The University Of Sheffield.

An Investigation into the Expression and the Role of the CRTC1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma

Esra B Amoura

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Oral & Maxillofacial Pathology
School of Clinical Dentistry

November 2020

This thesis is detected to the memory of my father

(Bashir Ali Amoura 1946-2002)

Acknowledgments

First and foremost, I would like to express my greatest gratitude to my supervisor Dr Lynne Bingle for her patience, guidance, encouragement and advice that she has constantly provided. I cannot thank her enough, under her supervision, I became a confident and independent researcher. She has not only supported me pursuing my research project, she supported my development and career and was there for me during my difficult times. I would also like to thank Professor Colin Bingle for his guidance and support throughout the project, his vision, sincerity and motivation have deeply inspired me. He taught me the methodology to carry out my research project. It was a great privilege and honour to be part of the Bingle's laboratory, and to work under their guidance.

I would like to extremely thank Professor Keith Hunter for his continuous help, support and valuable contributions throughout my PhD and master's degree. I am extremely grateful, and I cannot thank him enough for his supervision and guidance, helping with my career and development.

I would like to extend my sincerest thanks and appreciation to Dr Vivian Wagner for her endless support and encouragement during my PhD.

I would like to say thanks to those valuable people who helped and encouraged me to pursue my laboratory skills, technical staff: Mrs Brenka McCabe, Mr Jason Heath, Dr Benjamin Durham, Mrs Hayley Stanhope, Mrs Kirsty Franklin and Mr David Thompson; PhD colleagues: Asma, Naeima, Anita, Cher, Amy, Basma, Maryam, Lamis, Rawan and Sven.

I would like to thank the Bingle group for their support, help and the great times we had during my studies (Rachel, Zulaiha, Priyanka, Hannah, Miraj, Renata, Zeyad, Fawaz, Irene and Abdulaziz).

A special thanks to Professor Paul Speight and Dr Daniel Brierley for helping with normal salivary gland and tumour images.

A special word of thanks also goes to my husband Anas, and daughters Lina, Noor and Sarah, my mother and siblings who always provide me with energy, support and enthusiasm, I could not have done this without you. Finally, I am grateful to Tripoli University and The Ministry of Higher Education, Libya for providing financial support and fully funding this PhD project.

Oral Presentations

- BSODR, Leeds UK (3rd-5th September 2019) Presented an Oral talk “The CRTC1-MAML2Fusion Protein- an innocent bystander or tumour promoter?
- “Investigating the Role of CRTC1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma” Gordon Research Seminar for Salivary Glands and Exocrine Biology, Texas, USA (2019).
- “CRTC1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma” North-East Postgraduate conference, Newcastle, UK (2018).

Poster Presentations

- “Investigating the Role of CRTC1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma” SInFoNiA Conference, University of Sheffield, 21st March 2019.
- “Investigating the Role of CRTC1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma” Gordon Research Seminar for Salivary Glands and Exocrine Biology, Texas, USA (2019).
- “The Role of CRTC1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma” UK Society for Extracellular Vesicles (UK-EV), Sheffield, UK (2018).
- “CRTC1-MAML2 Fusion in Mucoepidermoid Carcinoma” The British Association for Cancer Research (BACR) Student Conference, London, UK (2018).
- “The Role of MECT1-MAML2 Fusion Protein in Mucoepidermoid Carcinoma” International Association for Dental Research (IADR), 96th General Session, London, UK (2018).
- “Towards understanding the Role of MECT1-MAML2 Fusion Protein Mucoepidermoid Carcinoma” North-East Postgraduate conference, Newcastle, UK (2017).

Manuscript

- **Accepted for publication in Biomedicines**

Vivian Petersen Wagner, Manoela Domingues Martins, Esra B Amoura, Virgilio Gonzales Zanella, Rafael Roesler, Caroline Brunetto de Farias, Colin Bingle, Pablo Agustin Vargas, Lynne Bingle “TrkB-targeted therapy for Mucoepidermoid Carcinoma” November 2020.

Abstract

Introduction

Mucoepidermoid carcinoma (MEC) is the most common salivary gland cancer and the majority of cases harbour a gene translocation between the CRTC1 and MAML2 genes resulting in a novel oncogenic fusion protein. Some MEC cases are difficult to diagnose and this, along with the fact that the method currently used to detect the fusion gene (MAML2), a break apart probe and fluorescence *in situ* hybridization (FISH) is complicated and does not provide information about the fusion event, highlights the need for a test which can routinely be performed on FFPE tissues. The role of the fusion protein in tumour development and growth has not been fully elucidated and its molecular pathogenesis remains elusive. The aim of this study was to develop a specific mRNA BaseScope assay to detect the CRTC1-MAML2 fusion in MEC FFPE samples and to develop an *in vitro* model to investigate the role of the fusion protein in MEC pathogenesis.

Methods

A retrospective cohort of 29 MEC samples were retrieved from the tissue bank with ethical approval, clinicopathological data was reviewed, thorough histological analysis was conducted and a chromogenic BaseScope assay, using a probe targeting the novel exon-exon junction in the CRTC1-MAML2 fusion transcript, were performed on FFPE tumour specimens. An epitope tagged CRTC1-MAML2 expression clone was generated by amplifying the translocated genes from fusion-expressing MEC cell lines. The CRTC1-MAML2 expression clone was transfected into primary salivary gland cells (sublingual and submandibular), isolated from human explanted tissues and into mammalian cell lines, to study the down-stream effects of novel fusion protein expression. I also developed three-dimensional (3D) organoid

models of normal human salivary gland cells and CRTC1-MAML2 fusion transfected cells to validate the monolayer cell culture results.

Results

Using the BaseScope assay, we detected the RNA transcript signal for the CRTC1-MAML2 junction in known fusion-positive cells but not in known fusion-negative cells. We also detected distinct fusion events in 34% of MEC cases, however, only a small proportion of the tumour cells harboured the translocation. A flag-tagged mammalian expression construct was generated and verified by overlapping sequencing. Primary salivary gland cells and mammalian cell lines were successfully transfected, and protein expression subsequently detected; the generation of a stable cell line was not successful. The exogenous expression of the CRTC1-MAML2 in transiently transfected HuSL cells led to alteration of CREB target gene expression. The CRTC1-MAML2 fusion influenced NCI-H647 cell migration, NCI-H647 and HEK293 cells survival and invasion, but it does not play any role in cell proliferation. The 3D organoid models of normal, transfected salivary gland cells and MEC cells have been developed and our preliminary histological analysis indicated variations in the organisation of the cells across the established models.

Conclusion

In this study, we have demonstrated for the first time, that the BaseScope assay accurately detects CRTC1-MAML2 fusion transcripts. This could provide an alternative chromogenic technique, for use in routine clinical labs, to aid accurate diagnosis of MEC after validation of specificity and sensitivity against currently used techniques. The CRTC1-MAML2 fusion construct provides a valuable tool for studying the downstream effects of de novo expression on cell behaviour. Developing 3D models will allow us to further elucidate the role of the fusion protein in tumourigenesis.

Contents

List of Figures

1. Chapter 1

2. Chapter 2

3. Chapter 3

4. Chapter 4

5. Chapter 5

6. Chapter 6

7. Chapter 7

List of Tables

1. Chapter 1

2. Chapter 2

3. Chapter 3

Abbreviations

+ve	Positive
-ve	Negative
×	Times
%	Percentage
α- SMA	alpha smooth muscle actin
C°	Celsius
∆	Delta
µg	Microgram
µl	Microlitre
µm	Micrometre
2D	Two-dimensional
3D	Three-dimensional
AB	Antibody
ACC	Adenoid cystic carcinoma
AFIP	Armed Forces Institute of Pathology
AKT2	Serine/threonine kinase 2
ALI	Air-Liquid-interface
AMP	Amplification
ANOVA	Analysis of variance
AP1	Activated protein 1, transcriptional factor
AREG	Amphiregulin
ATF1	Activating transcription factor 1
ATF3	Activating transcription factor 3
ATP	Adenosine triphosphate
B2M	Beta-2-Microglobulin
BCL2	B-cell lymphoma gene
BGH	Bovine Growth Hormone
BIRC2	Baculoviral IAP Repeat Containing 2
bp	Base Pair
BSA	Bovine serum albumin

BrdU	5-bromo-2’-deoxyuridine
C-	Carboxyl group
C1M2	CRTC1-MAML2 fusion
cAMP	cyclic adenosine monophosphate
CBD	CREB binding domain
CCDH	Charles Clifford Dental Hospital
CD34	Cluster of differentiation 34
CDKN2A	cyclin-dependent kinase inhibitor 2A
cDNA	complementary DNA
CFTR	Cystic fibrosis transmembrane regulator gene
CGH	Comparative genomic hybridization
CK5	Cytokeratin 5
CK7	Cytokeratin 7
C-Kit	Stem cell factor
CLDN7	Claudin-7
cm	Centimetre (s)
CMV	Cytomegalovirus
CO_2	Carbon dioxide
COS7	Fibroblast-like cells derived from monkey kidney tissue
CRTC1	CREB Regulator Transcriptional Coactivator 1
CRTC3	CREB Regulator Transcriptional Coactivator 3
CREB	cAMP response element binding mediated transcription
CRISPR	Clustered regularly interspaced short palindromic repeats
CSL	CBF1, Suppressor of Hairless, Lag-1
Ctrl	Control
CT	Cycle threshold
$CuSO_4$	Copper sulfate
CUX1	Homeobox protein cut-like 1
DAB	3,3’-Diaminobenzidine
DAPB	4-hydroxy-tetrahydrodipicolinate reductase Bacillus
DAPI	4', 6-diamidino-2-phenylindole
DDH_2O	Deionised distilled water

DFS	Disease free survival
DMBT1	Deletion in malignant brain tumour1
DMEM	Dulbecco's Modified Eagles Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPX	Diethylphenyl-zanthine
DSS	Disease free survival
DTT	Dithiothreitol
DUSP1	Dual Specificity Phosphatase 1
ECL	Enhanced luminol-based chemiluminescent substrate
ECM	Extracellular matrix
E. coli	Escherichia Coli
EDTA	Ethylene-diamine-tetra-acetic acid
EdU	5-ethynyl-2'-deoxyuridine
EGF	Epidermal growth factor
eGFP	Enhanced green fluorescence protein
EGFR	Epidermal growth factor receptor
EHS	Engelbreth-Holm-Swarm mouse sarcoma
EMT	Epithelial mesenchymal transition
EST	Expressed sequence tag
ETS	Erythroblast Transformation Specific
ETV6	ETS Variant Transcription factor 6
EV	Empty Vector
EWSR	Ewing sarcoma RNA binding protein
F	Forward
FACS	Fluorescence activated cell sorting
FBS	Foetal bovine serum
FFPE	Formalin Fixed Paraffin embedded
FGF	Fibroblast growth factor
FISH	Fluorescence *in situ* Hybridization
FITC	Fluorescent isothiocyanate
g	gram

g	Gravity
G418	Geneticin
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gDNA	Genomic DNA
GFP	Green fluorescent protein
H&E	Haematoxylin and Eosin
H4	Histone 4 gene
HBS	HEPES Buffer Saline
HDACs	Histone deacetylases
HEPES	Hydroxyethyl-piperazine ethane sulfonic acid
HES1	Basic helix-loop-helix transcriptional repressors
HES5	Basic helix-loop-helix transcriptional repressors
HG	High glucose
HIF-1 α	Hypoxia-inducible factor 1-alpha
HMG	High mobility group
HMGA2	High mobility group A2
HPG	Human parotid gland cells
HRP	Horseradish peroxidase
HuSL	Normal human salivary glands
h	Hour (s)
H_2O_2	Hydrogen peroxide
H_2O	hydrogen dioxide (Water)
HybEZ	Hybridization
ID	Identification
IDA	Industrial methylated spirit (denatured alcohol) 99%
IF	Immunofluorescence
IGF	Insulin-like growth factor
IHC	Immunohistochemistry
iPSC	Induced pluripotent stem cells
IPTG	Isopropyl B-D-1-thiogalactopyranoside
IREs	Internal ribosomal entry site
K	Kilo

KDa	Kilo Dalton
KGM	Keratinocytes growth medium
L	Litre
LB	Luria-Bertani media
LG	Low glucose
L-G	L- Glutamine
LINC00473	Long Intergenic Non-Protein Coding RNA 473
LKB1	Tumour suppressor gene (serine-threonine kinase 11)
M	Molar
MAD1L1	Mitotic Arrest Deficient Like 1
MAML2	Master mind-like gene2
MEC	Mucoepidermoid carcinoma
MECT1	Mucoepidermoid carcinoma translocation gene1
MG132	Proteasome inhibitor
min	minute (s)
mL	millilitre
mM	Millimolar
MMP	Matrix metalloproteases
mRNA	Messenger RNA
MUC1	Membrane-bound mucins1
MUC2	Membrane-bound mucins2
MUC4	Membrane bound mucins4
MUC5AC	Membrane bound mucins5AC
MUC5B	Membrane bound mucins5B
MYB	**My**elo**b**lastosis
MYC	**My**elo**c**ytomatosis
N-	amino-terminus
Nanog	Stem cell transcriptional factor
NBD	Notch binding domain
NCBI	National Centre for Biotechnology Information
NCI	National Cancer Institute
NDRG1	N-Myc Downstream-Regulated Gene 1

NFIB	Nuclear factor I B
ng	Nanogram
NGS	Normal goat serum
NIH3T3	NIH Swiss mouse embryo, Three-day transfer
nm	Nanometre
NR4A2	Nuclear receptor 4 related A2
NREC	NHS Research Ethics Committee
NTRK1	Neurotrophic Receptor Tyrosine Kinase 1
OAZ1	Ornithine Decarboxylase Antizyme 1
OS	Overall survival
OSCC	Oral squamous cell carcinoma
P	Phosphorylation
P	Passage number
p	Short arm of the chromosome
P63	Transforming relating tumour protein 63
P/S	Penicillin/ Streptomycin
PA	Pleomorphic adenoma
PAGE	Polyacrylamide gel electrophoresis
PAS-D	Periodic acid–Schiff–diastase
PAS	Periodic-Acid-Schiff
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDE4B	Phosphodiesterase 4B
pH	Potentiometric hydrogen ion concentration
PI	Protease inhibitor
PKA	Protein kinase A
PLAG1	Pleomorphic adenoma gene1
Pou6f2	POU Class 6 Homeobox 2
PPIB	Peptidylprolyl isomerase B
Pts	Point (s)
PTC	Papillary thyroid carcinoma

pUC19	Plasmid University of California
PVDF	Polyvinylidene fluoride
qPCR	Quantitative polymerase chain reaction
q	Long arm of the chromosome
R	Reverse
RET	Oncoproteins (**R**earranged **d**uring **t**ransfection)
RIPA	Radioimmunoprecipitation assay buffer
RISH	RNA In Situ Hybridization
RK3E	Rat Kidney epithelial
RNA	Ribonucleic acid
ROCK	Rho kinase inhibitors
RPM	Revolutions per minute
RS1	R-spondin-1
RT	Room temperature
RT-PCR	Reverse transcriptase polymerase chain reaction
SA-β-Gal	Senescence associated b galactosidase
SCC	Squamous cell carcinoma
SCDM	Santa Casa de Misericordia Hospital
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SEM	Scanning electron microscopy
SEM	Standard error of the mean
SGC	Salivary glands cancer
SGM	Salivary glands medium
SGT	Salivary glands tumour
sh	Short Hairpin
SIgA	Salivary immunoglobulin A
siRNA	Small interfering RNA
SLG	Sublingual gland cells
SMG	Submandibular gland cells
SOC	Super Optimal broth with Catabolite repression
STC1	Stanniocalcin 1

t	Translocation
T	Time point
TAD	Transcriptional activation domain
TAE	Tris base, acetic acid and EDTA
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline-tween
TE	Trypsin EDTA
TFE3	Transcriptional factor E3
TFEB	Transcriptional factor EB
TFF1	Trefoil factor 1
TGFβ-1	Transforming growth factor beta-1
TK	Tyrosine protein Kinase
TNF α	Tumour necrosis factor alpha
TORC1	Transducer of regulated CREB activity1
TP53	Tumour suppressor gene
Tris	Trisaminomethane
Tris HCl	Tris Hydrochloride
TSP-1	Thrombospondin 1
U	Unit
U6	RNU6B
UM-HMC2	University of Michigan-Human mucoepidermoid carcinoma 2
UTR	Untranslated region
UV	Ultraviolet
U2OS	Human osteosarcoma cells
V	Voltage
V	Volume
VEGF-A	Vascular endothelial growth factor A
W	Weight
WAMTP1	Warthin and mucoepidermoid tumour translocated partner 1
WHO	World Health Organisation
Wnt	Wingless-related integration site

1. Chapter 1 Literature Review

1.1 The normal salivary glands

The salivary glands are exocrine glandular tissues present in the oral cavity and divided into major and minor glands. In humans, the salivary glands are composed of three symmetrical pairs of major salivary glands: parotid, submandibular (previously known as submaxillary) and sublingual (figure1.1). There are also numerous, irregular, scattered minor salivary glands.

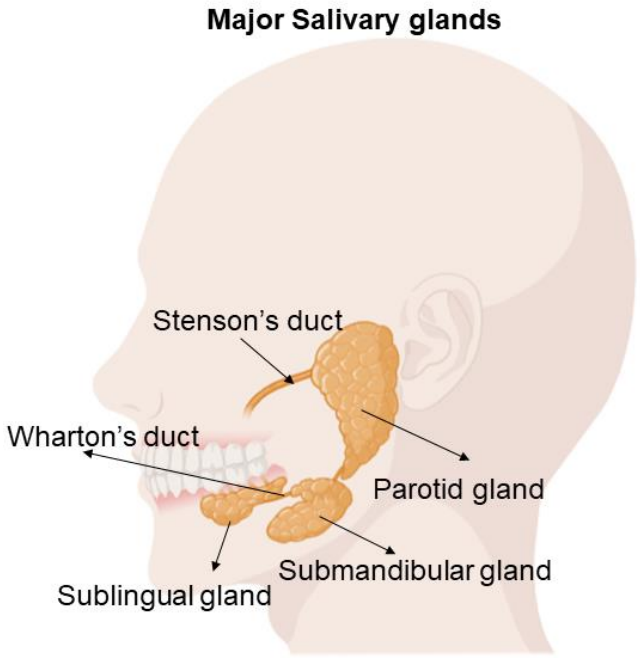


Figure 1.1 Major salivary glands, the anatomical locations of the parotid, submandibular and sublingual gland. Image created in BioRender.com.

A major function of salivary glands is to keep the oral cavity moist through the secretion of saliva, 80% of which is secreted by the major glands, whereas the minor salivary glands are responsible for the final 20%. Among the major glands, the submandibular glands are responsible for secreting 70% of the saliva, while the parotid glands secrete 25%, and the remaining 5% is secreted by the sublingual

glands. Saliva is important for lubricating the oral mucosa and helping with speech, mastication, and swallowing. Saliva is crucial in food digestion via the secretion of specific digestive enzymes such as amylase and other proteins like mucin and glycogen. Saliva is also responsible for maintaining the normal oral flora, as well as playing an essential role in the prevention of dental caries and periodontal diseases by secreting the antimicrobial agent salivary immunoglobulin A (SIg A) and lysozymes.

The basic structures of all salivary glands are similar, however, there are some differences in secretions, functions and histology. The salivary gland unit consists of acinar cells, intercalated ducts, striated ducts, and excretory ducts (figure 1.2). These units are arranged in two sets of cells: luminal cells, which include acinar and ductal cells, and abluminal cells. The abluminal cells are known as myoepithelial cells and they surround the acinus and the intercalated ducts at the periphery and the basal cells on the proximal side of excretory ducts. There are two types of acini, serous and mucous, and the composition of saliva differs between the two types. The serous acini contain amylase, whereas the mucous acini secrete sialomucins. The saliva is stored in granules inside the epithelial cells of the acini. The proteins and electrolytes inside these acini are dissolved in water with which they are mainly filled. The secretion of saliva increases due to stimuli such as eating and chewing and controlled mainly by the autonomous nervous system (Iorgulescu, 2009). However, even without a stimulus the salivary glands constitutively secrete some oral fluid.

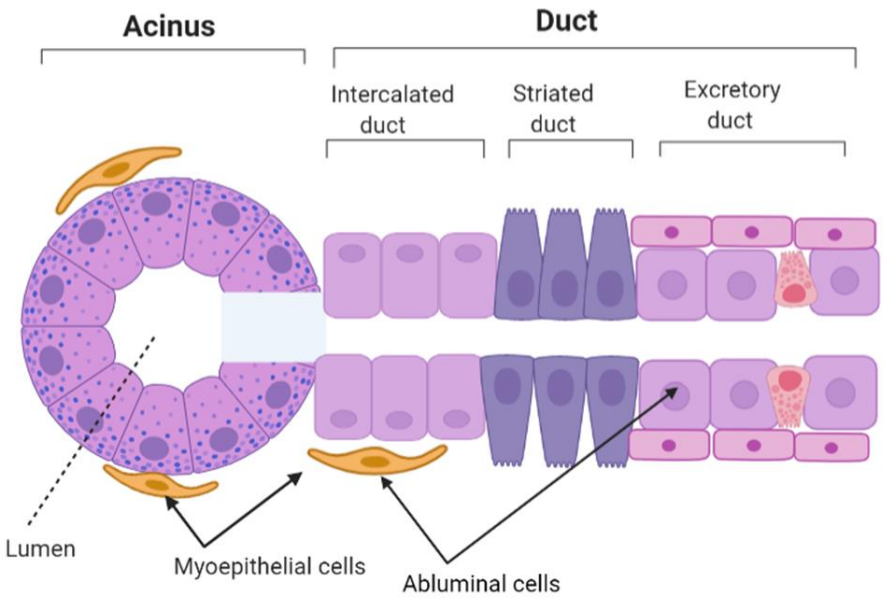


Figure 1.2 General salivary gland unit: Acinus, intercalated duct, striated duct and the main excretory duct. Arrows indicate Myoepithelial cells. Image created in BioRender.com.

1.1.1 Major salivary glands

- **Parotid glands**

The parotid gland is the largest gland in the oral region in adults. Its weight ranges from 15-30 g. It lies anterior to the ear and inferior to the mandible. It is triangular in shape, contained in a fine capsule and divided into a superficial upper lobe and deep lower lobe by the facial nerve. The secretions of the parotid gland flow through a main duct orifice called the Stenson's duct. It is 7cm in length, and opens into the oral cavity on the buccal mucosa opposite to the maxillary second molars.

The majority of salivary gland neoplasms occur in the superficial lobe and despite the vital structures and important nerves and veins in this area, they are still surgically convenient to access.

- **Histology of parotid glands**

Parotid glands histologically consist of epithelial secretory tissue and protective mesenchyme. A fine uniform fibrous capsule surrounds and protects the gland. The parenchymal tissue separates and divides the gland into lobules. A main collecting duct branches into smaller inter-lobular ducts, which branch into smaller intra-lobular striated ducts. The striated ducts divide into intercalated ducts, which connect into the terminal secretory system of the acini, figures 1.2 and 1.3.

The parotid glands have an entirely serous secretion. The serous acini are composed of triangular shaped cells surrounding the central small lumen, with a prominent and round nucleus at the base. The cytoplasm is basophilic due to the presence of zymogen secretary granules: rough endoplasmic reticulum studded with ribosomes which mainly secrete amylase, largely explains the granular appearance. These secretory zymogen granules stain positive for Periodic-Acid-Schiff (PAS) and resist diastase digestion. In contrast, they are negative for mucicarmine and Alcian-blue, which explains the watery nature of the serous secretion. The connective tissue consists of plasma cells, that produce immunoglobulin A (IgA), blood vessels and nerves.

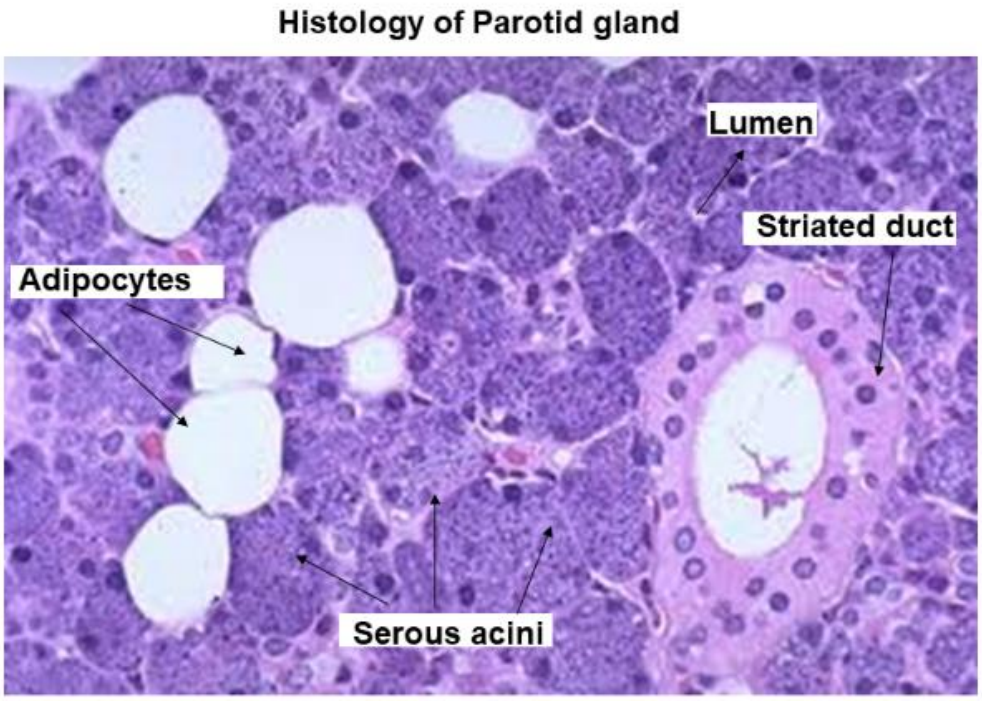


Figure 1.3 Normal parotid gland glandular structure: normal parotid glands with serous acini organised in lobules. Serous acini contain dark granules secreting watery fluid with amylase draining to the ducts and secreting into the oral cavity.

The myoepithelial cells are present at the basal membrane of the acinar cells and intercalated ducts. They are so called because they share histologic features with both epithelial and smooth muscle cells. Myoepithelial cells are flat, stellate cells with cellular processes that extend to surround the acinar cells and intercalated ducts (figure 1.2). The important feature of the cytoplasmic filaments is that they contract to aid saliva secretion. The myoepithelial cells are rarely seen in routine haematoxylin and eosin (H&E) sections, due to an artefact during the fixation process; special staining for alpha smooth muscle actin (α-SMA) is required to detect them. Myoepithelial cells play a significant role in many salivary gland neoplasms and are believed to be responsible for the morphological variability among these tumours (Martinez-Madrigal and Micheau, 1989). This belief is due to the fact that these cells

share both epithelial and mesenchymal characteristics which can change under some neoplastic conditions.

As mentioned above, there are three types of secretory ducts, the first being the intercalated duct, which is longer and quite distinct in the parotid glands (although can sometimes be difficult to distinguish histological sections), shorter in the submandibular glands and poorly developed in the sublingual glands. Intercalated ducts are lined with simple cuboidal cells that have abundant cytoplasm and regular round nuclei.

The striated duct is a very active in modification of the final ductal secretion, with a diameter is six times larger than the acini cells. It is lined with columnar cells with vertical striations in the cytoplasm. These striations represent basal folds, microvilli, and can be seen under high power with a light microscope. The striated duct cells react more with acid in the H&E, due to the fact that they contain large numbers of mitochondria which have ionic pumping activity to regulate the salivary fluids. The cytoplasm is eosinophilic with centralised and round nuclei.

Finally, the excretory ducts are collecting ducts lined with pseudo-stratified columnar epithelium with low-basal cells scattered in the basement membrane and infrequently with goblet like mucous cells that stain positive with PAS. As the excretory ducts merge into the oral epithelia to secrete the saliva, the lined epithelial changes to a stratified squamous epithelium.

Around the parotid gland are small nodules of lymphoid tissue, sebaceous glands or small clusters of sebaceous cells and adipose tissue. Sebaceous glands may be

present in 10 to 42% of normal parotid glands and secrete a waxy, oily substance (Sebum) that helps to prevent the loss of saliva secretion (Ellis and Auclair, 2008).

- **Submandibular glands**

The second largest salivary gland is the submandibular gland, with weights ranging from 7 to 15 g and located in the submandibular triangle of the neck. The submandibular gland is similar to the parotid gland having one excretory duct, the Wharton's duct. This duct extends about 5 cm along the lingual sulcus with its opening adjacent to the lingual fraenum in the floor of the mouth. The two openings of the duct are located a few milometers from each other.

- **Histology of submandibular glands**

Submandibular glands are surrounded by a fine fibrous capsule. The architecture of the submandibular gland is similar to that of the parotid glands, consisting of acini, intralobular ducts (intercalated and striated), interlobular ducts and main excretory ducts.

The submandibular gland acini are distinct from their parotid counterparts, as they are mixed, with serous acini in the majority, but around 5-10% of acini are mucous secreting. Serous acini sometimes appear, in histological sections, to surround the mucous acini to form a crescent shaped cap at the periphery, called a serous demi-lune (figure 1.4); this is considered by some to be an artefact of the fixation process. The cytoplasm in mucous cells is distinctly basophilic with spherical nuclei. These cells are positive for PAS, Mucicarmine and Alcian blue, due to the mucus production by the acinar cells (sialomucin acids).

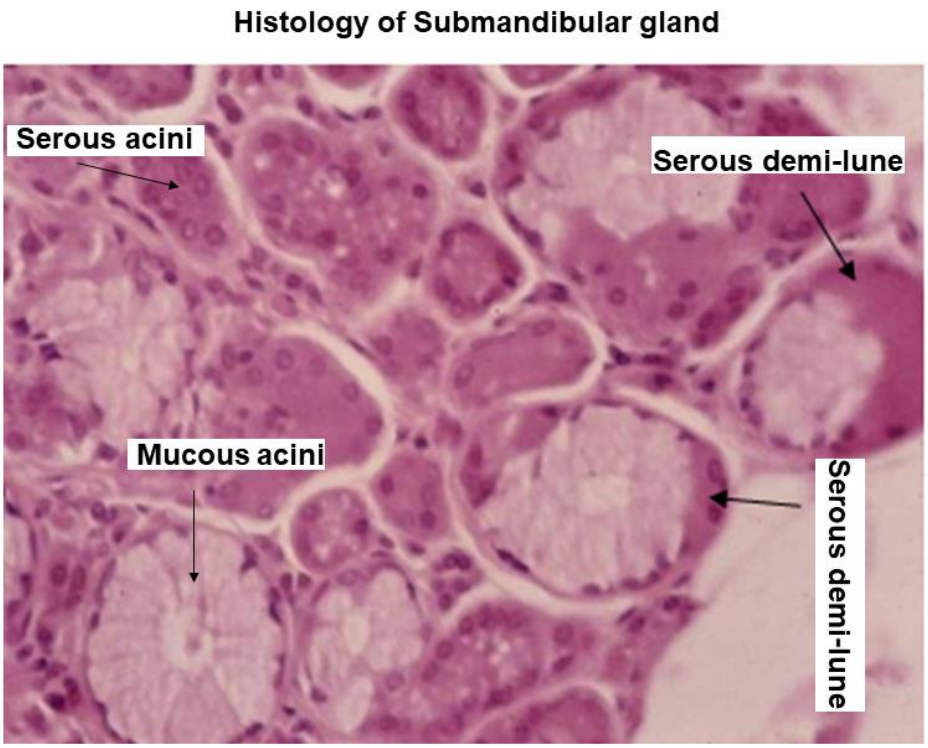


Figure 1.4 Normal submandibular gland structure: with both serous acini and mucous acini containing pale coloured mucin. Black arrows indicate the serous cells surrounding the mucous cells (serous demi-lune). (Image from Dr Daniel J Brierley).

The intercalated duct in the submandibular gland is shorter and less apparent than in the parotid gland, whereas the striated ducts are more evident. Again, the myoepithelial cell is not routinely visible in H&E staining, due to an artefact resulting from the fixation process. Unlike the parotid glands, lymphoid tissue and peripheral nerves are not present in submandibular glands, which could explain the rarity of lymphatic spread of malignant salivary gland neoplasms that develop in this gland. Sebaceous glands are rare in submandibular glands being present in only 5-6% of submandibular glandular structures (Ellis and Auclair, 2008).

- **Sublingual glands**

The sublingual gland is the smallest of the major salivary glands, weighing on average only two to three grams. It is located between the floor of the mouth and the mylohyoid muscle. The sublingual gland, unlike the other major glands, has several excretory ducts of which the main and largest duct, the Bartholin's duct, opens into the mouth through the ducts of the submandibular glands. The other smaller ducts are called Rivinus ducts and open directly into the oral cavity.

- **Histology of sublingual glands**

Sublingual glands are poorly encapsulated and their lobular structures are less organised than in other major salivary glands. The acini in sublingual glands are mixed, with a mucous predominance (figure 1.5). The ends of these secretory units often appear elongated instead of the spherical-like acini found in the parotid and submandibular glands. Both types of intra-lobular ducts, intercalated and striated, are shorter in sublingual glands than in other major glands. The lymphoid tissues and sebaceous glands are not present in the sublingual gland.

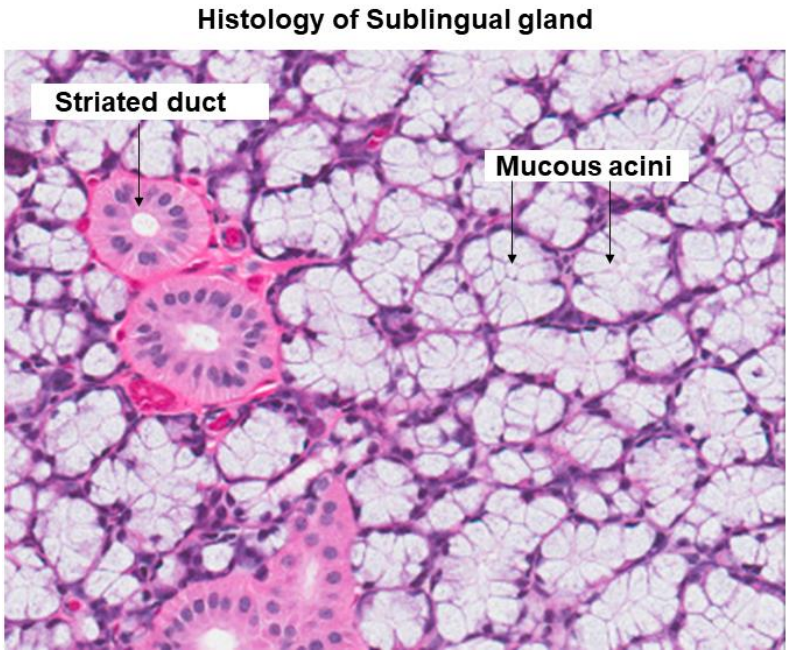


Figure 1.5 Normal sublingual gland structure: mucous acini with pale coloured mucin.

1.1.2 Minor salivary glands

400-800 minor salivary glands are distributed in the oral cavity (Speight and Barrett, 2009) and are located in the mucosa and submucosal layer in the cheeks, lips, ventral, dorsal and lateral surfaces of the tongue, hard and soft palate, and retromolar area. However, the gingivae and anterior part of the hard palate are lacking in, or free from, minor salivary glands; salivary glands tumours are, therefore, unlikely to develop at these anatomical sites.

Minor salivary glands are generally un-encapsulated (figure 1.6), a feature which allows neoplasms in minor salivary glands to metastasise rapidly and grow invasively. The secretion in minor glands is mainly mucous, with the exception of the lingual minor glands, where it is entirely serous (Serous glands of Von Ebner).

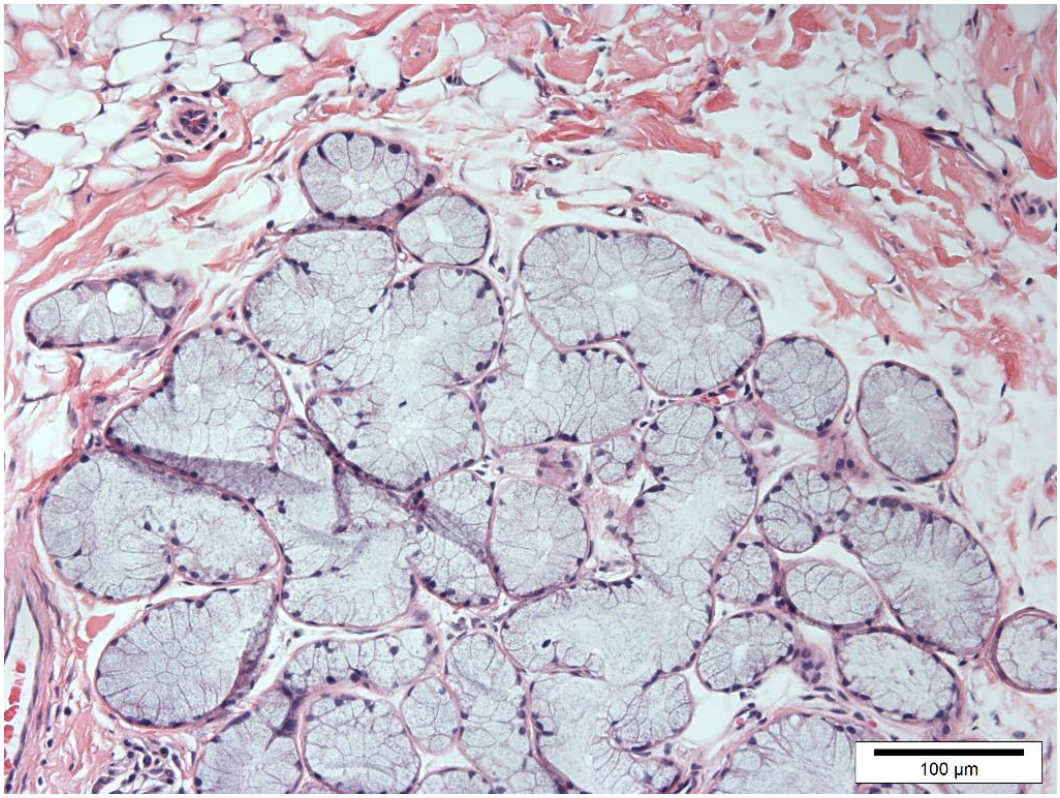


Figure 1.6 Normal minor salivary glands: aggregates of un-encapsulated secretory acini and ducts organised in the submucosa layer of the oral cavity (not, however, in the anterior part of the hard palate and gingivae). Minor salivary gland secretion is mainly mucous.

