0030029	0.001023	474	16	0.033755	actin filament-based process
GO:0030334	0.00109	532	17	0.031955	regulation of cell migration
GO:0065007	0.001122	6460	84	0.013003	biological regulation
GO:0097435	0.001122	426	15	0.035211	supramolecular fiber organization
GO:0009987	0.001163	8473	98	0.011566	cellular process
GO:0050794	0.001163	5770	78	0.013518	regulation of cellular process
GO:0080134	0.001163	966	24	0.024845	regulation of response to stress
GO:0051384	0.001163	86	7	0.081395	response to glucocorticoid
GO:0010466	0.001163	118	8	0.067797	negative regulation of peptidase activity
GO:0052547	0.001332	240	11	0.045833	regulation of peptidase activity
GO:0040012	0.001505	609	18	0.029557	regulation of locomotion
GO:0048731	0.00153	2677	46	0.017183	system development
GO:0051270	0.00153	611	18	0.02946	regulation of cellular component movement
GO:0061041	0.001645	92	7	0.076087	regulation of wound healing
GO:0031960	0.001645	92	7	0.076087	response to corticosteroid
GO:0008283	0.001691	1134	26	0.022928	cell population proliferation
GO:0007010	0.001691	866	22	0.025404	cytoskeleton organization
GO:0034599	0.001691	205	10	0.04878	cellular response to oxidative stress
GO:0001667	0.001921	253	11	0.043478	ameboidal-type cell migration
GO:2000145	0.002001	569	17	0.029877	regulation of cell motility
GO:0034330	0.002056	211	10	0.047393	cell junction organization
GO:0001885	0.002138	41	5	0.121951	endothelial cell development
GO:0031323	0.002243	3626	56	0.015444	regulation of cellular metabolic process
GO:0019222	0.00227	3912	59	0.015082	regulation of metabolic process
GO:0007154	0.002381	3450	54	0.015652	cell communication

GO:0050789	0.002425	6114	80	0.013085	regulation of biological process
GO:0051171	0.002702	3375	53	0.015704	regulation of nitrogen compound metabolic process
GO:0007166	0.00272	1690	33	0.019527	cell surface receptor signaling pathway
GO:0048608	0.00288	268	11	0.041045	reproductive structure development
GO:0060255	0.003079	3578	55	0.015372	regulation of macromolecule metabolic process
GO:0014070	0.003107	596	17	0.028523	response to organic cyclic compound
GO:0061458	0.003107	271	11	0.04059	reproductive system development
GO:0019221	0.003172	426	14	0.032864	cytokine-mediated signaling pathway
GO:0001890	0.003272	106	7	0.066038	placenta development
GO:0051239	0.003852	1804	34	0.018847	regulation of multicellular organismal process
GO:0048513	0.003881	1964	36	0.01833	animal organ development
GO:0071495	0.004147	867	21	0.024221	cellular response to endogenous stimulus
GO:0045861	0.004147	190	9	0.047368	negative regulation of proteolysis
GO:1903201	0.004245	49	5	0.102041	regulation of oxidative stress-induced cell death
GO:2000026	0.004257	1215	26	0.021399	regulation of multicellular organismal development
GO:0051346	0.004681	239	10	0.041841	negative regulation of hydrolase activity
GO:0016043	0.004712	3926	58	0.014773	cellular component organization
GO:1901698	0.005636	694	18	0.025937	response to nitrogen compound
GO:0033993	0.006316	577	16	0.02773	response to lipid
GO:0051241	0.007471	711	18	0.025316	negative regulation of multicellular organismal process
GO:0031175	0.007499	587	16	0.027257	neuron projection development
GO:0035239	0.007499	527	15	0.028463	tube morphogenesis
GO:0001525	0.007637	306	11	0.035948	angiogenesis
GO:0010243	0.007637	651	17	0.026114	response to organonitrogen compound
GO:0044092	0.007747	653	17	0.026034	negative regulation of molecular function
GO:1900407	0.007747	57	5	0.087719	regulation of cellular response to oxidative stress
GO:0045682	0.007747	57	5	0.087719	regulation of epidermis development
GO:0007163	0.007748	166	8	0.048193	establishment or maintenance of cell polarity
GO:0051098	0.007912	259	10	0.03861	regulation of binding
GO:0007015	0.007912	259	10	0.03861	actin filament organization

GO:0000380	0.008107	58	5	0.086207	alternative mRNA splicing, via spliceosome
GO:0048522	0.008107	3263	50	0.015323	positive regulation of cellular process
GO:0048514	0.008357	366	12	0.032787	blood vessel morphogenesis
GO:0043065	0.008357	421	13	0.030879	positive regulation of apoptotic process
GO:0008065	0.008357	3	2	0.666667	establishment of blood-nerve barrier
GO:0043068	0.008505	424	13	0.03066	positive regulation of programmed cell death
GO:0001568	0.008505	423	13	0.030733	blood vessel development
GO:0031324	0.008505	1586	30	0.018916	negative regulation of cellular metabolic process
GO:0001889	0.008505	92	6	0.065217	liver development
GO:0032269	0.008505	665	17	0.025564	negative regulation of cellular protein metabolic process
GO:0022603	0.008505	664	17	0.025602	regulation of anatomical structure morphogenesis
GO:0051099	0.008669	130	7	0.053846	positive regulation of binding
GO:0031032	0.009014	131	7	0.053435	actomyosin structure organization
GO:0061008	0.009245	94	6	0.06383	hepaticobiliary system development
GO:0036473	0.009245	61	5	0.081967	cell death in response to oxidative stress
GO:0006357	0.010222	1379	27	0.019579	regulation of transcription by RNA polymerase II
GO:0045595	0.010222	1086	23	0.021179	regulation of cell differentiation
GO:0009725	0.010501	618	16	0.02589	response to hormone
GO:0033043	0.010501	880	20	0.022727	regulation of organelle organization
GO:0051128	0.010501	1614	30	0.018587	regulation of cellular component organization
GO:0065009	0.010501	1770	32	0.018079	regulation of molecular function
GO:0051130	0.010836	817	19	0.023256	positive regulation of cellular component organization
GO:0051172	0.010836	1466	28	0.0191	negative regulation of nitrogen compound metabolic process
GO:0071840	0.010836	4072	58	0.014244	cellular component organization or biogenesis
GO:1902882	0.010836	64	5	0.078125	regulation of response to oxidative stress
GO:0034329	0.010836	180	8	0.044444	cell junction assembly
GO:0045765	0.011337	182	8	0.043956	regulation of angiogenesis
GO:0001944	0.011337	443	13	0.029345	vasculature development
GO:0003006	0.011337	386	12	0.031088	developmental process involved in reproduction
GO:0023052	0.011733	3427	51	0.014882	signaling
GO:0032956	0.011981	231	9	0.038961	regulation of actin cytoskeleton organization
GO:0030162	0.01204	447	13	0.029083	regulation of proteolysis

GO:1904019	0.012459	67	5	0.074627	epithelial cell apoptotic process
GO:0072358	0.012459	449	13	0.028953	cardiovascular system development
GO:0032268	0.012946	1643	30	0.018259	regulation of cellular protein metabolic process
GO:1904798	0.013167	4	2	0.5	positive regulation of core promoter binding
GO:0010605	0.013184	1725	31	0.017971	negative regulation of macromolecule metabolic process
GO:0048511	0.01324	188	8	0.042553	rhythmic process
GO:0051248	0.013508	705	17	0.024113	negative regulation of protein metabolic process
GO:0007596	0.013508	189	8	0.042328	blood coagulation
GO:0061061	0.013508	397	12	0.030227	muscle structure development
GO:0048518	0.013757	3644	53	0.014544	positive regulation of biological process
GO:0048666	0.013984	643	16	0.024883	neuron development
GO:0043086	0.014005	458	13	0.028384	negative regulation of catalytic activity
GO:0007599	0.014116	192	8	0.041667	hemostasis
GO:0010942	0.014116	460	13	0.028261	positive regulation of cell death
GO:2000116	0.014116	147	7	0.047619	regulation of cysteine-type endopeptidase activity
GO:0007165	0.014116	3192	48	0.015038	signal transduction
GO:0050817	0.014116	192	8	0.041667	coagulation
GO:0010468	0.014116	2569	41	0.01596	regulation of gene expression
GO:0009651	0.014472	18	3	0.166667	response to salt stress
GO:0001709	0.014472	18	3	0.166667	cell fate determination
GO:0035295	0.014472	648	16	0.024691	tube development
GO:0071353	0.01696	19	3	0.157895	cellular response to interleukin-4
GO:1901701	0.017265	726	17	0.023416	cellular response to oxygen-containing compound
GO:0042493	0.017849	598	15	0.025084	response to drug
GO:0097190	0.017975	415	12	0.028916	apoptotic signaling pathway
GO:1904796	0.018951	5	2	0.4	regulation of core promoter binding
GO:1902177	0.018951	5	2	0.4	positive regulation of oxidative stress-induced intrinsic apoptotic signaling pathway
GO:0051017	0.019208	114	6	0.052632	actin filament bundle assembly
GO:0032970	0.019624	254	9	0.035433	regulation of actin filament-based process
GO:0061572	0.019758	115	6	0.052174	actin filament bundle organization
GO:0031099	0.019758	115	6	0.052174	regeneration
GO:1903036	0.019758	45	4	0.088889	positive regulation of response to wounding
GO:0048545	0.020281	256	9	0.035156	response to steroid hormone
GO:1901342	0.02036	206	8	0.038835	regulation of vasculature development