1.2 Oral cancer

Oral cancers are malignant tumours that involve the structures of the oral cavity. Oral squamous cell carcinoma (OSCC) represents more than 90% of oral cancers (Johnson *et al*., 2011) and is the sixth most common neoplasm worldwide, with more than 500,000 new cases diagnosed annually (Matta and Ralhan, 2009). Despite advances in the management of OSCC, the 50% overall survival rate at 5-year has not improved during the past 3 decades (Sasahira *et al*., 2014).

OSCC can develop at any site in the mouth, although there are specific locations such as the tongue and floor of the mouth where OSCC is most likely to occur. The main

causative factors linked to OSCC are tobacco smoking and alcohol consumption. Surgical biopsy is the only way to provide an accurate diagnosis of OSCC and to determine the grade and stage which are crucial in deciding upon post-surgical treatment and clinical outcome. In recent years, there has been increasing interest in investigating numerous biological and molecular markers for diagnosis and prognosis of OSCC. However, routine clinical examination with comprehensive histopathological investigation of surgical specimens is still the gold standard for providing precise management and predicting the prognosis.

1.2.1 Salivary Gland Tumours

The overall annual incidence of salivary gland tumours is less than 5 per 100,000 in the Western world (Speight and Barrett, 2020) with around 720 cases every year in the United Kingdom (https://www.cancerresearchuk.org/about-cancer/salivary-gland-cancer/about, accessed in October 2020). According to the latest histological classification of the World Health Organisation (WHO) 2017, there are more than forty types of salivary glands tumours, 23 of which are malignant (table 1.1) (El-Naggar *et al*., 2017). Almost 80% of salivary gland neoplasms occur in the parotid glands and 15% arise in the submandibular glands. Only 5% of salivary gland cancers develop in the sublingual and minor salivary glands (Bell and Hanna, 2012). Malignant tumours of salivary glands are a rare form of oral cancer representing only 6% of all head and neck malignancies and less than 0.5% of all cancers in the body (Barnes, 2005; Lin *et al*., 2018).

The occurrence of salivary gland cancer (SGC) is truly site dependent as twenty per cent of parotid gland tumours are malignant, nearly half of submandibular gland neoplasms are malignant but more than 60% of sublingual and minor gland tumours are malignant (Guzzo *et al*., 2010).

Salivary gland carcinomas are a highly heterogeneous and clinically diverse group, which despite their rarity are life-threatening and exhibit various biological behaviours. This complexity of salivary gland tumours renders accurate diagnosis challenging even for a highly skilled pathologist.

The exact causes of salivary gland cancer remain unclear, however, several risk factors have been identified, including exposure to radiation, low-nutrition diet, history of childhood benign salivary gland tumour, infection and genetics (Guzzo *et al*., 2010). Nearly 80% of salivary gland tumours are benign. Pleomorphic adenoma (PA) is the most common salivary gland tumour, while in contrast, mucoepidermoid carcinoma (MEC) is the most common malignant lesion, followed by adenoid cystic carcinoma (ACC).

This review will focus mainly on MEC with brief overviews of PA as it is the most common salivary gland tumour and ACC as it is the second most common salivary gland cancer.

Table 1.1 WHO Histological classification of tumours of the salivary glands (El-Naggar et al., 2017)

Malignant epithelial tumours	Benign epithelial tumours
Mucoepidermoid carcinoma	Pleomorphic adenoma
Adenoid cystic carcinoma	Myoepithelioma
Acinic Cell carcinoma	Basal cell adenoma
Polymorphous adenocarcinoma	Warthin tumour
Clear cell carcinoma	Oncocytoma
Basal cell adenocarcinoma	Canalicular adenoma
Intraductal carcinoma	Sebaceous adenoma
Adenocarcinoma, NOS	Lymphoaenoma
Salivary duct carcinoma	Ductal papillomas 3. Sebaceous 4. Non-sebaceous
Myoepithelial carcinoma	
Epithelial-myoepithelial carcinoma	Ductal papillomas 4. Inverted ductal papilloma 5. Intraductal papilloma 6. Sialadenoma papilliform
Carcinoma ex pleomorphic adenoma	
Secretory carcinoma	cystadenoma
Sebaceous adenocarcinoma	
Carcinosarcoma	
Poorly differentiated carcinoma 4. Small cell carcinoma 5. Large cell carcinoma 6. Undifferentiated carcinoma	**Haematolymphoid tumours**
Lymphoepithelial carcinoma	Hodgkin lymphoma
Squamous cell carcinoma	Diffuse large B-cell lymphoma
Oncocytic carcinoma	Extranodal marginal zone B-cell lymphoma
Uncertain malignant potential	
Sialoblastoma	**Secondary tumours**
	Soft tissue tumours
	Haemangioma

Abbreviations: NOS -not otherwise specified.

1.2.1.1 Pleomorphic adenoma

PA, a benign neoplasm, is the most common salivary gland tumour. The parotid gland is more likely to harbour PA, followed by the minor salivary glands. PA can happen at any age, with peak tendency between the fourth and sixth decades of life and is more common in women than men. There are significant variations in the histological features of these tumours. PA demonstrates a proliferation of mainly ductal and myoepithelial cells (spindle or plasmacytoid shape); in addition, it may also contain cuboidal, squamous, basaloid, oncocytic, and clear cells. All lie in myxoid or chondroid connective tissue stroma, surrounded by a capsule of varying thickness. These variations in histological characteristics make the diagnosis difficult in some cases.

1.2.1.2 Adenoid cystic carcinoma

ACC is the second most common salivary gland cancer and can also develop in any other exocrine glands such as breast, cervix, vulva, sino-nasal tract and tracheobronchial tree (Stenman, 2013). ACC is a slow-growing tumour, although it is an aggressive and life-threatening neoplasm because of the high metastasis rate. In cases of distance metastasis, the survival rate of ACC decreases from 10 years to less than 2-years (Laurie *et al*., 2011). Histologically, ACC is composed of ductal, basal and myoepithelial cells. Additionally, it can exhibit a tubular cribriform and/or solid growth pattern. The primary care for ACC patients is surgery and/or radiotherapy.

1.2.1.3 Mucoepidermoid carcinoma

MEC was initially reported by Volkmann in 1895, and then fully described in 1945 by Stewart *et al.*, who termed it “mucoepidermoid” and classified it as a tumour due to its slow growing it has been thought to have a benign behaviour (Stewart *et al*., 1945).

Later, in 1954, Foote and Frazell created the term “mucoepidermoid carcinoma” and established this as a malignant tumour (Foote and Frazell, 1954). MEC is the most common salivary gland cancer in adults, children and adolescents, and can develop in major and minor salivary glands. It accounts for more than 30% of all salivary gland cancers (Jones *et al*., 2008). MEC most commonly develop in the parotid gland but have also been documented in submandibular and minor salivary glands, more frequently in the palate, retromolar area, cheek, tongue and lips (Speight and Barrett, 2009).

1.2.1.3.1 Histo-pathology of MEC

MEC is a malignant epithelial tumour. Its histopathology mainly comprises three types of cells, including mucous, intermediate and epidermoid cells, which are considered to be the basic cell types of MEC (figure 1.7). Mucous-producing cells usually line the cystic spaces and may be tall columnar or goblet-shaped with obvious cytoplasmic mucin or be more subtle, requiring special stains (PAS/ PASD) to identify the mucin. Intermediate cells, the most frequently seen cell type, concentrate at the tumour periphery and are small with a high nuclear-to-cytoplasmic ratio and oval hyperchromatic nuclei. Epidermoid cells are polygonal with well-defined borders and abundant eosinophilic cytoplasm. In addition to these three cell types, MEC may also include clear, oncocytic or columnar cells (figure 1.8). Changes in the basic cell types of the MEC to oncocytic or clear are thought to be due to differentiation of the intermediate cell type (García *et al*., 2011).

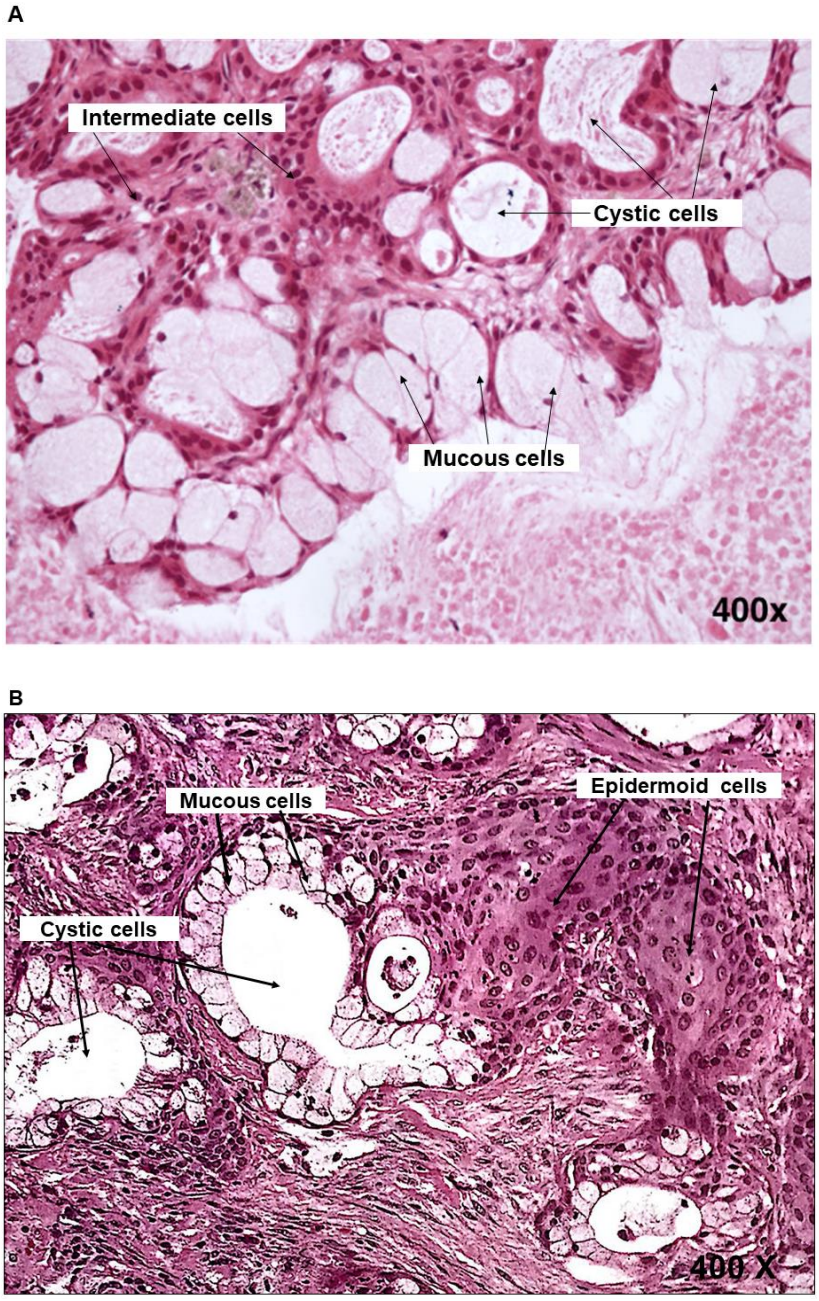


Figure 1.7 Mucoepidermoid carcinoma (H&E) showing variable cyst size and composed of the basic three cell types: mucous, epidermoid and intermediate.

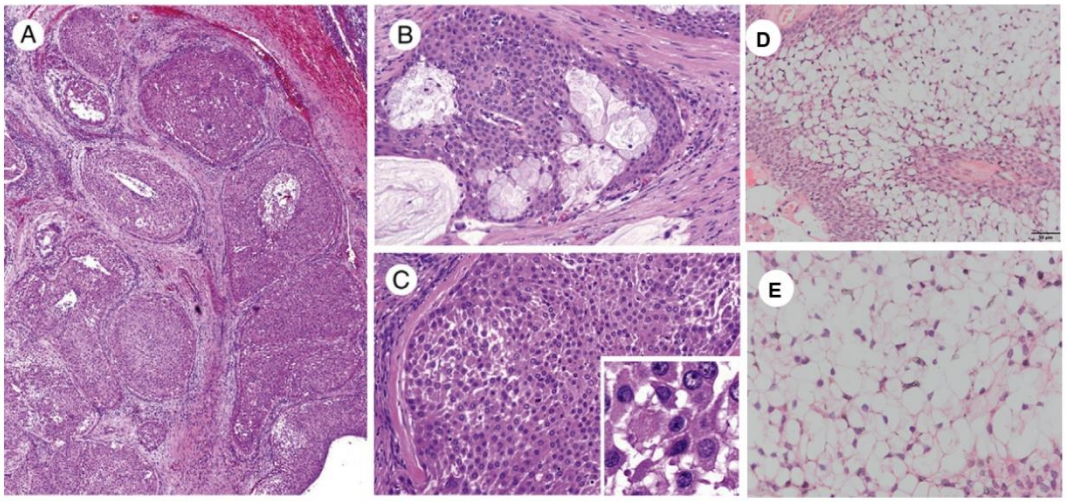


Figure 1.8 Variants of MEC: A-C) Oncocytic MEC showing nests of cells with micro-cystic formation. **D-E)** Clear cell MEC, a mixture of epidermoid cells with clear cell differentiation across the tumour. Images **A-C** were taken with permission from Joaquin *et al.*, 2011.

The most common presentation of MEC is as macro or micro multi-cystic spaces lined by epidermoid or ductal cells, and mucous-secreting cells, with the presence of solid sheets or areas that mainly contain intermediate cells. The cystic formation in MEC is closely linked to low-grade tumours, with more than 70% of low-grade tumours having a higher proportion of mucous cells. In contrast, high-grade tumours typically exhibit a solid pattern with the presence of intermediate and epidermoid cells. In addition, high-grade tumours exhibit cellular atypia, necrosis, perineural invasion and an increase in mitotic rate. Further features that may be seen in rare cases of MEC include focal keratinisation and squamous carcinoma-like areas. Infiltration of lymphocytes, meanwhile, is mainly seen in cases of MEC in parotid or minor glands.

In terms of the pathogenesis of MEC, it is thought to originate from scattered mucous cells in excretory ducts. However, there is still no evidence that the MEC originated from a single cell type. In normal salivary gland tissues, these mucous cells are scant and not uniformly existent in the lining of major excretory ducts. The frequency of

these cells tends to decrease along the intercalated ducts and striated ducts and they have not been seen within intralobular ducts. Under certain inflammatory conditions, such as chronic interstitial sialitis, the number of mucous cells increases, which may confirm the origin of the MEC (Foote and Frazell, 1954).

1.2.1.3.2 Mucoepidermoid carcinoma grades

MEC is graded as low (well differentiated), intermediate or high (poorly differentiated; figure 1.9). At present, there are several grading systems. The most universally used grading systems are those of the Armed Forces Institute of Pathology (AFIP), Brandwein, and the modified Healey system (Seethala, 2009) (table 1.2). MEC is the only salivary gland cancer to have point-based (quantitative) grading systems: AFIP and Brandwein, with the AFIP grading system agreed on and recommended by WHO since 2005. However, the utility of these grading systems in the clinic is still restricted, largely due to the high heterogeneity within and between different types. In addition, the AFIP scheme tends to under-grade the tumours, whereas the Brandwein system over-grades them (Seethala *et al*., 2008).

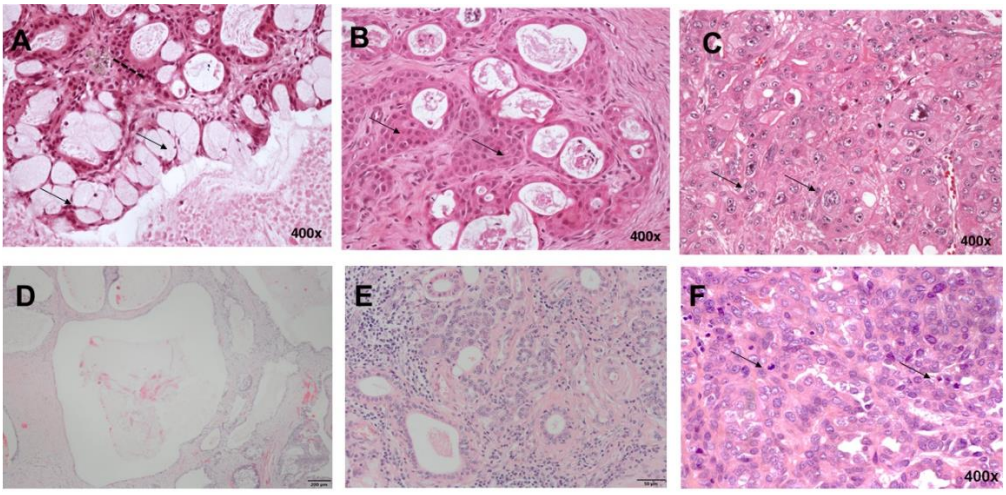


Figure 1.9 Haematoxylin and Eosin staining of MEC cases. Low grade MEC consists of large proportions of mucous-secreting cells (arrows), intermediate cells (dashed lines) and large cyst formations **(A and D)**. Intermediate grade MEC showing small size cyst with a mixture of intermediate cells, epidermoid cells (arrows) and lower proportions of mucous cells **(B and E)**. High-grade MEC showing solid growth of epidermoid and intermediate cells with lack or little presence of mucous cells, lack of microcytic formation and the presence of cellular atypia **(**arrows in **C)** and mitotic figures **(**arrows in **F)**.

Several studies have linked the grading system in MEC carcinoma with prognosis, claiming that low-grade MEC cases have good prognosis and high-grade cases have poor clinical outcomes. One of the main obstacles is the intermediate grade, which does not seem to have any prognostic value as this grade can behave either indolently or aggressively.

1.2.1.3.3 Mucoepidermoid carcinoma immunohistochemical markers

Different markers, in various ratios, are expressed by MECs, including several classes of membrane-bound mucins (MUC1, MUC4, MUC5AC and MUC5B). The expression of MUC1 is closely associated with high histological grade and poor prognosis such as high recurrence rate, metastasis and short survival rates, whereas

MUC4 expression is linked to histologically low-grade carcinoma and better prognosis. A recent study was conducted by Robinson and colleagues (2020) to investigate, using immunohistochemistry, the expression pattern of MUC1, MUC2, MUC4, and MUC5AC in MEC and their correlation with the tumour's grade, based on 15 cases each of low, intermediate and high grade. Their results indicated that the expression of MUC1 was significantly associated with high-grade tumours, whereas MUC4 expression was linked to low-grade tumours. MUC5AC expression was not correlated with any tumour grade and MUC2 was only expressed in a single case and so could be used as an indicative marker to exclude MEC in diagnosis. Cytokeratin 14 and p63 are expressed by epidermoid and basal cells, while strong expression of proliferating cell nuclear antigen (PCNA) and p53 is related to high-histological grade MEC (Namboodiripad, 2014). Also, intense staining of some inflammatory cytokines occurs in MEC, such as tumour necrosis factor alpha (TNF α) and deletion in malignant brain tumours 1(DMBT1), which are thought to act as tumour suppressor genes (Namboodiripad, 2014).

Table 1.2 Comparison of Grading Systems for Mucoepidermoid Carcinoma. (Seethala, 2009)

Modified Healey Qualitative	AFIP Point based	Brandwein Point based
Low grade • **Macrocysts, microcysts, transition with excretory ducts** • **Differentiated Mucin producing Epidermoid Cells, often in a 1:1 ration; minimal to moderate intermediate cell population** • **Daughter cyst proliferation from large cysts** • **Broad-front, often circumscribed invasion** • **Pools of extravasated mucin with stromal reaction** **Intermediate grade** • **No macrocysts, few microcysts, solid nests of cells** • **Large duct not conspicuous** • **Slight to moderate pleomorphism, few mitoses, prominent nuclei and nucleoli** • **Invasive quality, usually well-defined and uncircumscribed** • **Chronic inflammation at periphery, fibrosis separates nests of cells and groups of nests** **High grade** • **No macrocysts, predominantly solid but may be nearly all glandular** • **Cell constituents range from poorly differentiated to recognizable epidermoid and intermediate to ductal type adenocarcinoma** • **Considerable pleomorphism, easily found mitoses** • **Unquestionable soft tissue, perineural and intravascular invasion** • **Chronic inflammation less prominent, desmoplasia of stroma may outline invasive clusters**	Intracystic component <20% = 2pts Neural invasion present = 2pts Necrosis present = 3pts Mitosis (4or more per 10 HPF) = 3pts Anaplasia = 4pts **Low grade** = 0-4 pts **Intermediate grade** = 5-6 pts **High grade** = 7-14 pts	Intracystic component <25% = 2pts Tumour invades in small nests and islands = 2pts Pronounced nuclear atypia = 2pts Lymphatic and/or vascular invasion = 3pts Bony invasion = 3pts >4mitoses per 10 HPF = 3pts Perineural spread = 3pts Necrosis = 3pts **Low grade** = 0 pts **Intermediate grade** = 2-3 pts **High grade** = 4 or more pts

Abbreviation: pts = points

1.2.1.3.4 Management of MEC

The first line treatment for MEC patients is surgical resection, with or without radiotherapy in cases of close margin, nodal invasion or perineural involvement. Treatment in cases with local advanced lesions, recurrence of tumour or distant metastases is currently limited, largely symptomatic, with no targeted therapy currently being offered. The five-year survival rate ranges between 86% and 22 % for low-grade and high-grade tumours respectively (Sood *et al*., 2016). Although the low-grade tumour is usually indolent, it can occasionally be aggressive (Ghosh-Laskar *et al*., 2011). Perineural and lymphovascular invasion is not unusual in intermediate and high grade MECs and the incidence of lymphatic metastasis is 40% in intermediate and high-grade tumours (Sood *et al*., 2016). Therefore, according to National Multidisciplinary Guidelines, UK, in the management of salivary gland disease, the histological grade of MEC is an important factor that should be considered in outcome and treatment plans. Despite the importance of MEC, there remains a paucity of evidence on the limitations of systematic treatment and this situation will only be improved by gaining a full understanding of the molecular pathology of this carcinoma.

1.3 Fusion proteins

Fusion proteins are created by joining two or more genes which originally coded for separate proteins (Fredriksson and Bülow, 2001). Chromosomal rearrangement, the switching of genes between two or more chromosomes, results in disruption of the biological function for which these genes were originally encoded (Aman, 1999).

At present, at least 400 fusion oncogenes have been reported in human neoplasms, representing about 20% of human carcinomas (Brenner and Chinnaiyan, 2009). The fusion oncogenes can originate from, or be encoded by, different factors including: 1)

Transcription factors; 2) Transcription regulators; 3) Tyrosine kinases receptors (Bell and El-Naggar, 2013). They have the potential to be involved in tumourigenesis via one of two actions: upregulation of one of the genes resulting in deregulation of the cell cycle or fusion of two genes at the breakpoint and creation of a chimeric oncogene. The rearrangement could occur between the promoter regions of the two genes or between the promoter region of one gene and the non-coding segment of the other gene. Also, either could have the functional domain or normal protein (Aman, 1999).

The first report of a fusion protein in relation to a solid tumour was in papillary thyroid carcinoma (PTC), known as H4-RET (histone 4-RET proto-oncogene) and involved the tyrosine kinase inhibitor and neurotrophic receptor kinase1 (NTRK1). Subsequently a fusion protein in renal cell carcinoma was found to involve the transcriptional factors 3 and B (TFE3 and TFEB), (Argani and Ladanyi, 2003). Recent evidence suggests that these chromosomal rearrangements are involved in the initiation and early stage of these tumours (Stenman, 2005).

Multiple studies have investigated the association between salivary gland tumours and specific, non-random, recurrent chromosomal translocations (table 1.3). The consequence of these translocations is the creation of fusion proteins that are thought to be involved in the pathogenic and molecular pathology of the tumours. As previously stated in this review, more than forty types of morphologically diverse salivary gland tumours have been recognised by the World Health Organisation (WHO, 2017; El-Naggar *et al*., 2017). Such heterogeneity has, however, caused difficulties for pathologists in diagnosing and studying tumours. Therefore, studying the chromosomal disarrangements could play an important role in addressing the issue of inconsistent diagnosis of salivary gland neoplasms.

This review focuses mainly on the CRTC1-MAML2 fusion but highlights briefly the PLAG1, HMGA2 and MYB-NFIB fusions as they are the most frequent translocations found in PA and ACC.

Table 1.3 Fusion oncogenes in salivary gland neoplasms

Tumour	Cell of origin	Site of tumour	Pathogenic translocation	Fusion oncogene
Pleomorphic adenoma	Epithelial and myoepithelial	Salivary gland	t(8;12)(q12;q14-15)	PLAG1-HMGA2
Mammary analogue secretory carcinoma	Epithelial	Salivary gland	t(12; 15) (p13; q25)	ETV6-NTRK3
Hyalinizing clear cell carcinoma	Epithelial	Salivary gland and oral cavity	t(12; 22) (q13; q12)	EWSR-ATF1
Adenoid cystic carcinoma	Epithelial and myoepithelial	Salivary gland, breast	t(6; 9)(q22–23; p23–24)	MYB-NFIB
Mucoepidermoid carcinoma	Exocrine ducts	Salivary gland, lung, breast, skin, oesophagus, pancreas, thyroid and cervix.	t(11; 19) (q21-22; p13)	CRTC1/3-MAML2

Abbreviations: t= translocation, q= chromosome long arm, p= chromosome short arm.

1.3.1 Pleomorphic adenoma

During the last few decades, comprehensive cytogenetic studies have been conducted to detect the specific fusion proteins associated with Pleomorphic adenoma cases.

PLAG1: Pleomorphic Adenoma Gene 1 was the first chromosomal rearrangement detected in a benign tumour (Mark *et al*., 1980). PLAG1 is a member of the PLAG (zinc finger protein) family, which includes PLAG1, PLAGL1 and PLAGL2. PLAG1 and PLAGL2 are architecturally alike and share an oncogenic role based on their ability to transform NIH3T3 cells. PLAG1 is a transcriptional factor located on chromosome 8q12 and is expressed in a limited number of foetal tissues during development including the liver, kidney and lung (Stenman, 2005). It has been reported to inhibit tumour growth and thus act as a tumour suppressor gene. It is also considered an important factor for adenoma carriers and as a target gene in PA.

HMGA2: is a member of the high mobility group (HMG) and is also known as HMGIC. Its involvement has been identified in a number of different mesenchymal benign tumours such as lipoma, fibroadenoma, uterine leiomyoma, hamartomas, angiomyxoma and other soft tissue chondromas. HMGA2 is located on chromosome 12q14-15 and expression is necessary for development of embryonic tissues. Expression becomes low or down-regulated in certain adult tissues, including lung, kidney and normal salivary gland tissue.

PLAG1 and HMGA2 fusion proteins are specific to PA and have not been documented in any other salivary gland neoplasms. The molecular mechanism behind the fusion needs to be determined in order to allow their use as clinical biomarkers.

1.3.2 Adenoid cystic carcinoma

MYB-NFIB fusion: ACC is associated with a specific recurrent translocation that fuses the MYB oncogene to the NFIB transcriptional factor gene.

MYB is one of the transcriptional regulator family of proteins and has three domains; the N-terminal DNA-Binding domain, a transcription activator domain situated in the middle, and the C-terminal transcriptional suppressor domain. MYB has a crucial function in cell growth and differentiation and it is up-regulated during cell proliferation and division, whereas it is expressed at low level in fully differentiated cells. In ACC an increase occurs in the expression of MYB-NFIB fusion protein. The transcriptional suppressed domain of MYB is exchanged with single or multiple terminal exons of NFIB. As a result of this fusion, the DNA binding domain and the transcriptional activated domain remain active and may activate various MYB target genes such as BCL2, KIT, CD34, BIRC2, MYC and MAD1L1, resulting in development of a neoplasm (Stenman, 2013).

Data from several studies suggest that no fewer than 80-90% of ACC are positive for MYB-NFIB, while even in fusion negative cases, there is still an increase in MYB expression (Stenman, 2013). To date, a number of studies have confirmed that the MYB-NFIB fusion has not been documented in any other salivary gland tumour, which clearly indicates its specificity in ACC.

Expression of MYB-NFIB can be detected by reverse-transcription polymerase chain reaction (RT-PCR) and Fluorescence in situ hybridization (FISH) analysis or by Immunohistochemical (IHC) staining of MYB proteins. More research needs to focus on targeting MYB and its up-regulated genes to generate novel therapies and enhance the survival rates and outcome for patients suffering from this carcinoma.

1.3.3 Mucoepidermoid Carcinoma

1.3.3.1 CRTC1-MAML2 fusion

Several cytogenetic analyses have indicated that more than 50% of salivary gland MEC are associated with a CRTC1-MAML2 fusion protein (Anzick *et al*., 2010, Shinomiya *et al*., 2016). In addition to the salivary glands, CRTC1-MAML2 fusion has been identified in MEC of other tissues including lung, breast, cervix, thyroid and skin (Amelio *et al*., 2014). The CRTC1-MAML2 fusion is expressed by all three cell types of MEC: epidermoid, mucous and intermediate cells (Namboodiripad, 2014). In 1994 Nordkvist *et al*. were the first to report the presence of CRTC1-MAML2 fusion protein in MEC (Nordkvist *et al*., 1994), followed in 2003 by Tonon *et al*. who fully described this translocation (Tonon *et al*., 2003).

The CRTC1-MAML2 fusion is a nuclear protein and is composed of exon one of the CRTC1 gene fused to exons 2-5 of MAML2. CRTC1 is a CREB Regulator Transcriptional Coactivator, also known as MECT1 (Mucoepidermoid Carcinoma Translocated gene 1), or TORC1 (Transducer of Regulated CREB Activity 1), and previously known as WAMTP1 (Warthin and Mucoepidermoid Tumour Translocated Partner 1).

CRTC1 gene is located at 19p13 and encodes for a 75 KDa protein that has the N-terminal CREB (cAMP cyclic Adenosine Monophosphate Response Element Binding mediated transcription) binding domain (CBD) and the C-terminal transcriptional activation domain (TAD; figure 1.10). CRTC1 is one of three members of the CRTC family (CRTC1, CRTC2 and CRTC3) and CRTC1 shows 32% homology with the CRTC2 and CRTC3 isoforms.

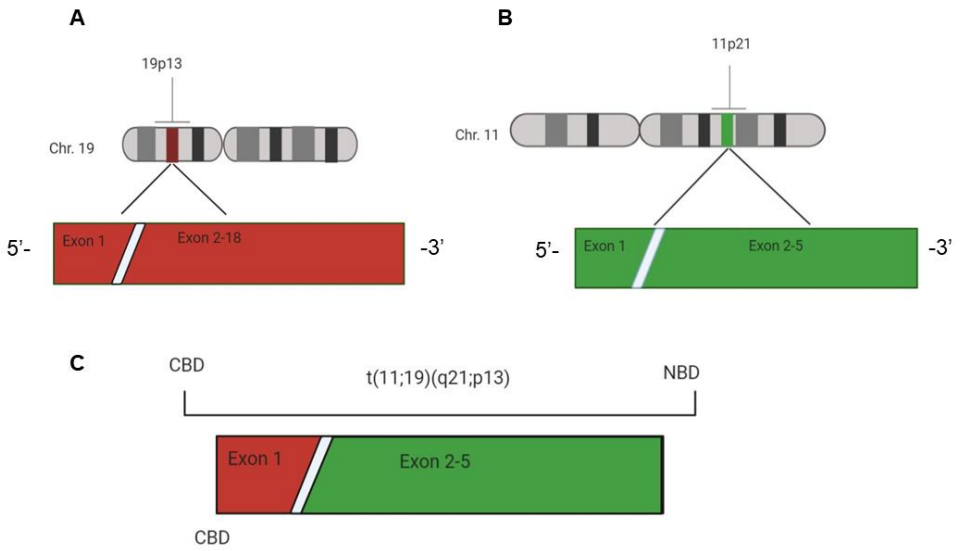


Figure 1.10 Schematic illustration of oncogenic fusion transcripts in mucoepidermoid carcinoma: functional domains in the encoded proteins are CBD, CREB-binding domain; TAD transactivation domain; and NBD, Notch binding domain. Figure adapted from Stenman *et al*., (2014).

CRTC function is essential for the activation of CREB in the cAMP signalling pathway (figure 1.11). CRTC1 is normally expressed in the cytoplasm where it is phosphorylated by AMP-related kinase and inactive, however, an increase in the concentration of calcium or cyclic adenosine monophosphates (cAMP) intracellularly, stimulates the dephosphorylation of CRTC1 through the activation of calcineurin and protein kinase A respectively. This induces the translocation of CRTC1 to the nucleus to bind with the transcriptional partner (CREB) and induce the transcription of CREB target genes (Altarejos and Montminy, 2011).

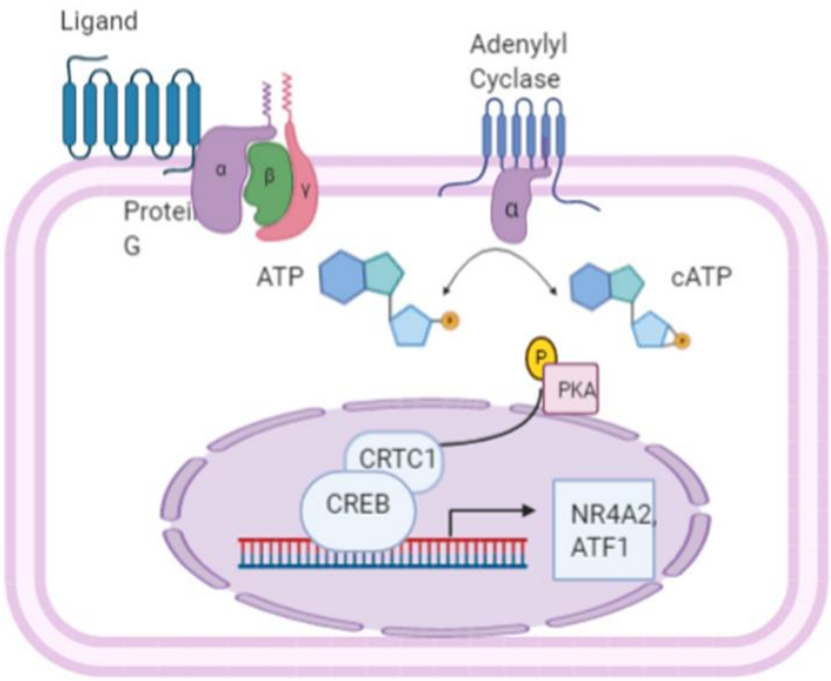


Figure 1.11 cAMP signalling pathway. Image created in BioRender.com.

CRTC1 also plays essential roles in metabolic activity, aging and cancer development (Coxon *et al*., 2005, Komiya *et al*., 2006). The expression pattern of the CRTC1 transcript seems to vary across tissues, but it is mostly expressed in the brain with expression in normal salivary gland being modest (figure 1.12). The metabolic function of CRTC1 is mainly related to its tissue expression, thus its high expression in the brain contributes to various functions including memory, neuronal plasticity, dendritic growth and suprachiasmatic circadian clock (Schumacher *et al*., 2016). CRTC1 protein was found in the submandibular salivary glands of mice during initiation of the morphogenesis of epithelial tissue, but disappeared during acinar maturation returning during the early stage of tumour formation (Jaskoll *et al*., 2011).

Besides its metabolic functions, there is now data inferring that CRTC1 is an emerging promoter of tumorigenesis in different tissues. Reports have suggested that

constitutive activation of CRTC1 is required to stimulate LKB1 loss-driven carcinogenesis in epithelia such as lung and oesophagus (Feng *et al*., 2012, Gu *et al*., 2012) and enhances the activity of human and mouse colonic tumours (Schumacher *et al*., 2016).

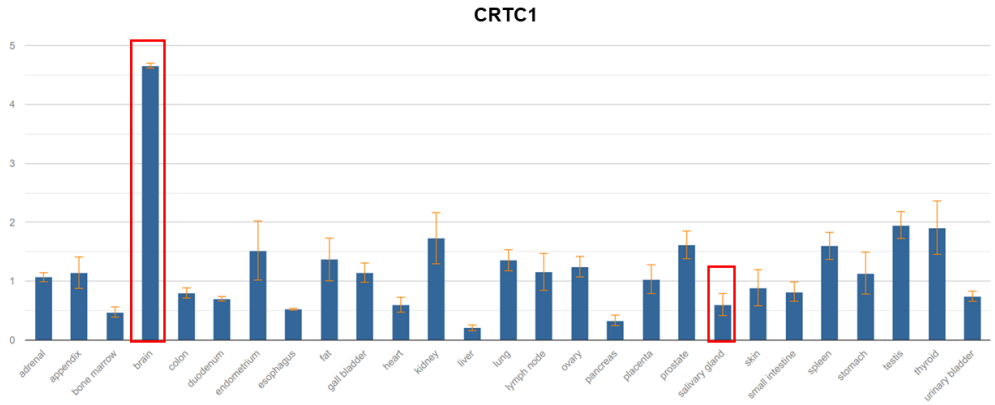


Figure 1.12 Expression level of the CRTC1 RNA in normal human tissue. RNA-sequence was performed on tissue samples from 95 human individuals representing 27 different tissues in order to determine tissue-specificity of all protein-coding genes. Data analysed in April (2018) and accessed in October 2020. Data from National Centre for Biotechnology Information (NCBI website).
https://www.ncbi.nlm.nih.gov/gene/23373/?report=full_report

Some studies have reported the presence of CRTC3 as an alternative fusion partner to the MAML2 gene. Data from Fehr and colleagues (2008) demonstrated 71% CRTC1-MAML2 fusion in 66 MEC cases with only one case showing a CRTC3-MAML2 fusion gene. Nakayama *et al.* (2009) showed 34% of 101 MEC cases studied had a CRTC1-MAML2 fusion and 6% a CRTC3-MAML2 fusion. Moreover, a further study, by Okumura *et al*. (2011), detected CRTC1-MAML2 fusion in 34% and CRTC3-MAML2 fusion in 5% of the 111 MEC studied. Based on these studies, we suggest

that the CRTC3 fusion partner is a rarely occurring event in MEC; this study, therefore, focused only on CRTC1 as a fusion partner with MAML2.

MAML2 is a member of the master mind-like family, functioning as a transcriptional co-activator in the Notch signalling pathway (figure 1.13). MAML2 is located at 11q21 and encodes for a 125 KDa nuclear protein that has the N-terminal of Notch binding domain (NBD) and the C-terminal of the transcriptional activation domain (TAD; figure 1.10). Again, the MAML2 transcript is expressed in a wide number of tissues, with high expression in the brain, placenta, spleen, lung and fat tissue, but modest in salivary gland tissue (figure 1.14).