GO:0022604	0.020451	310	10	0.032258	regulation of cell morphogenesis
GO:0009892	0.020567	1870	32	0.017112	negative regulation of metabolic process
GO:0008285	0.020664	425	12	0.028235	negative regulation of cell population proliferation
GO:1902175	0.020866	21	3	0.142857	regulation of oxidative stress-induced intrinsic apoptotic signaling pathway
GO:0070670	0.020866	21	3	0.142857	response to interleukin-4
GO:0030336	0.021183	161	7	0.043478	negative regulation of cell migration
GO:0051493	0.02123	369	11	0.02981	regulation of cytoskeleton organization
GO:0032989	0.022648	684	16	0.023392	cellular component morphogenesis
GO:0050790	0.022825	1408	26	0.018466	regulation of catalytic activity
GO:0051336	0.023027	753	17	0.022576	regulation of hydrolase activity
GO:0045214	0.023243	22	3	0.136364	sarcomere organization
GO:0006366	0.023741	1491	27	0.018109	transcription by RNA polymerase II
GO:1905049	0.02523	6	2	0.333333	negative regulation of metallopeptidase activity
GO:0090594	0.02523	6	2	0.333333	inflammatory response to wounding
GO:0000003	0.025415	762	17	0.02231	reproduction
GO:0022414	0.025415	762	17	0.02231	reproductive process
GO:0010837	0.025739	23	3	0.130435	regulation of keratinocyte proliferation
GO:2000146	0.026365	169	7	0.04142	negative regulation of cell motility
GO:0071396	0.026783	383	11	0.028721	cellular response to lipid
GO:0071407	0.026783	383	11	0.028721	cellular response to organic cyclic compound
GO:0051246	0.027623	1750	30	0.017143	regulation of protein metabolic process
GO:1903203	0.027944	24	3	0.125	regulation of oxidative stress-induced neuron death
GO:1903707	0.027944	86	5	0.05814	negative regulation of hemopoiesis
GO:0010638	0.027944	446	12	0.026906	positive regulation of organelle organization
GO:0051173	0.027944	1998	33	0.016517	positive regulation of nitrogen compound metabolic process
GO:0051179	0.027944	3875	54	0.013935	localization
GO:0120036	0.028034	913	19	0.020811	plasma membrane bounded cell projection organization
GO:0031325	0.029026	2089	34	0.016276	positive regulation of cellular metabolic process
GO:0050878	0.029337	276	9	0.032609	regulation of body fluid levels
GO:0031330	0.0294	174	7	0.04023	negative regulation of cellular catabolic process
GO:0032879	0.029647	1602	28	0.017478	regulation of localization

GO:0030182	0.029832	779	17	0.021823	neuron differentiation
GO:1901890	0.030039	25	3	0.12	positive regulation of cell junction assembly
GO:0043154	0.030039	53	4	0.075472	negative regulation of cysteine-type endopeptidase activity involved in apoptotic process
GO:1903393	0.030039	25	3	0.12	positive regulation of adherens junction organization
GO:0045616	0.030039	25	3	0.12	regulation of keratinocyte differentiation
GO:0036475	0.030039	25	3	0.12	neuron death in response to oxidative stress
GO:0110053	0.030241	176	7	0.039773	regulation of actin filament organization
GO:0030856	0.030589	89	5	0.05618	regulation of epithelial cell differentiation
GO:0007167	0.030859	648	15	0.023148	enzyme linked receptor protein signaling pathway
GO:0051045	0.03086	7	2	0.285714	negative regulation of membrane protein ectodomain proteolysis
GO:0045597	0.031037	583	14	0.024014	positive regulation of cell differentiation
GO:0051240	0.031252	1000	20	0.02	positive regulation of multicellular organismal process
GO:0050678	0.031252	228	8	0.035088	regulation of epithelial cell proliferation
GO:0072359	0.031572	651	15	0.023041	circulatory system development
GO:0010604	0.031888	2110	34	0.016114	positive regulation of macromolecule metabolic process
GO:0048583	0.032328	2454	38	0.015485	regulation of response to stimulus
GO:0065008	0.033266	2286	36	0.015748	regulation of biological quality
GO:0022008	0.033384	935	19	0.020321	neurogenesis
GO:0030030	0.033649	936	19	0.020299	cell projection organization
GO:1902905	0.034323	135	6	0.044444	positive regulation of supramolecular fiber organization
GO:0080135	0.034323	527	13	0.024668	regulation of cellular response to stress
GO:2000117	0.034323	56	4	0.071429	negative regulation of cysteine-type endopeptidase activity
GO:0055002	0.034758	93	5	0.053763	striated muscle cell development
GO:0043281	0.035244	136	6	0.044118	regulation of cysteine-type endopeptidase activity involved in apoptotic process
GO:0010035	0.037083	348	10	0.028736	response to inorganic substance
GO:0014745	0.037211	8	2	0.25	negative regulation of muscle adaptation

GO:0048585	0.037211	1022	20	0.019569	negative regulation of response to stimulus
GO:0048699	0.037211	877	18	0.020525	generation of neurons
GO:0099173	0.037211	95	5	0.052632	postsynapse organization
GO:0061028	0.037211	28	3	0.107143	establishment of endothelial barrier
GO:0031340	0.037211	8	2	0.25	positive regulation of vesicle fusion
GO:2000008	0.037211	28	3	0.107143	regulation of protein localization to cell surface
GO:0043484	0.037211	95	5	0.052632	regulation of RNA splicing
GO:1902903	0.038314	239	8	0.033473	regulation of supramolecular fiber organization
GO:0008150	0.03894	9139	99	0.010833	biological_process
GO:0009605	0.039256	1489	26	0.017461	response to external stimulus
GO:0009968	0.039583	812	17	0.020936	negative regulation of signal transduction
GO:0010631	0.03961	189	7	0.037037	epithelial cell migration
GO:0051252	0.039935	2150	34	0.015814	regulation of RNA metabolic process
GO:0043616	0.040022	29	3	0.103448	keratinocyte proliferation
GO:0090132	0.040311	190	7	0.036842	epithelium migration
GO:0051651	0.04058	60	4	0.066667	maintenance of location in cell
GO:0051591	0.04058	60	4	0.066667	response to cAMP
GO:0051271	0.041022	191	7	0.036649	negative regulation of cellular component movement
GO:0051254	0.04144	1111	21	0.018902	positive regulation of RNA metabolic process
GO:0007399	0.041689	1342	24	0.017884	nervous system development
GO:0090130	0.042963	193	7	0.036269	tissue migration
GO:0090185	0.044299	9	2	0.222222	negative regulation of kidney development
GO:1990403	0.044299	9	2	0.222222	embryonic brain development
GO:0071287	0.044299	9	2	0.222222	cellular response to manganese ion
GO:0055001	0.044306	101	5	0.049505	muscle cell development
GO:0051492	0.044306	62	4	0.064516	regulation of stress fiber assembly
GO:0000902	0.045477	620	14	0.022581	cell morphogenesis
GO:0090199	0.045477	31	3	0.096774	regulation of release of cytochrome c from mitochondria
GO:0009636	0.045477	305	9	0.029508	response to toxic substance
GO:0051973	0.045477	31	3	0.096774	positive regulation of telomerase activity
GO:0040013	0.045477	196	7	0.035714	negative regulation of locomotion
GO:0051093	0.046913	556	13	0.023381	negative regulation of developmental process
GO:0040007	0.048096	625	14	0.0224	growth

GO:1903076	0.048096	64	4	0.0625	regulation of protein localization to plasma membrane
GO:0045766	0.048216	104	5	0.048077	positive regulation of angiogenesis
GO:0071320	0.048216	32	3	0.09375	cellular response to cAMP
GO:1900408	0.048216	32	3	0.09375	negative regulation of cellular response to oxidative stress
GO:0008631	0.048216	32	3	0.09375	intrinsic apoptotic signaling pathway in response to oxidative stress
GO:0070268	0.048216	32	3	0.09375	cornification
GO:1903202	0.048216	32	3	0.09375	negative regulation of oxidative stress-induced cell death
GO:0009893	0.049166	2276	35	0.015378	positive regulation of metabolic process
GO:2001236	0.049363	105	5	0.047619	regulation of extrinsic apoptotic signaling pathway
GO:0033554	0.049591	1448	25	0.017265	cellular response to stress
GO:0097237	0.049592	151	6	0.039735	cellular response to toxic substance
GO:0098974	0.049739	10	2	0.2	postsynaptic actin cytoskeleton organization
GO:0001765	0.049739	10	2	0.2	membrane raft assembly
GO:0002683	0.049739	257	8	0.031128	negative regulation of immune system process
GO:1990440	0.049739	10	2	0.2	positive regulation of transcription from RNA polymerase II promoter in response to endoplasmic reticulum stress
GO:0035722	0.049739	33	3	0.090909	interleukin-12-mediated signaling pathway
GO:0110020	0.049739	66	4	0.060606	regulation of actomyosin structure organization
GO:0006887	0.049739	563	13	0.023091	exocytosis
GO:1902883	0.049739	33	3	0.090909	negative regulation of response to oxidative stress
GO:0045935	0.049739	1217	22	0.018077	positive regulation of nucleobase-containing compound metabolic process
GO:0061614	0.049739	33	3	0.090909	pri-miRNA transcription by RNA polymerase II
GO:0045603	0.049739	10	2	0.2	positive regulation of endothelial cell differentiation
GO:1905331	0.049739	10	2	0.2	negative regulation of morphogenesis of an epithelium
GO:0099188	0.049739	10	2	0.2	postsynaptic cytoskeleton organization
Luminal duct cells					
native	**p_value**	**term_size**	**intersection_size**	**recall**	**name**
GO:0070887	1.58E-08	1947	50	0.025681	cellular response to chemical stimulus

GO:0012501	1.58E-08	1247	39	0.031275	programmed cell death
GO:0042221	1.58E-08	2528	57	0.022547	response to chemical
GO:0008219	1.69E-08	1324	40	0.030211	cell death
GO:0009888	2.43E-08	1162	37	0.031842	tissue development
GO:0032502	2.84E-08	3495	67	0.01917	developmental process
GO:0048856	4.18E-08	3254	64	0.019668	anatomical structure development
GO:0080134	4.32E-08	966	33	0.034161	regulation of response to stress
GO:0048731	5.41E-08	2677	57	0.021292	system development
GO:0006950	7.61E-08	2381	53	0.02226	response to stress
GO:0010033	7.61E-08	1998	48	0.024024	response to organic substance
GO:0006915	1E-07	1194	36	0.030151	apoptotic process
GO:0006928	1.79E-07	1160	35	0.030172	movement of cell or subcellular component
GO:0050896	1.89E-07	4834	78	0.016136	response to stimulus
GO:0009653	2.01E-07	1567	41	0.026165	anatomical structure morphogenesis
GO:0032501	2.01E-07	3804	68	0.017876	multicellular organismal process
GO:0051704	3.37E-07	1597	41	0.025673	multi-organism process
GO:0048513	3.56E-07	1964	46	0.023422	animal organ development
GO:0071345	4.23E-07	639	25	0.039124	cellular response to cytokine stimulus
GO:0048869	4.86E-07	2379	51	0.021438	cellular developmental process
GO:0009605	4.86E-07	1489	39	0.026192	response to external stimulus
GO:0071310	5.11E-07	1635	41	0.025076	cellular response to organic substance
GO:0034097	5.11E-07	703	26	0.036984	response to cytokine
GO:0044092	5.14E-07	653	25	0.038285	negative regulation of molecular function
GO:0007275	5.14E-07	2983	58	0.019444	multicellular organism development
GO:0006952	5.15E-07	818	28	0.03423	defense response
GO:0030855	5.15E-07	372	19	0.051075	epithelial cell differentiation
GO:0042981	5.87E-07	939	30	0.031949	regulation of apoptotic process
GO:0030154	6.25E-07	2258	49	0.021701	cell differentiation
GO:0043067	8.35E-07	956	30	0.031381	regulation of programmed cell death
GO:0006955	8.86E-07	1081	32	0.029602	immune response
GO:0051674	1.04E-06	910	29	0.031868	localization of cell
GO:0048870	1.04E-06	910	29	0.031868	cell motility
GO:0043086	2.11E-06	458	20	0.043668	negative regulation of catalytic activity
GO:0097190	2.31E-06	415	19	0.045783	apoptotic signaling pathway
GO:0048523	2.67E-06	2872	55	0.01915	negative regulation of cellular process
GO:0048646	2.67E-06	669	24	0.035874	anatomical structure formation involved in morphogenesis
GO:0016477	2.67E-06	838	27	0.03222	cell migration
GO:0048583	2.67E-06	2454	50	0.020375	regulation of response to stimulus

GO:0050793	2.76E-06	1547	38	0.024564	regulation of developmental process
GO:0001775	3.06E-06	789	26	0.032953	cell activation
GO:0065009	3.06E-06	1770	41	0.023164	regulation of molecular function
GO:0010941	3.18E-06	1030	30	0.029126	regulation of cell death
GO:0040011	3.18E-06	1029	30	0.029155	locomotion
GO:0007155	3.18E-06	737	25	0.033921	cell adhesion
GO:0009913	3.36E-06	125	11	0.088	epidermal cell differentiation
GO:0007154	3.42E-06	3450	61	0.017681	cell communication
GO:0022610	3.42E-06	742	25	0.033693	biological adhesion
GO:0007165	4.09E-06	3192	58	0.01817	signal transduction
GO:0009611	4.27E-06	396	18	0.045455	response to wounding
GO:0048584	4.27E-06	1379	35	0.025381	positive regulation of response to stimulus
GO:0008544	4.27E-06	194	13	0.06701	epidermis development
GO:0051239	4.27E-06	1804	41	0.022727	regulation of multicellular organismal process
GO:0043066	4.56E-06	543	21	0.038674	negative regulation of apoptotic process
GO:0051248	4.88E-06	705	24	0.034043	negative regulation of protein metabolic process
GO:0051716	5.57E-06	4082	67	0.016414	cellular response to stimulus
GO:0009607	5.64E-06	768	25	0.032552	response to biotic stimulus
GO:0051346	6.2E-06	239	14	0.058577	negative regulation of hydrolase activity
GO:0023052	6.2E-06	3427	60	0.017508	signaling
GO:0043069	6.2E-06	556	21	0.03777	negative regulation of programmed cell death
GO:0007166	6.2E-06	1690	39	0.023077	cell surface receptor signaling pathway
GO:0032269	6.38E-06	665	23	0.034586	negative regulation of cellular protein metabolic process
GO:0060548	6.38E-06	611	22	0.036007	negative regulation of cell death
GO:2000026	7.01E-06	1215	32	0.026337	regulation of multicellular organismal development
GO:0048585	7.01E-06	1022	29	0.028376	negative regulation of response to stimulus
GO:0002376	7.07E-06	1631	38	0.023299	immune system process
GO:0010951	8.62E-06	112	10	0.089286	negative regulation of endopeptidase activity
GO:0097193	9.3E-06	212	13	0.061321	intrinsic apoptotic signaling pathway
GO:0002252	9.45E-06	683	23	0.033675	immune effector process
GO:0019221	9.63E-06	426	18	0.042254	cytokine-mediated signaling pathway
GO:0080135	1.03E-05	527	20	0.037951	regulation of cellular response to stress
GO:0043207	1.17E-05	750	24	0.032	response to external biotic stimulus
GO:0051707	1.17E-05	750	24	0.032	response to other organism

GO:0010466	1.25E-05	118	10	0.084746	negative regulation of peptidase activity
GO:0048519	1.25E-05	3230	57	0.017647	negative regulation of biological process
GO:0032879	1.25E-05	1602	37	0.023096	regulation of localization
GO:0052548	1.33E-05	221	13	0.058824	regulation of endopeptidase activity
GO:0001525	1.58E-05	306	15	0.04902	angiogenesis
GO:0061061	1.63E-05	397	17	0.042821	muscle structure development
GO:1901700	2.07E-05	1022	28	0.027397	response to oxygen-containing compound
GO:0060429	2.07E-05	720	23	0.031944	epithelium development
GO:0051270	2.18E-05	611	21	0.03437	regulation of cellular component movement
GO:0032268	2.18E-05	1643	37	0.02252	regulation of cellular protein metabolic process
GO:0042119	2.31E-05	317	15	0.047319	neutrophil activation
GO:0036230	2.38E-05	318	15	0.04717	granulocyte activation
GO:0006954	2.44E-05	364	16	0.043956	inflammatory response
GO:0048514	2.6E-05	366	16	0.043716	blood vessel morphogenesis
GO:0051179	2.7E-05	3875	63	0.016258	localization
GO:0045087	2.76E-05	465	18	0.03871	innate immune response
GO:0055001	2.83E-05	101	9	0.089109	muscle cell development
GO:0052547	2.9E-05	240	13	0.054167	regulation of peptidase activity
GO:0098609	2.98E-05	419	17	0.040573	cell-cell adhesion
GO:0097529	2.98E-05	102	9	0.088235	myeloid leukocyte migration
GO:0071621	2.99E-05	52	7	0.134615	granulocyte chemotaxis
GO:0051246	3.15E-05	1750	38	0.021714	regulation of protein metabolic process
GO:0045766	3.4E-05	104	9	0.086538	positive regulation of angiogenesis
GO:0097435	3.56E-05	426	17	0.039906	supramolecular fiber organization
GO:0051240	3.91E-05	1000	27	0.027	positive regulation of multicellular organismal process
GO:0030334	3.91E-05	532	19	0.035714	regulation of cell migration
GO:0042060	3.91E-05	335	15	0.044776	wound healing
GO:0045321	3.92E-05	698	22	0.031519	leukocyte activation
GO:0042127	3.92E-05	939	26	0.027689	regulation of cell population proliferation
GO:0050790	4.03E-05	1408	33	0.023438	regulation of catalytic activity
GO:0032940	4.45E-05	884	25	0.028281	secretion by cell
GO:0045055	4.45E-05	486	18	0.037037	regulated exocytosis
GO:0014070	4.85E-05	596	20	0.033557	response to organic cyclic compound
GO:0002274	5.24E-05	392	16	0.040816	myeloid leukocyte activation
GO:1903034	5.25E-05	111	9	0.081081	regulation of response to wounding
GO:0140352	6.17E-05	902	25	0.027716	export from cell

GO:0031324	6.44E-05	1586	35	0.022068	negative regulation of cellular metabolic process
GO:0040012	6.47E-05	609	20	0.032841	regulation of locomotion
GO:0098542	6.75E-05	556	19	0.034173	defense response to other organism
GO:0031099	6.76E-05	115	9	0.078261	regeneration
GO:1904018	7.2E-05	116	9	0.077586	positive regulation of vasculature development
GO:0043312	7.84E-05	312	14	0.044872	neutrophil degranulation
GO:0006887	7.84E-05	563	19	0.033748	exocytosis
GO:0002283	8.05E-05	313	14	0.044728	neutrophil activation involved in immune response
GO:0032101	8.43E-05	622	20	0.032154	regulation of response to external stimulus
GO:0048522	8.65E-05	3263	55	0.016856	positive regulation of cellular process
GO:0051172	8.65E-05	1466	33	0.02251	negative regulation of nitrogen compound metabolic process
GO:2000145	8.77E-05	569	19	0.033392	regulation of cell motility
GO:0040007	8.77E-05	625	20	0.032	growth
GO:0003008	8.86E-05	802	23	0.028678	system process
GO:0097530	9.56E-05	64	7	0.109375	granulocyte migration
GO:0002446	9.68E-05	320	14	0.04375	neutrophil mediated immunity
GO:0051094	9.68E-05	807	23	0.028501	positive regulation of developmental process
GO:0051336	0.00011	753	22	0.029216	regulation of hydrolase activity
GO:0055002	0.000112	93	8	0.086022	striated muscle cell development
GO:0030029	0.000115	474	17	0.035865	actin filament-based process
GO:0001568	0.000116	423	16	0.037825	blood vessel development
GO:0046903	0.000117	944	25	0.026483	secretion
GO:0050794	0.000124	5770	79	0.013692	regulation of cellular process
GO:1990000	0.000124	12	4	0.333333	amyloid fibril formation
GO:2001233	0.000124	283	13	0.045936	regulation of apoptotic signaling pathway
GO:0002366	0.000126	427	16	0.037471	leukocyte activation involved in immune response
GO:0001895	0.000139	26	5	0.192308	retina homeostasis
GO:0002263	0.000139	431	16	0.037123	cell activation involved in immune response
GO:0048468	0.000139	1223	29	0.023712	cell development
GO:0043299	0.000139	333	14	0.042042	leukocyte degranulation
GO:0072359	0.000142	651	20	0.030722	circulatory system development
GO:0065008	0.000142	2286	43	0.01881	regulation of biological quality
GO:0051241	0.000144	711	21	0.029536	negative regulation of multicellular organismal process
GO:0002275	0.000161	338	14	0.04142	myeloid cell activation involved in immune response
GO:0033554	0.000163	1448	32	0.022099	cellular response to stress