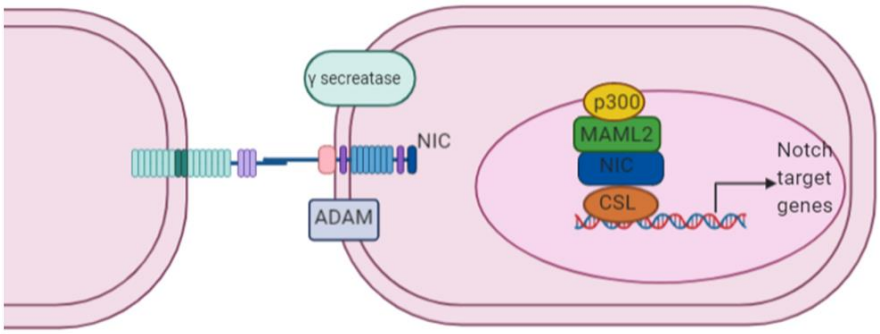


Figure 1.13 Notch signalling pathway. Image created in BioRender.com.

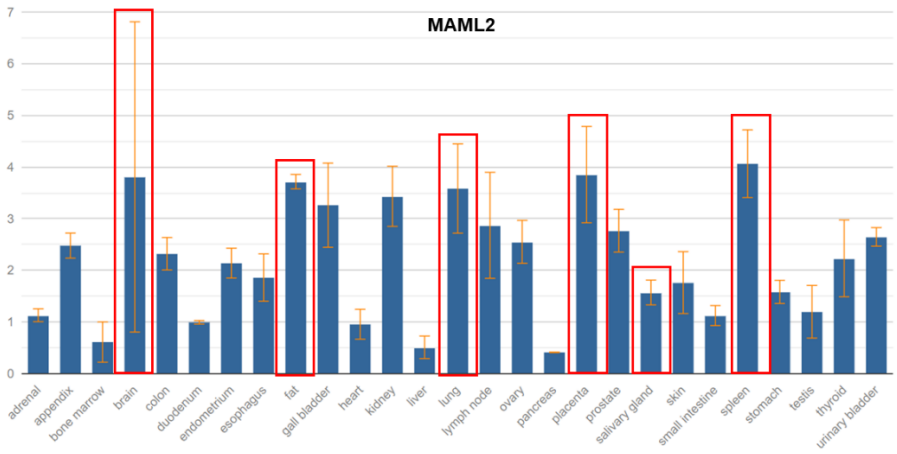


Figure 1.14 Expression level of the MAML2 RNA in normal human tissue. RNA-sequence was performed on tissue samples from 95 human individuals representing 27 different tissues in order to determine tissue-specificity of all protein-coding genes. Data analysed in April (2018) and accessed in October 2020. Data from National Centre for Biotechnology Information (NCBI website). https://www.ncbi.nlm.nih.gov/gene/23373/?report=full_report.

The breakpoint almost always takes place in the first intron of both genes (CRTC1 and MAML2; figure 1.15) (Chen *et al*., 2015). The individual CRTC1 and MAML2 genes do not have any oncogenic activity that leads to transformation of normal epithelial cells. The outcome of the fusion is the creation of an oncogenic protein that leads to replacement of the Notch Binding Domain in the MAML2 by the CREB binding Domain in CRTC1, (figure 1.10). The consequence of the translocation is fusion of 42 amino-acids of CRTC1 with 981 amino-acids of MAML2, which could lead to stimulation of both CREB and Notch signalling targets, resulting in disruption of cell cycle differentiation and regulation (Gupta *et al*., 2015).

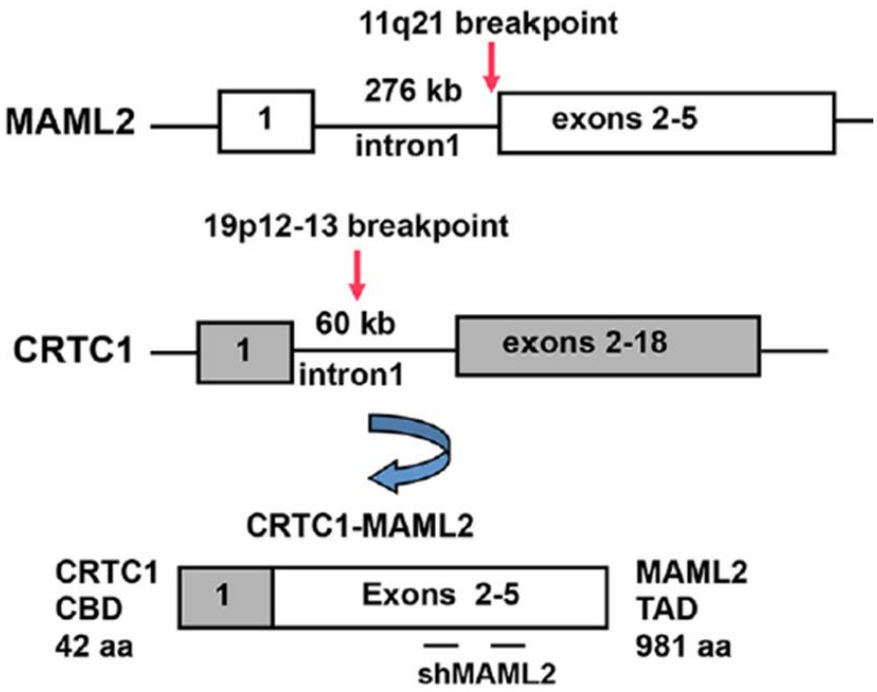


Figure 1.15 The CRTC1-MAML2 translocation: break apart point between intron 1 at 276 KB in MAML2 gene and intron 1 of CRTC1 at 60 Kb in CRTC1 gene. Copyright © 2015, Chen *et al*.

CRTC1-MAML2 fusion functions as a transcriptional factor that is activated mainly through alterations in gene expression. The CRTC1-MAML2 fusion oncogene consistently stimulates the CREB-signalling pathway through interaction of the CRTC1-MAML2 with CREB via CRTC1 CBD-domain. Also, it stimulates the CREB-mediated transcription via the MAML2 TAD-domain (Wu *et al*., 2005). However, recent studies have demonstrated the CRTC1-MAML2 fusion to develop CREB-independent activities via interaction with different coactivator factors including AP1 and MYC (Canettieri *et al*., 2009, Amelio *et al*., 2014). This could confirm that CRTC1-MAML2 leads to tumourigenesis through a number of different gene regulating signals.

A new gain-of-function interaction between CRTC1-MAML2 fusion and Myelocytomatosis (MYC) oncogene was shown to occur due to the interaction between CREB transcription and MYC, resulting in cellular transformation in epithelia (Amelio *et al*., 2014). This interaction could lead to development of a promising new anticancer therapy. However, the exact mechanism of the interaction between the fusion protein and MYC oncogene needs to be defined.

Chen *et al*. (2015) analysed the gene expression profile of CRTC1-MAML2 fusion in human MEC cells as well as knocking-down the CRTC1-MAML2 fusion. They found that the CRTC1-MAML2 fusion stimulated transcriptional activity in MEC cells and regulated differential gene expression in both CREB-dependent and -independent pathways (Chen *et al*., 2015).

Amphiregulin (AREG) is an epidermal growth factor receptor ligand (EGFR), which is a transmembrane protein that binds to EGFR and activates the signalling pathway. AREG is target gene in the cAMP/CREB pathway that is reported to be over-expressed in fusion positive MEC cases and suppressed in fusion negative cases (Shinomiya *et al*., 2016). Shinomiya *et al.* studied 23 fusion-positive MEC and found that there was a significant correlation between the presence of CRTC1-MAML2 and AREG and this in turn was associated with a favourable prognosis. However, they found no correlation between fusion positive cases and expression of EGFR (Shinomiya *et al*., 2016).

Chen *et al*. (2014) have suggested that expression of AREG in MEC could be related to its tumourigenesis and may have a crucial role to play in the initiation and maintenance of fusion positive MEC cancer cells via the AREG-EGFR signalling pathway (Chen *et al*., 2014). AREG expression in MEC positive cases, detected by

IHC, could prove a useful prognostic marker as FISH and RT-PCR are expensive and time consuming for routine use in the clinic. Further systematic research is needed to establish convenient prognostic biomarkers.

Kang *et al*. (2017) conducted a whole-exome sequence of 18 MEC salivary glands and found that TP53 (Tumour suppression gene) was the most commonly mutated gene and that the mutation was present only in high and intermediate grades. In addition, cases with the TP53 mutation had more types of mutations, such as Pou6f2, than those without; this, accordingly, could demonstrate the association between these mutations and advanced histological grades. In the same study, the CRTC1-MAML2 fusion was detected in 88% of MEC cases and thus it was concluded that TP53 and CRTC1-MAML2 fusion could together be an early major cause of MEC (Kang *et al*., 2017). Nevertheless, such a conclusion is hard to apply clinically as a prognostic outcome and more research is needed to establish this.

1.3.3.2 CRTC1-MAML2 fusion and grading

Several studies have indicated that the fusion protein is strongly associated with low and intermediate grades of MEC patients and less often linked to high-grade MEC patients (Seethala *et al*., 2010). Surprisingly, high-grade MEC cases frequently show loss of the CDKN2A tumour suppressor gene. The latest genome-wide array-CGH research suggests that MEC could be subdivided into the following categories: 1. Fusion positive MEC cases of low and intermediate grades; 2. Fusion positive MEC cases of high grade with or without CDKN2A; 3. Fusion negative MEC cases of high grade with multiple genomic imbalances (Stenman *et al*., 2014). However, such sub classification lacks the inclusion of some of the fusion negative low and intermediate cases. In addition, classification of fusion negative high grade MEC needs to be further refined.

1.3.3.3 Detecting the expression of CRTC1-MAML2 fusion

The gold standards for detecting the expression of CRTC1-MAML2 fusion are either the use of RT-PCR to detect the transcript or FISH for the break apart MAML2 translocated partner at a DNA level. The detection of the MAML2 translocation in MEC cases using FISH could assist with the correct diagnosis, particularly in certain cases of clear cell or oncocytic variants. As an example, the study by Garcia *et al.* (2011) successfully distinguished the MEC oncocytic variant in 14 cases from other oncocytic lesions through detection of the CRTC1-MAML2 fusion. In cases of low-grade MEC FISH could allow accurate differentiation from other cystic conditions and the existence of a MAML2 translocation could help to exclude metastatic squamous cell carcinoma or salivary duct carcinoma from MEC (García *et al*., 2011). Although the FISH technique has been widely applied in FFPE tissue samples, much uncertainty still exists about its' usefulness in clinics. Additionally, a failure rate of about 40% has been reported (Just *et al*., 2012). Preservation of fresh frozen fixed tissue samples is needed to perform RT-PCR.

1.3.3.4 CRTC1-MAML2 fusion and prognosis

A much-debated question is whether CRTC1-MAML2 fusion positive cases do better in terms of having long-term disease-free survival. This is reflected in discrepancies in the literature regarding the correlation between fusion positive cases and favourable prognosis. Table 1.4 summarises the data from a variety of independent studies carried out from 2006-2020.

As mentioned previously, the CRTC1-MAML2 fusion protein is associated with low and intermediate grades, and this in turn has been linked with better survival of up to 10 years compared to the fusion negative 1.6 years (Behboudi *et al*., 2006). Most studies favour the correlation between fusion positive cases and good prognosis.

However, the number of cases studied varied considerably, making it hard to achieve an adequate comparison. In addition, fusion could occur in a high-grade tumour and this might not produce the expected prognosis for the patient.

A recent systematic review was conducted to examine the prognostic value of the CRTC1-MAML2 fusion in MEC. Seven out of the ten studies included in this review were significantly associated with a favourable clinical outcome (Pérez-de-Oliveira *et al*., 2019). However, only five studies out of the ten were included in the quantitative analysis.

All of the studies reviewed so far have limitations. Firstly, all are retrospective, with analysis of archive stored cancer tissue, so the quality of the tissue (RNA and proteins) differs from case to case, unless frozen tissue with better quality RNA and proteins is used. There is also a lack of clarity regarding the analytical and technical approaches of each study, as the rarity of salivary gland carcinomas and the lack of an evidence-based study could lead to discrepancies. Moreover, not all of these studies used the same quantitative analytical approaches to examine the correlation between the fusion positive cases and good prognosis. Finally, since four of these studies were from the same centre (Department of Pathology, Nagoya City University Graduate School of Medical Sciences, and Department of Maxillofacial Surgery, Japan), we believe that there could be replication of the cases included (Miyabe *et al*., 2009, Okabe *et al*., 2006, Okumura *et al*., 2011, Okumura *et al*., 2020).

Table 1.4 Comparisons between the key studies investigating the correlation between fusion positive MEC and clinical significance

Scholars/group	Techniques	Number of fusion +ve/All, %	Prognosis implication
Behboudi *et al*., 2006	FISH & RT-PCR	16/29 (55.2%)	Useful
Okabe *et al*., 2006	RT-PCR	27/71 (38%)	Useful
Miyabe *et al*., 2009	RT-PCR	34/101 (33.6%)	Useful
Seethala *et al*., 2010	FISH & RT-PCR	36/55 (66%)	Useful
Okumura *et al*., 2011	RT-PCR	44/111 (39.7%)	Useful
Schwarz *et al*., 2011	FISH	25/40 (62.5%)	Not useful
Ilic Dimitrijevic *et al*., 2014	RT-PCR	8/20 (40%)	Useful
Saade *et al.,* 2016	FISH	50/90 (55.5%)	Not useful
Shinomiya *et al.,* 2016	RT-PCR	23/33 (69.7%)	Useful
Birkeland *et al.,* 2017	RT-PCR	53/90 (58.9%)	Not useful
Okumura *et al*., 2020	FISH & RT-PCR	90/153 (58.8%)	Useful

Abbreviations: +ve= positive, -ve = negative, % =per cent, No = number, RT-PCR= reverse-transcriptase polymerase chain reaction, FISH = Fluorescence *in situ* hybridization.

1.3.3.5 CRTC1-MAML2 fusion in non-salivary gland MEC

As mentioned previously, MEC can rarely develop in other parts in the body including the lung, thyroid, pleura, penis, lacrimal glands, breast, oesophagus, thymus, pancreas, tonsil and skin. Of note, the CRTC1-MAML2 fusion was also described in all of these non-salivary gland MECs (Bean *et al*., 2019, Costa *et al*., 2015, Eviatar and Hornblass, 1993, Gartrell *et al*., 2015, Goh *et al*., 2017, Hu *et al*., 2018, Liu *et al*., 2015, Moran and Suster, 2003, Roden *et al*., 2014, Stefanou *et al*., 2004). However, most of these studies were case reports with lack of in depth and comprehensive molecular analysis.

Interestingly, these carcinomas share the histological and cytogenetical features of their counterpart salivary MECs; hence, detecting the expression of the CRTC1-MAML2 fusion and understanding the exact role of the CRTC1-MAML2 fusion of salivary glands could assist the diagnosis and treatment of the infrequent MECs of non-salivary gland origin.

1.4 Summary

There is a variety of evidence in the literature, and summarised below, to confirm that the CRTC1-MAML2 fusion oncogene is implicated in the aetiology and pathology of MEC.

1) An *in vitro* study demonstrated that CRTC1-MAML2 fusion is involved in the initiation of clusters of epithelia in RK3E cells and a subsequent *in vivo* study of fusion transformed RK3E led to the development of a subcutaneous mass in immunocompromised mice (Chen *et al*., 2014); 2) Diminution of the CRTC1-MAML2 fusion protein in MEC cells, both *in vitro* and *in vivo,* led to a significant decrease in cell growth and survival (Chen *et al*., 2014). Most studies have only focused on the

effect of the fusion on cancer cell lines and immune-deficient, mouse xenograft models. No previous studies have investigated the effect of the CRTC1-MAML2 fusion protein on normal salivary gland cells.

The CRTC1-MAML2 oncoprotein could drive the tumourigenesis via interaction and co-activation of cAMP/ CREB signalling, leading to transcription of the CREB targets genes; simultaneously, with the absence of the Notch binding domain, it could cause the co-activation of the transcription of Notch target genes independently of the presence of CSL via the CREB binding domain. These transcribed genes might be implicated in cell proliferation, survival and migration.

Unlike the individual CRTC1 and MAML2 genes, the CRTC1-MAML2 fusion translocation results in a nuclear protein without any identified enzymatic activity; therefore, knowing the role of the CRTC1-MAML2 fusion protein and interfering with its downstream targets could be a promising approach for the development of new treatments.

1.5 Hypothesis

Previous data in the literature showed that the presence of the unique t(11;19)(21q-p13) translocation leads to the formation of the CRTC1-MAML2 fusion protein in MEC. We hypothesise that CRTC1-MAML2 translocation is associated with a specific molecular signature that results in specific phenotypic changes in the cell. Detection of this translocation and investigations into the function of this fusion protein would enable a better understanding of the biological effects of *de novo* expression, the pathogenesis of disease, and may have useful diagnostic and therapeutic applications.

1.5.1 Aims

The aims of this project are to investigate the expression of the CRTC1-MAML2 fusion gene in MEC histological tumour samples and the role of the CRTC1-MAML2 in the tumourigenesis of MEC.

1.5.2 Objectives

- Design a BaseScope probe to detect the expression of the CRTC1-MAML2 fusion transcript in formalin-fixed, paraffin-embedded tumour samples.
- Synthesise an *in vitro* tool, CRTC1-MAML2 fusion construct, from a MEC fusion positive cell line.
- Study the downstream transcriptional effects of the CRTC1-MAML2 fusion on normal human salivary gland cells.
- Study the downstream effect of the CRTC1-MAML2 fusion protein on cell behaviour using standard functional assays.

- Develop 3D organoid models of normal human primary salivary gland cells, CRTC1-MAML2 transfected normal human primary salivary gland cells and of established MEC cell lines.

2. Chapter 2 Materials and Methods

2.1 BaseScope Assay

2.1.1 Patient samples and clinical data

A total of 30 clinical cases, diagnosed with MEC of salivary glands were identified from the Charles Clifford Dental Hospital (CCDH), Sheffield, UK from 2007-2016 and also from Santa Casa de Misericordia Hospital, Porto Alegre, Brazil. Clinical data were collected from the patients' records in the Oral and Maxillofacial Pathology Department, University of Sheffield and Santa Casa de Misericordia Hospital. The parameters studied are summarised in table 2.1 and include age, sex, location and tumour pathological grades. All of the haematoxylin and eosin slides were reviewed in detail by two expert histopathologists: Professor Keith Hunter (Head, Academic Unit of Oral Pathology, University of Sheffield, UK) and Dr Vivian Petersen (Post-doctorate staff, Academic Unit of Oral and Maxillofacial Pathology, University of Sheffield, UK). The average time since fixation of the tissue blocks was 6.4 years (range 0–20 years). Histological grades of MEC cases were classified according to the AFIP, (see table 1.2) as recommended by WHO (Barnes *et al*., 2005) and it has high reproducibility (Auclair *et al*., 1992, Schwarz *et al*., 2011). Mrs Hayley Stanhope (histology laboratory technician) cut sections from the tissue blocks onto SuperFrost plus slides (Fisher Scientific, Loughborough, UK) in preparation for further BaseScope analysis.

Table 2.1 Demographic Parameters in MEC Retrospective Cohort

Mean Age*	**50.8 (Range: 14-90)**
Gender: Male: Female ratio	1:2
Anatomical site	
Parotid	15 (50%)
Palate	6 (20%)
Submandibular	4 (13.3%)
Buccal Mucosa	2 (6.7%)
Other	3 (10%)
Pathological grades	
Low	21 (70%)
Intermediate	4 (13.3%)
High	5 (16.7%)

*Age during disease presentation.

Cases were selected based on the adequate availability of the material. Cases with insufficient histological material or referred cases with unavailable histological material were excluded from this study.

2.1.2 Probe design

A specific 1ZZ probe named BA-Hs-CRTC1-MAML2-FJ, targeting nucleotides 143-174 of AY186997.1, was designed by Advanced Cell Diagnostics Inc. (Newark, CA) after the sequence and specifications had been provided. The probe was used on the FFPE sections to target the novel exon-exon junction in the CRTC1-MAML2 fusion transcript by using the BaseScope™ Detection Reagent Kit v2– RED according to the manufacturer's instructions (figure 2.1). All BaseScope reagents are summarised in table 2.2.

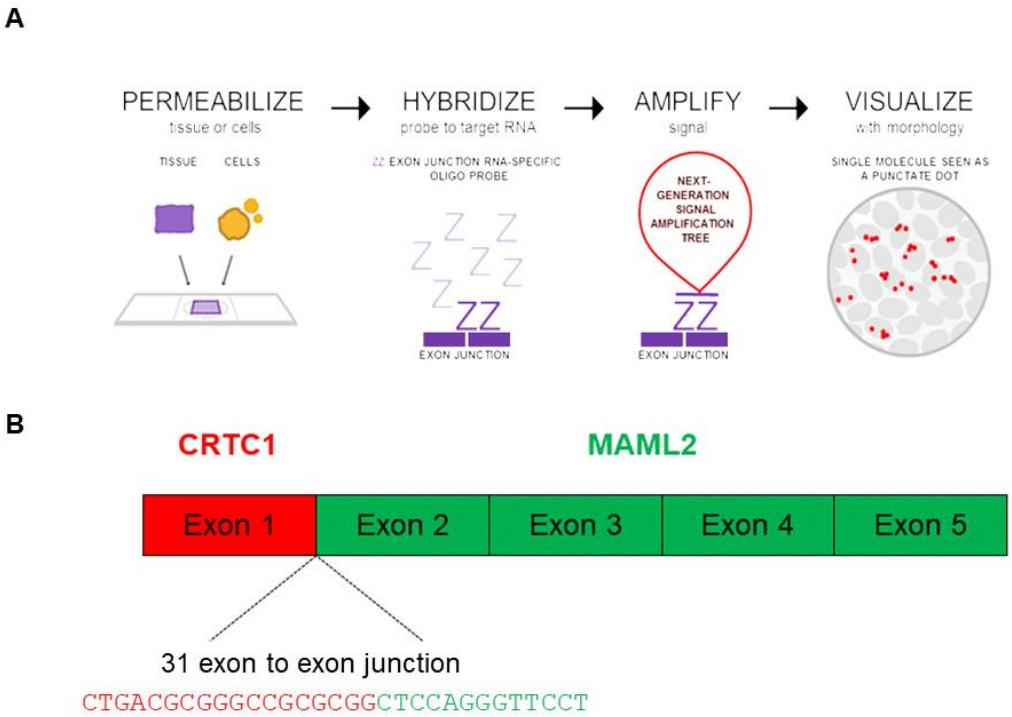


Figure 2.1 BaseScope Assay. Representative diagram of BaseScope assay workflow **(A)**. Representative diagram of BA-Hs-CRTC1-MAML2-FJ BaseScope assay probe design of 31 base pairs sequence cross exon1 of CRTC1 to exon 2 of MAML2 junction targeting 143-174 of AY186997.1 (**B).**

2.1.3 Sample preparation and pretreatment

Briefly, 4 µm sections were cut onto Superfrost plus slides (Thermo Scientific, New Hampshire, USA) and allowed to dry overnight at room temperature. Sections were baked at 60°C for one hour (h) before being deparaffinised in xylene (2 × 5 min) and ethanol (2 × 2 min), then dried by baking at 60°C for 15 min. Pretreat 1 (hydrogen peroxide) was applied for 10 min at RT, Pretreat 2 (target retrieval) for 30 min at 100°C followed by immersing briefly in distilled water for 15 seconds and drying in 100% ethanol for 5 min. Pretreat 3 (protease IV) was applied for 30 min (tissue sections) at 40°C or (protease III) for 15 min (cell pellets), with two rinses in distilled water between pretreatments.

Table 2.2 Reagents in BaseScope Assay

Materials	Supplier	Conditions	Procedure
Xylene	Fisher Scientific, UK	RT, Fume hood	2×, 5 min
Ethanol (100%)	Fisher Scientific, UK	RT, Fume hood	2×, 2 min
Pretreatment 1, RNAScope® Hydrogen Peroxide H_2O_2	Advanced Cell Diagnostics Inc. Newark, CA	RT, On bench	4-8 drops, 10 min
Pretreatment 2, RNAScope® Target Retrieval	Advanced Cell Diagnostics Inc. Newark, CA	100°C, Steamer	4-8 drops, 15 min for cell pellets, 30 min for tumour sections
Pretreatment 3, RNAScope® Protease III or IV	Advanced Cell Diagnostics Inc. Newark, CA	40°C, HybEZ™ Oven	4-8 drops, IV (tissue) 30 min or Protease III (cell pellets) 15 min.
BaseScope™ Target Probe Human (Hs)- -CRTC1-MAML2-FJ-1ZZ. Cat, No. 721321	Advanced Cell Diagnostics Inc. Newark, CA	40°C, HybEZ™ Oven	4-8 drops, 2 h
BaseScope™ Positive Control Probe Human (Hs)-PPIB-1ZZ, Probe targeting common housekeeping gene, Cat. No. 701041	Advanced Cell Diagnostics Inc. Newark, CA	40°C, HybEZ™ Oven	4-8 drops, 2 h
BaseScope™ Negative Control ProbeDapB-1ZZ, Probe targeting bacterial gene dapB, Cat. No. 701021	Advanced Cell Diagnostics Inc. Newark, CA	40°C, HybEZ™ Oven	4-8 drops, 2 h
RNAScope® Wash Buffer	Advanced Cell Diagnostics	RT, on bench using staining dish.	2×, 2 min

	Inc. Newark, CA		
Amplification kits, BaseScope™ v2 AMP 1-8	Advanced Cell Diagnostics Inc. Newark, CA	Varies, 40°C at HybEZ™ Oven or RT.	Varies, 60 min -15 min
Detection kit, BaseScope™ Fast RED-A and B, 1:60	Advanced Cell Diagnostics Inc. Newark, CA	RT not exposed to direct sunlight or UV light.	~120µl per section,10 min
Counter staining, Gill's Haematoxylin I, 50% Haematoxylin: DDH_2O	American Master Tech Scientific	RT, Fume Hood	2 min
Bluing reagent, Ammonium hydroxide, 28–30% (1N)	Sigma-Aldrich	RT, Fume Hood	Immersed for 2-3 times or till sections become blue.
Mounting reagent, BioCare EcoMount. Cat. No. 320409	BioCare Medical	RT, Fume Hood	1–2 drops

*Abbreviations: RT= Room Temperature, Cat. No= Catalogue Number.

2.1.4 Hybridize probe application and amplification

BaseScope target probe together with positive and negative probes were applied to slides placed on a slide holder (RNAScope® EZ-Batch™ Slide Holder) for 2 h at 40°C in a HybEZ™ oven, the slides were then placed in a Tissue-Tek® Slide Rack in a staining dish containing 1× wash buffer and washed 2× for 2 min with agitation. Subsequently, the slides were incubated with amplification BaseScope™ v2 (AMP) reagents 1 to 8 as indicated in table 2.3. The slides were rinsed with wash buffer (2 × 2 min) between AMP incubations. Finally, the slides were incubated with Fast Red BaseScope detection reagents for 10 min at room temperature in the dark. The slides were counterstained with 50% Gill's haematoxylin diluted in distilled water before drying for 15 min at 60° C in a drying oven and mounted using EcoMount permanent mounting medium (Vector Laboratories, Burlingame, CA).

Table 2.3 Amplification stages used in BaseScope assays with application conditions

Amplification Reagents	Application procedure
Amplification Hybridize 1 BaseScope (Amp 1)	40°C at HybEZ™ oven for 30 min
Amplification Hybridize 2 BaseScope (Amp 2)	40°C at HybEZ™ oven for 30 min
Amplification Hybridize 3 BaseScope (Amp 3)	40°C at HybEZ™ oven for 15 min
Amplification Hybridize 4 BaseScope (Amp 4)	40°C at HybEZ™ oven for 30 min
Amplification Hybridize 5 BaseScope (Amp 5)	40°C at HybEZ™ oven for 30 min
Amplification Hybridize 6 BaseScope (Amp 6)	40°C at HybEZ™ oven for 30 min
Amplification Hybridize 7 BaseScope (Amp 7)	RT on the bench for 1 Hour
Amplification Hybridize 8 BaseScope (Amp 8)	RT on the bench for 15 min

Abbreviations: min = minutes

2.1.5 Signals detection and quantification

Sections were examined under a standard bright field microscope (Olympus BX51) at 20-40× magnification. The positive control signal strength was assessed before the sections stained with the target probe were analysed. A case was considered positive when punctate red dots were visible in the nucleus and/or cytoplasm at 20-40× magnification; negative cases or negative control slides demonstrated no staining. A semi-quantitative assessment of the signal scoring was carried out as prescribed by the assay manufacturer (table 2.4).

Table 2.4 guideline for semi-quantitative assessment of BaseScope™ staining intensity

Staining Score	Microscopic
0	No staining
1	1 dot/cell (visible at 20–40× magnification)
2	2–3 dots/cell. No or very few dot clusters (visible at 20–40× magnification)
3	4–10 dots/cell. Less than 10% positive cells have dot clusters (visible at 20× magnification)
4	> 10 dots/cell. More than 10% positive cells have dot clusters (visible at 20× magnification)

2.2 Cell culture work

2.2.1 Cell culture reagents and procedure

All cell culture procedures were conducted in class II laminar hoods, under completely sterile techniques. Most of the routinely used cell culture supplements are listed in table 2.5 and were purchased from Sigma-Aldrich, Gillingham, UK, unless otherwise stated.

Table 2.5 Cell culture reagents

Reagent	Uses
Dulbecco's modified Eagle's medium (DMEM)	Routinely/ growing
L-glutamine	Routinely/ growing
Dulbecco's phosphate buffered saline (PBS)	Routinely/ washing
Trypsin/EDTA (ethylenediaminetetraacetic acid)	Routinely/ expanding
Foetal bovine serum (FBS)	Routinely/ growing
Nutrient mixture F-12 Ham	Routinely/ growing
Amphotericin B	Occasionally/ Therapeutic
Epidermal growth factor (EGF)	Routinely/ growing
Human insulin	Routinely/ growing
Hydrocortisone	Routinely/ growing
Adenine	Routinely/ growing
Penicillin/ streptomycin	Routinely/ growing
Dimethyl sulfoxide (DMSO)	Routinely/ freezing

2.2.2 Cell lines and primary cells

The cell lines and the primary cells used in this study are listed in tables 2.6 and 2.7. The primary cells were grown from explanted tissues obtained from human surgical biopsies of normal salivary glands and were isolated by Dr Zulaiha Rahman or Dr Lynne Bingle, (NREC Ethic approval number 13/NS/0120).

Table 2.6 List of Cell lines

Cell Line	Cell type	Site	Source
NCI-H647	Immortalised cells established from human adenosquamous carcinoma specimen, CRTC1-MAML2 translocation negative.	Lung	Prof. John Carmichael (University of Nottingham, UK)
NCI-H292	Immortalised cells established from human mucoepidermoid carcinoma specimen, CRTC1-MAML2 translocation positive.	Lung	Prof. John Carmichael (University of Nottingham, UK)
HEK293	Immortalised cells established from normal human embryonic epithelial cells.	Kidney	Prof. James Stewart (the Institute of Infection and Global Health, University of Liverpool, UK)
HA-R-Spondin1-Fc 293T	The 293T cell line is stably transfected to express murine Rspo1 with an N-terminal HA epitope tag and fused to a C-terminal murine IgG2a Fc fragment.	Kidney	AMSBIO (cat.no. 3710-001-01)
UM-HMC2	Immortalised cells established from human mucoepidermoid carcinoma specimen, CRTC1-MAML2 translocation positive.	Parotid Salivary glands	Prof. Jacques E. Nör (University of Michigan School of Dentistry, USA)
L-Wnt3a clone 5.5	Stable clone of human Wnt-A3 expression vector	Subcutaneous tissues, Mouse	Dr Hans Clevers, (Hubrecht Institute for Developmental Biology and Stem Cell Research, Netherlands).
A-253	Immortalised cells established from human mucoepidermoid carcinoma specimen, CRTC1-MAML2 translocation unknown.	Submandibular salivary gland	ATTC A253 (Cat.no ATCC® HTB41™)

Table 2.7 List of human primary salivary gland cells

Cell culture	Cell type	Source
HuSL	Human sublingual salivary gland	Dr Zulaiha A Rahman
SMG	Human submandibular salivary gland	Dr Lynne Bingle
SLG	Human sublingual salivary gland	Dr Lynne Bingle
HPG	Human parotid salivary gland	Dr Lynne Bingle

2.2.3 Culture Media

Keratinocyte growth medium (KGM) was used for normal salivary glands (HuSL, SMG, SLG and HPG) and consists of Low glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 23% (v/v) Ham's F12 (Sigma-Aldrich), 10% (v/v) Foetal bovine serum (FBS), 100 µg/ml penicillin and 100 U/ml streptomycin, 2 mM L-glutamine, 180 µM adenine (Sigma-Aldrich), 0.5 µg/ml hydrocortisone (Sigma-Aldrich), and 10 ng/ml epidermal growth factor (EGF; Sigma-Aldrich).

The UM-HMC-2 cell line was cultured in salivary gland medium (SGM), which consists of high glucose DMEM supplemented with 10% (v/v) FBS, 100 µg/ml penicillin and 100 U/ml streptomycin, 2 mM L-glutamine, 180 µM adenine (Sigma-Aldrich), 0.5 µg/ml hydrocortisone (Sigma-Aldrich), and 10 ng/ml epidermal growth factor (EGF; Sigma-Aldrich).

The A-253 cell line was cultured in McCoy's 5A medium modified supplemented with 10% (v/v) FBS, 100 µg/ml penicillin and 100 U/ml streptomycin, and 2 mM L-glutamine.

HA-R-Spondin1 (RS1) cell line was cultured in Opti-MEM Gibco® life techonologies™1× Medium with L-Glutamine and HEPES (Reduced serum, phenol red free) supplemented with 100 µg/ml penicillin and 100 U/ml streptomycin.

NCI-H647, NCI-H292, HEK293 and L-Wnt3a clone 5.5 cell lines were cultured in Low glucose DMEM supplemented with 10% (v/v) FBS, 100 µg/ml penicillin, 100 U/ml streptomycin, and 2 mM L-glutamine.

2.2.4 Routine culture and maintenance of cells

Cells were mostly grown as adherent monolayers in T75 flasks and incubated at 37°C with 5% CO_2. Cells were checked regularly and exhausted media was replaced with fresh, pre-warmed media every 2 to 3 days. Cells were routinely screened for Mycoplasma infections by a core facility service provided by School of Clinical Dentistry, The University of Sheffield.

2.2.5 Sub-culturing cells

Adherent cells were passaged when they had reached ≥ 80% confluency. Spent culture medium was removed, cells were washed twice with phosphate buffered saline (PBS, 2× 5 ml), and incubated for 2-6 minutes at 37°C with 5% CO_2 in prewarmed Trypsin/EDTA (0.05% trypsin/ 0.02% EDTA w/v) (TE, (Sigma-Aldrich) to disassociate the adherent cells. When the cells were detached a pre-warmed growth medium containing 10% (v/v) FBS (twice the volume of the trypsin) was added to inhibit enzyme activity. The solution was transferred to a centrifuge tube and centrifuged at 1000 rpm for 5 min. The supernatant was discarded, and the pellet re-suspended by adding the required amount of pre-warmed growth media. Cell counting (with a haemocytometer) was performed prior to seeding at different ratios depending on that required for downstream experiments and incubated at 37°C with 5% CO_2.

2.2.6 Cell Counting

When necessary, cells were counted using a haemocytometer. After trypsinising, the enzyme was de-activated by the addition of culture media containing 10% FBS. The cell suspension was transferred to a 15 ml tube and centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded, and the cell pellet re-suspended in 10 ml of fresh medium and mixed well. Cell viability was checked using Trypan Blue

(Invitrogen, Thermo Fisher Scientific); the cell suspension was mixed 1:1 (V: V) with trypan blue, 10 µl transferred to a haemocytometer, examined under an inverted microscope (Nikon, ECLIPSE TS100) and all viable cells in the four corner squares counted using a 10× objective. Dead cells stained blue whereas viable cells remained unstained. The number of cells per millilitre was calculated as indicated in the formula below.

$$\text{Concentration (cells / ml)} = \frac{Number\ \ of\ \ cells\ \times Dilution\ \ factor\ \ 10^4}{4\ (square\ )}$$

Total number of cells = Concentration (cells /ml) × Volume of the sample (ml)

2.2.7 Thawing cells

After removing the cryovial from liquid nitrogen, the cells were thawed immediately in a water bath at 37°C for <1 minute. The cells were then transferred to centrifuge tubes and suspended by adding 10 ml of pre-warmed required growth medium and centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded, and the cell pellet was re-suspended in the appropriate growth medium, seeded into a T75 flask and incubated at 37°C with 5% CO2. Cells were checked regularly on subsequent days and medium was changed to remove any unattached cells or debris.

2.2.8 Freezing cells

The adherent cells were detached on reaching ~80% confluency using the same protocol described above for sub-culturing the adherent cells (2.2.5). 1-2 $\times 10^6$ cells in normal growth media, supplemented with 10% dimethyl sulfoxide (DMSO) as a

cryoprotective agent, was added to each cryovial. Cells were initially stored in a cryo-freezing container for 24 hours at -80°C before transfer to liquid nitrogen.

2.2.9 Cell pellet preparation

When cells had reached 80% confluency, the used media was removed and the cells were washed twice with PBS, 5 ml of ice cooled PBS was added and cells were gently scraped off the bottom of the flask. The cell solution was centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded and the cell pellet re-suspended in RNA lysis buffer (Monarch® Total RNA Miniprep Kit, New England BioLabs®$_{inc}$) using RNase- free filtered tips for RNA extraction. For basic protein analysis, cell pellets were lysed directly in 2× SDS loading buffer (table 2.8) and stored at -20°C. For genomic DNA analysis (gDNA) cell pellets were resuspended in PBS and DNA extracted as described in 2.5.1.

Table 2.8 Recipe for 2× SDS preparation

Reagent	Weight/ volume	Supplier
Protease inhibitor	1 tablet	Roche
Dithiothreitol (DTT)	1 ml of 1 M DTT	Thermo Fisher Scientific
Sodium dodecyl sulfate (SDS)	1 ml of 20% SDS	Thermo Fisher Scientific
Glycerol	2 ml	Thermo Fisher Scientific
Tris-HCl	1.25 ml of 0.5 M pH 6.8	Sigma-Aldrich
Deionised H_2O	4.55 ml	-

2.2.10 Preparing cell clots

When cells reached 80% confluency the used media was removed and the cells were washed twice with PBS. The cells were then fixed with 10% formalin/ PBS solution

(1ml of 10% Formalin + 4 ml of 1 × PBS) for 5 min, washed with PBS, scraped from the culture surface and the cell suspension centrifuged for 5 min at 1000 rpm. The supernatant was discarded and the pellet resuspended in freshly prepared 200 µl of fibrinogen (10 mg/ml in PBS) and 200 µl of thrombin (50 unit/ ml). A clot formed rapidly after which it was sent for processing and sectioning prior to immunocytochemistry.

2.2.11 β-galactosidase (SA-β-Gal) assay

This assay was used to ensure that the primary salivary gland cells were mortal and would stop proliferating at a certain passage number. Six-well plates were seeded with forty thousand cells per well in normal growth media and left to adhere overnight. Senescence associated β Gal activity was detected using a Senescence Detection Kit (ab65351, Abcam, UK) following the manufacturer's guidance.

2.2.12 Immunofluorescence double staining of Human primary salivary gland cells

Twenty to forty thousand primary human salivary gland cells per well were seeded into 8-well culture chamber slides (x-well Tissue Culture Chambers, Sarstedt, Germany). After 24 hours, cell viability was checked, used medium removed and the cells washed three times for 5 minutes with ice cooled PBS. Cells were fixed with 4% paraformaldehyde in PBS (see appendix for fixative preparation) for 10 minutes at room temperature and washed two times for 5 minutes with PBS, permeabilised with 0.25% TRITON™ ×100 (Sigma-Aldrich, UK) in PBS for 5 minutes and washed two times for 5 minutes with PBS. Cells were blocked by incubating with 10% horse serum in PBS for 1 hour at room temperature. After blocking, all wells except the negative control were incubated with 1:100 Monoclonal Anti-Human Cytokeratin 5 (CK5, Clone XM26, Vector VP-C400) primary antibody diluted in 3% Bovine serum albumin (BSA) /PBS for 1 hour at room temperature. After primary antibody incubation, cells were

washed three times for 5 minutes with PBS and constant agitation. All wells were incubated with 1:1000 fluorescence conjugated Goat Anti-Mouse secondary antibody (Alexa Fluor 568 Goat anti-mouse Ab. Cat No- A11001. Red, Thermo Fisher Scientific, UK) for 1 hour at room temperature; from this point all procedures were conducted in the dark. Cells were washed three times for 5 minutes with PBS and constant agitation. Cells were blocked again as described above and incubated with 1:100 of the second primary antibody, Vimentin (MA5-11883, Thermo Fisher Scientific) diluted in 3% BSA/PBS for 1 hour at room temperature. After incubation, cells were washed three times for 5 minutes with PBS and constant agitation and incubated with a fluorescence conjugated Goat anti-Mouse IgG (H+L) Superclonal Secondary Antibody (1:1000, Alexa Fluor® 488 conjugate, Green, Thermo Fisher Scientific, UK) for 1 hour at room temperature. Cells were then washed three times for 5 minutes with PBS and constant agitation. The 8-well manifold was removed carefully at this step, leaving the slide to dry. One drop of ProLong ™ Gold antifade mountant with DAPI (Invitrogen) was applied to the slide, coverslipped, left to air dry for 24 hours at room temperature in the dark, then sealed with nail polish. Staining was examined using a Zeiss 880 AiryScan confocal fluorescence microscope (Carl Zeiss), using FITC at 492 nm absorption with an emission at 520 nm. The same settings were used with controls.

2.3 Cloning

2.3.1 Synthesis of a CRTC1-MAML2 fusion construct.

The CRTC1-MAML2 fusion comprises 1023 amino acids, consisting of exon 1 of the CRTC1 gene fused to exons 2 to 5 of the MAML2 gene. The strategy for the *in vitro* synthesis of this construct was carried out through eight different stages which were: **1.** The 5' end of the fusion transcript was amplified by RT-PCR from a fusion positive cell line, this fragment contained an internal *BamHI* site. **2.** The CRTC1-MAML2 5' end fragment was cloned into pCR2.1 TOPO and sequenced. **3.** Specific primers with *NheI* forward and in-frame FLAG tag were designed to amplify the fragment generated in stage 1 after it had been retrieved from step 2. **4.** The PCR product from stage 3 was cloned into PCR2.1 and sequenced. **5.** MAML2 is 1156 amino acids long and the fusion uses all of the sequence except the 171 amino acids in the first exon; to avoid any subsequent mutations being generated through PCR cloning, a full length Expressed Sequence Tag (EST) clone of MAML2 in PCR4-TOPO was purchased from Genomics-online under gene ID number (84441). The MAML2 fragment was cut out using *BamHI* (internal site in MAML2) and *EcoRI* in pCR4-TOPO (vector). **6.** The retrieved fragment from stage 5 was cloned into pCR3.1 previously prepared by double cutting with *BamHI* and *EcoRI*. **7.** The FLAG-tag 5' end fusion from stage 4 was cut with *NheI* and *BamHI* and retrieved. **8.** The *NheI/ BamHI* fragment from stage 7 was subcloned into the cut pCR3.1 MAML2 clone from stage 6 (figure 2.2).

The molecular genetic procedures were carried out through the following steps: designing PCR primers, restriction enzyme analysis and selection of vectors (table 2.11).

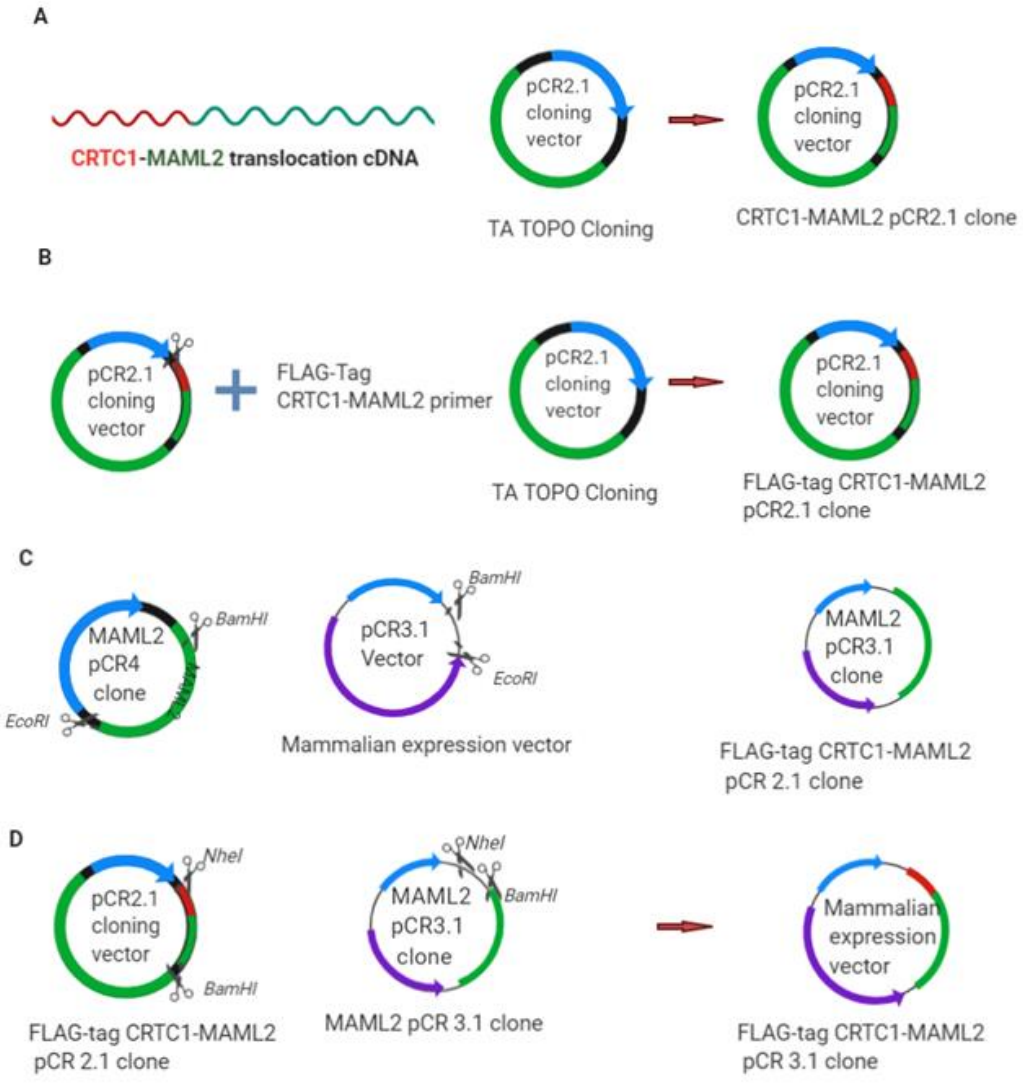


Figure 2.2 Gene cloning. Schematic diagram of full history cloning and different stages. Image created with bioRENDER.com.

2.3.2 Primer design and polymerase chain reaction

Primers were designed to amplify the 5' end of the human CRTC1-MAML2 translocation gene from the fusion positive cell lines, NCI-H292 and UM-HMC-2. A *NheI* enzyme restriction site, a Kozak consensus sequence and a FLAG-tag sequence (DYKDDDDK) were incorporated into the forward primer. A *NotI* or / *EcoRI* enzyme restriction site was incorporated into the reverse primer. The primer sequences, with the expected PCR product size, are summarised in table 2.10.

Primers amplified the desired cloning insert from the cDNA template via PCR as described in section 2.5.4. cDNA was derived from RT-PCR of extracted RNA from both fusion positive cell lines as described in sections 2.5.2 and 2.5.3. This fragment contained an internal *BamHI* site. All samples were amplified in triplicate with a positive (OAZ1) and negative control (master mix only, without primers) being used with each run. PCR thermal cycling is summarised in table 2.9.

Table 2.9 Description of PCR Thermal Cycle Thermal DreamTaq

Step	Temperature, C°	Time	Number of Cycles
Initial denaturation	95	3 min	1
Denaturation	95	30 s	35
Annealing	60	30 s	
Extension	72	1 min	
Final extension	72	10 min	1
Hold	4	Forever ∞	-

*Abbreviations, min: minutes, s: seconds.

Table 2.10 Primer sequences utilised for gene cloning

Gene transcript Human	Oligonucleotides Sequence 5'-3'	Product size (bp)
CRTC1-MAML2 (4)	F 5'ACA CAT TCT CCT GGC AAT GG'3	1,973
	R 3'CGT GGC ATC ATA GGG TTG TT'5	
CRTC1-MAML2 (5)	F 5'GGC GAG AAG ATG GCG ACT'3	699
	R 3'CAT GGG TAA GGA GGG CCT A'5	
CRTC1-MAML2 (6)	F 5'GAG GAG GAG GAG GAG GAG GT'3	833
	R 3'AGA GTT GGC CAT GGA GAA TG'5	
OZA1	F 5' CCA ACG ACA AGA CGA GGG ATT '3	164
	R 3' AGC GAA CTC CAG GAG AAC TG '5	
FLAG CRTC1-MAML2	F 5'TGCTAGCACCATCCACTACAAAGACGATGACGACAA GGCGACTTCGAACAA TCCGCGGAA	724
	R 3'CAT GGG TAA GGA GGG CCT A'5	

*Abbreviations, bp: Base Pair, F: forward, R: reverse.

Table 2.11 Description of the Plasmids used in the study

Vector	feature	Size	Use	Antibiotics resistance	Antibiotic selection	Source
pCR2.1	TOPO Cloning LacZα gene for blue/white screening	3931 bp	Cloning and sequence plasmid	ampicillin and kanamycin	Not Applicable	Thermo Fisher Scientific
pCR 4	TOPO Cloning LacZα gene for blue/white screening	3956 bp	Cloning and sequence plasmid	ampicillin and kanamycin	Not Applicable	Thermo Fisher Scientific
pCR3.1	CMV promotor	5060 bp	Mammalian expression Vector	ampicillin/ kanamycin	Geneticin G418	Professor Colin Bingle
pCDNA3.1+ IREsGFP	CMV promotor With GFP internal ribosome entry site	6805 bp	Mammalian expression Vector	ampicillin/ kanamycin	Neomycin/ Geneticin G418	Addgene
pEGFP-N1	CMV promotor	4733 bp	Reporter Vector	kanamycin	Neomycin	Professor Colin Bingle
pCDNA6	CMV promotor	5149 bp	Mammalian expression Vector	ampicillin	Blasticidin	Dr Devon Smith

Abbreviations, bp: base pairs, CMV: Cytomegalovirus, GFP: Green fluorescence protein

2.3.3 PCR2.1 TA TOPO Cloning

The amplified RT-PCR CRTC1-MAML2 product was cloned from the fusion positive cell lines into pCR2.1 TOPO vector by TA cloning (Invitrogen, Thermo Fisher Scientific) according to the manufacturer's instructions. The vector was added to the

DNA and incubated for 5 min at room temperature to prepare for bacterial transformation.

2.3.4 Plasmid transformation and selection of the positive colonies

The cloned plasmid was transformed into 50 µl of one shot TOP10F *E. coli* chemically competent cells (Invitrogen, Thermo Fisher Scientific), table 2.12. The competent cells were thawed on ice and incubated with the plasmids on ice for 30 min with intermittent gentle flicking, then heat shocked for 45 seconds in a 42°C water-bath. The mixture was returned to ice for 5 min and 200 µl of pre-warmed Super Optimal broth with Catabolite repression medium (New England Biolabs, UK) was added before incubating at 37°C for an hour on a rotary shaker at 225 rpm. 50-100 µl of transformed plasmids were added to LB Agar plates (Sigma Life Science, UK) with 50 µg/ml ampicillin, each plate having previously been spread with 75 µl of Isopropyl B-D-1-thiogalactopyranoside (IPTG) for blue/white colour colony selection. The plates were then incubated overnight at 37°C, including the pUC19 (positive control), competent cells only and plasmid free (negative control). White positive colonies were selected based on the criterion that the colony was well demarked and separated and expanded overnight at 37°C on a rotary shaker in 2 ml of LB broth (Sigma Life Science, UK) containing 50 µg/ml ampicillin.

Table 2.12 List of competent cells used in the study

Competent cells	Plasmid	Source
One Shot™ TOP10 chemically competent *E. coli*	pCR™4,pCR™2.1, pCR™3.1	Thermo Fisher Scientific
One Shot™ TOP10F chemically competent *E. coli*	pCR™4,pCR™2.1, pCR™3.1	Thermo Fisher Scientific
NEB® 5-alpha Competent *E. coli*	pRC™3.1, pCDNA 3.1+ IREsGFP	Thermo Fisher Scientific
NEB® pUC19	Control	Thermo Fisher Scientific

2.3.5 Isolation and preparation of plasmid DNA from *E. coli*

In order to extract the plasmid of interest from the *E. coli* a miniprep (ISOLATE II Plasmid Mini Kit, BIOLINE) procedure was performed. 1.5 ml of a selected and expanded colony was centrifuged for 30 seconds at 11,000×g. The supernatant was discarded and the bacteria pellets were lysed gently in resuspension buffer according to the manufacturer's instructions. The DNA was bound to a mini spin column (silica membrane), which was washed and dried. DNA elution was carried out with 50 µl of nuclease free H_2O and the samples were stored at -20°C. The isolated DNA was confirmed by overnight diagnostic restriction enzyme digest (*EcoRI* digestion) and by running on 1% agarose gel. Once the correct size was visualised, 10 µl of each purified DNA sample was sent for sequencing at the Genomic Core Facility of the Medical School, University of Sheffield. The primers used for sequences are described in table 2.13.

Table 2.13 Standard and specific overlapping primers used in the sequence

Primers	Oligonucleotides sequences	Supplier
T7 promoter, forward primer	5' TAATACGACTCACTATAGGG'3	Invitrogen
BGH terminator, reverse primer	3' TAGAAGGCACAGTCGAGG'5	Invitrogen
CRTC1-MAML2 F3S	5'CTTGCCTATGAGCCAAGCAC '3	Sigma-Aldrich (UK)
CRTC1-MAML2 F4S	5' TTGCCAAGCTGGCAGGAAGT '3	Sigma-Aldrich (UK)
CRTC1-MAML2 F2S	5' ACATTTGTCAAGGCCACCTC '3	Sigma-Aldrich (UK)
CRTC1-MAML2 F5S	5' AACTGCACCAAACTGGGCCT '3	Sigma-Aldrich (UK)
CRTC1-MAML2 R2S	5' AGGAATGGGATACCAAGTCT '3	Sigma-Aldrich (UK)
CRTC1-MAML2 R4S	5' AACAACCCTATGATGCCACG '3	Sigma-Aldrich (UK)

Abbreviations: BGH= Bovine growth hormones.

2.3.6 FLAG-Tagged CRTC1-MAML2 pCR2.1clone

The CRTC1-MAML2 pCR2.1 clone was linearised by *XhoI* digestion for use as a PCR template using the specific FLAG-tag primer described in section 2.3.2. A standard PCR reaction was used to amplify the 5' end of the fusion transcript and the PCR reaction was confirmed by separation of the products in a 1% TAE agarose gel. The FLAG tagged CRTC1-MAML2 PCR product was cloned into pCR2.1 TOPO and the clone transformed into One Shot™ TOP10F chemically competent *E. coli*. After selecting and expanding positive colonies, the DNA plasmid was isolated and purified as described previously. The isolated plasmids were digested with *EcoRI* and fragments of the correct size were sent for sequence analysis using T7 primers. The generated plasmid was double digested with *BamHI* and *NheI* and after digestion DNA was recovered and purified ready for sub cloning.

2.3.7 MAML2 pCR4-TOPO clone

A single clone of MAML2 in pCR4-TOPO was purchased from Sigma-Aldrich, UK based on the sequence verified full length expressed sequences tags (EST). The clone was regrown in ampicillin resistance LB agar plates. On the second day positive white colonies were selected and cultured in LB broth ready for plasmid isolation and preparation. Minipreps were used to extract the MAML2 pCR4-TOPO plasmid and assessed initially by diagnostic restriction enzyme digests *(EcoRI*, *BamHI/ HindIII*) and confirmed by sequencing in two directions using T7 as the forward sequence primer and BGH as the reverse sequence primer (table 2.13). In order to clone MAML2 into a mammalian pCR3.1 entry vector, the MAML2 pCR4-TOPO clone was cut with *BamHI* (in the MAML2 fragment) and *EcoRI* (in pCR4-TOPO). The digested DNA was electrophoresed on a 1% TAE-agarose gel, visualised under UV light, the required size band excised and the DNA purified using Wizard® SV Gel and PCR Clean-up System, Promega, USA following the manufacturer's instructions. The pure

DNA was visualised by electrophoresis and stored at -20°C for further mammalian entry vector cloning.

2.3.8 Subclone MAML2 into pCR3.1

Mammalian pCR3.1 entry vector (figure 2.3), the target vector for the CRTC1-MAML2 fusion was prepared by digesting with *BamHI/EcoRI*. After running the digested vector on a gel, the expected size band was excised, and the DNA was recovered and purified as previously described and stored at -20°C ready for ligation.

2.3.9 Ligation of MAML2 cut fragment into pCR3.1

Ligation was carried out to connect the insert DNA (MAML2) to the compatibly digested vector (pCR3.1) to obtain a recombinant plasmid; this was achieved using T4 DNA ligase enzyme (Promega, USA). In a sterile Eppendorf tube, the DNA insert was mixed with 10× ligase buffer, the ligase enzyme and the vector, which was added to the mixture using four different dilutions (1:5, 1:10, 1:20 and 1:40) and a fixed insert ratio. The final volume of the reaction was 10 µl. The ligation reaction was incubated at 4°C overnight, transformed into *E. coli* using One Shot™ TOP10F chemically competent cells and plated onto ampicillin agar plates. The positive colonies were selected and cultivated in LB broth with ampicillin (previously illustrated above under 2.3.4 section). Minipreps were performed to extract and purify the plasmid using the same kit as described above in section 2.3.5. The isolated DNA of interest was analysed using appropriate restriction enzymes, the fragment visualised to confirm the correct molecular size. Bands of the correct size were sequenced. When the clone sequences were confirmed, the clone was double digested with *NheI* and *BamHI* restriction enzymes. After digestion the required products were excised, and the DNA was purified as described above and stored at -20°C for further sub-cloning.

2.3.10 Sub-cloning FLAG CRTC1-MAML2 tagged into mammalian pCR3.1 entry vector

Serial dilutions of pCR3.1 plasmid were prepared (1:2, 1:4, 1:8 and 1:16) and T4 DNA ligase and 10× ligase buffer (Promega, USA) were used in the ligation. The mixture was incubated at 4˚C overnight. A self-ligation test for pCR3.1 vector (incubation of 1:2 diluted pCR3.1 with ligase enzyme and ligase buffer without DNA insert) confirmed the correct ligation. Ligation reactions were transformed into 50 µl of competent cells following the manufacturer's instructions. One Shot™ TOP10F chemically competent *E. coli* was used to transform the Flag CRTC1-MAML2 pCR3.1. The pUC19 DNA (New England Biolabs, UK) was used as positive control and competent cells only as negative control. All reactions were added to ampicillin agar plates. Only the positive colonies that met the criteria described previously (in section 2.3.4) were selected and cultured in LB broth with ampicillin. A high copy of the *E. coli* plasmids was extracted and purified as described previously. *EcoRI* digestion was used to detect the correct band size for all constructs. Internal primers (T7) were used to sequence and verify the complete fusion construct.

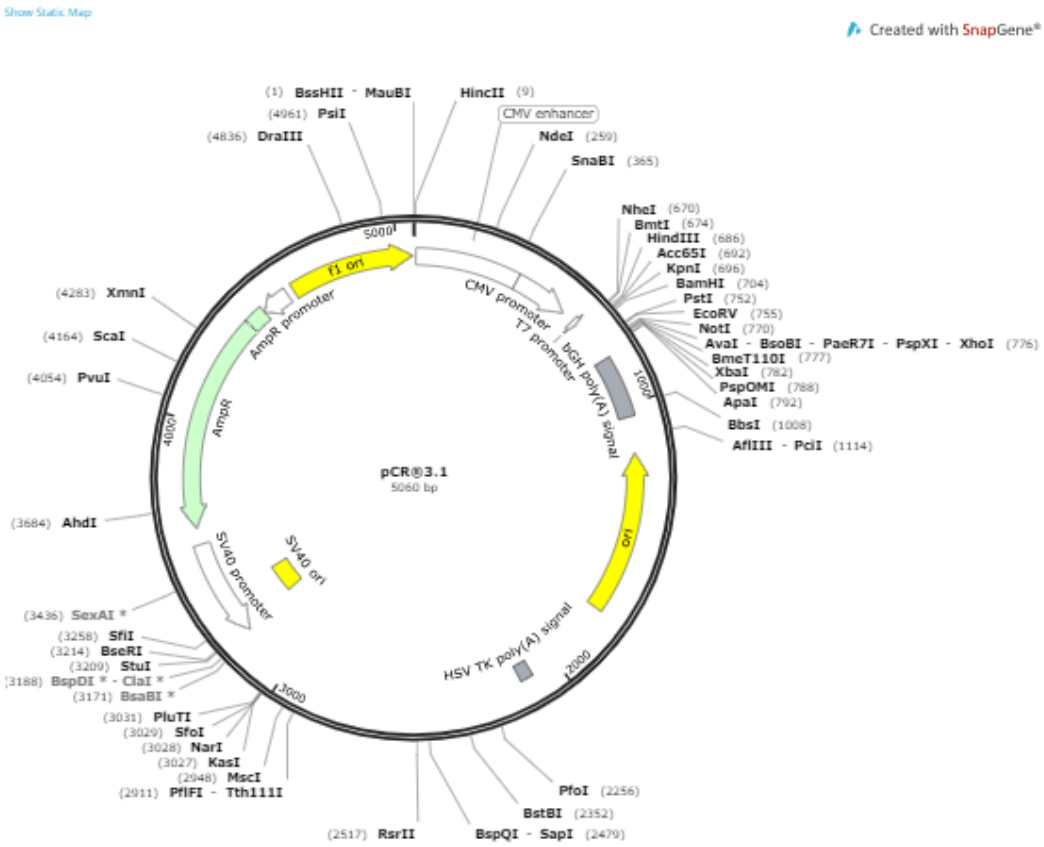


Figure 2.3 The pCR3.1 mammalian expression vector. A plasmid map of the target vector, pCR3.1. Labelled cassettes include the T7 promoter, CMV gene, BGH gene, ampicillin resistance gene and Neomycin/ G418 selection gene. Restriction sites are also shown. The plasmid vector is 5060 bp in size. Image is from Addgene.

2.3.11 FLAG Tagged CRTC1-MAML2 clone in pCDNA 3.1+IREs GFP vector

To facilitate the transfection experiments, the fusion construct was subcloned into an internal ribosomal entry site Green Fluorescent Protein mammalian expression vector. The FLAG Tagged CRTC1-MAML2 pCR3.1 clone was double cut using *NheI* and *NotI*, the fragment was retrieved and ligated into pCDNA 3.1+ IREs GFP (figure 2.4) plasmid purchased from Addgene and prepared by digesting with *NheI* and *NotI* to produce compatible sticky ends for the insertion of FLAG Tagged CRTC1-MAML2 fusion. A single digested plasmid was used as self-ligation control. The same above-

mentioned cloning procedures were carried out on this construct and the final confirmation by double cutting the clone with *BamHI* and *SpeI* restriction digestions.

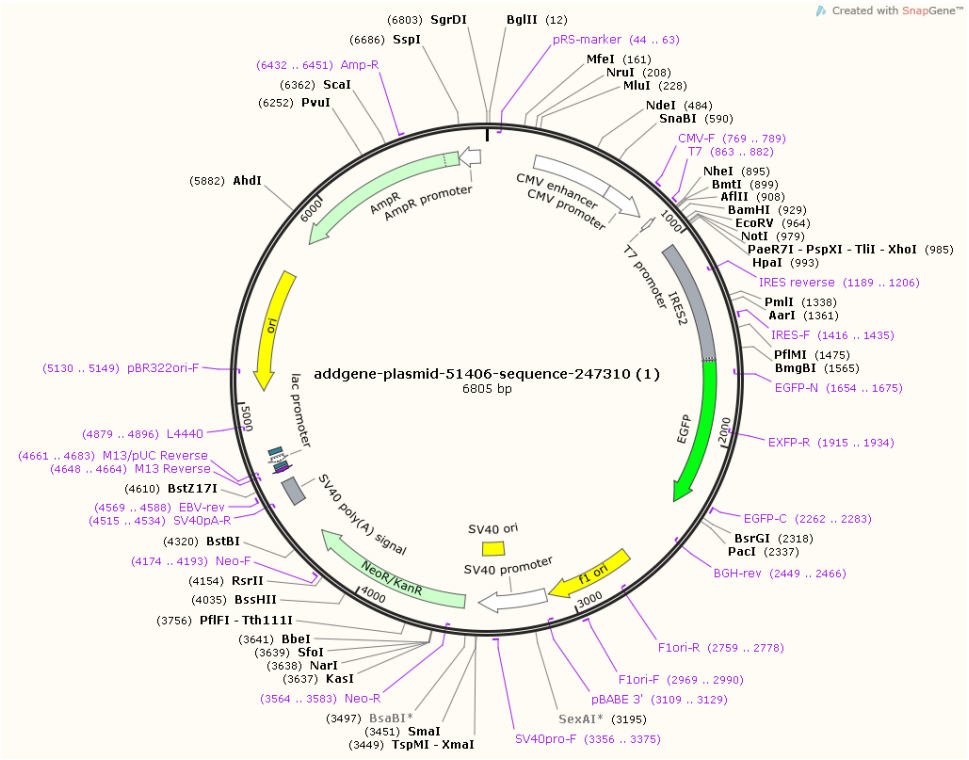


Figure 2.4 The pCDNA3.1+IREs GFP mammalian expression vector. A plasmid map of the target vector, pCDNA3.1+IREs GFP. Labelled cassettes include the T7 promoter, the Internal ribosomal site green fluorescence protein, CMV gene, ampicillin resistance gene and neomycin/ G418 selection gene. Restriction sites are also shown. The plasmid vector is 6805 bp in size. Purchased from Addgene plasmid # 51406; http://n2t.net/addgene:51406; RRID:Addgene 51406).

2.3.12 Final confirmation of the constructs

To verify cloning of the FLAG Tagged CRTC1-MAML2 pCR3.1, overlapping sequencing with specific primers was performed. The overlapping sequence, in both forward and reverse directions, was then read along the construct. A number of primers were designed in our laboratory, (table 2.12). The sequencing quality was examined and a blast analysis with NCBI nucleotide verified the sequence identity

and the alignment with sequence references. Snap Gene software was used to draw the plasmid map.

2.3.13 Plasmid isolation 'Midiprep'

In order to have a high DNA yield for long-time storage of the successfully sequenced FLAG tag CRTC1-MAML2 pCR3.1 clone, 2 ml of the previous transformed mix was cultivated in autoclaved LB broth supplemented with 50 mg/ml ampicillin and incubated overnight at 37ºC in a rotary shaker at 225 rpm. The next day, before plasmid extraction, a glycerol stock of the bacterial culture was prepared by adding 500 µl of autoclaved 50% glycerol solution to 500 µl of the bacterial culture in cryo-vials and stored at -80ºC. The cells were harvested by centrifuging the bacterial culture at 3.400× g for 10 minutes at 20ºC. The plasmid DNA was extracted and purified using ZymoPURE™ Plasmid Midiprep kit (ZymoPURE, USA). DNA concentration was measured with a Nanodrop spectrophotometer, and the accuracy was analysed by restriction digestion, *EcoRI* for the pCR3.1 construct and *BamHI* and *SpeI* for pCDNA3.1+IREs GFP construct, and agarose gel electrophoresis visualised under UV trans -illumination. Samples were stored at -20ºC.

2.4 Transfection

2.4.1 FugeneHD transient transfection

HuSL primary cells and NCI-H647 cells were cultured at approximately 50 $\times 10^4$ in 6-well plates in antibiotic-free DMEM (low-glucose),10% FBS and other essential growth additives for each cell type. The cells were incubated at 37 °C and 5% CO_2 for 24 hr. When the cells were approximately 80% confluent, they were transfected using FugeneHD transfection reagent (Promega). Ninety to ninety-nine microliters (dependent on the DNA concentration) of pre-warmed, room temperature Opti-

MEM™ I (Gibco) Reduced Serum Medium, no phenol, no additives, was mixed with 2 µg of DNA (FLAG tag CRTC1-MAML2 cloned in pCR3.1) to achieve 100 µl and pre-warmed, room temperature transfection reagent was added to the DNA media mixture at various ratios (3:1 -5:1 µl Transfection reagent: µg DNA), mixed well and incubated for 15 min at room temperature. During this incubation, the old media was aspirated, the cells washed with Opti-MEM media and 2 ml of Opti-MEM media was added to each well of the 6- well plates. 150 µl of the transfection solution was slowly added to each well and the plate incubated at 37°C with 5% CO_2. Co-transfection with 100 ng pEGFP-N1 reporter gene (ClonTech Laboratories) was carried out and an empty pCR3.1 vector used as transfection control. The cells were checked at 24 and 48 hr post-transfection for green fluorescence and cells were washed, lysed directly in two-× SDS lysis buffer and stored at -80°C for downstream analysis.

2.4.2 JetPRIME Transfection

HEK293, NCI-H647 and human primary salivary gland cells were seeded into 6-well plates at seeding density of 5×10^5 cells per well in antibiotic-free growth media and incubated overnight at 37°C and 5% CO_2. On the following day, the media was changed prior to transfection. Two wells were transfected with either FLAG-tag CRTC1-MAML2 pCR3.1 construct or FLAG-tag CRTC1-MAML2 pCDNA3.1+ IREs GFP construct. When the pCR3.1 construct was used it was co-transfected with the pEGFP-N1 reporter gene (ClonTech Laboratories). Two wells were transfected with empty vector without target DNA, and one well was used as a negative control to monitor the growth of the cells and the other was used as transfection reagent control to test the toxicity of the transfection reagent on cells without any DNA plasmids. Two µg of DNA was diluted in 200 µl of transfection buffer and 1:2 of DNA: transfection reagent ratio was used following an optimisation trial. The transfection mixture was incubated at room temperature for 10-15 minutes, and then 200 µl of the transfection

solution was added to the desired wells. After 4 hours the transfected media was replaced with fresh media and the cells were grown for a further 48 hours. When a larger scale of cells was needed for transfection (T25 or T75 flasks), the transfection protocol was adjusted appropriately. Cell viability and green fluorescence expression was checked using an inverted fluorescent microscope (Axiovert 200M, Zeiss). Transfection efficiency was determined either by fluorescence activated cell sorting (FACS) or microscopically. The cells were lysed directly in 2× SDS lysis buffer for downstream analysis.

2.4.3 Fluorescence activated cell sorting (FACS)

Transfected and control cells were washed twice with PBS, harvested using trypsin, and then resuspended in 500 µl of growth media for cell sorting (FACS) using a BD FACS Aria Ilu cell sorter (Research Core Facility, the Medical School, University of Sheffield, UK) based on GFP positive and GFP negative cells. Cell sorting was performed by Sue Clark (Flow Cytometry Core Facility technician). Un-transfected and empty vector transfected cells were used as controls to calibrate the machine, to set a baseline for the fluorescence gate and to eliminate any false positive GFP signals. After sorting, in addition to GFP positive cells, the GFP negative and empty vector cells were collected to serve as controls for all downstream analysis.

2.4.4 Cell expansion

Single cell colony selection was not possible with the primary human salivary gland cells (HuSL, HPG, SMG and SLG) as they did not grow in isolation. Dependent on the number of sorted cells, both GFP positive and negative cells were grown in smaller culture vessels, for example 48-well plates, in the presence of 10 µg/ml of the ROCK inhibitor Y-27632 (Abcam, Cambridge, UK), in order to promote cell growth

and temporarily inhibit cell senescence, as suggested by Chapman *et al*., 2010, Chapman *et al*., 2014, Koslow *et al*., 2019). Cell expansion continued until there was a sufficient number of cells for storage in liquid nitrogen. Y-27632 was removed before conducting functional analysis of transfected cells.

2.4.5 Kill curve

As the pCDNA3.1+IREsGFP and pCR3.1 plasmids carry Geneticin (G418) resistance, and in order to generate a stable cell line, a dose response experiment was performed on HuSL, NCI-H647 and HEK293 cells. Cells were grown in 24-well plates with antibiotic-free media to reach 70% confluence after 24 hours, as the G418 is most effective when cells are in active division. Growth media was replaced with fresh selective media containing varying concentrations of G418 disulphate salt (Sigma Aldrich, UK) in duplicate: 0 µg/ ml, 50 µg/ ml, 100 µg/ ml, 200 µg/ ml, 300 µg/ ml, 400 µg/ ml, 500 µg/ ml, 600 µg/ ml, 700 µg/ ml, 800 µg/ ml, 900 µg/ ml and 1 mg/ml. Selective media were maintained by replacing every 2 to 3 days for 10 days. Cells were checked regularly, and cell viability was determined under bright field microscopy to determine the optimal dose; this is the minimum concentration of antibiotic that killed the cells in 10 days.

2.4.6 Generation of stable cell lines

Once the optimum dose of antibiotic for each cell line was determined, transfection was carried out as described in section 2.4.2. Two controls were included in stable transfection experiments, one positive control and one negative control; the conditions were identical apart from the transfection event. After 48 hours cells were either sorted or exposed to selection media with optimum antibiotic concentration. The selection media were maintained for 10 days. The negative control was used to

assess the viability of non-transfected cells after exposure to antibiotics and the positive control to monitor the growth of transfected cells without exposure to antibiotics. All non-transfected cells were dead within 10 days of exposure to G418. Transfected cells survived and continued to grow and expand as described in section 2.4.4. Once stable cell lines were established the expression of the gene of interest could be analysed at protein and gene levels.

2.5 Gene expression analysis

2.5.1 Genomic DNA Extraction

Genomic DNA (gDNA) was extracted from cultured cells using the Wizard Genomic DNA Purification Kit (Promega, Southampton, UK) following the manufacturer's instructions. The quality and yield of the extracted DNA were analysed using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). DNA was stored at -20°C for future use.

2.5.2 RNA extraction

Total RNA was isolated from cultured cells using the Monarch® Total RNA Miniprep Kit (New England BioLabs UK), following the manufacturer's instructions and with the optional step of removing residual genomic DNase being carried out. The quality and yield of the RNA was examined using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). RNA was stored at -80°C for further analysis.

2.5.3 Reverse transcription

Complementary DNA (cDNA) was synthesised from 1 µg of total RNA using a High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) as per the manufacturer's protocol (table 2.14). cDNA was stored at -20°C for further application.

Table 2.14 Reverse transcription description of thermal cycle

Settings	Step 1	Step 2	Step 3	Step 4
Temperature	25°C	37°C	85°C	4°C
Time	10 minutes	120 minutes	5 minutes	∞

Abbreviations: ∞ forever

2.5.4 Polymerase chain reaction

Polymerase chain reaction (PCR) was performed using 1 µl of template cDNA, 1 µm forward and reverse oligonucleotide primers and 2× Dream Tag Green PCR Master Mix (Thermo Fisher Scientific) in a total reaction mixture of 50 µl. The primer pair sequences are summarised in Table 2.15. Annealing temperatures at 55°C, 60°C and 65°C were originally tested; 60°C was found to be optimal and used for all subsequent PCR reactions.

2.5.5 Analysis of the PCR products

PCR products were analysed on a 1 or 2% TAE- agarose gel (dependent on expected PCR product size) containing 0.2 µg / ml ethidium bromide. DNA loading buffer (6×, Bioline, UK) was added to the samples and 5 µl of 100 bp (PCRBIO Ladder IV) or 1 kb (GeneRuler™) DNA ladder were used as molecular weight markers. Gels were run in 1× TAE at 70 Volts for 90 minutes. An InGenius3 gel documentation system (Syngene) was used to visualise DNA bands.

Table 2.15 Primers sequence utilised in PCR analysis

Gene	Sequence	Product size	Supplier
E-cadherin	5'GACACCCGGGACAACGTTTA'3 3'GGGTCAGTATCAGCCGCTTT'5	293 bp	Sigma-Aldrich
Amylase	5'ACTTGTGGCAATGACTGGGT'3 3'TCTCCAGAAATGACATCACAGT'5	299 bp	Sigma-Aldrich
c-kit	5'GAAGGCTTCCGGATGCTCAG'3 3'AAAGGCTGGGGTAGGTAGGT'5	204 bp	Sigma-Aldrich
Nanog	5'CACCCAGCTGTGTGTACTCA'3 3'AAAGGCTGGGGTAGGTAGGT'5	221 bp	Sigma-Aldrich
CK5	5' CCAGGAGCTCATGAACACCA '3 3' GCTTCCACTGCTACCTCCG '5	235 bp	Sigma-Aldrich
GAPDH	5'TGATGACATCAAGAAGGTGGTGAAG'3 3'TCCTTGGAGGCCATGTGGGCCAT'5	221 bp	Sigma-Aldrich
hOAZ1	5' CCA ACG ACA AGA CGA GGG ATT '3 3' AGC GAA CTC CAG GAG AAC TG '5	164 bp	Sigma-Aldrich
MUC2	5' AACACAGTCCTGGTGGAAGG '3 3' CCTGGCACTTGGAGGAATAA 5'	474 bp	Sigma-Aldrich
MUC5AC	5' TGATCATCCAGCAGCAGGGCT '3 3' CCGAGCTCAGAGGACATATGG '5	409 bp	Sigma-Aldrich
MUC5B	5' CAACAGCCATGTGGACAACT '3 '3 CAACAGCCATGTGGACAACT '5	229 bp	Sigma-Aldrich

Abbreviation: bp = Base Pair.