GO:0022603	0.000185	664	20	0.03012	regulation of anatomical structure morphogenesis
GO:0001944	0.000186	443	16	0.036117	vasculature development
GO:0002682	0.000193	849	23	0.027091	regulation of immune system process
GO:0048518	0.000199	3644	58	0.015917	positive regulation of biological process
GO:0002444	0.000202	346	14	0.040462	myeloid leukocyte mediated immunity
GO:0030162	0.000203	447	16	0.035794	regulation of proteolysis
GO:0072358	0.000213	449	16	0.035635	cardiovascular system development
GO:0001817	0.000228	400	15	0.0375	regulation of cytokine production
GO:0065007	0.000239	6460	84	0.013003	biological regulation
GO:0051098	0.000239	259	12	0.046332	regulation of binding
GO:0050921	0.00029	78	7	0.089744	positive regulation of chemotaxis
GO:0045765	0.000314	182	10	0.054945	regulation of angiogenesis
GO:0043154	0.000323	53	6	0.113208	negative regulation of cysteine-type endopeptidase activity involved in apoptotic process
GO:0002685	0.000359	112	8	0.071429	regulation of leukocyte migration
GO:0002684	0.000359	581	18	0.030981	positive regulation of immune system process
GO:2000116	0.000359	147	9	0.061224	regulation of cysteine-type endopeptidase activity
GO:0071624	0.000366	16	4	0.25	positive regulation of granulocyte chemotaxis
GO:0035239	0.000373	527	17	0.032258	tube morphogenesis
GO:0030595	0.000375	113	8	0.070796	leukocyte chemotaxis
GO:0032370	0.000395	33	5	0.151515	positive regulation of lipid transport
GO:2000117	0.000425	56	6	0.107143	negative regulation of cysteine-type endopeptidase activity
GO:0060326	0.000427	151	9	0.059603	cell chemotaxis
GO:0045861	0.000427	190	10	0.052632	negative regulation of proteolysis
GO:0003012	0.000478	235	11	0.046809	muscle system process
GO:0050920	0.000527	119	8	0.067227	regulation of chemotaxis
GO:0030036	0.000537	434	15	0.034562	actin cytoskeleton organization
GO:0071407	0.000553	383	14	0.036554	cellular response to organic cyclic compound
GO:0072593	0.000621	159	9	0.056604	reactive oxygen species metabolic process
GO:0001816	0.000621	440	15	0.034091	cytokine production
GO:2001234	0.000649	160	9	0.05625	negative regulation of apoptotic signaling pathway
GO:0048857	0.000654	37	5	0.135135	neural nucleus development
GO:0009892	0.000654	1870	36	0.019251	negative regulation of metabolic process
GO:0050789	0.000682	6114	80	0.013085	regulation of biological process

GO:0036499	0.000689	19	4	0.210526	PERK-mediated unfolded protein response
GO:0008283	0.000702	1134	26	0.022928	cell population proliferation
GO:0007010	0.00071	866	22	0.025404	cytoskeleton organization
GO:0010605	0.000723	1725	34	0.01971	negative regulation of macromolecule metabolic process
GO:0023051	0.00073	2123	39	0.01837	regulation of signaling
GO:0002443	0.000736	449	15	0.033408	leukocyte mediated immunity
GO:0010648	0.000736	869	22	0.025316	negative regulation of cell communication
GO:0023057	0.000742	870	22	0.025287	negative regulation of signaling
GO:1901342	0.000769	206	10	0.048544	regulation of vasculature development
GO:0031323	0.000778	3626	56	0.015444	regulation of cellular metabolic process
GO:0032989	0.000779	684	19	0.027778	cellular component morphogenesis
GO:0001819	0.000781	251	11	0.043825	positive regulation of cytokine production
GO:0010035	0.000791	348	13	0.037356	response to inorganic substance
GO:0035966	0.000791	166	9	0.054217	response to topologically incorrect protein
GO:0051146	0.000825	167	9	0.053892	striated muscle cell differentiation
GO:0032102	0.000868	210	10	0.047619	negative regulation of response to external stimulus
GO:0051050	0.000868	571	17	0.029772	positive regulation of transport
GO:0045595	0.000894	1086	25	0.02302	regulation of cell differentiation
GO:0030593	0.000975	41	5	0.121951	neutrophil chemotaxis
GO:0033993	0.000975	577	17	0.029463	response to lipid
GO:0043270	0.000992	133	8	0.06015	positive regulation of ion transport
GO:0022414	0.000992	762	20	0.026247	reproductive process
GO:0000003	0.000992	762	20	0.026247	reproduction
GO:0009719	0.001057	1030	24	0.023301	response to endogenous stimulus
GO:0042246	0.001069	42	5	0.119048	tissue regeneration
GO:0045214	0.001113	22	4	0.181818	sarcomere organization
GO:0043281	0.00114	136	8	0.058824	regulation of cysteine-type endopeptidase activity involved in apoptotic process
GO:0080090	0.001164	3494	54	0.015455	regulation of primary metabolic process
GO:0032103	0.001195	365	13	0.035616	positive regulation of response to external stimulus
GO:0010646	0.001195	2098	38	0.018112	regulation of cell communication
GO:0048589	0.001215	419	14	0.033413	developmental growth
GO:0043065	0.001273	421	14	0.033254	positive regulation of apoptotic process

GO:0045185	0.001305	71	6	0.084507	maintenance of protein location
GO:0042692	0.001325	223	10	0.044843	muscle cell differentiation
GO:0043068	0.001355	424	14	0.033019	positive regulation of programmed cell death
GO:0050727	0.001391	181	9	0.049724	regulation of inflammatory response
GO:1903793	0.001497	24	4	0.166667	positive regulation of anion transport
GO:0071622	0.001497	24	4	0.166667	regulation of granulocyte chemotaxis
GO:1903672	0.001497	24	4	0.166667	positive regulation of sprouting angiogenesis
GO:1901701	0.001512	726	19	0.026171	cellular response to oxygen-containing compound
GO:0001890	0.001515	106	7	0.066038	placenta development
GO:1905954	0.00154	46	5	0.108696	positive regulation of lipid localization
GO:0044403	0.001572	605	17	0.028099	symbiotic process
GO:2000977	0.001676	2	2	1	regulation of forebrain neuron differentiation
GO:2000978	0.001676	2	2	1	negative regulation of forebrain neuron differentiation
GO:0061741	0.001676	2	2	1	chaperone-mediated protein transport involved in chaperone-mediated autophagy
GO:0006986	0.001797	147	8	0.054422	response to unfolded protein
GO:1990440	0.001854	10	3	0.3	positive regulation of transcription from RNA polymerase II promoter in response to endoplasmic reticulum stress
GO:0006935	0.001854	332	12	0.036145	chemotaxis
GO:0051171	0.001884	3375	52	0.015407	regulation of nitrogen compound metabolic process
GO:0044419	0.001891	616	17	0.027597	interspecies interaction between organisms
GO:0042330	0.001937	334	12	0.035928	taxis
GO:1905517	0.001948	26	4	0.153846	macrophage migration
GO:2001243	0.002012	78	6	0.076923	negative regulation of intrinsic apoptotic signaling pathway
GO:0071248	0.002016	112	7	0.0625	cellular response to metal ion
GO:0009968	0.002057	812	20	0.024631	negative regulation of signal transduction
GO:0060255	0.002105	3578	54	0.015092	regulation of macromolecule metabolic process
GO:1902903	0.002114	239	10	0.041841	regulation of supramolecular fiber organization
GO:0048568	0.002179	240	10	0.041667	embryonic organ development
GO:1990266	0.002345	51	5	0.098039	neutrophil migration

GO:0090201	0.002394	11	3	0.272727	negative regulation of release of cytochrome c from mitochondria
GO:0097191	0.002541	156	8	0.051282	extrinsic apoptotic signaling pathway
GO:0030901	0.002541	52	5	0.096154	midbrain development
GO:0010942	0.002822	460	14	0.030435	positive regulation of cell death
GO:0046677	0.002827	202	9	0.044554	response to antibiotic
GO:0044093	0.002918	1112	24	0.021583	positive regulation of molecular function
GO:0021762	0.00323	30	4	0.133333	substantia nigra development
GO:1903900	0.00323	86	6	0.069767	regulation of viral life cycle
GO:0035295	0.00323	648	17	0.026235	tube development
GO:0048709	0.00323	55	5	0.090909	oligodendrocyte differentiation
GO:0009636	0.003387	305	11	0.036066	response to toxic substance
GO:1903035	0.003475	56	5	0.089286	negative regulation of response to wounding
GO:0032368	0.003475	56	5	0.089286	regulation of lipid transport
GO:0009966	0.003478	1892	34	0.01797	regulation of signal transduction
GO:0002683	0.003562	257	10	0.038911	negative regulation of immune system process
GO:0060563	0.003587	31	4	0.129032	neuroepithelial cell differentiation
GO:0007568	0.003703	211	9	0.042654	aging
GO:2000379	0.003703	57	5	0.087719	positive regulation of reactive oxygen species metabolic process
GO:0045069	0.003703	57	5	0.087719	regulation of viral genome replication
GO:0046457	0.003773	13	3	0.230769	prostanoid biosynthetic process
GO:0001516	0.003773	13	3	0.230769	prostaglandin biosynthetic process
GO:0006979	0.003838	311	11	0.03537	response to oxidative stress
GO:2001242	0.00392	127	7	0.055118	regulation of intrinsic apoptotic signaling pathway
GO:0030216	0.00392	90	6	0.066667	keratinocyte differentiation
GO:0070268	0.003928	32	4	0.125	cornification
GO:0050778	0.00408	423	13	0.030733	positive regulation of immune response
GO:2000120	0.004204	3	2	0.666667	positive regulation of sodium-dependent phosphate transport
GO:0061740	0.004204	3	2	0.666667	protein targeting to lysosome involved in chaperone-mediated autophagy
GO:0051173	0.004224	1998	35	0.017518	positive regulation of nitrogen compound metabolic process
GO:0001889	0.004302	92	6	0.065217	liver development
GO:0061041	0.004302	92	6	0.065217	regulation of wound healing
GO:0090342	0.004321	33	4	0.121212	regulation of cell aging