2.5.6 Quantitative Polymerase Chain Reaction

PCR reactions were performed on cDNA samples using a rota-gene Q real-time PCR cycler (Qiagen, Manchester, UK) and either SYBR® Green or TaqMan Probe. For the CRTC1-MAML2 fusion transcript a custom-made TaqMan gene expression probe (Hs03024645_ft) was designed (Thermo Fisher Scientific, UK). SYBR Green primers, designed in house (Sigma-Aldrich, Gillingham, UK), are listed in table 2.17. Relative changes in the number of transcripts in each sample were determined by normalising for RNU6B (U6; Sigma-Aldrich, Gillingham, UK) for SYBR Green analysis or β-2 macroglobulin (B2M; Life Technologies) for TaqMan probe analysis. Each qPCR reaction was performed in triplicate with at least two biological repeats, using 1 µl of

cDNA template in a total reaction mixture of 10 µl. The standard thermal cycle consisted of 40 cycles including a melt curve analysis with SYBR Green (table 2.16).

Quantification of gene expression was calculated using the delta delta CT (2-ΔΔCT) method (Livak and Schmittgen, 2001). The cycle threshold is the cycle number that causes the fluorescent signal to reach a fixed threshold. The ΔCT value represents the difference in CT values between target genes and endogenous controls. The calculated 2-ΔΔCT values represent the relative fold change in gene expression between samples. Data will be presented as fold-change in target gene expression compared to the endogenous controls.

Table 2.16 Thermal cycle conditions for Real-Time PCR machine

Setting	Step 1	Step 2	Step 3	Cycle
Temperature	95°C	95°C	60°C	40
Time	10 minutes	10 Seconds	45 Seconds	

Table 2.17 Primers sequences utilised in qPCR analysis

Gene transcript	Oligonucleotides Sequence	Product size (bp)
ATF3	5'AAAAGAGGCGACGAGAAAGA'3	104
	3'ACTTTCCAGCTTCTCCGACT'3	
DUSP1	5' CTGGTTCAACGAGGCCATTG'3	114
	3' AGGTAAGCAAGGCAGATGGT'3	
NRA2	5' GTCTGATCAGTGCCCTCGT'3	119
	3' ATGCTGGGTGTCATCTCCAC'3	
STC1	5' CAGCTGCCCAATCACTTCTC'3	98
	5' AGGCTGTCTCTGATTGTGCT'3	
TFF-1	5' AATACCATCGACGTCCCTCC'3	112
	3' AGCTCTGGGACTAATCACCG'5	

*Abbreviations, bp: Base Pairs.

2.6 Protein Expression Analysis

2.6.1 Protein Extraction

Protein extraction was carried out as described in section 2.2.9, with the pellet being lysed in 10 ml of radio-immunoprecipitation assay (RIPA) buffer (ChemCruz®, Biotechnology, USA) containing EDTA free protease inhibitor cocktail (Roche, Basel, Switzerland). The cell lysis solution was incubated with constant agitation at 4°C for 30 minutes, and then centrifuged at 12,000 rpm for 20 minutes at 4°C. The supernatant was gently aspirated into a fresh tube and stored at -80°C.

2.6.2 Protein Quantification assay

Total protein concentration was calculated using the bicinchoninic acid assay (BCA; Thermo Fischer Scientific, Cambridge, UK) according to the manufacturer's instructions. BSA (Thermo Fisher Scientific) was used to produce the standard curve, the POLAR star Galaxy Spectrophotometer (BMG LabTech) was used to measure the absorbance of the protein samples and standards at 570 nm.

2.6.3 Protein Sample Preparation

CRTC1-MAML2 fusion protein is a non-secreted nuclear protein; to enable detection of the expressed recombinant protein from transfected cells, cells were lysed in 2× SDS buffer and homogenised using a QIAshredder (Qiagen, UK) to ensure removal of gDNA and insoluble debris. Briefly, the cell lysate was transferred to a QIAshredder spin column in a 2 ml collection tube, samples were homogenised by centrifuging at 16,000 g for 3 minutes after which the column was discarded and the sample was saved for further use. Samples were sonicated for 10 seconds, and then placed on ice for 5 minutes before use. Other protein samples (not CRTC1-MAML2 fusion protein) lysed in RIPA buffer, were mixed 1:1 (V: V) with 2× SDS buffer. All protein samples were heated at 100°C for 10 minutes prior to electrophoresis.

2.6.4 Protein Sample Separation

CRTC1-MAML2 fusion recombinant is a large molecular weight protein and so 20 µl of sample was separated on a 3-8% gradient Tris-Acetate precast protein gel (NuPAGE™, Invitrogen™) alongside 5 µl of pre-stained broad-range protein standards (Prime-Step, BioLegend, Inc) using 1× Tris-Acetate SDS running buffer (NuPAGE™, Invitrogen™). Gels were initially run for 15 minutes at 130 Volts; when

all of the loaded samples had reached the end of the stacking gels the power was changed to 150 Volts.

2.6.5 Protein Transfer

After separation by electrophoresis proteins were transferred to a Polyvinylidene Difluoride (PVDF) membrane (0.45 µm pore size; Invitrogen™). Wet transfer was carried out to ensure transfer of the larger molecular weight proteins. Before assembling the gel-membrane sandwiches, the membrane was soaked in 100% methanol for 15 seconds and then rinsed by soaking in distilled water for 5 minutes. The membrane was then soaked in transfer buffer (table 2.18) for 10 minutes. The assembled gel-membrane sandwich was placed in a gel tank (XCell SureLock Mini-Cell Electrophoresis System) filled with buffer and 30 V applied for 3 hours. In some instances, when the target protein had a lower molecular weight, an alternative protein transfer method was used (Semi-dry transfer) using the Trans-Blot® Turbo™ Transfer System (Bio-Rad, Deeside, UK). Transfer conditions were 25 V, 2.5 A for 10 minutes. To confirm the transfer of proteins the membrane was stained with Ponceau Solution (Sigma-Aldrich, Gillingham, UK) for 5 minutes on a shaker. The membrane was then washed with distilled water and the gel stained with Instant Blue™ staining solution (Expedeon, Abcam) for 15 minutes at room temperature to ensure that the proteins had completely transferred to the membrane.

Table 2.18 Preparation of 1× NuPAGE® Transfer buffer for wet transfer

Reagent	Volume or percentage
NuPAGE® Transfer Buffer (20×)	50 ml
Methanol 100%	50 ml
Deionized Water	900 ml
SDS	0.1%

*Abbreviations, ml: millilitres, g: grams, SDS: Sodium Dodecyl Sulphate.

2.6.6 Blocking and Antibody Incubation

The membrane was blocked with 5% milk powder (Marvel) dissolved in 1× 10 mM Tris buffered saline (1×TBS) for one hour at room temperature on a rocking platform. The membrane was then briefly rinsed with 1×TBS containing 0.1% Tween (v/v) (1×TBS-Tween 20). Primary antibodies (table 2.19) were diluted in 5% milk powder (Marvel) in 1× TBS-Tween 20. Membranes were incubated with diluted antibodies overnight at 4°C on a rocking platform, washed with 1×TBS-Tween 20 three times for 10 minutes and then incubated with horseradish peroxidase conjugated (HRP) secondary antibodies (table 2.19) diluted in 5% milk 1× TBS-Tween 20 for one hour at room temperature on a rocking platform. Membranes were washed with 1×TBS-Tween 20 three times for 10 minutes.

Table 2.19 Antibodies used in Western blot analysis

Primary Antibody	Manufacturer/	Description	Dilution	Secondary Antibody

	catalogue number			
Anti-FLAG	Sigma-Aldrich/ F3165	Mouse monoclonal to FLAG-tagged proteins	1: 1000	Anti-mouse HRP Cell signalling 1:3000
Anti-MAML2	Cell Signalling / #4618	Rabbit Polyclonal to human MAML2 protein	1:1000	Anti-rabbit IgG, HRP Cell signalling 1:3000
Anti-β Actin	Cell Signalling / #3700	Mouse Monoclonal to human β-Actin protein	1: 10,000	Anti-mouse HRP Cell signalling 1:3000

Abbreviation: HRP= Horseradish peroxidase

2.6.7 Protein Detection

Protein was detected using enhanced chemiluminescence (ECL) Western blotting substrate (Thermo Fisher Scientific) following the manufacturer's instructions. Signals were visualised by exposure to x-ray film (CL-XPosure™ Film, Thermo Fisher Scientific, Cambridge, UK) and the film was developed with a Compact ×4 Developer (Xograph Imaging Systems). Protein was alternatively detected using a Li-Cor C-Digit Western Blot Scanner and Image Studio Software. When re-probing with other antibodies, the membranes were washed with 1× TBS-Tween and stripped using 10 ml of stripping buffer (Thermo Fischer Scientific, Cambridge, UK) on a shaker for 13 minutes at room temperature. The membranes then were washed for 10 min twice with 1× TBS-Tween. This procedure was only conducted once per membrane.

2.6.8 Immunofluorescence staining for transfected cells

NCI-H647 cells were grown in 8-well culture chamber slides (x-well Tissue Culture Chambers, Sarstedt, Germany) at a seeding density of 20,000 to 40,000 cells for 24 hours. Cells were transfected with FLAG-Tagged CRTC1-MAML2 pCDNA3.1+IREsGFP or pCDNA3.1+IREsGFP empty vector as described in 2.4.2. After 48 hours, cells were checked for viability, the old medium removed, and the cells washed with ice-cooled PBS three times for 5 minutes. Cells were then fixed with 4% paraformaldehyde in PBS (see appendix for fixative preparation) for 10 minutes at room temperature. Fixed cells were washed twice with PBS for 5 minutes. The cells were permeabilised by incubating with 0.25% TRITON™ ×100 (Sigma-Aldrich, UK) in PBS for 5 minutes and then washed twice with PBS for 5 minutes. Cells were blocked with 10% BSA in PBS for 30 minutes at 37°C after which all cells except the negative control chamber were incubated with 1:1000 Monoclonal anti-FLAG M2 (Sigma-Aldrich) primary antibody diluted in 3% BSA/PBS overnight at 4°C. On the second day, cells were washed three times with PBS for 5 minutes with constant agitation. All cells were incubated with fluorescent conjugated Goat Anti-Mouse secondary antibody (Alexa Fluor 488, Thermo Fisher Scientific, UK) for 45 minutes at 37°C, with the chamber being wrapped in aluminium foil and kept in the dark. Cells were then washed three times with PBS for 5 minutes with constant agitation. The 8-chamber manifold was removed carefully, and the slide left to dry. One drop of ProLong ™ Gold antifade mountant with DAPI (Invitrogen) was applied to the slide, coverslipped and left to air dry for 24 hours at room temperature in the dark prior to sealing with nail polish. Staining was examined using Zeiss 880 AiryScan confocal fluorescence microscope (Carl Zeiss), using FITC at 492 nm absorption with an emission at 520 nm. The same settings were used with the controls.

2.7 Functional Assays

2.7.1 Proliferation assay

EdU assay (iFluor 488, ab219801) was used to detect and quantify cell proliferation in live cells, using both flow cytometry and fluorescence microscopy. DNA proliferation was assessed by directly measuring DNA synthesis. EdU assay was performed on CRTC1-MAML2 transfected cells and mock transfected (empty vector) cells to determine the effect of the fusion protein on cell proliferation. 1.5 $\times 10^5$ cells were seeded into a 24- well plate with 500 µl growth media to reach 70% confluency within 24 hours. Cells were then incubated for 4 hours at 37°C and 5% CO_2 with 1× EdU solution (20 µm) by replacing the 500 µl of culture media with 250 µl of 2× EdU solution plus 250 µl of normal culture media. At various time points, the media containing EdU solution was aspirated and cells fixed with 500 µl of 4% formaldehyde in PBS for 15 minutes at room temperature, protected from light. The cells then were washed twice with 500 µl wash buffer (3% BSA solution in PBS), permeabilised with 500 µl of 1× Triton ×-100 in PBS for 20 minutes at room temperature, washed twice with 500 µl of washing buffer and then treated with EdU reaction mix (table 2.20) for 30 minutes at room temperature, protected from light but on a rocking platform to ensure the reaction cocktail was distributed evenly. After 30 minutes the EdU mix was aspirated and the cells were washed with 500 µl washing buffer and then with 500 µl of PBS. Cells were then stained with 5 µg/ ml of Hoechst (Hoechst 33342, Thermo Fisher Scientific, UK) solution in PBS (1:2000) for 30 minutes at room temperature, protected from light, washed twice with 250 µl PBS, kept in PBS prior to analysis using a ZEISS AxioCam MRm2 fluorescent light microscope (Carl Zeiss, UK) and imaged using Image AxioVision Rel.4.8 image software at 20 ×magnification power.

Table 2.20 EdU reaction mix solution

Component	Reaction mix
TBS	21.4 ml
$CuSO_4$	1 ml
iFluor 488 azide	62 µl
EdU additive solution	2.5 ml
Total volume	25 ml

EdU assay data analysis was completed after coding and mixing the images to ensure unbiased results. Cells with Hoechst staining were counted using Image J with an auto cell counter plug in. At least 3 images per well of transfected and control wells were used and the mean reading recorded. Green fluorescent cells were counted, and the percentage of proliferating cells calculated per well. The assay was performed in triplicate with each cell condition. Data was analysed with Graph Pad prism and Unpaired t test was used to compare the differences between the transfected cells and controls. Flow cytometry data was analysed by a BD™ LSR II cytometer (Research Core Facility, the Medical School, University of Sheffield, UK) with established appropriate FSC vs SSC gates to exclude debris and cell aggregates. Gate EdU-positive cells based on iFluor 488 intensity (FL1 channel) or, using mean fluorescent intensity, determine fold changes between transfected and control cells. Cells excited using a 488 nm laser and data analysed using FACSDiva 8.0.1. The flow analysis was performed by Sue Clark (Flow Cytometry Core Facility technician).

2.7.2 Migration assay

A scratch (wound) assay was used to study the migration of the CRTC1-MAML2 fusion transfected cells. Wound closure was studied by taking snapshot pictures at different time points: T0, T8, T16, T24, T32, T40, and T48. Based on the width of the wound, the migration distance and the speed of the cell migration was

calculated. Cells were seeded into a 6-well plate and after reaching 100% confluency, 2 µg/ ml of Mitomycin C was added to the cells for two hours to inhibit cell proliferation. The optimum dose of Mitomycin was previously determined in our lab (Dr Vivian Wagner). The cells were then washed with PBS and images captured of the pre-wounded confluent cell monolayer. A wound was made in the centre of the well using a p-200 tip applying the same pressure across the wound for all wounds. Wounding was confirmed using an inverted microscope and the cells were washed twice with PBS to remove detached cells. Images captured at this point were assigned as time zero (T0; with the wound being 100% open) with two pictures per well and 4 wells per condition. Cells were examined every 8 hours, the same two points in each wound were imaged with the same light conditions, using the same microscope, same magnification (10×) and all images saved as tiff files of the same size (figure 2.5). The experiment was carried out in triplicate and Image J software was used to analyse migration, with wound area being defined using freehand selection. At each time point the area free of migrating cells was calculated and expressed as a percentage of the initial scratched area. Data was analysed on Graph Pad prism; a two-way ANOVA test was used to determine the effect of two nominal predicated variables on a continuous outcome variable.

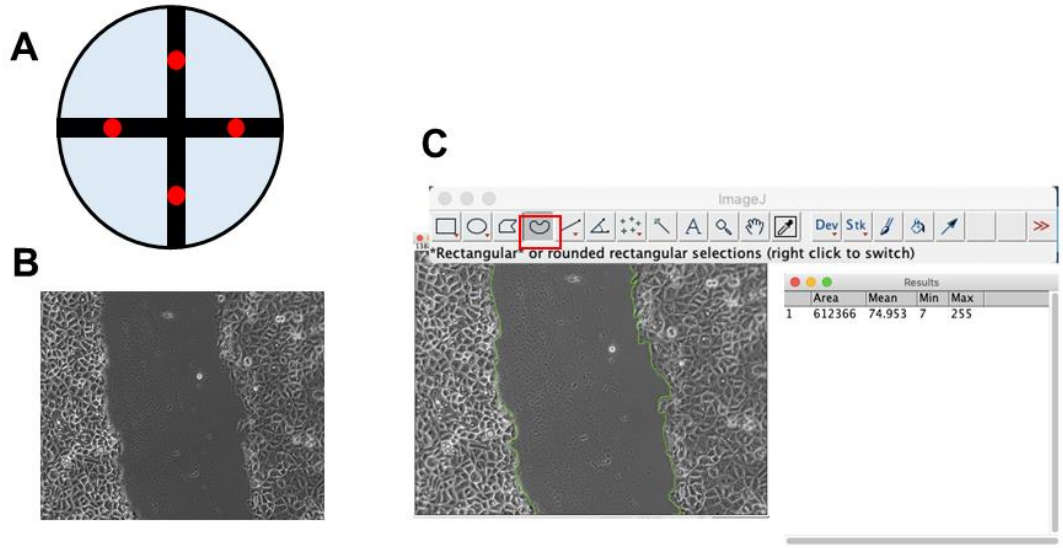


Figure 2.5 Migration assay. The technique involves basic steps of making a linear thin scratch "wound" (creating a gap) in a confluent cell monolayer (A); data acquisition was through microscopic image capturing two points per wound (Red dots) (A) and gap measurement at each time point, T0 (B); and data analysis using Image J software and selection and measuring of wound area (C).

2.7.3 Invasion assay

A trans-well invasion assay was used to study the effect of CRTC1-MAML2 fusion protein expression on cell behaviour and ability to invade the extracellular matrix (ECM) substitute. The assay was performed on 24-well plates using inserts with 8 µm pores (Falcon, Loughborough, UK). Two coating materials were used in the invasion assay; Fibronectin and pre-coated Matrigel trans wells. Inserts were coated with 10 µg/ ml fibronectin in PBS (Sigma-Aldrich), incubated for 24 hours at 4°C, washed twice with sterile PBS and left to dry in a tissue culture hood for 24 hours at room temperature. Inserts were stored in 100 mm plates sealed with parafilm at 4°C prior to use. Alternatively, precoated Corning® BioCoat™ Matrigel® Invasion Chambers (Catalog No. 354480, 354481) were used. 1×10^4 CRTC1-MAML2 fusion transfected and control cells were seeded onto the fibronectin / Matrigel-coated inserts in 400 µl of normal growth media supplemented with 10% FBS, 700 µl of the same media was

placed in the bottom chamber and the plate incubated at 37°C and 5% CO_2 for 4 hours. After 4 hours the media in the transwell chamber was changed to growth media supplemented with 1% FBS; the media under the insert remained the same to act as chemoattractant to the cells and incubated at 37°C and 5% CO_2 for 48 hours. On day two, all of the non-invaded cells were removed with a cotton bud. Cells were fixed with 500 µl of 100% cool-methanol and incubated for 6 minutes at -20°C, the cells were then stained with Harris Haematoxylin (Thermo-Scientific) and Shandon Eosin Y (Thermo-Scientific) as described in table 2.21. After staining the insert, membranes were cut out using a scalpel blade and put onto a glass slide with invading cells facing up. The slides were mounted with mounting media (EcoMount) and sealed with cover slips (figure 2.6).

Table 2.21 Haematoxylin and Eosin staining schedule for invaded cells

Step	Time
PBS	5 min
PBS	5 min
Haematoxylin	2 min
Distilled water	5 min
Distilled water	5 min
Eosin	2 min
Distilled water	5 min
Distilled water	5 min

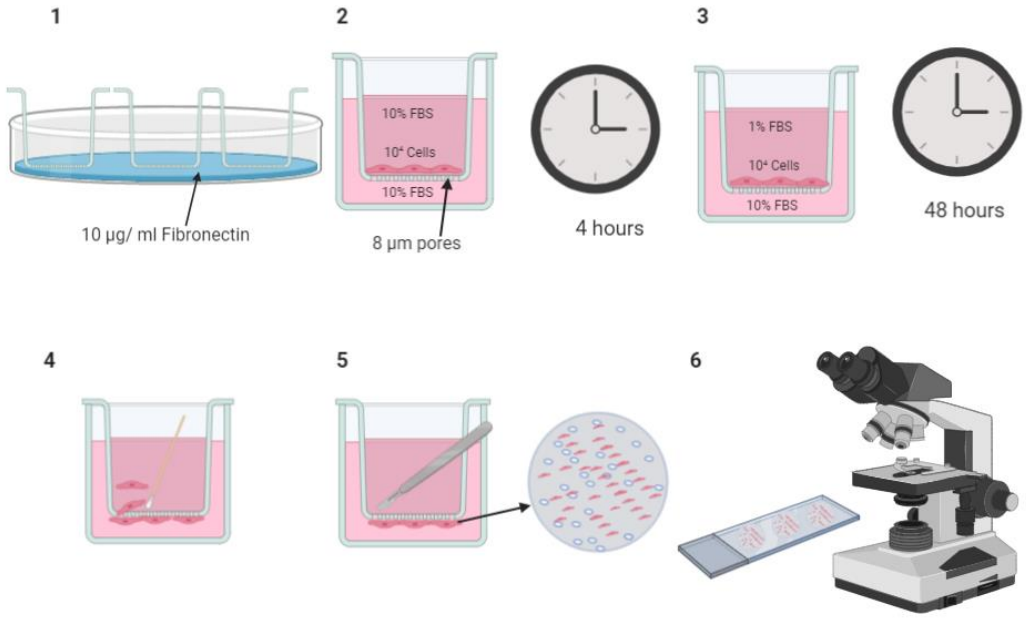


Figure 2.6 Trans well invasion assay diagram CRTC1-MAML2 transfected, mock transfected and non-transfected HEK293 cells were seed in 24 well plate trans well inserts previously coated with 10 µg/ ml fibronectin (1-2) (PET membrane, 8.0 µm, Falcon, Loughborough, UK) in 10% serum (2), Cells were left to settle for 4 hours and then media was changed to 1% FBS and left to migrate for 48 hours (3) to media containing 10% FBS. Following 48 hours, non-invaded cells were removed with cotton bud (4). The invaded cells were fixed, stained, and the membrane was cut (5). Membrane was mounted on glass slides and analysis (6). Diagram was created using bio RENDER.com software.

Quantification of invaded cells was carried out blind to ensure unbiased results, using the cell counter function on Image J software. Experiments were performed on at least six replicates, three images of each membrane were captured at 400× magnification using a light microscope (OLYMPUS BX51), the invaded cells were counted, the average score calculated for each membrane and the values used for statistical analysis. The pre-coated Matrigel inserts were finally chosen over the fibronectin coated inserts, because of the quality of the images and ability to differentiate between the cells and pores in the inserts. Data was presented on Graph

Pad prism; an Un-paired two-tailed t test was used to compare the average between the two groups.

2.7.4 Clonogenic Assay

This assay was carried out to investigate the ability of single cells to grow into colonies, to undergo unlimited cell division and whether the CRTC1-MAML2 fusion protein influenced cell survival. CRTC1-MAML2 and empty vector transfected cells were seeded per well in a 6-well plate (600 per well) in normal growth media and the cells allowed to grow for 10 days. Cells were washed twice for 5 min with PBS and fixed with 100% Methanol mixed with 100% acetic acid (7:1) and left for 5 min at room temperature. Fixative was removed and 0.2% of Crystal Violet staining solution added to each well for 2 min. Staining solution was removed, the plate placed under running tap water for 3 min and then left to dry. Colonies were counted by scanning the plate and using ImageJ to analyse the number of cells and size of the colonies. Colonies with cell number less than 50 were not included in the analysis (Franken *et al*., 2006). The assay was carried out in triplicate.

2.8 Three-Dimensional Models

Human primary salivary gland cells at low passage number (0-4) were used to develop 3D organoids cultured on inserts at an air-liquid interface (ALI). Explanted tissue samples were obtained from surgical biopsies which had been previously processed by Dr Zulaiha Rahman (previous PhD student), Dr Lynne Bingle and Miss Rachel Furmidge (undergraduate placement student) under ethical approval number 13/NS/0120. Human primary salivary gland cells transfected with CRTC1-MAML2 fusion construct were used to grow 3D organoid fusion positive models and MEC cell lines (NCI-H292, UM-HMC 2, and A-253) to grow the 3D models of cancer cell lines.

2.8.1 Developing 3D Organoid model of salivary glands

The organoid models were prepared from normal cell cultures, maintained in KGM medium. When the cells had reached 80% confluency, they were trypsinised, counted as described in section 2.2.6 and resuspended with freshly prepared DMEM/F12 media, supplemented with 10% R-Spondin, 50% Wnt3a conditioned media and other growth factors (table 2.22). Y-27632 (Sigma) was added to the cells for the first three days only.

8×10^5 cells in 100 µl cell media was mixed gently using a cooled pipette tip, without producing bubbles, with 200 µl of ice-cooled basement membrane extract (BME) Matrigel™; this mixture was sufficient for four inserts (figure 2.7). The cell/matrix suspension was seeded into a pre-warmed 12-well plate with 0.4 µm pore, PET transwell inserts (FALCON®). 500 µl of fresh media was added to the cells and 1500 µl was placed under the insert before being incubated at 37°C with 5% CO_2. After four days the culture media was removed from the insert so that the organoids were now at an ALI and the media under the insert replenished; this was counted as day zero. The exhausted media was replaced every 2-4 days and the model inspected daily by inverted bright field microscopy. Growth was continued for up to two weeks (figure 2.8).

2.8.2 Developing 3D Organoid models of CRTC1-MAML2 transfected salivary glands cells

The organoid model was prepared from monolayer cells transfected with FLAG-Tagged CRTC1-MAML2 pCR3.1 construct. The method described in section 2.8.1 was used to establish 3D transfected cell models.

2.8.3 Developing 3D models of Mucoepidermoid Carcinoma cell lines

MEC 3D organoid modes were developed from monolayers of A-253, UM-MEC-2 and NCI-H292 MEC cell lines, maintained in growth culture media specific to each cell type (Table 2.23). The method described in section 2.8.1 was used to establish 3D MEC cell models (figure 2.9).

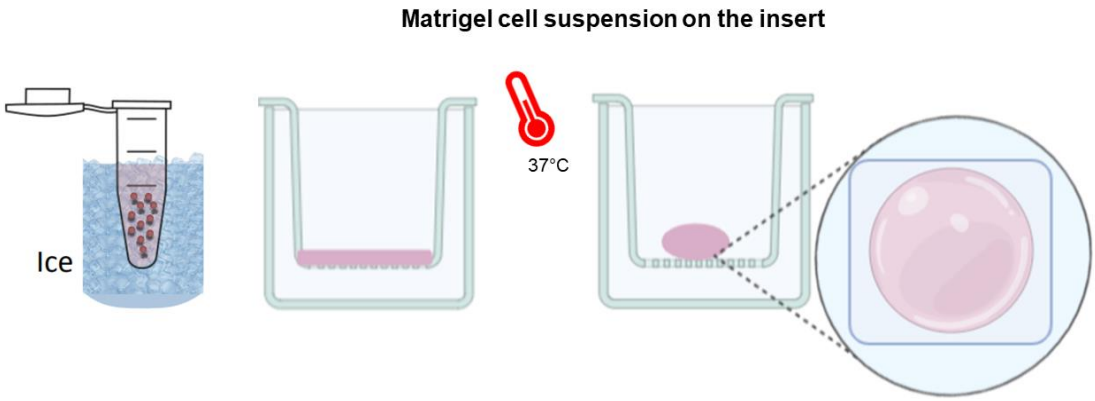


Figure 2.7 Schematic diagram of seeding the cells with the Matrigel. Cells were mixed with Ice-cooled Matrigel using a cooled pipette tip. Cells-ECM suspension were then seeded on the middle of a pre-warm 12-well plate insert and incubated for 30 min at 37°C with 5% CO_2 to allow the Matrigel to solidify. Diagram was created using bio-RENDER software.

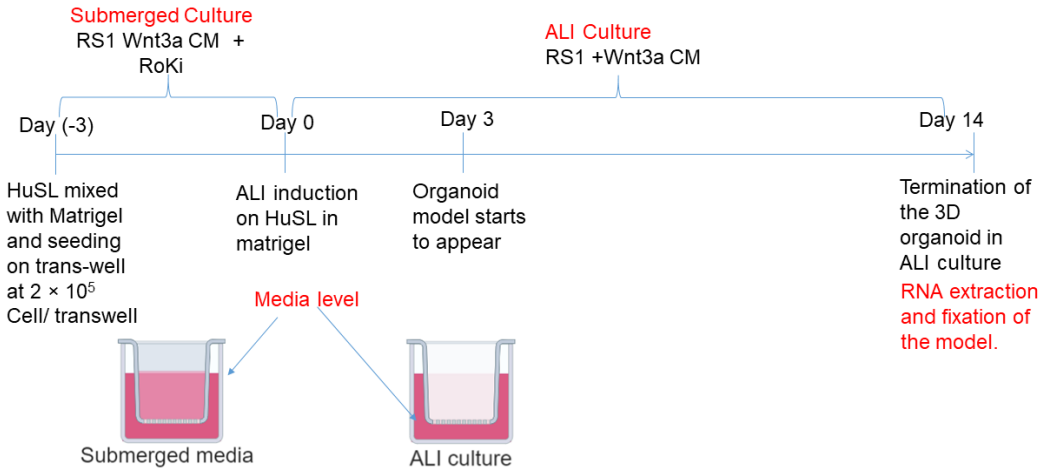


Figure 2.8 Developing 3D organoid model of salivary glands Human primary salivary gland cells (mixed populations of epithelium and connective tissue cells) with low passage number (0-4) were mixed with Matrigel and grown in freshly prepared media which mainly depended on the presence of R-Spondin and Wnt3a on insert at air-liquid interface (ALI). After

one to three days of growing the organoid structure was noticed and after 14 days was successfully developed. Diagram was created using bio RENDER software.

Table 2.22 Compositions of the freshly prepared media used to grow the organoid 3D Model

Component	Final concentration	Supplier
Dulbecco's modified Eagle's Medium (DMEM) low glucose	(3:1) DMEM: F12	Sigma-Aldrich®
Nutrient mixture F12 HAM	(1:3) F12: DMEM	Sigma-Aldrich®
Penicillin/Streptomycin	100i.u./ml penicillin and 100µg/ml streptomycin	Sigma-Aldrich®
L-Glutamine	1%	Sigma-Aldrich®
N-2 supplements	1%	Gibco®
Epidermal Growth Factor (EGF)	20 ng/ml	Sigma-Aldrich®
Rh-FGF2 fibroblast growth factor recombinant human	10 ng/ml	R&D SYSTEMS
Human Insulin	10µg/ml	Sigma-Aldrich®
Dexamethasone	1 µM	Sigma-Aldrich®
*Y-27632 (RHOKi) Protein Kinase Inhibitor	10 µM	Sigma-Aldrich®
R-Spondin1 (4645-R5)	10%	Conditioned media
Wnt3a (50.6-WN	50%	Conditioned media

*Note: Y-27632 additive was only added to the first preparation of the media

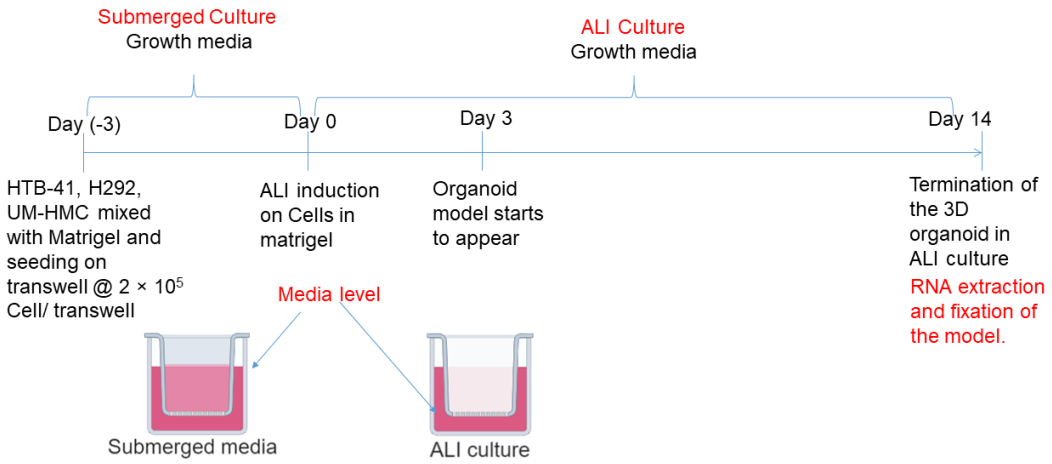


Figure 2.9 Mucoepidermoid carcinoma cell lines 3D model. Human mucoepidermoid carcinoma cell lines (A-253, NCI-H292, UM-HMC2) were mixed with Matrigel and grown in culture media on insert at air-liquid interface (ALI). After one to three days of growing the organoid structure was noticed and after 14 days was successfully developed. Diagram was created using bio-RENDER software.

Table 2.23 Compositions of the media used to grow the MEC organoid 3D Model

Cell type	Growth Medium	Supplements	Supplier
A-253	McCoy's 5A medium modified	10% FBS, 1% L-G, 1% P/S	Sigma-Aldrich
UM-HMC-2	High glucose Dulbecco's modified Eagle's medium (DMEM)	10% FBS, 1% P/S, 1%L-G, 180 µM adenine, 0.5 µg /ml hydrocortisone 10 ng/ml EGF	Sigma-Aldrich
NCI-H292	Low glucose Dulbecco's modified Eagle's medium (DMEM)	10% FBS, 1% L-G, 1% P/S	Sigma-Aldrich

2.8.4 Fixation, processing and sectioning of 3-D models

On day 14, the 3D models were inspected under microscopy to confirm development. Models were fixed in 10% neutral buffered formalin (NBF) for 24 hrs, processed and sections prepared for H&E and IHC staining. For immunofluorescence staining, wells were fixed in 100% ice-cold methanol for 20-30 minutes at -20°C, washed three times with 200 µl PBS and stored in PBS at 4°C.

Samples were processed overnight using a Citadel 2000 (SHANDON) benchtop tissue processor. The processed samples were embedded perpendicular to the bottom surface of the mould in molten paraffin wax using a LEICA EG1150 H embedding station so that they could be cut in the correct orientation. 5 µm sections were mounted onto SuperFrost Plus slides and left to air dry. Before use the slides were baked at 60°C for at least 30 minutes, to ensure adherence of the sections to the slides.

For Haematoxylin and Eosin staining, an automated linear stainer (LEICA ST 4040) was used. Slides were mounted with DPX mounting media, visualised and imaged with Cell^D software (Olympus soft imaging solutions, GmbH, Münster, Germany).

2.8.5 Periodic Acid Schiff staining of three-dimensional organoid models

Periodic Acid Schiff staining was performed on 3D MEC cell models. Slides were deparaffinised, hydrated, washed with distilled water, immersed in 0.5% periodic acid solution for 5 minutes and washed three times with distilled water. Slides were then placed in Schiff reagent for 15 minutes until they became light pink and were then washed in lukewarm tap water for 5 minutes to a dark pink colour. Slides were counterstained with Harris Haematoxylin for 1 minute and washed with tap water for 5 minutes. Slides were dehydrated with alcohol and coverslipped using DPX mountant media.

2.8.6 Immunofluorescence staining of 3D organoid models

Models were fixed with 4% paraformaldehyde, washed twice with 200 µl of PBS, permeabilised with 0.2% Triton ×-100 in PBS for 30 minutes at room temperature and washed twice with PBS. Phalloidin conjugated green fluorescent solution (Abcam®) was added to the models and incubated overnight at 4°C in the dark. After aspirating the Phalloidin solution models were washed three times with PBS and DAPI was added (1:1000 dilution in PBS) for 30 minutes in the dark. Models were washed three times with PBS and the membrane removed from the plastic support, mounted on a glass slide (model facing up) in DAPI-FREE ProlongGold antifade mountant and kept in the dark at room temperature for 24 hours. The slides were sealed with nail polish and stored at 4°C to be examined under a fluorescence microscope.

2.9 Online Resources

All online software tools used in this study are summarised in table 2.24.

Table 2.24 online analysis tools used in the study

Analysis Tool	Purpose	Software URL Link
NCBI VecScreen	Screen a Sequence for Vector Contamination	https://www.ncbi.nlm.nih.gov/tools/vecscreen/
NCBI blastn	Search query sequence with nucleotide database	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome
NCBI Blastx	Search protein databases using a translated nucleotide query	https://blast.ncbi.nlm.nih.gov/Blast.cgi?LINK_LOC=blasthome&PAGE_TYPE=BlastSearch&PROGRAM=blastx
NEBcutter V2.0	Search for Restriction Enzyme Mapping	http://nc2.neb.com/NEBcutter2/
NCBI' ORF finder	Finding Open reading frames in unknown nucleotide sequence	https://www.ncbi.nlm.nih.gov/orffinder/
Clustal W	Multiple sequence alignment	https://www.ebi.ac.uk/Tools/msa/clustalo/
Primer3	Primer design tool	http://bioinfo.ut.ee/primer3-0.4.0/
Clustal Omega	Multiple sequence alignment	https://www.ebi.ac.uk/Tools/msa/clustalo/

2.10 Statistics

GraphPad Prism 8 software was used to determine statistical significance. An unpaired t-test was used to compare two groups. One-way analysis of variance was used to analyse more than two groups. A two-way ANOVA test was used to determine the effect of two nominal predicated variables on a continuous outcome variable. A *P*-value of < 0.05 was considered to be statistically significant. The number of biological repeats is shown as 'N='. Technical repeats are represented as 'n='.

Asterisks denote statistical significance (*P <0.05; **P <0.01; ***P <0.001; ****P <0.0001, and NS P >0.05).

2.11 Ethical Consideration

Ethics approval for use of clinical data of MEC cases and fixed tissues was in accordance with NREC (Ref number: 05/Q2305/127) and Santa Rita Hospital of the Santa Casa de Misericordia, Porto Alegre, Brazil (Human Research Ethics Committee approval: 74754317.5.0000.5335). Isolation and the use of the primary salivary gland cells were in accordance with NREC (Ethical approval number 13/NS/0120).