GO:0051099	0.004365	130	7	0.053846	positive regulation of binding
GO:0051247	0.004421	1079	23	0.021316	positive regulation of protein metabolic process
GO:0051651	0.004424	60	5	0.083333	maintenance of location in cell
GO:0048871	0.004459	267	10	0.037453	multicellular organismal homeostasis
GO:0031032	0.004459	131	7	0.053435	actomyosin structure organization
GO:0071241	0.004459	131	7	0.053435	cellular response to inorganic substance
GO:0031325	0.004459	2089	36	0.017233	positive regulation of cellular metabolic process
GO:0032309	0.004459	14	3	0.214286	icosanoid secretion
GO:0008150	0.004472	9139	99	0.010833	biological_process
GO:0048608	0.004502	268	10	0.037313	reproductive structure development
GO:0032504	0.004594	431	13	0.030162	multicellular organism reproduction
GO:0032270	0.004602	1014	22	0.021696	positive regulation of cellular protein metabolic process
GO:0061008	0.004602	94	6	0.06383	hepaticobiliary system development
GO:0006446	0.004602	61	5	0.081967	regulation of translational initiation
GO:0050918	0.004603	34	4	0.117647	positive chemotaxis
GO:0006508	0.004742	1088	23	0.02114	proteolysis
GO:0061458	0.004782	271	10	0.0369	reproductive system development
GO:1902904	0.004782	95	6	0.063158	negative regulation of supramolecular fiber organization
GO:0043542	0.004922	134	7	0.052239	endothelial cell migration
GO:0050900	0.004937	223	9	0.040359	leukocyte migration
GO:0031424	0.005028	35	4	0.114286	keratinization
GO:0031399	0.005028	1167	24	0.020566	regulation of protein modification process
GO:0009893	0.005028	2276	38	0.016696	positive regulation of metabolic process
GO:0030239	0.005028	35	4	0.114286	myofibril assembly
GO:0010604	0.005078	2110	36	0.017062	positive regulation of macromolecule metabolic process
GO:0048732	0.005078	274	10	0.036496	gland development
GO:0090023	0.005122	15	3	0.2	positive regulation of neutrophil chemotaxis
GO:1901571	0.005122	15	3	0.2	fatty acid derivative transport
GO:0071715	0.005122	15	3	0.2	icosanoid transport
GO:0019222	0.005185	3912	56	0.014315	regulation of metabolic process
GO:0071396	0.005237	383	12	0.031332	cellular response to lipid
GO:0001836	0.005432	36	4	0.111111	release of cytochrome c from mitochondria
GO:0051272	0.005746	332	11	0.033133	positive regulation of cellular component movement

GO:0021782	0.005775	65	5	0.076923	glial cell development
GO:0019730	0.005985	37	4	0.108108	antimicrobial humoral response
GO:0031579	0.006114	16	3	0.1875	membrane raft organization
GO:0010927	0.006118	66	5	0.075758	cellular component assembly involved in morphogenesis
GO:0002688	0.006118	66	5	0.075758	regulation of leukocyte chemotaxis
GO:0070527	0.006552	38	4	0.105263	platelet aggregation
GO:0007517	0.006558	234	9	0.038462	muscle organ development
GO:0043903	0.00677	143	7	0.048951	regulation of symbiosis, encompassing mutualism through parasitism
GO:0043900	0.006931	455	13	0.028571	regulation of multi-organism process
GO:0040017	0.006962	341	11	0.032258	positive regulation of locomotion
GO:0071211	0.007049	4	2	0.5	protein targeting to vacuole involved in autophagy
GO:1905907	0.007049	4	2	0.5	negative regulation of amyloid fibril formation
GO:0010631	0.007049	189	8	0.042328	epithelial cell migration
GO:0007596	0.007049	189	8	0.042328	blood coagulation
GO:0090022	0.007092	17	3	0.176471	regulation of neutrophil chemotaxis
GO:1903038	0.007232	69	5	0.072464	negative regulation of leukocyte cell-cell adhesion
GO:0090132	0.007232	190	8	0.042105	epithelium migration
GO:0051259	0.007282	290	10	0.034483	protein complex oligomerization
GO:0022408	0.007282	105	6	0.057143	negative regulation of cell-cell adhesion
GO:2001236	0.007282	105	6	0.057143	regulation of extrinsic apoptotic signaling pathway
GO:0051249	0.007308	239	9	0.037657	regulation of lymphocyte activation
GO:0006984	0.007558	40	4	0.1	ER-nucleus signaling pathway
GO:0050817	0.007558	192	8	0.041667	coagulation
GO:0007599	0.007558	192	8	0.041667	hemostasis
GO:0050866	0.007558	106	6	0.056604	negative regulation of cell activation
GO:0090130	0.007797	193	8	0.041451	tissue migration
GO:0030155	0.00786	405	12	0.02963	regulation of cell adhesion
GO:1902532	0.007983	349	11	0.031519	negative regulation of intracellular signal transduction
GO:0061844	0.008067	18	3	0.166667	antimicrobial humoral immune response mediated by antimicrobial peptide
GO:1900120	0.008067	18	3	0.166667	regulation of receptor binding
GO:0016192	0.008244	1293	25	0.019335	vesicle-mediated transport
GO:0009987	0.008275	8473	95	0.011212	cellular process
GO:0050728	0.008324	72	5	0.069444	negative regulation of inflammatory response

GO:0002687	0.008324	72	5	0.069444	positive regulation of leukocyte migration
GO:0000302	0.008324	150	7	0.046667	response to reactive oxygen species
GO:0097237	0.008587	151	7	0.046358	cellular response to toxic substance
GO:0050865	0.008729	299	10	0.033445	regulation of cell activation
GO:0071674	0.008729	42	4	0.095238	mononuclear cell migration
GO:0050776	0.008838	533	14	0.026266	regulation of immune response
GO:0022008	0.008937	935	20	0.02139	neurogenesis
GO:0051128	0.00898	1614	29	0.017968	regulation of cellular component organization
GO:0042493	0.009044	598	15	0.025084	response to drug
GO:0048246	0.009126	19	3	0.157895	macrophage chemotaxis
GO:1902624	0.009126	19	3	0.157895	positive regulation of neutrophil migration
GO:0007569	0.009162	74	5	0.067568	cell aging
GO:0009617	0.00942	303	10	0.033003	response to bacterium
GO:2000377	0.009894	113	6	0.053097	regulation of reactive oxygen species metabolic process
GO:0001667	0.010123	253	9	0.035573	ameboidal-type cell migration
GO:1905952	0.010226	76	5	0.065789	regulation of lipid localization
GO:0032970	0.010266	254	9	0.035433	regulation of actin filament-based process
GO:0032310	0.010266	5	2	0.4	prostaglandin secretion
GO:2000118	0.010266	5	2	0.4	regulation of sodium-dependent phosphate transport
GO:0032306	0.010266	5	2	0.4	regulation of prostaglandin secretion
GO:0001894	0.010266	114	6	0.052632	tissue homeostasis
GO:0032308	0.010266	5	2	0.4	positive regulation of prostaglandin secretion
GO:0042491	0.010272	20	3	0.15	inner ear auditory receptor cell differentiation
GO:1905521	0.010272	20	3	0.15	regulation of macrophage migration
GO:0034599	0.010528	205	8	0.039024	cellular response to oxidative stress
GO:0019079	0.01055	77	5	0.064935	viral genome replication
GO:1903670	0.010637	45	4	0.088889	regulation of sprouting angiogenesis
GO:0061077	0.010637	45	4	0.088889	chaperone-mediated protein folding
GO:0022604	0.01072	310	10	0.032258	regulation of cell morphogenesis
GO:0015698	0.011062	78	5	0.064103	inorganic anion transport
GO:0009887	0.011123	614	15	0.02443	animal organ morphogenesis
GO:0048609	0.011404	428	12	0.028037	multicellular organismal reproductive process
GO:0006690	0.011404	46	4	0.086957	icosanoid metabolic process

GO:0006693	0.011435	21	3	0.142857	prostaglandin metabolic process
GO:0043393	0.011435	161	7	0.043478	regulation of protein binding
GO:0043534	0.011435	79	5	0.063291	blood vessel endothelial cell migration
GO:0006692	0.011435	21	3	0.142857	prostanoid metabolic process
GO:1902175	0.011435	21	3	0.142857	regulation of oxidative stress-induced intrinsic apoptotic signaling pathway
GO:0014003	0.011435	21	3	0.142857	oligodendrocyte development
GO:0034620	0.011539	118	6	0.050847	cellular response to unfolded protein
GO:0045936	0.011559	371	11	0.02965	negative regulation of phosphate metabolic process
GO:0030335	0.011685	315	10	0.031746	positive regulation of cell migration
GO:0010563	0.011751	372	11	0.02957	negative regulation of phosphorus metabolic process
GO:0000902	0.011848	620	15	0.024194	cell morphogenesis
GO:0006810	0.01193	3026	45	0.014871	transport
GO:0051093	0.01193	556	14	0.02518	negative regulation of developmental process
GO:0009266	0.012331	164	7	0.042683	response to temperature stimulus
GO:0010628	0.012578	1267	24	0.018942	positive regulation of gene expression
GO:0042063	0.012655	165	7	0.042424	gliogenesis
GO:0045429	0.012655	22	3	0.136364	positive regulation of nitric oxide biosynthetic process
GO:0036003	0.012655	22	3	0.136364	positive regulation of transcription from RNA polymerase II promoter in response to stress
GO:0007162	0.012655	165	7	0.042424	negative regulation of cell adhesion
GO:0060119	0.012655	22	3	0.136364	inner ear receptor cell development
GO:0044070	0.01267	48	4	0.083333	regulation of anion transport
GO:0043588	0.012934	166	7	0.042169	skin development
GO:0051250	0.012934	82	5	0.060976	negative regulation of lymphocyte activation
GO:0009408	0.013115	122	6	0.04918	response to heat
GO:0044703	0.013338	501	13	0.025948	multi-organism reproductive process
GO:1903201	0.013483	49	4	0.081633	regulation of oxidative stress-induced cell death
GO:0034109	0.013483	49	4	0.081633	homotypic cell-cell adhesion
GO:0032305	0.01362	6	2	0.333333	positive regulation of icosanoid secretion
GO:0010966	0.01362	6	2	0.333333	regulation of phosphate transport
GO:0015732	0.01362	6	2	0.333333	prostaglandin transport

GO:0044341	0.01362	6	2	0.333333	sodium-dependent phosphate transport
GO:1905906	0.01362	6	2	0.333333	regulation of amyloid fibril formation
GO:0042989	0.01362	6	2	0.333333	sequestering of actin monomers
GO:1902622	0.01374	23	3	0.130435	regulation of neutrophil migration
GO:2000147	0.01374	325	10	0.030769	positive regulation of cell motility
GO:0010001	0.01374	124	6	0.048387	glial cell differentiation
GO:0007219	0.01374	124	6	0.048387	Notch signaling pathway
GO:0060047	0.01374	124	6	0.048387	heart contraction
GO:0044706	0.01374	124	6	0.048387	multi-multicellular organism process
GO:1904407	0.01374	23	3	0.130435	positive regulation of nitric oxide metabolic process
GO:0035556	0.013959	1758	30	0.017065	intracellular signal transduction
GO:0032507	0.013985	50	4	0.08	maintenance of protein location in cell
GO:0016032	0.013985	570	14	0.024561	viral process
GO:0003006	0.014501	386	11	0.028497	developmental process involved in reproduction
GO:0002690	0.01498	51	4	0.078431	positive regulation of leukocyte chemotaxis
GO:0035051	0.015074	86	5	0.05814	cardiocyte differentiation
GO:1903707	0.015074	86	5	0.05814	negative regulation of hemopoiesis
GO:0002694	0.015238	275	9	0.032727	regulation of leukocyte activation
GO:0050878	0.015586	276	9	0.032609	regulation of body fluid levels
GO:0006413	0.015688	128	6	0.046875	translational initiation
GO:1900046	0.015688	52	4	0.076923	regulation of hemostasis
GO:0050863	0.015688	174	7	0.04023	regulation of T cell activation
GO:0030193	0.015688	52	4	0.076923	regulation of blood coagulation
GO:0043462	0.015688	52	4	0.076923	regulation of ATPase activity
GO:0006936	0.015688	174	7	0.04023	muscle contraction
GO:0031333	0.015688	87	5	0.057471	negative regulation of protein complex assembly
GO:0008585	0.0168	53	4	0.075472	female gonad development
GO:0035315	0.016805	25	3	0.12	hair cell differentiation
GO:0043403	0.016805	25	3	0.12	skeletal muscle tissue regeneration
GO:0014706	0.016951	227	8	0.035242	striated muscle tissue development
GO:0030856	0.016951	89	5	0.05618	regulation of epithelial cell differentiation
GO:0010243	0.016951	651	15	0.023041	response to organonitrogen compound
GO:0003015	0.017163	131	6	0.045802	heart process
GO:2000193	0.017504	7	2	0.285714	positive regulation of fatty acid transport

GO:0032303	0.017504	7	2	0.285714	regulation of icosanoid secretion
GO:0007567	0.017504	7	2	0.285714	parturition
GO:0070555	0.017552	132	6	0.045455	response to interleukin-1
GO:0050792	0.017552	132	6	0.045455	regulation of viral process
GO:0090257	0.017552	132	6	0.045455	regulation of muscle system process
GO:0051291	0.017552	54	4	0.074074	protein heterooligomerization
GO:0051234	0.01781	3100	45	0.014516	establishment of localization
GO:0002695	0.018256	91	5	0.054945	negative regulation of leukocyte activation
GO:0046456	0.018262	26	3	0.115385	icosanoid biosynthetic process
GO:0032956	0.018319	231	8	0.034632	regulation of actin cytoskeleton organization
GO:0071900	0.018689	343	10	0.029155	regulation of protein serine/threonine kinase activity
GO:0035967	0.018689	134	6	0.044776	cellular response to topologically incorrect protein
GO:0051129	0.018991	465	12	0.025806	negative regulation of cellular component organization
GO:0019220	0.019254	1092	21	0.019231	regulation of phosphate metabolic process
GO:1902905	0.019266	135	6	0.044444	positive regulation of supramolecular fiber organization
GO:0002253	0.019326	345	10	0.028986	activation of immune response
GO:1903706	0.019351	288	9	0.03125	regulation of hemopoiesis
GO:0046545	0.019409	56	4	0.071429	development of primary female sexual characteristics
GO:0050818	0.019409	56	4	0.071429	regulation of coagulation
GO:0051174	0.01943	1094	21	0.019196	regulation of phosphorus metabolic process
GO:0007159	0.019473	183	7	0.038251	leukocyte cell-cell adhesion
GO:0006897	0.019478	346	10	0.028902	endocytosis
GO:0010259	0.019726	27	3	0.111111	multicellular organism aging
GO:0090049	0.019726	27	3	0.111111	regulation of cell migration involved in sprouting angiogenesis
GO:0001932	0.020105	877	18	0.020525	regulation of protein phosphorylation
GO:0051049	0.020226	1024	20	0.019531	regulation of transport
GO:1900407	0.020355	57	4	0.070175	regulation of cellular response to oxidative stress
GO:0050877	0.020478	409	11	0.026895	nervous system process
GO:0031400	0.020824	410	11	0.026829	negative regulation of protein modification process
GO:0006869	0.020924	186	7	0.037634	lipid transport
GO:0042592	0.020964	1103	21	0.019039	homeostatic process
GO:0060537	0.021006	238	8	0.033613	muscle tissue development
GO:1900122	0.02123	8	2	0.25	positive regulation of receptor binding

GO:0031340	0.02123	8	2	0.25	positive regulation of vesicle fusion
GO:0048146	0.02123	28	3	0.107143	positive regulation of fibroblast proliferation
GO:0070169	0.02123	28	3	0.107143	positive regulation of biomineral tissue development
GO:0110151	0.02123	28	3	0.107143	positive regulation of biomineralization
GO:0002548	0.02123	28	3	0.107143	monocyte chemotaxis
GO:0014745	0.02123	8	2	0.25	negative regulation of muscle adaptation
GO:0016338	0.02123	8	2	0.25	calcium-independent cell-cell adhesion via plasma membrane cell-adhesion molecules
GO:0061635	0.02123	8	2	0.25	regulation of protein complex stability
GO:0043217	0.02123	8	2	0.25	myelin maintenance
GO:0042542	0.021373	96	5	0.052083	response to hydrogen peroxide
GO:0009628	0.021525	814	17	0.020885	response to abiotic stimulus
GO:0071216	0.021667	140	6	0.042857	cellular response to biotic stimulus
GO:0007399	0.022087	1342	24	0.017884	nervous system development
GO:1903428	0.023026	29	3	0.103448	positive regulation of reactive oxygen species biosynthetic process
GO:0071675	0.023026	29	3	0.103448	regulation of mononuclear cell migration
GO:0070098	0.023026	29	3	0.103448	chemokine-mediated signaling pathway
GO:0001892	0.02339	60	4	0.066667	embryonic placenta development
GO:0034341	0.023827	99	5	0.050505	response to interferon-gamma
GO:0023056	0.023827	1119	21	0.018767	positive regulation of signaling
GO:0032496	0.023827	192	7	0.036458	response to lipopolysaccharide
GO:0031214	0.023827	99	5	0.050505	biomineral tissue development
GO:0110148	0.023827	99	5	0.050505	biomineralization
GO:0043085	0.024061	897	18	0.020067	positive regulation of catalytic activity
GO:0043535	0.024386	61	4	0.065574	regulation of blood vessel endothelial cell migration
GO:0036473	0.024386	61	4	0.065574	cell death in response to oxidative stress
GO:0090066	0.024386	302	9	0.029801	regulation of anatomical structure size
GO:0009725	0.024689	618	14	0.022654	response to hormone
GO:0060113	0.024732	30	3	0.1	inner ear receptor cell differentiation
GO:0042326	0.024732	303	9	0.029703	negative regulation of phosphorylation
GO:0008285	0.025571	425	11	0.025882	negative regulation of cell population proliferation

GO:0046597	0.025605	9	2	0.222222	negative regulation of viral entry into host cell
GO:0090343	0.025605	9	2	0.222222	positive regulation of cell aging
GO:1903897	0.025605	9	2	0.222222	regulation of PERK-mediated unfolded protein response
GO:0071287	0.025605	9	2	0.222222	cellular response to manganese ion
GO:1900102	0.025605	9	2	0.222222	negative regulation of endoplasmic reticulum unfolded protein response
GO:0072321	0.025605	9	2	0.222222	chaperone-mediated protein transport
GO:0070661	0.025605	146	6	0.041096	leukocyte proliferation
GO:0009615	0.025729	196	7	0.035714	response to virus
GO:0043269	0.025867	306	9	0.029412	regulation of ion transport
GO:0042325	0.025941	980	19	0.019388	regulation of phosphorylation
GO:0050670	0.026035	102	5	0.04902	regulation of lymphocyte proliferation
GO:0030968	0.026035	102	5	0.04902	endoplasmic reticulum unfolded protein response
GO:0051928	0.026169	63	4	0.063492	positive regulation of calcium ion transport
GO:0002237	0.026169	197	7	0.035533	response to molecule of bacterial origin
GO:1901698	0.026169	694	15	0.021614	response to nitrogen compound
GO:0046660	0.026169	63	4	0.063492	female sex differentiation
GO:0090199	0.026169	31	3	0.096774	regulation of release of cytochrome c from mitochondria
GO:0050868	0.026169	63	4	0.063492	negative regulation of T cell activation
GO:0008625	0.026169	63	4	0.063492	extrinsic apoptotic signaling pathway via death domain receptors
GO:0006636	0.026169	31	3	0.096774	unsaturated fatty acid biosynthetic process
GO:0010632	0.026386	148	6	0.040541	regulation of epithelial cell migration
GO:0032944	0.026626	103	5	0.048544	regulation of mononuclear cell proliferation
GO:0016049	0.027299	310	9	0.029032	cell growth
GO:1903076	0.027299	64	4	0.0625	regulation of protein localization to plasma membrane
GO:1902882	0.027299	64	4	0.0625	regulation of response to oxidative stress
GO:0035690	0.027299	253	8	0.031621	cellular response to drug
GO:1903409	0.027299	64	4	0.0625	reactive oxygen species biosynthetic process
GO:0010594	0.027434	104	5	0.048077	regulation of endothelial cell migration
GO:0006457	0.027756	150	6	0.04	protein folding