Chapter 3
In Situ Detection of the CRTC1-MAML2 Translocation Expression in Mucoepidermoid Carcinoma

3. Chapter 3

3.1 Introduction

The heterogeneity of salivary gland neoplasms, within and between histological types, presents a major challenge for accurate diagnosis (Speight and Barratt, 2020). As indicated in chapter 1, several cytogenetic analyses have indicated that more than 50% of salivary gland MEC are associated with the CRTC1-MAML2 translocation (Anzick *et al*., 2010, Shinomiya *et al*., 2016). In addition to the salivary glands, the CRTC1-MAML2 fusion has also been identified in MEC in other tissues including lung, breast, cervix, thyroid and skin (Amelio *et al*., 2014).

Determination of fusion gene expression in MEC cases could assist in diagnosis, especially in more difficult cases, or for the exclusion of other entities. An alternative fusion of MAML2 with CRTC3 has been documented in the literature but with relatively low abundance, as discussed in chapter 1, and so this study will focus only on the CRTC1-MAML2 translocation (Fehr *et al*., 2008, Nakayama *et al*., 2009). Currently expression of CRTC1-MAML2 fusion is detected either by RT-PCR or FISH (Bell and Hanna, 2012) (figure 3.1), however, there are weaknesses associated with these techniques. For example, the Fluorescence *In Situ* Hybridization (FISH) MAML2 DNA break-apart technique is complicated, does not provide specific information about the fusion event or the partner gene involved and there is considerable uncertainty about its application in the clinic. Furthermore, RT-PCR only measures the transcript in solution, without any histological details. Thus, an alternative and accurate diagnostic test that can be routinely performed on Formalin Fixed Paraffin Embedded (FFPE) tissues is needed.

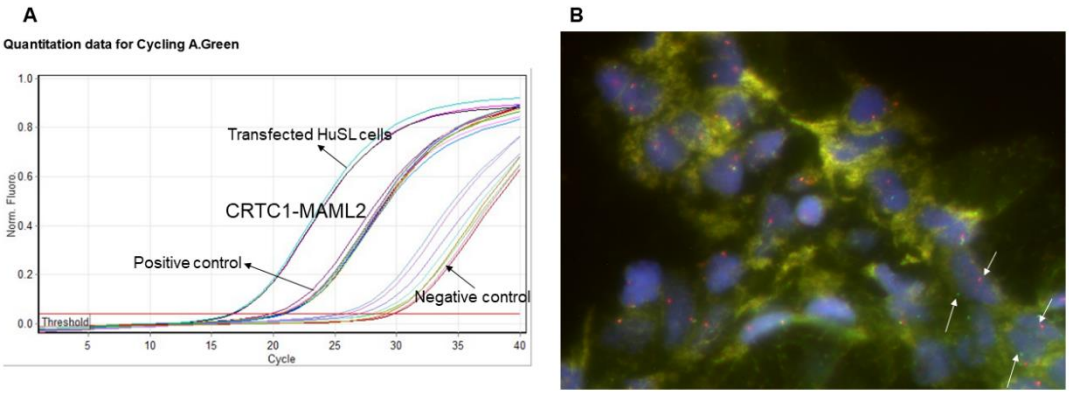


Figure 3.1 Current techniques used to detect the CRTC1-MAML2 translocation. A) Real-time RT-PCR shows amplification of CRTC1-MAML2 fusion product. **B)** Fluorescence *in Situ* Hybridisation (FISH) for MAML2 rearrangements (arrows) indicate break-apart separation of the red and green signals within the cells.

In the present study, we have developed an RNA *In Situ* Hybridisation (RISH) detection method targeting the fusion transcript junction. RNA probes are single-stranded and offer several advantages over DNA probes including improved signal or hybridization blots. We applied this technique to detect and quantify expression of the CRTC1-MAML2 transcript, using a specific probe straddling the junction between exon 1 of CRTC1 and exon 2 of MAML2. Following validation of the detection of junction-specific CRTC1-MAMl2 transcripts in cell lines and in FFPE specimens from MEC patients, we quantified CRTC1-MAML2 transcript expression in positive cases. We present, for the first time, the visualisation of CRTC1-MAML2 fusion transcripts in morphologically intact tissue and demonstrate, again for the first time, a specific and quantifiable CRTC1-MAML2 RNA *in situ* hybridisation test for detection of the fusion in MEC patients.

3.2 Aim and Objectives

The aim of this study was to detect the expression of CRTC1-MAML2 translocation in MEC and I achieved this by;

- Designing a novel BaseScope probe targeting the exon-exon junction in the CRTC1-MAML2 fusion transcript.
- Identifying specific cell types harbouring the translocation.
- Determining expression levels of the transcript and quantifying the positive signals.

3.3 Methods

The relevant materials and methods utilised in this chapter are detailed in chapter 2, under heading 2.1 as indicated below:

- Patient samples and clinical data is presented in section 2.1.1.
- Probe design, sample preparation and pretreatment were performed as described in section 2.1.2 and 2.1.3.
- Probe Hybridization and amplification were performed as described in section 2.1.4.
- BaseScope assay signal detection and analysis were assessed using a semi-quantitative histological methodology based on ACD scoring criteria.

3.4 Results

3.4.1 Histopathological findings

A retrospective cohort of 30 MEC cases from 2007- 2016 was examined in this study. The clinicopathological characteristics of the patient cohort are presented in figure 3.2. The histological features and classification of the cases were reviewed with H&E sections; all histological findings are summarised in table 3.1. The tumours consisted predominantly of epidermoid cells, intermediate cells and mucous secreting cells of variable proportion, however, six (20%) of the 30 cases were clear-cell variants of MEC (figure 3.4). The most common histological type was cystic, 18 of 30 (60%), varying between macro and micro cysts. Of these cases, 10 (33.3%) of 30 demonstrated solid growth patterns. Two (6.6%) of the 30 cases had a mix of solid and cystic components. Twenty-one of the cases (70%) were low grade, 13.3% (4/30) were intermediate grade and 16.7% (5/30) were high grade (figure 3.3). Necrosis and perineural invasion were seen in 3/30 (10 %) and 4/30 (13.3%) of the cases, respectively. Notably, none of the cases demonstrated an angiolymphatic invasion. The mitotic rate was low, with most of the cases showing a 0 mitotic rate per 10 high magnification fields. Ten cases showed mitosis of 1 to 2 per 10 high power fields and only two cases showed mitosis of 5 to 6 per 10 high power fields.

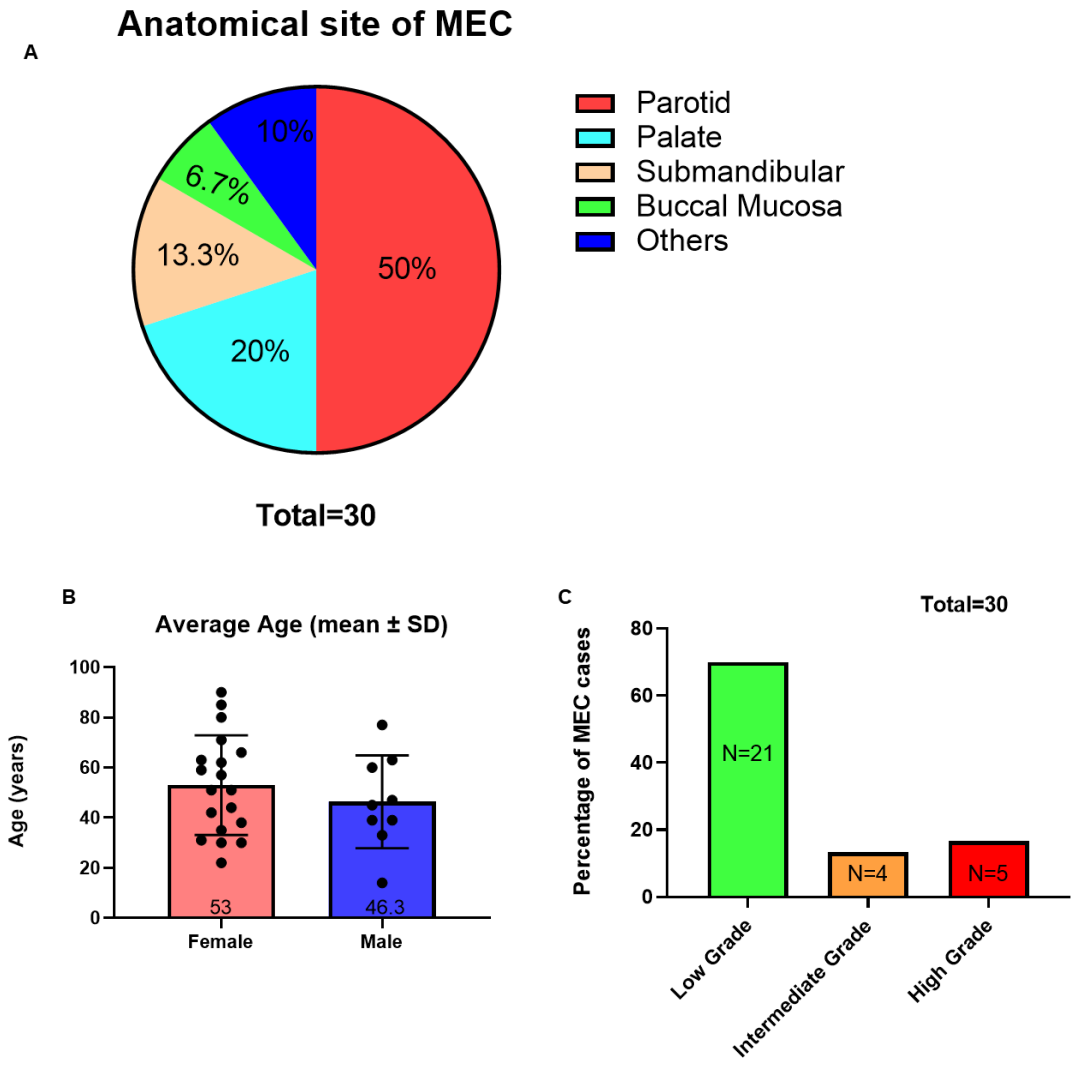
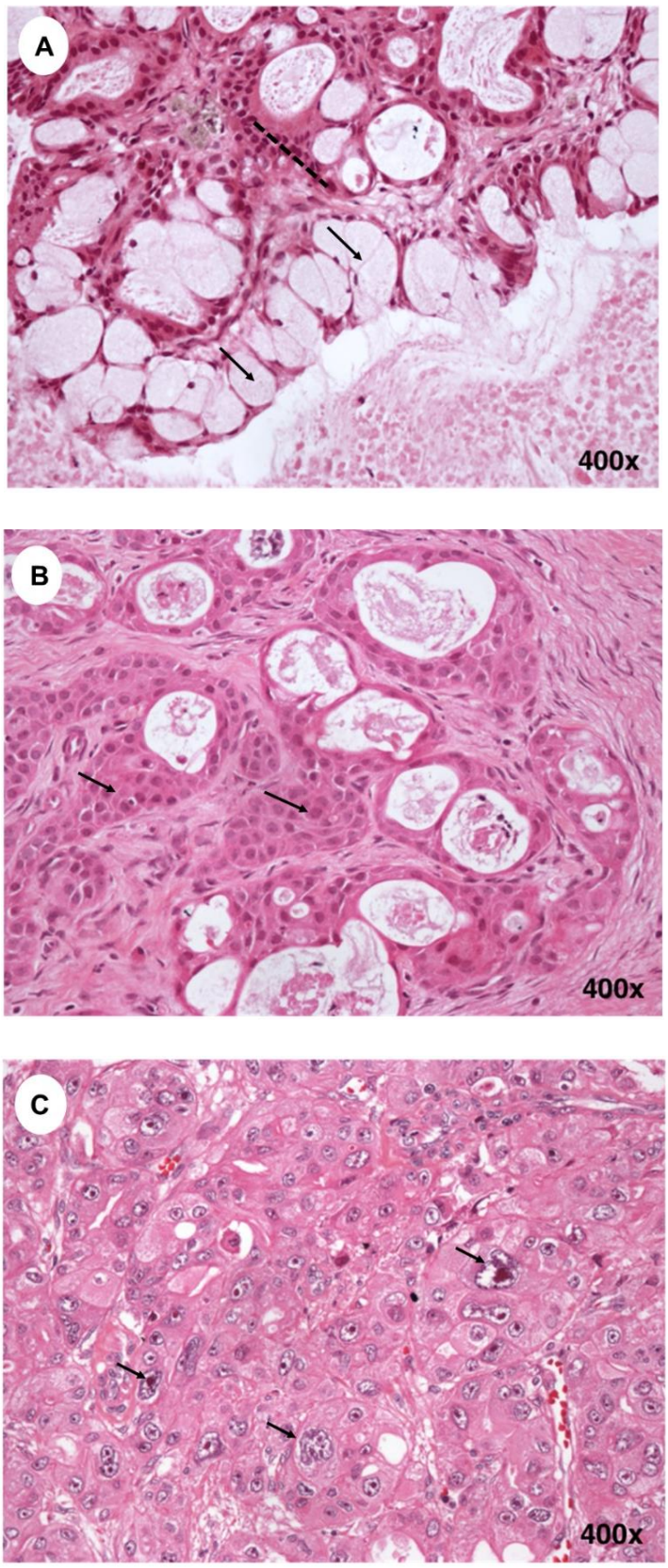
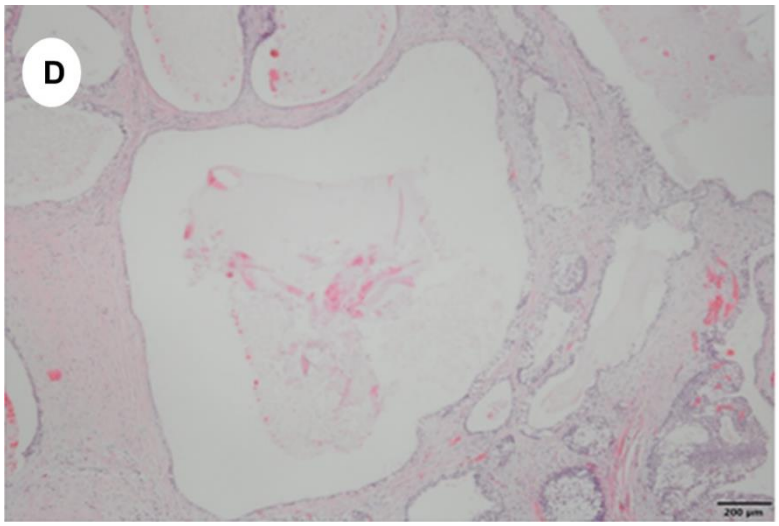
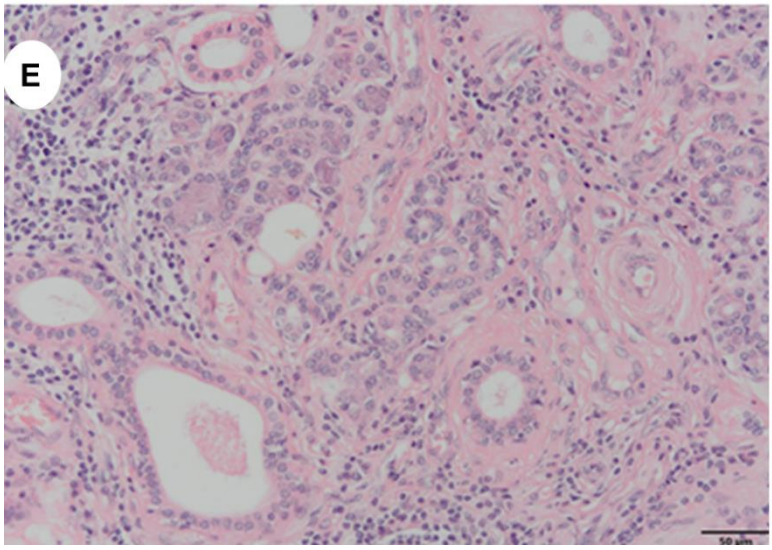


Figure 3.2 The anatomical site, average age and histological grades of MEC cohort. A) Anatomical location of MEC displays that most of the carcinomas (50%) were developed in the Parotid (major salivary glands), followed by the palate (20%, minor salivary glands). 13.3% of MEC cases arose from the submandibular glands (major salivary glands) and 6.7% of MEC cases initiated in the buccal mucosa (minor salivary glands). **B)** Average age of MEC cases (Mean ± SD) shows a broad age range of female patients compared to male cases. **(C)** Histological grading of MEC cases indicates that 70% of the cases were low grade, 13.3% were intermediate grade and 16.7% were high grade.


A
400x
B
400x
C
400x


D


E

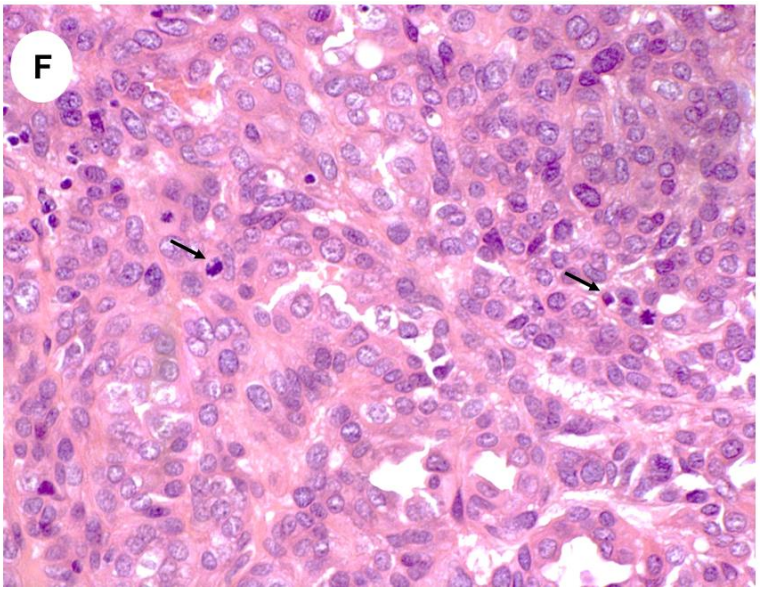

F

Figure 3.3 Haematoxylin and Eosin staining of MEC. Low grade MEC consists of large proportions of mucous-secreting cells (arrows), intermediate cells (dashed lines) and large cyst formations **(A and D)**. Intermediate grade MEC showing small size cyst with a mixture of intermediate cells, epidermoid cells (arrows) and lower proportions of mucous cells **(B and E)**. High-grade MEC showing solid growth of epidermoid and intermediate cells with lack or little presence of mucous cells, lack of microcytic formation and the presence of cellular atypia **(**arrows in **C)** and mitotic figures **(**arrows in **F)**.

Table 3.1 Histopathological findings of MEC cases including histological type, grade and the presence of other histological parameters

Case	Histologic type	Histologic grade (AFIP)	Necrosis	Angio-lymphatic invasion	Perineural invasion	Mitotic count/10 high-power field
1	Cystic	Low	No	No	No	0
2	Cystic	Low	No	No	No	2
3	Clear cell	Low	No	No	No	2
4	Cystic	Low	No	No	No	1
5	Solid clear cell	Low	No	No	No	0
6	Cystic	Low	No	No	No	0
7	Solid	Low	No	No	No	0
8	Solid clear cell	Low	No	No	No	1
9	Cystic	Low	No	No	No	0
10	Cystic	Low	No	No	No	0
11	Solid	Intermediate	Yes	No	No	1
12	Solid Clear	Intermediate	No	No	No	1
13	Solid/ cystic/ clear	Low	No	No	No	0
14	Solid	Low	No	No	No	0
15	Clear	Intermediate	No	No	Yes	0
16	Solid microcystic	Low	No	N0	No	1
17	Cystic	Low	No	No	No	0
18	Solid	High	No	No`	Yes	5
19	Cystic	Low	No	No	No	0
20	Cystic	Low	No	No	No	0
21	Cystic	Low	No	No	No	0
22	Cystic	Intermediate	No	No	Yes	0
23	Cystic	Low	No	No	No	0
24	Cystic	Low	No	No	No	0
25	Cystic	Intermediate	No	No	No	0
26	Cystic	High	No	No	No	2
27	Cystic	High	No	No	No	6
28	Cystic	High	Yes	No	No	2
29	Solid clear cell	High	No	No	No	0
30	Solid	High	Yes	No	No	1

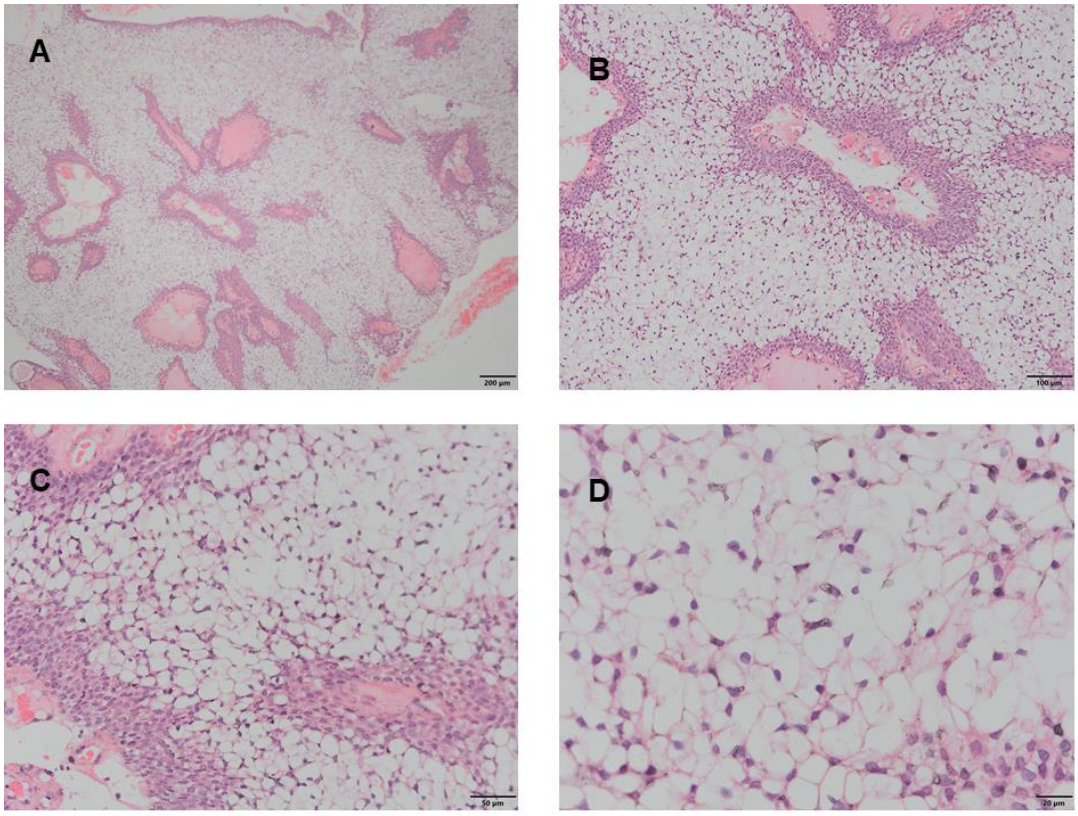
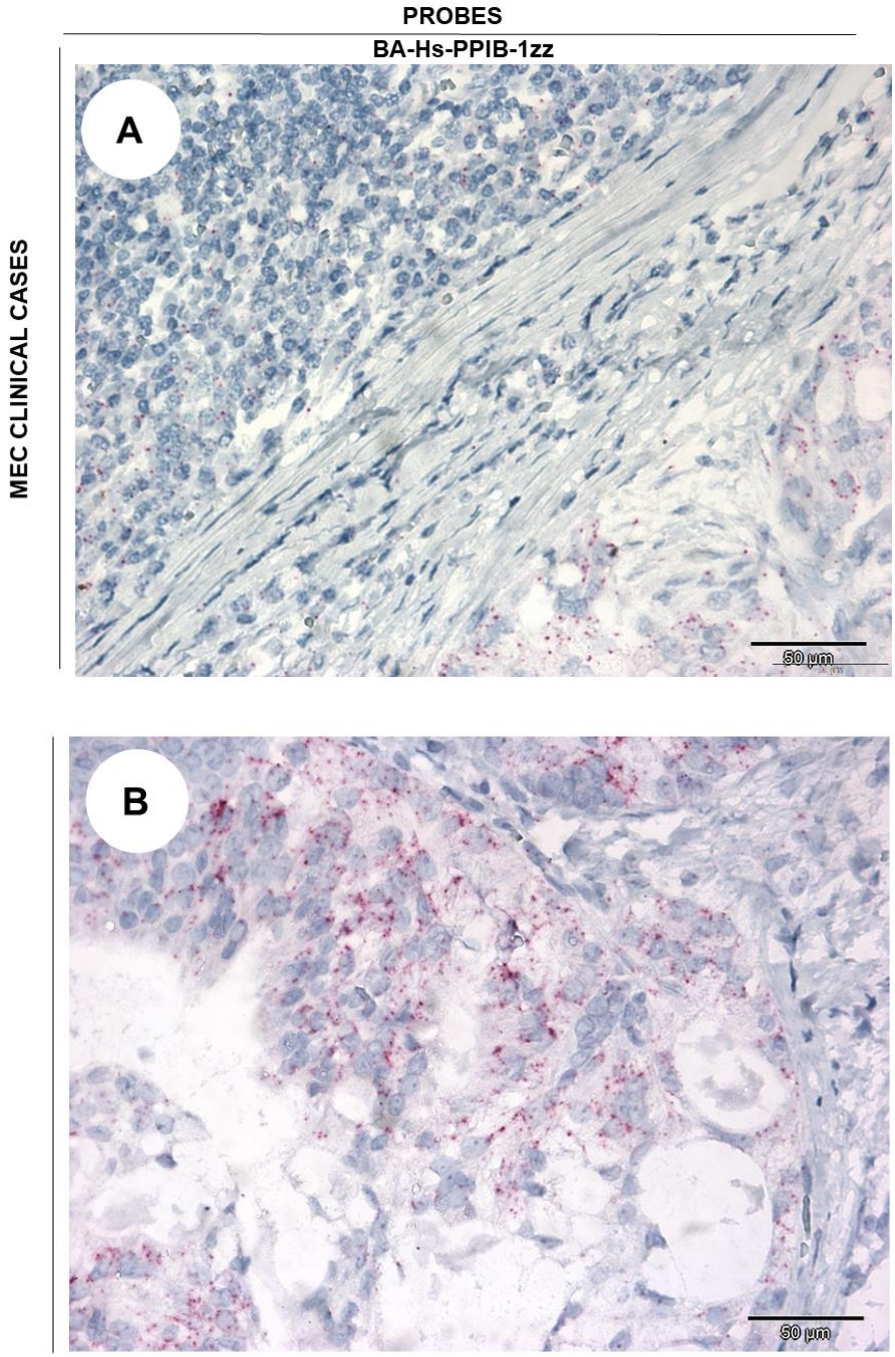
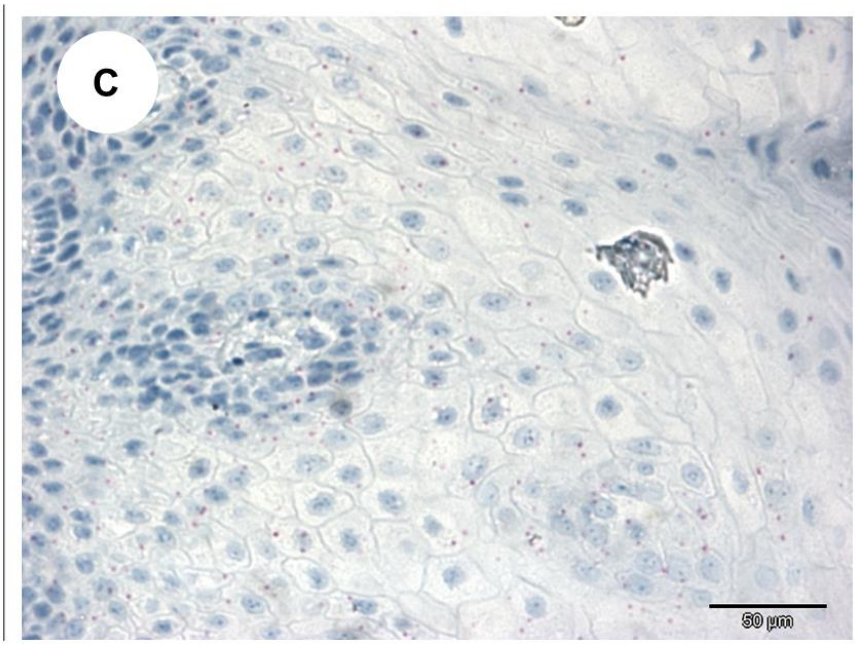


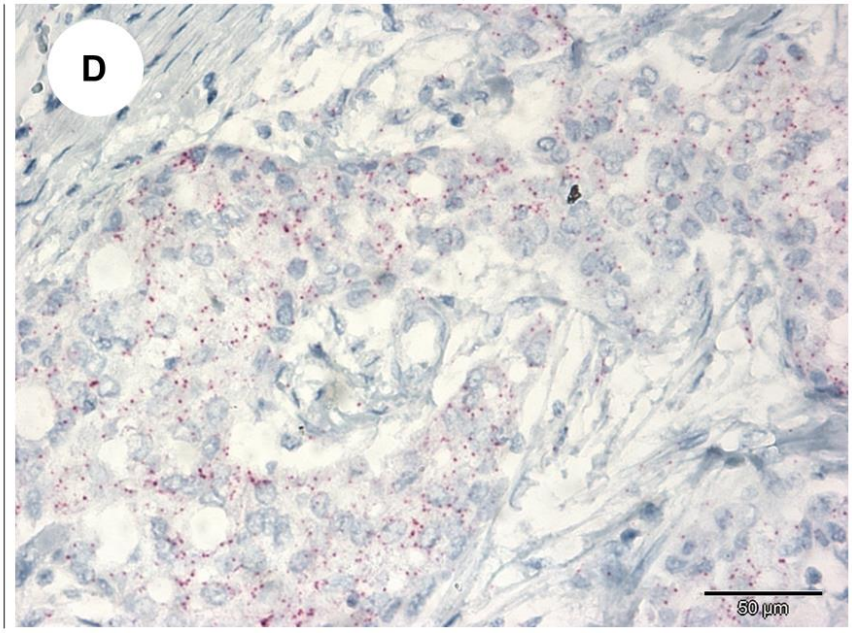
Figure 3.4 Clear cell MEC. Large cystic cavity lined with mucous and epidermoid cells with clear cell differentiation across the tumour.

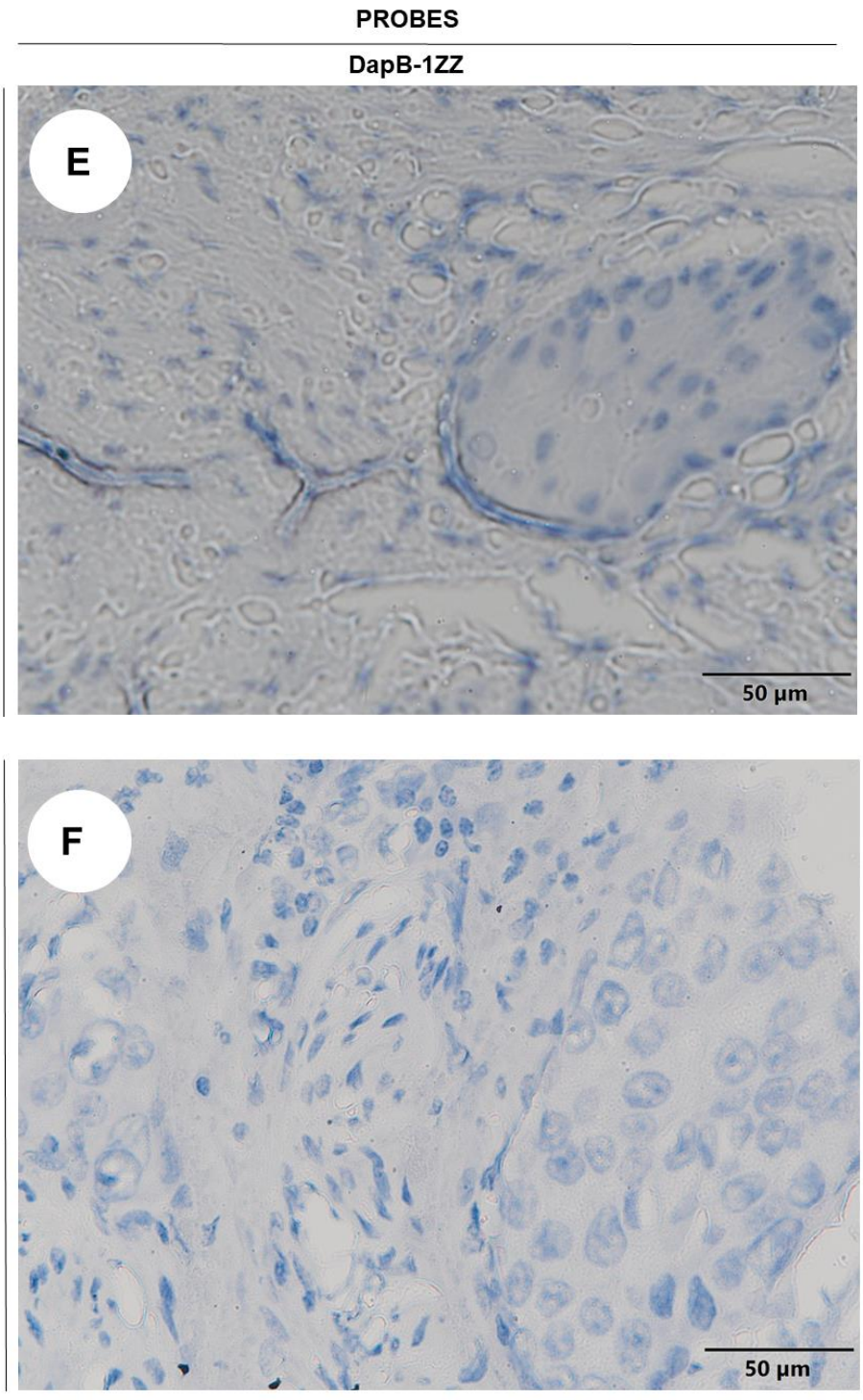
3.4.2 Assessment of the integrity of the mRNA in MEC cases

Before applying the fusion BaseScope assay to the MEC cases, the integrity of the mRNA in the FFPE samples was examined using the probe against the housekeeping gene (PPIB), BA-Hs-PPIB-1zz, and the bacterial gene probe, DapB-1ZZ. Positive signals were detected as punctate red dots in the cytoplasm or nucleus (figure 3.5). Across the cases examined, an average of 95.6% showed positive signals with the positive control probe (BA-Hs-PPIB-1zz) and none were positive for the negative control probe (DapB) (figure 3.5). One sample was excluded from the study as it did not meet the mRNA quality criteria (no signal with the BA-Hs-PPIB-1zz positive control probe). Notably, the positive signals were detected in normal tissues, lining epithelia, tumour islands and associated stroma (figure 3.5). However, the expression level varied, possibly due to variability in the degree of mRNA degradation, which may have occurred during sample fixation and processing, and the inter-tumour heterogeneity of the expression level of BA-Hs-PPIB-1zz mRNA.