GO:1903202	0.027864	32	3	0.09375	negative regulation of oxidative stress-induced cell death
GO:0071320	0.027864	32	3	0.09375	cellular response to cAMP
GO:1900408	0.027864	32	3	0.09375	negative regulation of cellular response to oxidative stress
GO:0008631	0.027864	32	3	0.09375	intrinsic apoptotic signaling pathway in response to oxidative stress
GO:0015850	0.028257	105	5	0.047619	organic hydroxy compound transport
GO:0010595	0.028434	65	4	0.061538	positive regulation of endothelial cell migration
GO:0031347	0.028438	435	11	0.025287	regulation of defense response
GO:0048545	0.028597	256	8	0.03125	response to steroid hormone
GO:0002757	0.028989	314	9	0.028662	immune response-activating signal transduction
GO:1903963	0.029213	10	2	0.2	arachidonate transport
GO:0010468	0.029213	2569	38	0.014792	regulation of gene expression
GO:0042490	0.029213	33	3	0.090909	mechanoreceptor differentiation
GO:0007009	0.029213	66	4	0.060606	plasma membrane organization
GO:0035722	0.029213	33	3	0.090909	interleukin-12-mediated signaling pathway
GO:0048569	0.029213	10	2	0.2	post-embryonic animal organ development
GO:1902074	0.029213	10	2	0.2	response to salt
GO:2001237	0.029213	66	4	0.060606	negative regulation of extrinsic apoptotic signaling pathway
GO:0099188	0.029213	10	2	0.2	postsynaptic cytoskeleton organization
GO:1902883	0.029213	33	3	0.090909	negative regulation of response to oxidative stress
GO:0050482	0.029213	10	2	0.2	arachidonic acid secretion
GO:0045620	0.029213	33	3	0.090909	negative regulation of lymphocyte differentiation
GO:0001765	0.029213	10	2	0.2	membrane raft assembly
GO:0110020	0.029213	66	4	0.060606	regulation of actomyosin structure organization
GO:1900047	0.029213	33	3	0.090909	negative regulation of hemostasis
GO:0030195	0.029213	33	3	0.090909	negative regulation of blood coagulation
GO:0098974	0.029213	10	2	0.2	postsynaptic actin cytoskeleton organization
GO:0030182	0.029491	779	16	0.020539	neuron differentiation
GO:0007015	0.029563	259	8	0.030888	actin filament organization
GO:0035270	0.030387	67	4	0.059701	endocrine system development
GO:0070663	0.030427	108	5	0.046296	regulation of leukocyte proliferation
GO:0043549	0.030461	573	13	0.022688	regulation of kinase activity

GO:0002697	0.030584	206	7	0.033981	regulation of immune effector process
GO:0051100	0.031247	109	5	0.045872	negative regulation of binding
GO:0071349	0.031247	34	3	0.088235	cellular response to interleukin-12
GO:1900024	0.031247	34	3	0.088235	regulation of substrate adhesion-dependent cell spreading
GO:0042743	0.031247	34	3	0.088235	hydrogen peroxide metabolic process
GO:2000045	0.031247	109	5	0.045872	regulation of G1/S transition of mitotic cell cycle
GO:0000904	0.031335	444	11	0.024775	cell morphogenesis involved in differentiation
GO:0002065	0.031473	68	4	0.058824	columnar/cuboidal epithelial cell differentiation
GO:0006898	0.032214	157	6	0.038217	receptor-mediated endocytosis
GO:0019058	0.032556	209	7	0.033493	viral life cycle
GO:0045446	0.032982	69	4	0.057971	endothelial cell differentiation
GO:0032330	0.033389	35	3	0.085714	regulation of chondrocyte differentiation
GO:0070671	0.033389	35	3	0.085714	response to interleukin-12
GO:0043583	0.033389	111	5	0.045045	ear development
GO:0006071	0.033652	11	2	0.181818	glycerol metabolic process
GO:0050860	0.033652	11	2	0.181818	negative regulation of T cell receptor signaling pathway
GO:0009415	0.033652	11	2	0.181818	response to water
GO:2000341	0.033652	11	2	0.181818	regulation of chemokine (C-X-C motif) ligand 2 production
GO:0097284	0.033652	11	2	0.181818	hepatocyte apoptotic process
GO:0010876	0.033652	211	7	0.033175	lipid localization
GO:0070293	0.033652	11	2	0.181818	renal absorption
GO:0045597	0.03386	583	13	0.022298	positive regulation of cell differentiation
GO:0043409	0.033961	112	5	0.044643	negative regulation of MAPK cascade
GO:0050663	0.033961	112	5	0.044643	cytokine secretion
GO:1903037	0.034329	160	6	0.0375	regulation of leukocyte cell-cell adhesion
GO:0045428	0.035356	36	3	0.083333	regulation of nitric oxide biosynthetic process
GO:0050819	0.035356	36	3	0.083333	negative regulation of coagulation
GO:0031532	0.035387	71	4	0.056338	actin cytoskeleton reorganization
GO:0002040	0.035387	71	4	0.056338	sprouting angiogenesis
GO:0051017	0.036151	114	5	0.04386	actin filament bundle assembly
GO:0071347	0.036151	114	5	0.04386	cellular response to interleukin-1
GO:0071417	0.036534	392	10	0.02551	cellular response to organonitrogen compound
GO:0043500	0.036933	72	4	0.055556	muscle adaptation

GO:0061572	0.03731	115	5	0.043478	actin filament bundle organization
GO:0002042	0.037428	37	3	0.081081	cell migration involved in sprouting angiogenesis
GO:0048260	0.037428	37	3	0.081081	positive regulation of receptor-mediated endocytosis
GO:0045859	0.037428	524	12	0.022901	regulation of protein kinase activity
GO:1990869	0.037428	37	3	0.081081	cellular response to chemokine
GO:1990868	0.037428	37	3	0.081081	response to chemokine
GO:0002764	0.037975	333	9	0.027027	immune response-regulating signaling pathway
GO:0016043	0.038135	3926	52	0.013245	cellular component organization
GO:0061684	0.038135	12	2	0.166667	chaperone-mediated autophagy
GO:0060117	0.038135	12	2	0.166667	auditory receptor cell development
GO:0061024	0.038135	526	12	0.022814	membrane organization
GO:0050858	0.038135	12	2	0.166667	negative regulation of antigen receptor-mediated signaling pathway
GO:1904994	0.038135	12	2	0.166667	regulation of leukocyte adhesion to vascular endothelial cell
GO:0035994	0.038135	12	2	0.166667	response to muscle stretch
GO:0044849	0.038135	12	2	0.166667	estrous cycle
GO:2000191	0.038135	12	2	0.166667	regulation of fatty acid transport
GO:0072567	0.038135	12	2	0.166667	chemokine (C-X-C motif) ligand 2 production
GO:0030325	0.038135	12	2	0.166667	adrenal gland development
GO:0010647	0.038335	1110	20	0.018018	positive regulation of cell communication
GO:0071222	0.038848	117	5	0.042735	cellular response to lipopolysaccharide
GO:0046683	0.039158	74	4	0.054054	response to organophosphorus
GO:0043536	0.039158	38	3	0.078947	positive regulation of blood vessel endothelial cell migration
GO:1903573	0.039158	38	3	0.078947	negative regulation of response to endoplasmic reticulum stress
GO:0032781	0.039158	38	3	0.078947	positive regulation of ATPase activity
GO:0001933	0.039652	277	8	0.028881	negative regulation of protein phosphorylation
GO:0071356	0.039793	167	6	0.035928	cellular response to tumor necrosis factor
GO:0031348	0.039793	118	5	0.042373	negative regulation of defense response
GO:0032092	0.040584	75	4	0.053333	positive regulation of protein binding
GO:0042129	0.040584	75	4	0.053333	regulation of T cell proliferation

GO:0051101	0.040584	75	4	0.053333	regulation of DNA binding
GO:1902806	0.040818	119	5	0.042017	regulation of cell cycle G1/S phase transition
GO:0030308	0.040818	119	5	0.042017	negative regulation of cell growth
GO:0010212	0.040818	119	5	0.042017	response to ionizing radiation
GO:0034976	0.040878	222	7	0.031532	response to endoplasmic reticulum stress
GO:0010038	0.042801	224	7	0.03125	response to metal ion
GO:0010460	0.043117	13	2	0.153846	positive regulation of heart rate
GO:0016485	0.043117	121	5	0.041322	protein processing
GO:0030049	0.043117	13	2	0.153846	muscle filament sliding
GO:0050926	0.043117	13	2	0.153846	regulation of positive chemotaxis
GO:0043496	0.043117	13	2	0.153846	regulation of protein homodimerization activity
GO:0050927	0.043117	13	2	0.153846	positive regulation of positive chemotaxis
GO:0033275	0.043117	13	2	0.153846	actin-myosin filament sliding
GO:0010823	0.043432	40	3	0.075	negative regulation of mitochondrion organization
GO:0042698	0.043432	40	3	0.075	ovulation cycle
GO:0050777	0.043432	77	4	0.051948	negative regulation of immune response
GO:1903078	0.043432	40	3	0.075	positive regulation of protein localization to plasma membrane
GO:1902041	0.043432	40	3	0.075	regulation of extrinsic apoptotic signaling pathway via death domain receptors
GO:0045604	0.043432	40	3	0.075	regulation of epidermal cell differentiation
GO:0048678	0.043432	40	3	0.075	response to axon injury
GO:0071219	0.044052	122	5	0.040984	cellular response to molecule of bacterial origin
GO:0022407	0.044766	227	7	0.030837	regulation of cell-cell adhesion
GO:0003158	0.04496	78	4	0.051282	endothelium development
GO:0045596	0.045448	410	10	0.02439	negative regulation of cell differentiation
GO:0007423	0.0456	286	8	0.027972	sensory organ development
GO:0006809	0.04602	41	3	0.073171	nitric oxide biosynthetic process
GO:0010660	0.04602	41	3	0.073171	regulation of muscle cell apoptotic process
GO:1905475	0.04635	124	5	0.040323	regulation of protein localization to membrane
GO:0030330	0.04635	79	4	0.050633	DNA damage response, signal transduction by p53 class mediator
GO:0008637	0.04635	79	4	0.050633	apoptotic mitochondrial changes
GO:1904375	0.04635	79	4	0.050633	regulation of protein localization to cell periphery