PROBES
BA-Hs-PPIB-1zz
MEC CLINICAL CASES
A
50 µm
B
50 µm


C
50 µm


D
50 µm


PROBES
DapB-1ZZ
E
50 µm
F
50 µm

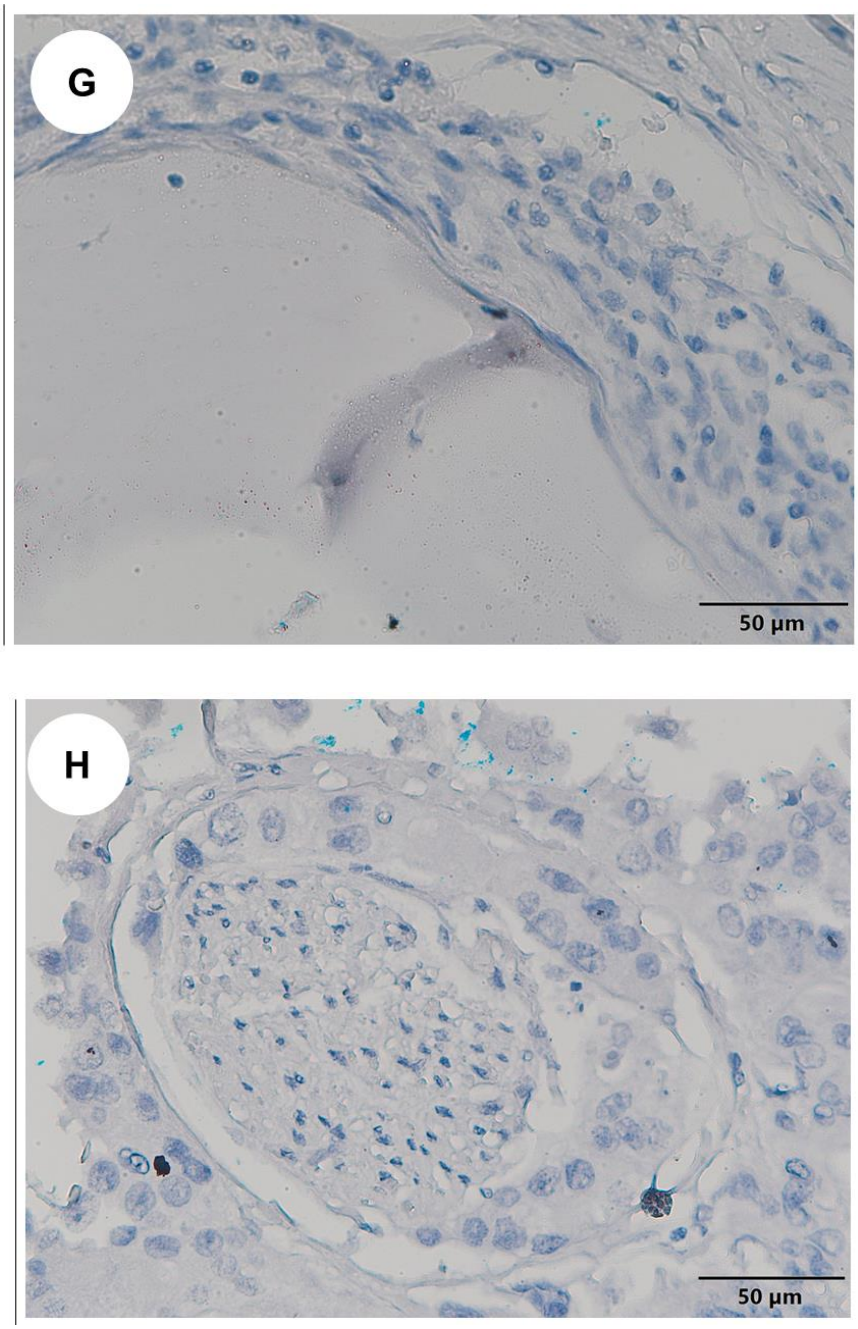
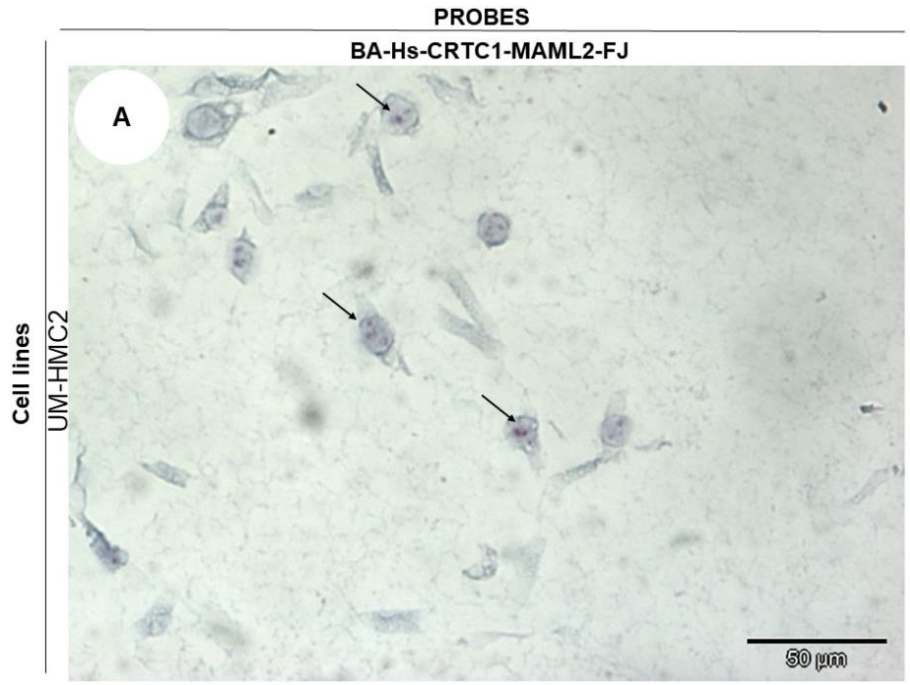
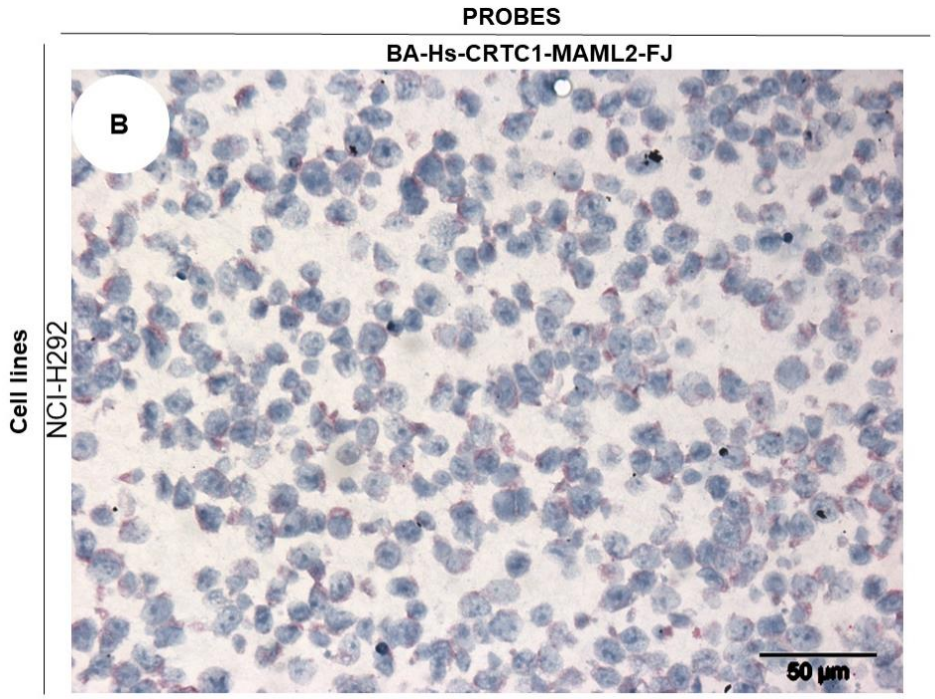


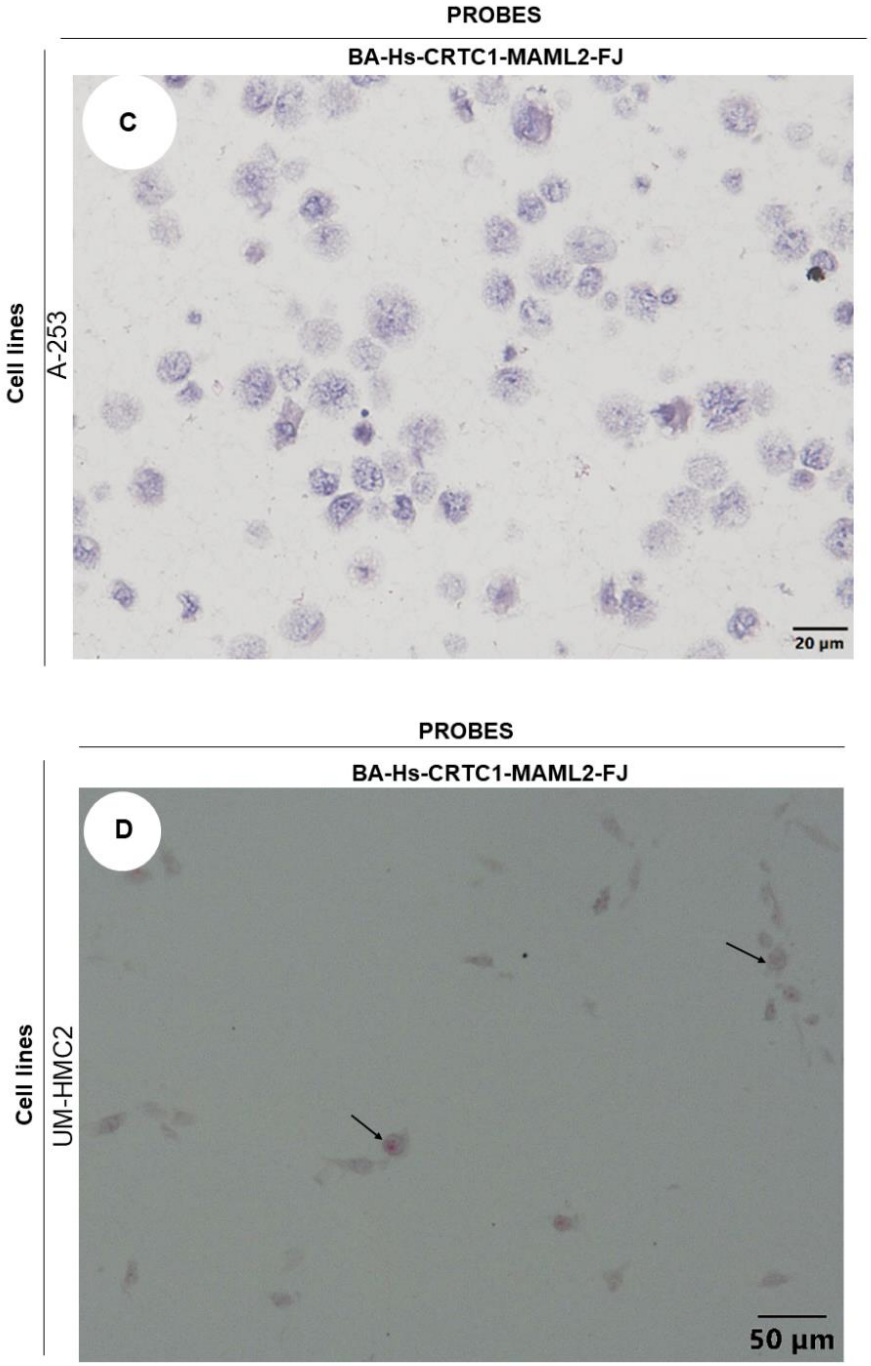
Figure 3.5 Validation of the BaseScope assay in tumour samples. Representative images of BaseScope assay performed on FFPE MEC samples using the BA-Hs-PPIB-1zz RNA quality control probes to examine the integrity of the mRNA among the clinical samples before conducting the assay using the target probe and DapB-1ZZ negative control probe. The integrity of the mRNA was preserved in the majority of the clinical samples, with the expression of the control probe showing distinct sharp red dots in tumour areas **(A, B and D)**, stroma cells **(B)** and normal lining epithelium **(C)**. Images **(E-H)** show no positive signals for the negative control probe (DapB).

3.4.3 Validation of CRTC1-MAML2 1ZZ probe in human cell lines

To validate the specificity of the BaseScope assay probe, well-characterised human MEC cell lines known to be fusion positive (NCI-H292 and UM-HMC2) and an unknown fusion status MEC cell line (A-253) were used. To validate the assay workflow, human HeLa cell pellets were used as a control with each run to confirm that the assays were working efficiently as was recommended and provided by the manufacturer. All assays included a probe targeting the fusion sequence (BA-Hs-CRTC1-MAML2-FJ), together with positive and negative RNA quality controls (bacterial mRNA DapB, positive control: housekeeping gene Hs-PPIB-1ZZ). As shown in figure 3.6 the BaseScope signals were readily distinguishable as punctate red dots within cells, either in the cytoplasm or in the nucleus. Across the cell lines tested, both of the known fusion positive MEC cell lines were positive but no positive signals were detected in the A-253 cell line.


PROBES
BA-Hs-CRTC1-MAML2-FJ
A
Cell lines
UM-HMC2
50 µm


PROBES
BA-Hs-CRTC1-MAML2-FJ
B
Cell lines
NCI-H292
50 µm


PROBES
BA-Hs-CRTC1-MAML2-FJ
Cell lines
A-253
C
20 µm
PROBES
BA-Hs-CRTC1-MAML2-FJ
Cell lines
UM-HMC2
D
50 µm

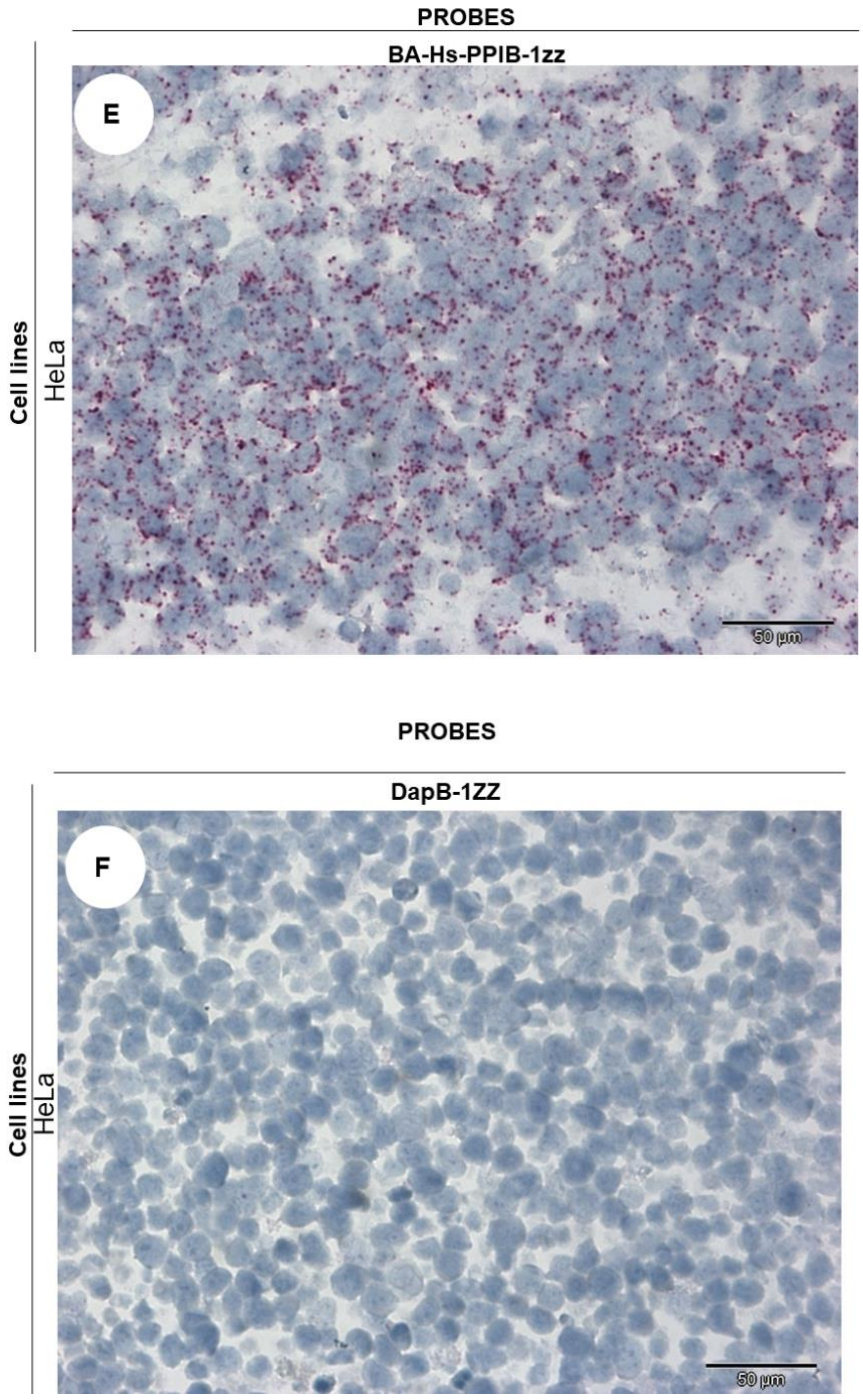


Figure 3.6 Validation of the BaseScope assay in the cell lines. The exon junction assay was performed on the MEC cell lines (UM-HMC2, NCIH292 and A-253) and non-MEC cells (HeLa) prepared as FFPE cell pellets using BA-Hs-CRTC1-MAML2-FJ probe for the fusion positive cell images **(A-D),** the positive control probe, BA-Hs-PPIB-1zz **(E)**, and a negative control probe, DapB-1ZZ **(F)** for HeLa cells.

3.4.4 Application of CRTC1-MAML2 1ZZ probe in human tumours

BaseScope probe technology was used on archived MEC tumour cases. Overall, 34% of the cases in this retrospective group were shown to harbour the CRTC1-MAML2 fusion (table 3.2), with 38% of the low-grade tumours being positive, 25% of the intermediate grade and 20% of the high-grade cases. When present, the CRTC1-MAML2 fusion was limited to the tumour cells but noted in all cell types: mucous cells, intermediate cells and epidermoid cells (figure 3.7). Notably, we detected the fusion transcript in clear cell variants of MEC (figure 3.8). No positive staining was seen in the normal tissue or surrounding stroma, which suggests that the probe is very specific to CRTC1-MAML2 fusion events in MEC. No signals were detected in any normal salivary glands (major and minor). Of the negative cases, no unique morphological features were noted. Similarly, the expression level of the CRTC1-MAML2 transcript was variable among the cases, with expression levels varying among the tumour cells within and between tumours.

Table 3.2 Detection of the CRTC1-MAML2 fusion transcript in MEC

Case number	BA-Hs-CRTC1-MAML2-FJ Probe	CRTC1-MAML2 Fusion positive cells
CCDH-01	Positive Score +2	Intermediate and mucous cells
CCDH-02	Positive Score 1	Intermediate cells
CCDH-03	Positive Score +2	Clear cells
CCDH-04	Negative	-
CCDH-05	Positive Score 1	Epidermoid cells
CCDH-06	Positive Score 1	Intermediate cells
CCDH-07	Positive Score 1	Intermediate cells
CCDH-08	Negative	-
CCDH-09	Negative	-
CCDH-10	Positive Score 1	Intermediate cells
SCDM-11	Negative	-
SCDM-12	Positive Score 1	Epidermoid cells
SCDM -13	Negative	-
SCDM -14	Negative	-
SCDM -15	Negative	-
SCDM -16	Negative	-
SCDM -17	Positive Score 1	Epidermoid cells
SCDM -18	Positive Score 3	Intermediate cells
SCDM -19	Negative	-
SCDM -20	Negative	-
SCDM -21	Negative	-
SCDM -22	Negative	-
SCDM -23	Negative	-
SCDM -24	Negative	-
SCDM -25	Negative	-
SCDM -26	Negative	-
SCDM -27	Negative	-
SCDM -28	Negative	-
SCDM -28	Negative	-
SCDM -29	Negative	-

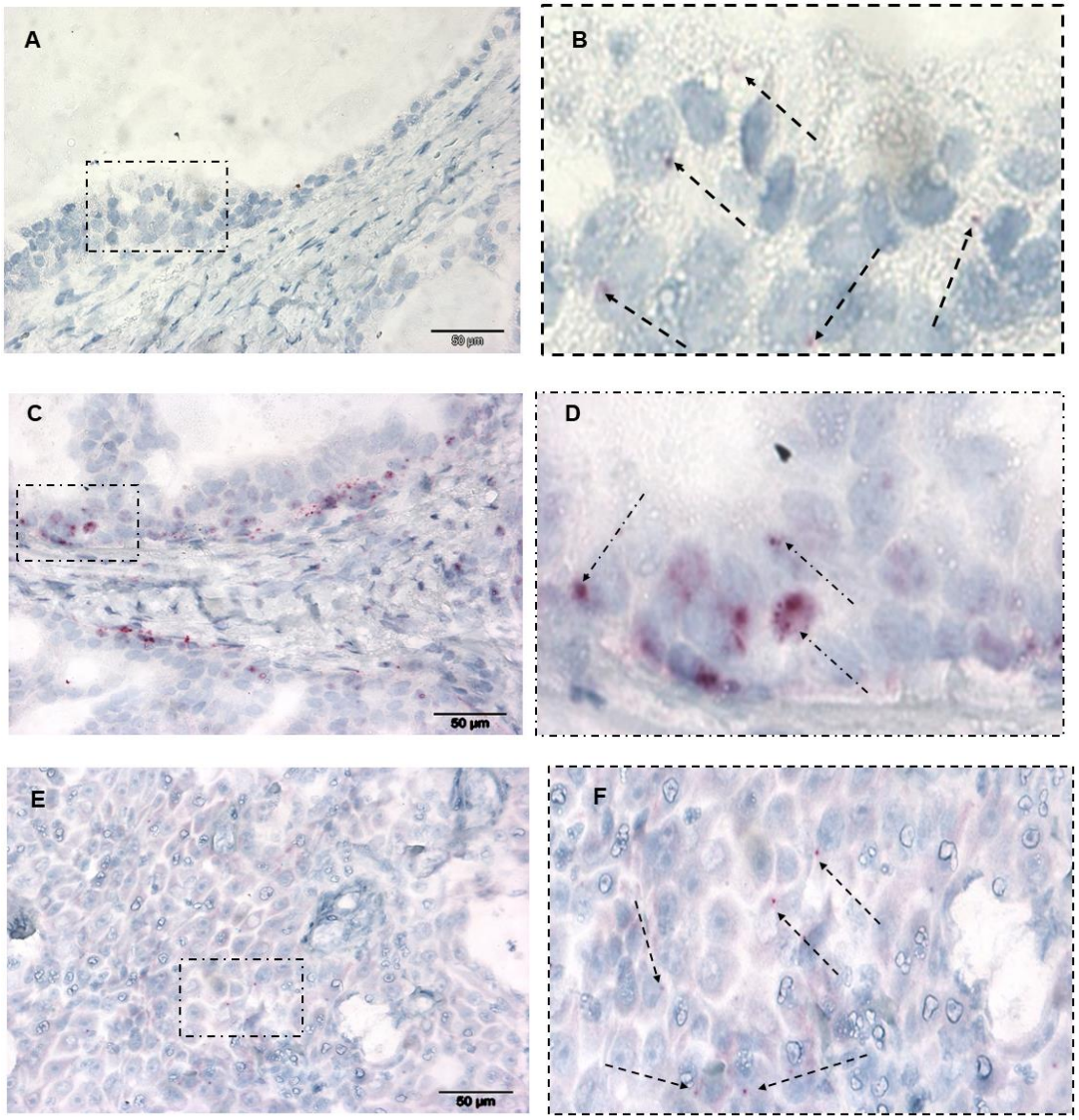


Figure 3.7 Detection of the CRTC1-MAML2 fusion transcript in MEC. Representative microscopic images of the BaseScope exon junction detection assay, performed on FFPE MEC using the BA-Hs-CRTC1-MAML2-FJ. Expression of the fusion transcript was detected in the three cell types: mucous secreting cells **(A, B)**, Intermediate cells **(C, D)**, and epidermoid cells **(E, F)**.

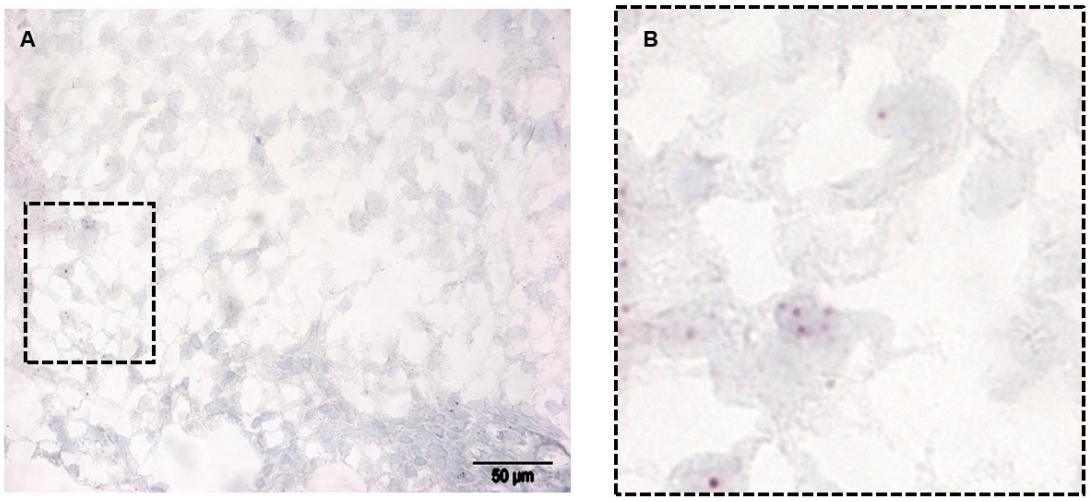


Figure 3.8 Detection of the CRTC1-MAML2 fusion transcript in MEC. Representative microscopic images of the BaseScope exon junction detection assay performed on FFPE MEC BA-Hs-CRTC1-MAML2-FJ probe. Expression of the fusion transcript was noted in the clear cell variant **(A)** and higher power **(B)**.

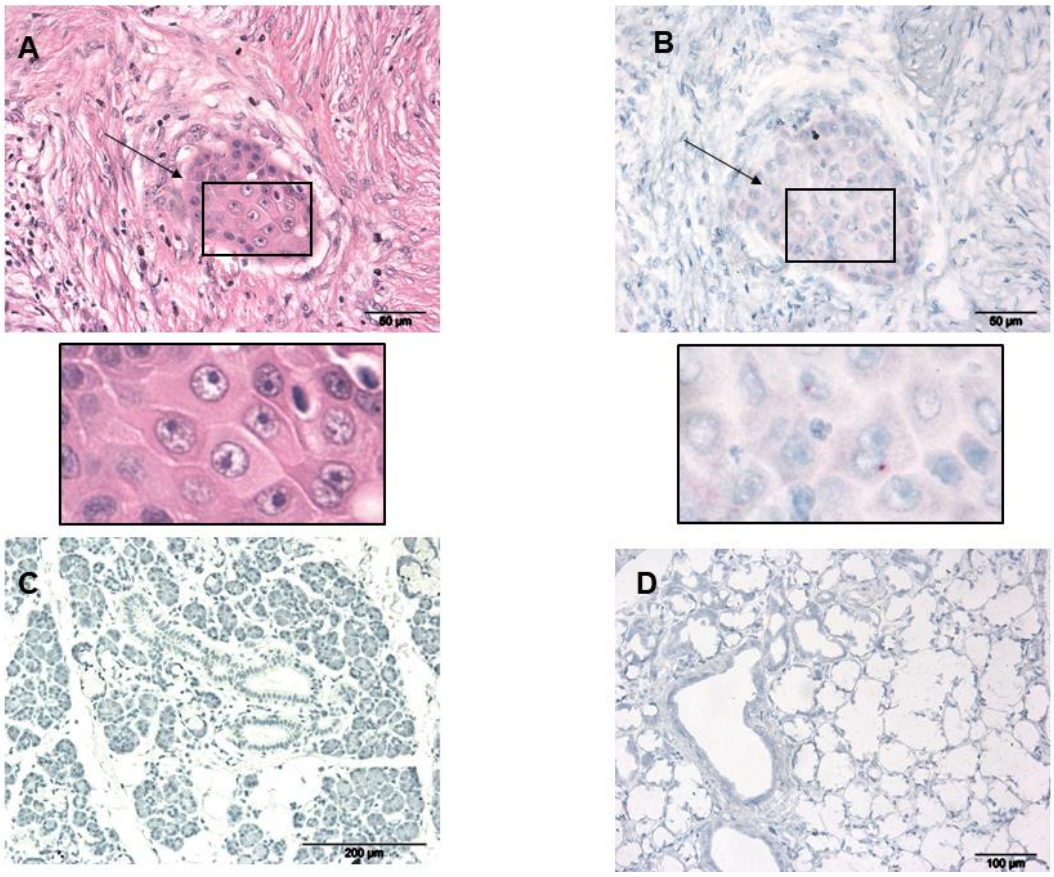
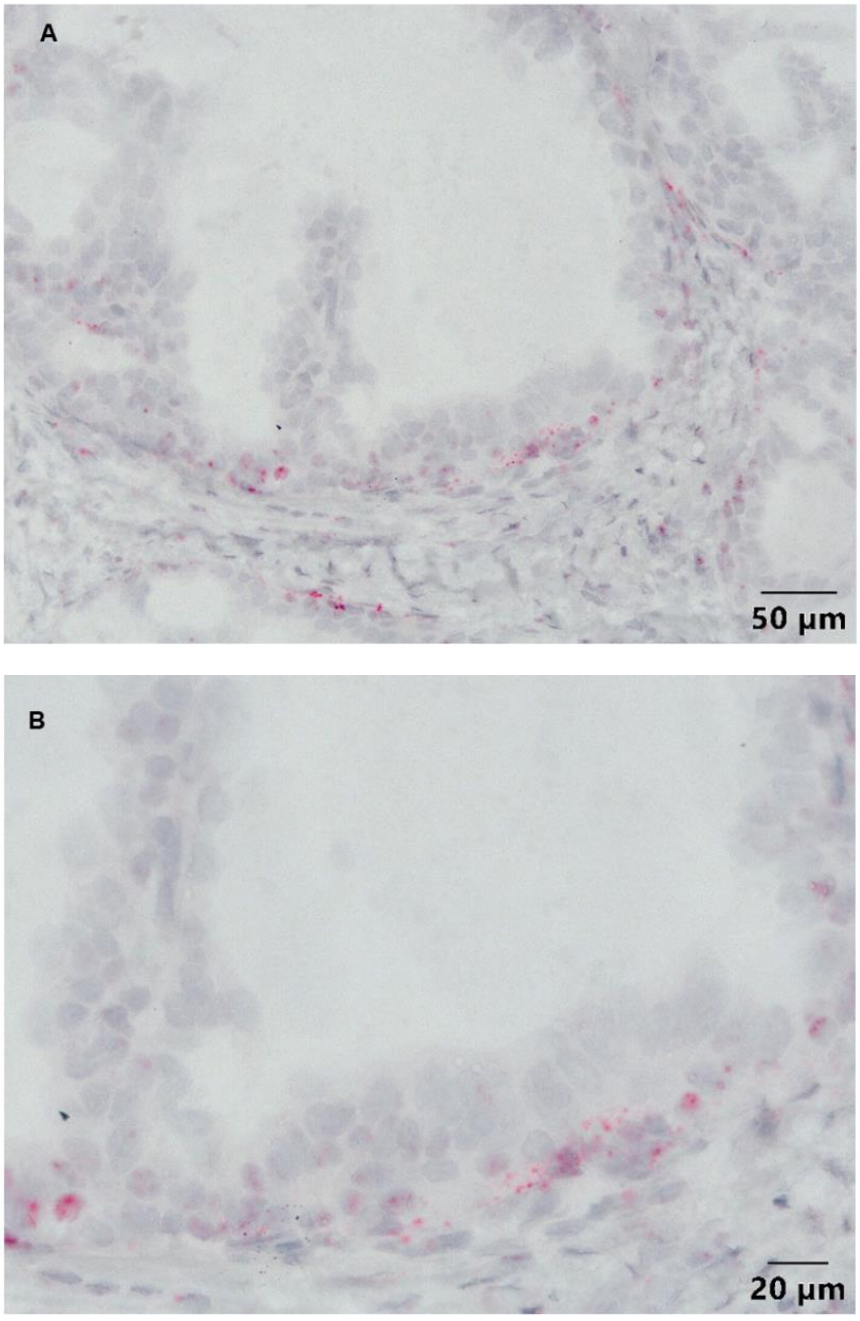


Figure 3.9 Specificity of BaseScope BA-Hs-CRTC1-MAML2-FJ probe to MEC areas. Representative microscopic images of MEC stained with standard haematoxylin and eosin showing an island of tumour cells lying in the stroma **(A)** and the same case with the BaseScope assay indicating positive signals were limited to tumour cells, being negative in the surrounding stroma **(B)**. Representative microscopic images of BaseScope assay using BA-Hs-CRTC1-MAML2-FJ probe for normal major and minor salivary glands showing no signals **(C & D)** respectively.

3.4.5 Specific architectural expression of CRTC1-MAML2 transcript in MEC cases

Two MEC cases in our cohort presented an interesting pattern of CRTC1-MAML2 expression and require further explanation. In both cases the transcript signal was not equally expressed in all tumour cells. In the first case, an intermediate-grade MEC, intense signals for the CRTC1-MAML2 fusion gene were noted in the external layers of neoplastic islands in direct contact with the surrounding stroma (figure 3.10). In the second case, a high grade MEC, the expression of the CRTC1-MAML2 fusion transcript was intense in the tumour cells surrounding the nerve (figure 3.10-3.11). Notably, the custom probe showed no non-specific signals and robustly detected the specific architectural expression pattern among these two cases.


A
50 µm
B
20 µm

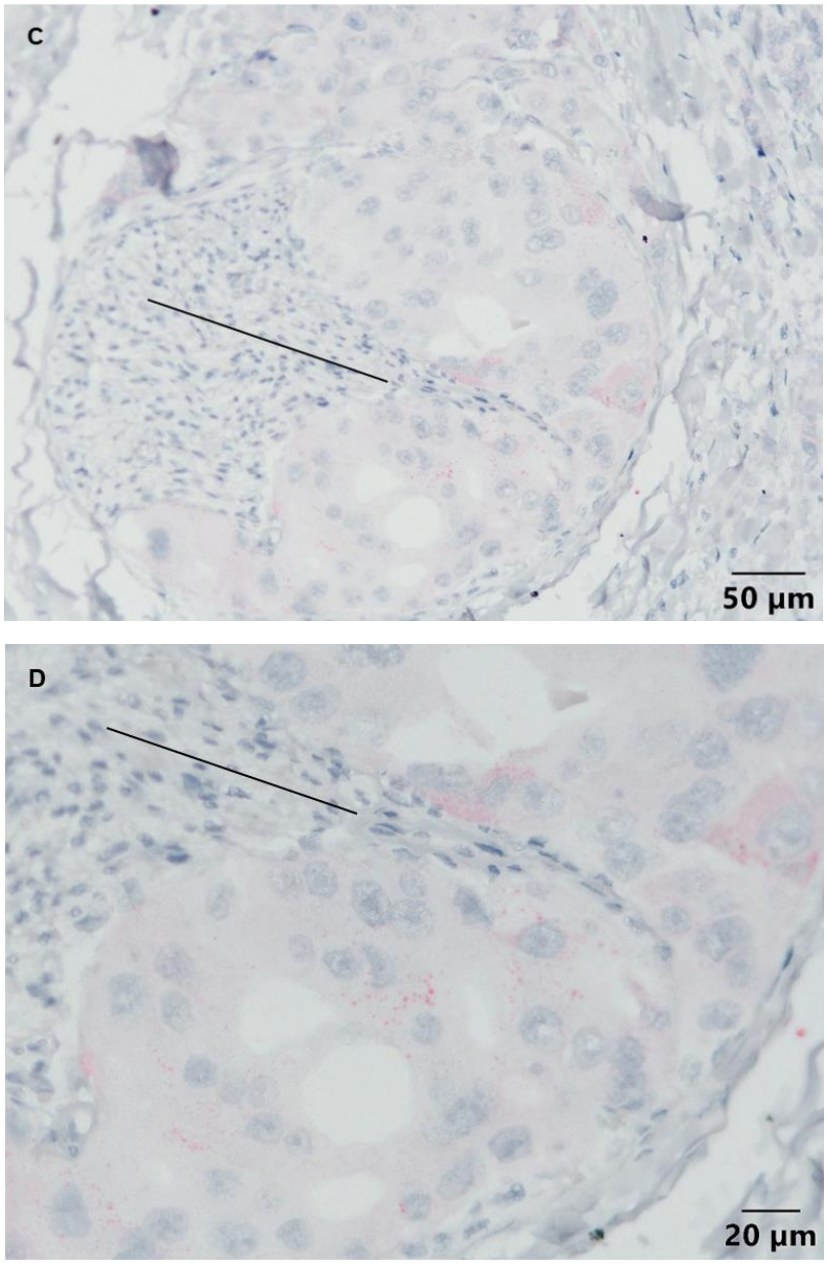


Figure 3.10 Specific architectural expression of the CRTC1-MAML2 transcript in MEC. Representative microscopic images of BaseScope assay using custom probe BA-Hs-CRTC1-MAML2-FJ targeting exon to exon junction showing specific architectural expression of CRTC1-MAML2 mRNA. A higher percentage of positive tumour cells was noted in the external layers of the neoplastic islands, in contact with the stroma **(A)** and at higher magnification power **(B)**. Also, showing high expression of CRTC1-MAML2 fusion transcript in tumour cells invading a nerve (black lines) in **(C)** and at higher magnification power **(D)**.

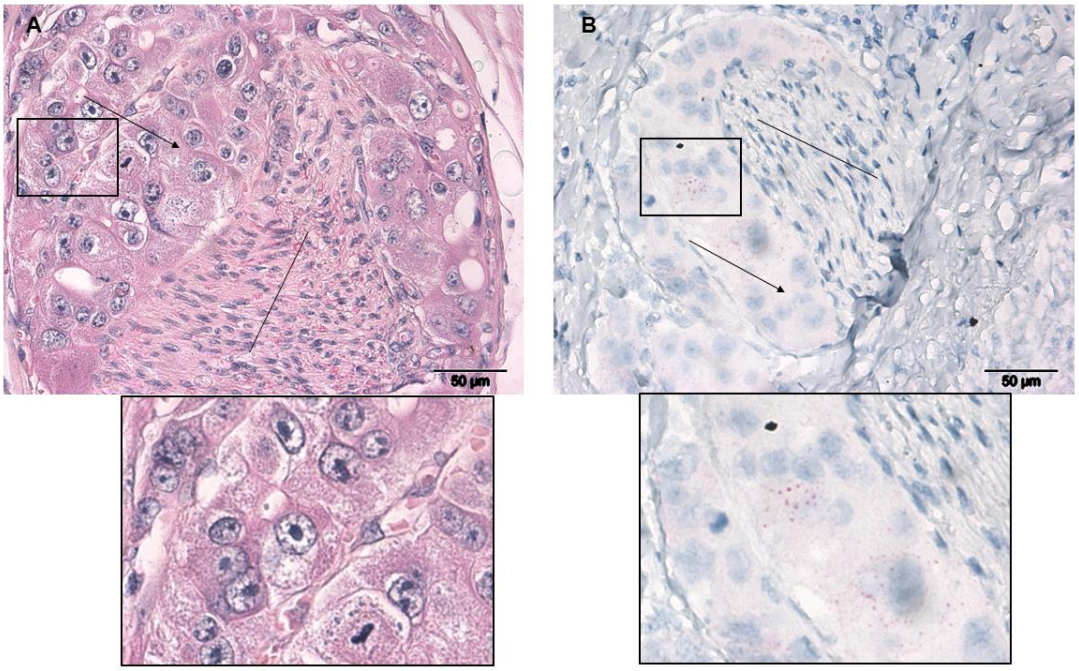
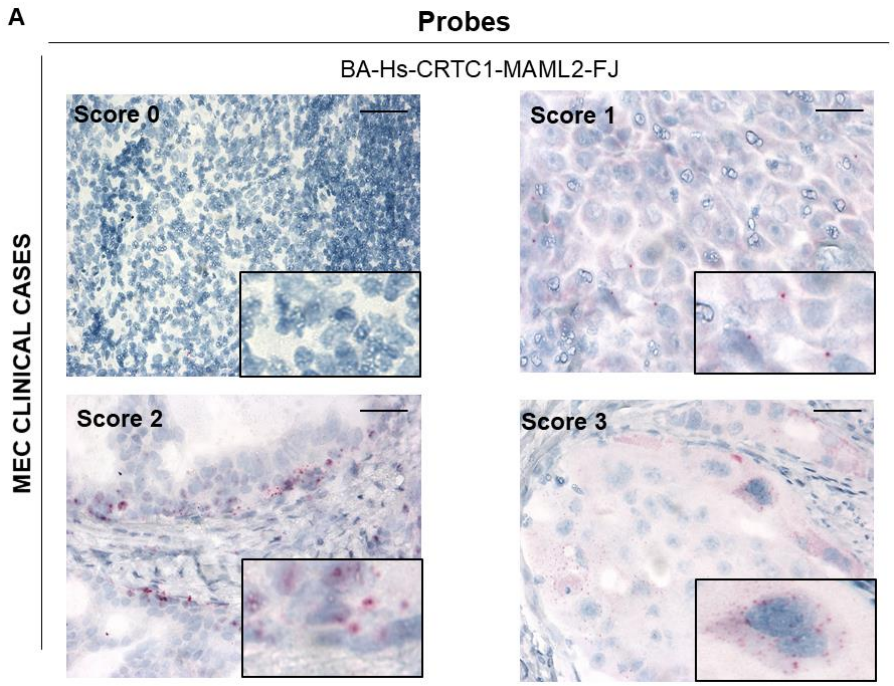


Figure 3.11 Specific architectural expression of the CRTC1-MAML2 transcript in MEC. Representative microscopic images of high grade MEC case stained with haematoxylin and eosin showing tumour cells (black arrow) surrounding the nerve tissue (black line) **(A)**. Representative microscopic images of BaseScope assay using BA-Hs-CRTC1-MAML2-FJ probe showing specific architectural expression of fusion expression in tumour cells (black arrow) invading the nerve (black line) **(B)**.