GO:0014074	0.04635	79	4	0.050633	response to purine-containing compound
GO:0090344	0.047367	14	2	0.142857	negative regulation of cell aging
GO:0072575	0.047367	14	2	0.142857	epithelial cell proliferation involved in liver morphogenesis
GO:0006959	0.047367	80	4	0.05	humoral immune response
GO:0006817	0.047367	14	2	0.142857	phosphate ion transport
GO:0071677	0.047367	14	2	0.142857	positive regulation of mononuclear cell migration
GO:0072574	0.047367	14	2	0.142857	hepatocyte proliferation
GO:0090050	0.047367	14	2	0.142857	positive regulation of cell migration involved in sprouting angiogenesis
GO:0008406	0.047367	125	5	0.04	gonad development
GO:0019400	0.047367	14	2	0.142857	alditol metabolic process
GO:0043558	0.047367	14	2	0.142857	regulation of translational initiation in response to stress
GO:0030260	0.047367	80	4	0.05	entry into host cell
GO:1902176	0.047367	14	2	0.142857	negative regulation of oxidative stress-induced intrinsic apoptotic signaling pathway
GO:0032892	0.047367	14	2	0.142857	positive regulation of organic acid transport
GO:0044409	0.047367	80	4	0.05	entry into host
GO:0006468	0.047367	1220	21	0.017213	protein phosphorylation
GO:0070613	0.047438	42	3	0.071429	regulation of protein processing
GO:0055013	0.047438	42	3	0.071429	cardiac muscle cell development
GO:0002027	0.047438	42	3	0.071429	regulation of heart rate
GO:0060291	0.047438	42	3	0.071429	long-term synaptic potentiation
GO:0120036	0.048409	913	17	0.01862	plasma membrane bounded cell projection organization
GO:0031401	0.048409	764	15	0.019634	positive regulation of protein modification process
GO:0106015	0.049382	1	1	1	negative regulation of inflammatory response to wounding
GO:0048685	0.049382	1	1	1	negative regulation of collateral sprouting of intact axon in response to injury
GO:0048683	0.049382	1	1	1	regulation of collateral sprouting of intact axon in response to injury
GO:1900138	0.049382	1	1	1	negative regulation of phospholipase A2 activity
GO:0051623	0.049382	1	1	1	positive regulation of norepinephrine uptake
GO:0048673	0.049382	1	1	1	collateral sprouting of intact axon in response to injury
GO:1901328	0.049382	1	1	1	response to cytochalasin B

GO:0002434	0.049382	1	1	1	immune complex clearance
GO:0098657	0.049382	418	10	0.023923	import into cell
GO:0010042	0.049382	15	2	0.133333	response to manganese ion
GO:0044858	0.049382	1	1	1	plasma membrane raft polarization
GO:0006937	0.049382	84	4	0.047619	regulation of muscle contraction
GO:0060122	0.049382	15	2	0.133333	inner ear receptor cell stereocilium organization
GO:0060164	0.049382	1	1	1	regulation of timing of neuron differentiation
GO:1902310	0.049382	1	1	1	positive regulation of peptidyl-serine dephosphorylation
GO:0045977	0.049382	1	1	1	positive regulation of mitotic cell cycle, embryonic
GO:0060284	0.049382	560	12	0.021429	regulation of cell development
GO:0060298	0.049382	1	1	1	positive regulation of sarcomere organization
GO:1902847	0.049382	1	1	1	regulation of neuronal signal transduction
GO:0045137	0.049382	129	5	0.03876	development of primary sexual characteristics
GO:0042554	0.049382	15	2	0.133333	superoxide anion generation
GO:1901699	0.049382	424	10	0.023585	cellular response to nitrogen compound
GO:0044856	0.049382	1	1	1	plasma membrane raft localization
GO:0060723	0.049382	1	1	1	regulation of cell proliferation involved in embryonic placenta development
GO:0060722	0.049382	1	1	1	cell proliferation involved in embryonic placenta development
GO:0010758	0.049382	15	2	0.133333	regulation of macrophage chemotaxis
GO:0060584	0.049382	1	1	1	regulation of prostaglandin-endoperoxide synthase activity
GO:0060585	0.049382	1	1	1	positive regulation of prostaglandin-endoperoxide synthase activity
GO:0090323	0.049382	1	1	1	prostaglandin secretion involved in immune response
GO:0044117	0.049382	1	1	1	growth of symbiont in host
GO:0044110	0.049382	1	1	1	growth involved in symbiotic interaction
GO:1903317	0.049382	44	3	0.068182	regulation of protein maturation
GO:0060720	0.049382	1	1	1	spongiotrophoblast cell proliferation
GO:0090214	0.049382	1	1	1	spongiotrophoblast layer developmental growth
GO:1902998	0.049382	1	1	1	positive regulation of neurofibrillary tangle assembly
GO:0030421	0.049382	1	1	1	defecation

GO:0043179	0.049382	1	1	1	rhythmic excitation
GO:0043112	0.049382	129	5	0.03876	receptor metabolic process
GO:0014740	0.049382	1	1	1	negative regulation of muscle hyperplasia
GO:0060904	0.049382	1	1	1	regulation of protein folding in endoplasmic reticulum
GO:1903901	0.049382	44	3	0.068182	negative regulation of viral life cycle
GO:1903906	0.049382	1	1	1	regulation of plasma membrane raft polarization
GO:1903916	0.049382	1	1	1	regulation of endoplasmic reticulum stress-induced eIF2 alpha dephosphorylation
GO:1903917	0.049382	1	1	1	positive regulation of endoplasmic reticulum stress-induced eIF2 alpha dephosphorylation
GO:0090074	0.049382	1	1	1	negative regulation of protein homodimerization activity
GO:0044855	0.049382	1	1	1	plasma membrane raft distribution
GO:1904313	0.049382	1	1	1	response to methamphetamine hydrochloride
GO:0036497	0.049382	1	1	1	eIF2alpha dephosphorylation in response to endoplasmic reticulum stress
GO:1904576	0.049382	1	1	1	response to tunicamycin
GO:1905895	0.049382	1	1	1	negative regulation of cellular response to tunicamycin
GO:1904577	0.049382	1	1	1	cellular response to tunicamycin
GO:1904764	0.049382	1	1	1	chaperone-mediated autophagy translocation complex disassembly
GO:0042313	0.049382	1	1	1	protein kinase C deactivation
GO:1904997	0.049382	1	1	1	regulation of leukocyte adhesion to arterial endothelial cell
GO:1904998	0.049382	1	1	1	negative regulation of leukocyte adhesion to arterial endothelial cell
GO:1905457	0.049382	1	1	1	negative regulation of lymphoid progenitor cell differentiation
GO:0061035	0.049382	44	3	0.068182	regulation of cartilage development
GO:0061078	0.049382	1	1	1	positive regulation of prostaglandin secretion involved in immune response
GO:0021577	0.049382	1	1	1	hindbrain structural organization
GO:0061102	0.049382	1	1	1	stomach neuroendocrine cell differentiation

GO:1905892	0.049382	1	1	1	negative regulation of cellular response to thapsigargin
GO:1905894	0.049382	1	1	1	regulation of cellular response to tunicamycin
GO:1905891	0.049382	1	1	1	regulation of cellular response to thapsigargin
GO:0061105	0.049382	1	1	1	regulation of stomach neuroendocrine cell differentiation
GO:0061106	0.049382	1	1	1	negative regulation of stomach neuroendocrine cell differentiation
GO:0021558	0.049382	1	1	1	trochlear nerve development
GO:0060721	0.049382	1	1	1	regulation of spongiotrophoblast cell proliferation
GO:1905933	0.049382	1	1	1	regulation of cell fate determination
GO:1905934	0.049382	1	1	1	negative regulation of cell fate determination
GO:0072749	0.049382	1	1	1	cellular response to cytochalasin B
GO:0021768	0.049382	1	1	1	nucleus accumbens development
GO:1990832	0.049382	1	1	1	slow axonal transport
GO:0070373	0.049382	43	3	0.069767	negative regulation of ERK1 and ERK2 cascade
GO:0021589	0.049382	1	1	1	cerebellum structural organization
GO:0061395	0.049382	1	1	1	positive regulation of transcription from RNA polymerase II promoter in response to arsenic-containing substance
GO:0034612	0.049382	181	6	0.033149	response to tumor necrosis factor
GO:0034391	0.049382	15	2	0.133333	regulation of smooth muscle cell apoptotic process
GO:0034390	0.049382	15	2	0.133333	smooth muscle cell apoptotic process
GO:0072576	0.049382	15	2	0.133333	liver morphogenesis
GO:0003065	0.049382	1	1	1	positive regulation of heart rate by epinephrine
GO:0032717	0.049382	15	2	0.133333	negative regulation of interleukin-8 production
GO:0006710	0.049382	1	1	1	androgen catabolic process
GO:0032286	0.049382	1	1	1	central nervous system myelin maintenance
GO:2000226	0.049382	1	1	1	regulation of pancreatic A cell differentiation
GO:2000227	0.049382	1	1	1	negative regulation of pancreatic A cell differentiation
GO:0030185	0.049382	1	1	1	nitric oxide transport
GO:0061757	0.049382	1	1	1	leukocyte adhesion to arterial endothelial cell

GO:0031583	0.049382	1	1	1	phospholipase D-activating G protein-coupled receptor signaling pathway
GO:2000520	0.049382	1	1	1	regulation of immunological synapse formation
GO:2000521	0.049382	1	1	1	negative regulation of immunological synapse formation
GO:2000866	0.049382	1	1	1	positive regulation of estradiol secretion
GO:2000974	0.049382	1	1	1	negative regulation of pro-B cell differentiation
GO:0062094	0.049382	1	1	1	stomach development
GO:2001189	0.049382	1	1	1	negative regulation of T cell activation via T cell receptor contact with antigen bound to MHC molecule on antigen presenting cell
GO:0050801	0.049382	423	10	0.023641	ion homeostasis
GO:0010657	0.049786	45	3	0.066667	muscle cell apoptotic process
GO:0070059	0.049786	45	3	0.066667	intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress
GO:0046209	0.049786	45	3	0.066667	nitric oxide metabolic process
GO:0015908	0.049786	45	3	0.066667	fatty acid transport
GO:1903036	0.049786	45	3	0.066667	positive regulation of response to wounding