3.4.6 Quantification of CRTC1-MAML2 fusion expression in MEC cases

Having established the BaseScope assay to detect the exon-to-exon junction of the CRTC1-MAML2 fusion it was possible to quantify expression of the fusion transcript *in situ*. BaseScope technology is semi-quantitative, with each of the dot signals corresponding to a single mRNA transcript. Analysis was carried out as prescribed by the assay manufacturer based on a quantitative assessment of the signal. A scoring system from 0-4 was used by three pathologists to assess the transcriptional activity of the rearrangement genes in MEC cases (table 2.4). As demonstrated earlier the expression pattern was heterogenous among the cases and the majority of cases scored 1 (1 dot/cell). Three cases scored above 1, two scored 2 (2–3 dots/cell and none or very few dot clusters) and one scored 3 (4–10 dots/cell), surprisingly, this last case was one of the high-grade cases (figure 3.12).



B

Case number	Score
CCDH-01	+2
CCDH-02	+1
CCDH-03	+2
CCDH-05	+1
CCDH-06	+1
CCDH-07	+1
CCDH-10	+1
SCDM-12	+1
SCDM-17	+1
SCDM-18	+3

C

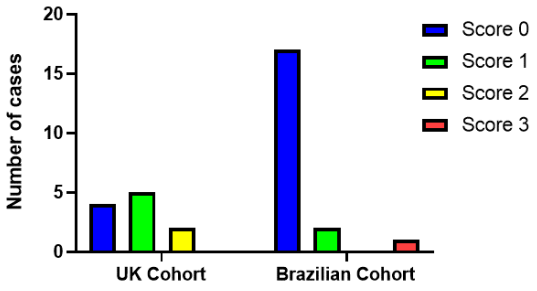


Figure 3.12 BaseScope assay semi-quantitative scoring system for MEC cases. (A) Representative microscopic images of mRNA *in situ* hybridisation using custom probe BA-Hs-CRTC1-MAML2-FJ targeting CRTC1-MAML2 fusion with higher magnification (inset). Scale bar at 50 µm. **(B)** Table (3.3) showing the quantification analysis of the CRTC1-MAML2 junction specific BaseScope assay of the UK and Brazilian cohorts. **(C)** Graph showing the number of cases in two MEC tumour tissue cohorts utilised in this study (UK cohort comprises 10 cases and Brazilian cohort consists of 19 cases). 70% (7/10) of the UK cohort harboured the CRTC1-MAML2 fusion, with five of the positive cases scoring 1 and two scoring 2. 16% of the cases from the Brazilian cohort were fusion positive and of these two cases scored 1 and one case scored 3.

3.5 Discussion

Improvements in the molecular technology used for diagnosis allow a better understanding of pathological conditions and are assisting in the classification of patients and in the development of targeted therapies. BaseScope technology enables highly specific and sensitive detection and visualisation of RNA targets with sensitivity of one nucleotide difference, *in situ*, with spatial mapping and a morphological context, under a brightfield light microscope. This study has described, for the first time, the application of the 1ZZ BaseScope *in situ* hybridisation technology to screen and quantify CRTC1-MAML2 fusion gene expression in FFPE human samples using a custom-made probe. Our analysis demonstrated that the high specificity of the BaseScope probe allowed precise detection of the CRTC1-MAML2 fusion gene, which was not limited to a specific grade or cell type as it was detected in all grades and all cell types, including the clear cell variant.

Prior to this study little was known about the expression pattern and architectural distribution of the CRTC1-MAML2 rearrangements *in situ.* We have also reported, for the first time, that the fusion transcript was detected in a substantial number of cells infiltrating the nerve and also in cells invading the surrounding stroma. Notably, the fusion event was not found in all of the cells constituting the tumour regardless of the histological type. Another interesting finding is that we noted a likely natural (biological) variation in gene expression which suggests that not all cells necessarily express the fusion transcript at detectable levels, despite carrying the translocated gene; genetic elements could be controlling such an effect.

The utility of using CRTC1-MAML2 fusion expression in the diagnosis of MEC, especially in difficult cases, has not been a strong focus in the literature. The detection of CRTC1-MAML2 translocation in MEC cases could assist with more accurate

diagnosis, particularly in certain clear cell cases or oncocytic variants. As an example, the study by Garcia *et al*. (2011) successfully distinguished more than 70% of MEC oncocytic variants from other oncocytic salivary gland lesions, which histologically overlapped, by detection of the CRTC1-MAML2 fusion. Also, in cases of low-grade MEC this assay could achieve more accurate differentiation from other cystic conditions and the demonstration of CRTC1-MAML2 translocation could help to exclude metastatic squamous cell carcinoma or salivary duct carcinoma (García *et al*., 2011). Thus, detecting the CRTC1-MAML2 fusion could significantly support the diagnosis of MEC. It could be argued that not all MEC cases carry the CRTC1-MAML2 translocation; nevertheless, with the positive cases it is important to segregate cases based on their molecular genetics to further understand the pathogenesis, their clinical implications, and provide targeted therapy to fusion positive MEC with advanced disease stage.

Gene rearrangements can act as oncogenic drivers across many cancer types (Stenman, 2013), hence, they are often considered ideal targets for drug development. Key studies have suggested that the CRTC1-MAML2 fusion could be implicated in the initiation and the progression of the disease (Amelio *et al*., 2014, Coxon *et al*., 2005, Komiya *et al*., 2006, O'Neill, 2009, Tonon *et al*., 2003, Wu *et al*., 2005) which are discussed in more detail in chapters 1 and 4. With the absence of targeted therapy and limited treatment options available for MEC patients, developing a targeted therapy interfering with signalling pathways associated with the CRTC1-MAML2 fusion could help with the management of fusion positive cases, potentially leading to successful treatment of the disease. Therefore, developing an accurate technique that would assist with the identification of these cases should be considered crucial.

Prior studies detected the CRTC1-MAML2 translocation in 34% (34/101), 38% (27/71), 59% (53/90), 61% (14/23), 81% (18/22) and 55% (16/29) of cases using RT-PCR and FISH (Anzick *et al*., 2010, Behboudi *et al*., 2006, Birkeland *et al*., 2017, Miyabe *et al*., 2009, Okabe *et al*., 2006, Tirado *et al*., 2007). In the present study, using the BaseScope assay, 34% (10/29) of cases were CRTC1-MAML2 fusion positive. BaseScope measures the mRNA in FFPE, and thus the instability and degradation of RNA in the tumour samples could result in lower levels of detection. The integrity and validity of RNA in our cohort was examined prior to fusion gene expression analysis suggesting that the BaseScope technique is comparable to other assessment techniques used in previously published studies. The use of frozen sections to perform RT-PCR in some previous studies resulted in a completely preserved quality of the RNA it is, however, often not possible to obtain frozen tumour samples.

The BaseScope assay is a new technique, established early in 2016 by Advanced Cell Diagnostics, and is based on the RNAScope assay which was launched in 2011. The BaseScope assay is an ultra-sensitive technique, uses a double ZZ pair probe which is only 30-50 base pairs in comparison to the ≥ 300 nucleotide probe used with RNAScope. BaseScope also uses more amplification steps, which allows for detection of exon-exon junctions, short target/splice variants, highly homologous sequences and point mutations, with single cell detection sensitivity in a broad range of tissues, samples and species. Several reports have shown the successful application of the BaseScope technology. For example, it was used to detect specific, novel RNA mutations in colorectal cancer samples, and it allowed the precise spatial mapping and morphological context to be determined (Baker *et al*., 2017). BaseScope has also been shown to visualise and quantify the novel junction of the androgen

receptor splice variant 7 (AR-V7) in morphologically intact cells of metastatic castration-resistance prostate cancer (Zhu *et al*., 2018).

The BaseScope assay technique involves a range of steps before the application of the probe, to optimise and enhance the detection, which include treatment with H_2O_2 to quench any endogenous peroxidase activity that might interfere with signal detection in the final step. Additionally, it involves antigen retrieval steps, which were optimised in this study to retrieve our targets in MEC tissue samples as well as cell pellets. Antigen retrieval required extra time for tissue samples in comparison to cell pellets as the tissue sections are thicker and had been fixed for longer than the cell pellets. Furthermore, treatment of the sections with protease was necessary to release the RNA from the samples by digesting the protein crosslinks which developed during the fixation process. The heating steps carried out in the hybridization oven help to irreversibly digest the crosslinks.

RNA-based detection tools for fusion genes are considered superior as DNA-based methods could be hampered by large intronic regions (Davies *et al*., 2018). The use of FISH to identify the translocation is limited as it only performs break-apart detection and does not detect the fusion sequence. For example, CRTC1-MAML2 detection by FISH involves only a probe for the MAML2 gene; it is not designed to detect the fusion partner. In contrast, our BaseScope probe (BA-Hs-CRTC1-MAML2-FJ) was specifically designed to cross the fusion transcript between exon 1 of the CRTC1 and exon 2 of MAML2 fusion partner and thus we are confident that this provides a more reliable and trusted detection technique.

Another benefit of RNA-based detection method over a DNA-based method is that it not only provides information regarding fusion status, but it can also allow determination of expression levels. The BaseScope assay allowed, for the first time, the accurate and precise identification of which cell type harbours the translocation (figures 3.7 and 3.8). Previous studies have suggested that the three major cell types found in MEC, epidermoid, mucous and intermediate, stained positively for the fusion (Behboudi *et al*., 2006); however, in the absence of convincing data to support this claim, it was hard to draw such a conclusion. In this study, it was clear exactly which cell types carried the fusion transcript, and our data also included a MEC with clear cell variant. The intermediate and epidermoid cell types were predominately harbouring the translocation among the positive cases. One case showed translocation in both mucous cells and intermediate cells. Therefore, our data is not only in line with previous results, but it is also further supported by accurate visualisation and detection of the translocation at a single-cell level in the intact histological samples.

The advantages of using *in situ* RNA based detection methods over protein detection techniques such as IHC are that IHC requires a specific and sensitive antibody to detect the antigen of interest, availability of which is not always possible. Moreover, developing an antibody is time consuming, expensive, laborious, requires validation and could vary from batch to batch. A prior study conducted by Behboudi and co-workers in (2006) involved the development of a polyclonal antibody for the CRTC1-MAML2 fusion for IHC, however, this antibody has yet to be made commercially available or validated for use in research and clinical laboratories. Added to these critical issues, detection of secretory proteins could be problematic and hard to localise in tissue, whereas mRNA is localised intracellularly, rendering it easy to identify. Furthermore, proteins and RNA levels do not always correlate (Maier *et al*.,

2009), therefore, measuring the mRNA level may indicate the biological status of the transcriptome, which may vary from its protein counterpart. Lastly, an RNA based technique could be the only option for detecting non-coding RNAs.

Detection of the CRTC1-MAML2 transcript in all cell-types of MEC, including the cell variants in our study (clear cell variant; figure 3.8) and in a previously published study (Oncocytic cell variant; (García *et al*., 2011)), suggests that there is not an original single cell type that harbours the translocation and drives tumourigenesis in MEC. More importantly, the fusion event seems to be expressed across all cell types, which make it hard to target one cell type to interfere with tumour growth and progression in *in vitro* studies and preclinical trials.

Several studies have indicated that the fusion protein is associated with low and intermediate grades of MEC but less frequently linked to high-grade MEC patients (Seethala *et al*., 2010). Interestingly, in my study the fusion event was detected in a high-grade case with neural invasion which clearly supports previous observations (Anzick *et al.*, 2010, Seethala *et al*., 2010) identifying high grade fusion positive MEC lethal cases. Another study has confirmed that the CRTC1-MAML2 fusion may not be correlated with disease severity and prognosis, as no statistically significant differences were noted between survival rate and fusion status (Birkeland *et al*., 2017). In our series, the presence of fusion positive cells which invaded the nerve in the high-grade case, suggested the translocation could play a role in tumour progression, cell invasion and metastasis. However, this was only detected in one case and further work is needed to support this claim.

A recent systematic review (Pérez-de-Oliveira *et al*., 2020) conducted to investigate the association between the CRTC1-MAML2 translocation and patient disease-free survival (DFS), disease-specific survival (DSS) and overall survival (OS) among 10 retrospective studies, concluded that the presence of the CRTC1-MAML2 fusion may act as useful prognostic tool in MEC cases. Further studies are required, however, to support this conclusion by performing prospective clinical studies, with appropriate statistical tests, including more intermediate and high-grade cases and allowing a minimum five-year follow up of the MEC cases.

This study demonstrated that the expression of the CRTC1-MAML2 fusion transcript was low in most of the cases, and that not all neoplastic cells harboured the translocation. Of the positive cases only three cases showed higher expression patterns, but even these showed that most of the cells (regardless of cell type) were negative. Two possible explanations for this low detection are that the percentage of the translocated cells may vary from case to case, as noted in cases with direct contact with nerve or at the external layer next to the stroma. Thus, even within fusion positive cases the expression pattern varies. The other possible explanation is that the BaseScope assay is only detecting the expression level of the CRTC1-MAML2 mRNA when it is active, mature and transcribed in the cells and shows the biological status of the mRNA during at the point of fixation (Sharova *et al*., 2009, Vogel and Marcotte, 2012). Nonetheless, the BaseScope assay is superior to other RNA assays and can detect the expression of low abundance transcripts in tissue.

The tumour microenvironment is composed of heterogeneous cell populations; cancer stem cells, the most highly tumorigenic cells within the tumour bulk, representing a very small percentage (Bao *et al*., 2013) yet they are recognised as more often being associated with drug resistance and tumour relapse/metastasis.

Thus, determining that only a small percentage of cells are fusion positive does not mean this is not an important event, it only shows it is very specific and that the consequences are still poorly understood. Also, it remains to be evaluated/understood how the positive cells influence negative cells within the tumour microenvironment. It might be that the presence of even a small percentage of fusion positive cells might have influence on the entire environment by controlling the initiation and progression of the tumour via specific signalling as discussed in chapters 1 and 4.

Our finding that all the major cell types (epidermoid, mucous and intermediate) can harbour the translocation is an encouraging finding, supporting the potential for a positive effect of any drugs that target this molecular event. A specific cell type does not need to be targeted, rather disruption of fusion positive cells, and perhaps through this we can create an imbalance/instability of the tumour microenvironment. Many studies have shown that targeting one specific molecular event might lead to successful target/personalized therapy. For example, drugs that usually “kill” more cancer stem cells are often not as efficient in the tumour bulk and vice-versa. Hence combined therapy is often suggested based on this rationale. MEC treatment by targeting the CRTC1-MAML2 fusion could be enhanced through use in association with drugs inhibiting other molecular events.

A recent study by Chen and colleagues (2018) using RNA *in situ* hybridization analysis demonstrated that there was predominant nuclear localization of the LINC00473 in two CRTC1-MAML2 fusion positive MEC cell lines (NCI-H292 and NCI-H3118), human primary fusion positive MEC and xenograft tumours but not in fusion negative primary MEC tumours and non-MEC fusion negative cell line (HPA-1); the

expression pattern or cell type were not confirmed. LINC00473 is one of the long non-coding RNAs (lncRNA) found to be downregulated when the CRTC1-MAML2 fusion expression was depleted in MEC cell lines. Two lentiviral based shRNAs, shM2-1 and shM2-3, that target the respective 3′-untranslated region (UTR) and the TAD region of CRTC1-MAML2 fusion and MAML2 were used to deplete fusion/MAML2 expression in fusion-positive MEC cells and MAML2 in fusion-negative cells, which results in reduction *in vitro* and *in vivo* in cell proliferation and survival (Chen *et al*., 2018).

Our results suggest that the BaseScope assay provides a robust method for CRTC1-MAML2 detection, which could have a potential application in clinics. Less optimisation is needed for the BaseScope than with the other methods and the amplification steps are universal. Furthermore, the result of the BaseScope assay is a chromogenic stain instead of a fluorescence dye, and so does not require specialist microscopes for analysis, can be stored for long periods of time and can be viewed on multiple occasions without any subsequent fading problems. BaseScope could, therefore, provide a much-needed aid for diagnosis and could assist in the development of targeted therapy when identifying which cells harbour the translocation.

Limitations of my study include the small sample size and the fact that most of the cohort were low grade tumours. The sample size was limited due to difficulty in obtaining an adequate number of MEC treated cases and the poor representation of intermediate and high grade tumours was largely due to the nature of the disease as previous studies have suggested that more than 50% of cases are low grade (Janjua *et al*., 2012). The BaseScope assay also does have some limitations in contrast to

other techniques, for example, the CRTC1-MAML2 mRNA must be preserved at a detectable level in FFPE samples. Fixation time and method may result in poor quality RNA, particularly in older specimens which were not processed adequately. Screening of the quality of the mRNA is, therefore, critical prior to application of the target probe. Furthermore, we did not generate CRTC1-MAML2 translocation genomic sequencing data of our MEC cohort to match our transcript data and to confirm the specificity and sensitivity of our technique. Such an analysis is required to exclude the presence of false- negative results.

3.5.1 Conclusion

The Base Scope assay accurately detects the CRTC1-MAML2 fusion transcripts and could provide an alternative diagnostic technique which can be easily applied in clinical labs if the sensitivity and specificity of this technique are validated. The CRTC1-MAML2 fusion transcript was detected in all the grades; however, only a low number of cells in a tumour harboured the translocation. The CRTC1-MAML2 transcript was also detected in the tumour cells with direct contact with stroma and nerve invasion.

3.5.2 Future work

3.5.2.1 Real-time PCR validation

In order to evaluate the efficacy of the BaseScope technology for detection of the CRTC1-MAML2 transcripts in tissue, real-time PCR transcript validation should be performed. Total RNA was extracted from FFPE MEC specimens and four 7µm sections were cut from the tissue blocks; RNA was extracted using an miRNeasy FFPE Kit (QAIAGEN®) according to the manufacturer's guidelines and the quality and yield of RNA were determined by Nanodrop and paraformaldehyde gel

electrophoresis. Specific primers to reverse transcribe the RNA were designed in house and purchased from Sigma, UK. Sadly, time restraints meant that RT-PCR analysis could not be carried out for specific detection of the fusion segment of cDNA created, but in future the results from such an assay should be combined with the results acquired from the BaseScope assay, and the sensitivity of the BaseScope assay assessed and recorded.

3.5.2.2 Molecular understanding of the CRTC1-MAML2 fusion

To develop a full picture of CRTC1-MAML2 expression in MEC, additional molecular *in vitro* studies will be needed to understand the role of the fusion protein in the tumorigenesis of MEC, such as performing gene expression studies and studying the cell behaviour upon expression of the fusion protein.

Chapter 4
Synthesis and Analysis of the CRTC1-MAML2 Fusion Construct

4. Chapter 4

4.1 Introduction

Previous studies have utilised human CRTC1-MAML2 constructs in *in vitro* and *in vivo* analyses in order to further understand the role of the fusion protein in the development and progression of salivary gland MEC (Amelio *et al*., 2014, Behboudi *et al*., 2006, Coxon *et al*., 2005, Komiya *et al*., 2006, Tonon *et al*., 2003, Wu *et al*., 2005). To date, however, the true function of the CRTC1-MAML2 fusion has yet to be fully elucidated.

Importantly, in this chapter I will demonstrate optimisation of *de novo* gene and protein expression in mammalian cell lines, but also in normal primary human salivary gland cells isolated from tissue explants. Custom-made TaqMan probes, specifically designed to target the fusion region, along with FLAG and MAML2 antibodies, were used to compare endogenous MAML2 expression with that related to the fusion gene. I also validated key study results, previously carried out using HeLa cells and MEC cells (Chen *et al*., 2015, Wu *et al*., 2005), with normal primary human salivary gland cells.

4.2 Aim and Objectives

- Generate a human CRTC1-MAML2 fusion construct in a mammalian expression vector, transfect into normal primary salivary gland cells and also relevant cell lines, detect gene and protein expression and colocalise the protein expression in the cells.

- Study the downstream effects following the introduction of the CRTC-MAML2 fusion gene into normal primary human salivary gland cells.

4.3 Materials and Methods

The relevant materials and methods utilised in this chapter are detailed in chapter 2, under headings 2.2, 2.3, 2.4, 2.5, 2.6 as indicated below:

- Synthesis of CRTC1-MAML2 fusion construct was performed as described in section 2.3.1.
- Primer design and polymerase chain reaction were performed as described in section 2.3.2.
- PCR2.1 TA TOPO cloning was performed as described in section 2.3.3.
- Plasmid transformation and selection of positive colonies were performed as described in section 2.3.4.
- Isolation and preparation of plasmid DNA from *E. coli* were performed as described in section 2.3.5.
- TOPO cloning and subcloning into mammalian expression vectors including sequences and plasmid midi preps were carried out as described in sections 2.3.6-2.3.13.
- The human primary salivary gland cells and cell lines were cultured as outlined in sections 2.2.1-2.2.3.
- Transfection experiments and method optimisations were performed as described in sections 2.4.1-2.4.6.
- Gene expression analysis was carried out as described in sections 2.5.1-2.5.6.
- Protein analysis experiments were performed as described in sections 2.6.1-2.6.6.

4.4 Results

4.4.1 Construction of CRTC1-MAML2 fusion construct

Human CRTC1-MAML2 fusion constructs were synthesised using conventional PCR cloning techniques. In order to produce the construct, CRTC1-MAML2 cDNA was generated using RNA extracted from fusion positive cell lines (NCI-H292 and UM-HMC-2). A range of gene specific oligonucleotides were used to amplify the cDNA. The PCR results were visualised on a 1% agarose gel which indicated positive results (figure 4.1). The cDNA amplified with the primers F6 R5 was chosen for TA TOPO-cloning as no primer dimers were created during the amplification process; this fragment contained exon 1 of CRTC1 and exon 2 of MAML2.

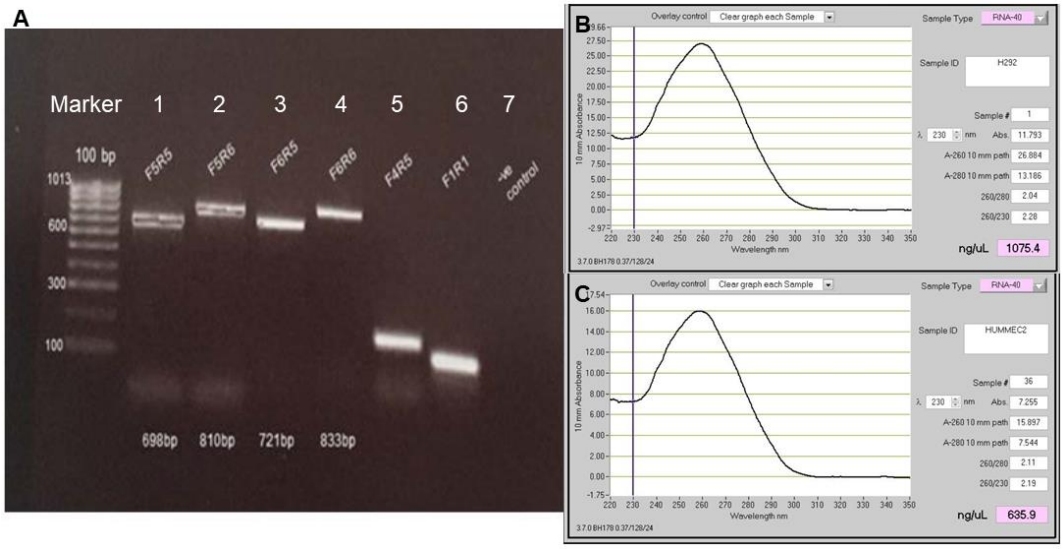


Figure 4.1 Amplification of the CRTC1-MAML2 cDNA A) Resolution of PCR products for cloning of CRTC1-MAML2 construct. RNA from a fusion positive cell line (NCI-H292) was reverse-transcribed into cDNA. cDNA was amplified using a range of specific primers. The PCR products were run on 1% agarose gel, with a 100 bp molecular DNA marker and ethidium-bromide. Lane 1-5 indicates the CRTC1-MAML2 amplicon bands corresponded to the expected size of the different specific primers, housekeeping gene (OAZ1) in lane 6, and the negative control in lane 7. **B and C)** Quality and concentration of NCI-H292 and UM-HMC2 RNA extracted: Nanodrop Spectrophotometer measurement of the RNA indicates good quality and quantity of the RNAs.

Primers were designed to span the 5' coding region of the fusion gene. The primers were also designed to incorporate a FLAG-Tag sequence at the N terminus of the gene just after the start codon and enzyme restriction sites upstream and downstream of the coding DNA. PCR was used to amplify target cDNA encoding the fusion region. The resultant PCR products were confirmed to be of the expected size by DNA gel electrophoresis (figure 4.2-A). The PCR products were cloned into pCR2.1 vector (figure 4.2-B), verified initially through restriction digests indicating that the insert fragment had been cloned in the correct orientation, and then confirmed by T7 forward sequence and blast analysis. The sequence results indicated a 100% identity to the subject.

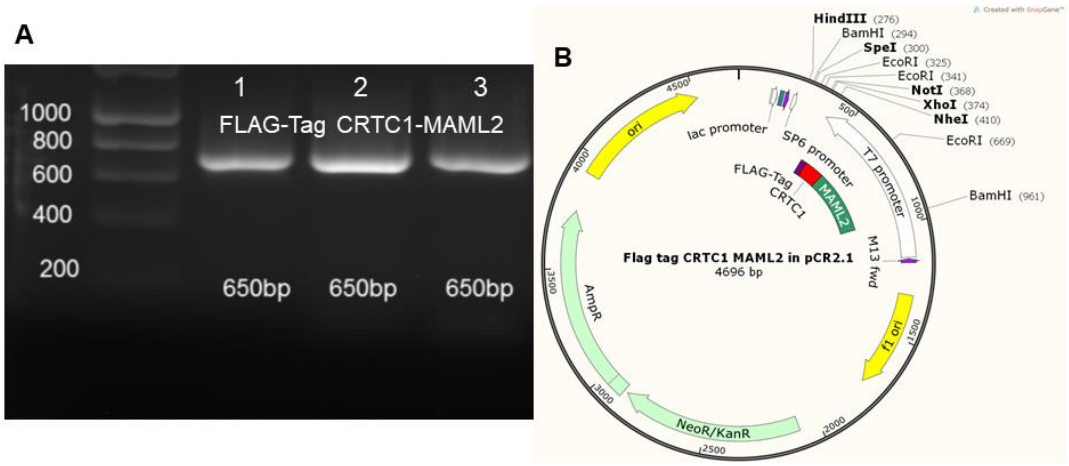


Figure 4.2 Generation of FLAG-Tag CRTC1-MAML2 Clone. A) Resolution of PCR products amplified the 5' coding region of the fusion gene using specific primers with incorporated FLAG-Tag sequences at the start of the coding region. The PCR products were separated on 1% agarose gel and visualised with ethidium bromide staining and products of the expected size of 650 bp were detected in lane 1 to 3. **B)** Representative map of the clone indicating the FLAG-Tag CRTC1-MAML2 fragment (violet, red and green colours) cloned into pCR2.1 for further mammalian expression sub cloning. Image created with SnapGene software.

In order to sub-clone the FLAG-Tag CRTC1-MAML2 fusion into a mammalian entry vector, the fragment was double digested with *NheI* and *BamHI*. The digestion was

separated on a 1% agarose gel, the fragment of FLAG-Tag CRTC1-MAML2, at the expected size of 650 bp, was excised and the DNA extracted and purified.

MAML2 is 1156 amino acids long and the fusion gene uses the sequence of all but the first exon (171 amino acids). In order to obtain full length MAML2 cDNA without any mutations, a MAML2 pCR4 clone was purchased, regrown and the plasmids extracted and purified. Diagnostic specific endonuclease digestion indicated bands of the expected size (figure 4.3). The clone was confirmed by two directional sequencing with results of blast analysis demonstrating 100% identity of the clone to the subject.

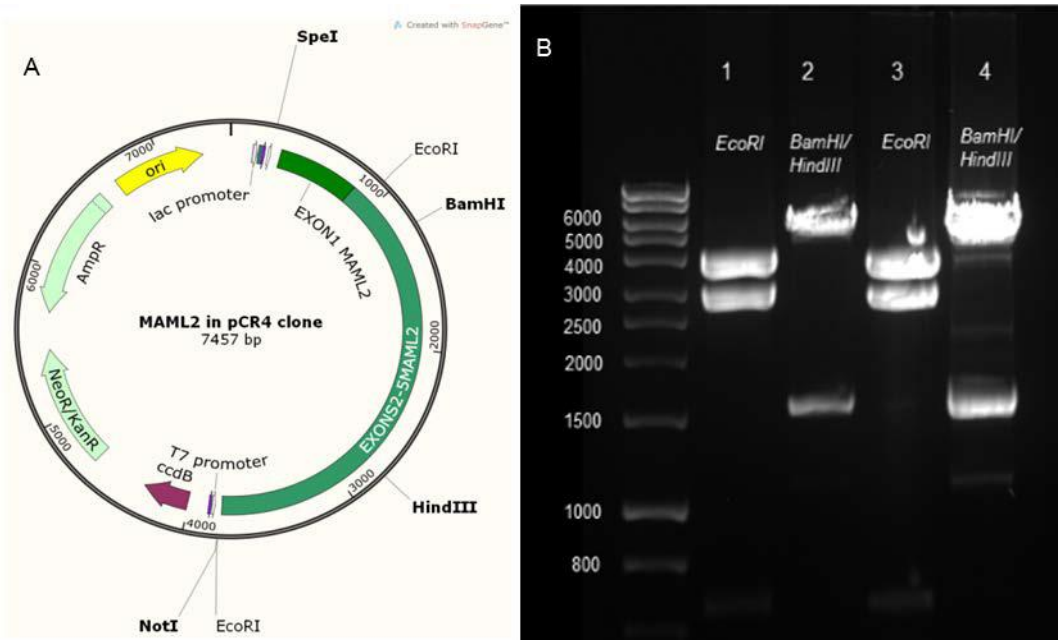
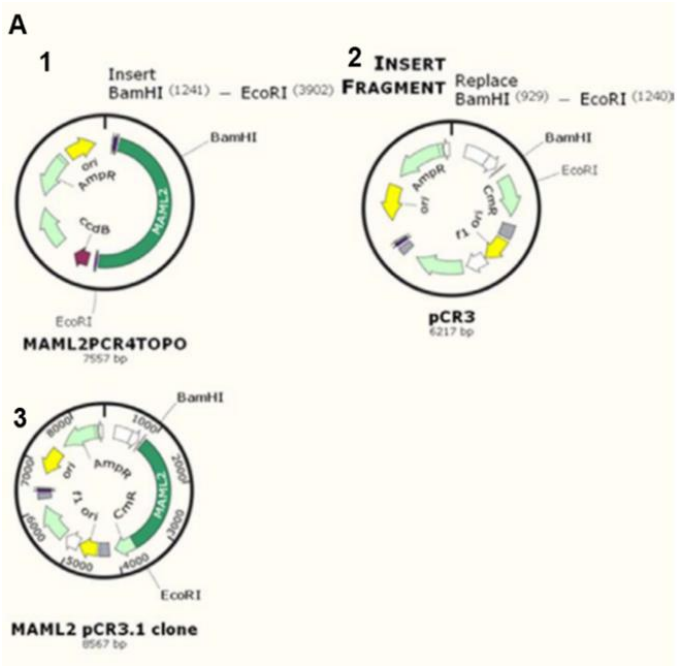


Figure 4. 3 MAML2 pCR4 clone: A) Representative map of the clone demonstrating the full expression sequence of MAML2 fragment (green colour) cloned into pCR4 vector. The map also shows the main restriction enzyme sites; image created with SnapGene software. **B)** Diagnostic digestion of MAML2 pCR4 clone: MAML2 pCR4 clone was regrown and the positive colonies were selected and cultivated. The plasmid was successfully extracted and verified with specific endonuclease digest. Both duplicates indicated the expected band size. Lanes 1 and 3 indicate *EcoRI* digestion of the clone resulting in two close bands: at 3,900 bp pCR4 vector and at 2,900 bp, MAML2 *EcoRI* digest fragment. Similarly, lanes 2 and 4

represent *BamHI* and *HindIII* double cutting generating two separate bands: vector at 5,900 bp and MAML2 fragment at 1,500 bp.

In the translocation event the first exon of MAML2 is replaced by exon 1 of CRTC1. To generate the same fusion sequence, we started with the full-length coding region of MAML2, which has an internal *BamHI* site. The MAML2 fragment was subjected to endonuclease restriction enzyme digestion with *BamHI* (in MAML2) and *EcoRI* (in the pCR4 TOPO vector) (figure 4.3-A). The released fragment was subcloned into a pCR3.1 mammalian expression vector (figure 4.4). The clone was verified with restriction enzyme digestion, further confirmed by T7 forward sequencing and blast analysis, indicating 100% identity of the clone to the subject.



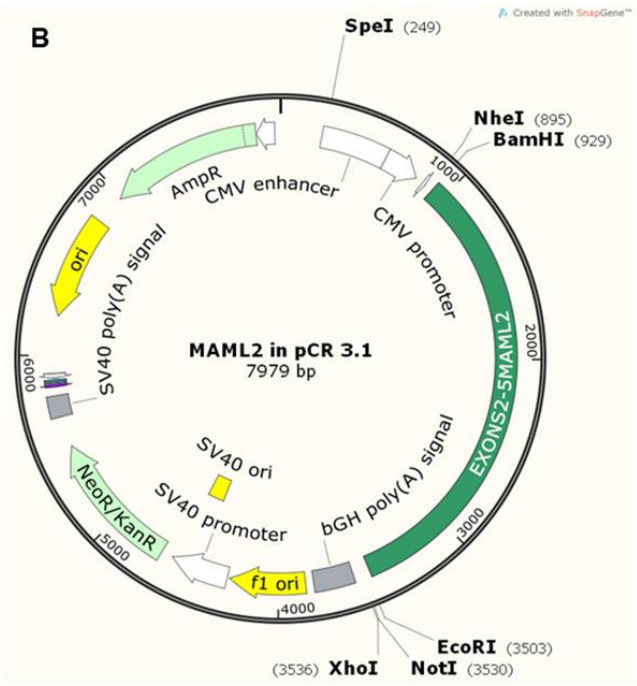


Figure 4.4 Sub cloning of MAML2 into pCR3.1: A) History map illustrating the sub cloning steps. The full length of MAML2 cloned into pCR4 vector with cleavage at *BamHI* and *EcoRI* sites and sub cloned into pCR3.1 which was previously digested with *BamHI* and *EcoRI*. **B)** A representative map of MAML2 pCR3.1 clone demonstrating the fusion sequence Tag of MAML2 fragment (green colour) cloned into a mammalian expression vector (pCR3.1). The map also shows the main restriction enzyme sites ready for the final sub-cloning step (Images created with SnapGene software).

The final cloning step was to ligate the FLAG-Tag CRTC1-MAML2 fragment into the MAML2 pCR3.1 clone. T4 DNA ligase enzyme was used to ligate the fragments, which were then transformed into competent cells and spread on ampicillin resistant agar plates. The positive colonies were selected and cultivated. Minipreps were used to extract plasmids. Restriction enzyme digestion initially confirmed cloning in the correct direction. 2900 bp bands were detected as predicted (figure 4.5).

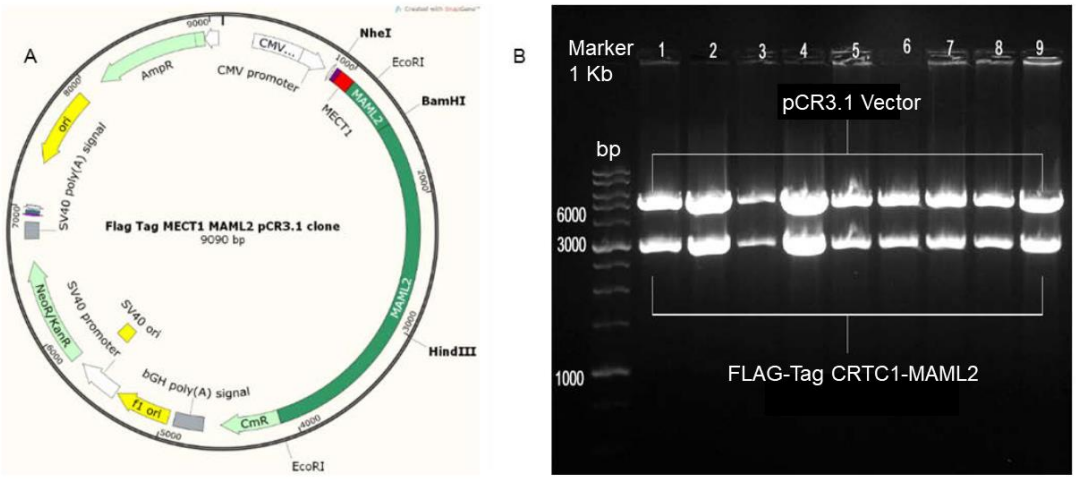


Figure 4.5 FLAG-Tag CRTC1-MAML2 fusion in pCR3.1: **A)** Map representing the final cloning construct of the fusion using *NheI* and *BamHI*, image created with SnapGene software. **B)** Restriction enzyme digestion: a sequence specific diagnostic *EcoRI* digestion of the fusion construct confirmed the expected band size released at 2900 bp from the pCR3.1 vector at the top of the gel. Different vector to insert ratios, with fixed insert, in the range 2:1 to 16:1 were used in the sub cloning. All 16 replicates positively indicated the same size, which was also confirmed by the T7 forward, R5 Reverse sequences and characterised using overlapping sequencing.

4.4.2 CRTC1-MAML2 pCR3.1 clone Characterisation

The successful cloning of the CRTC1-MAML2 fusion into a mammalian expression vector was confirmed by applying two directional sequencing, using a standard (T7) forward primer and (R5) reverse primer, as used during the cloning process. Good quality sequence data identified the gene of interest. A blast analysis was done to

align the results to the sequence reference/ subject, confirming no errors in sequence identity had been introduced through PCR reactions and 100% identity of the clone to the subject.

Further validation was performed by applying overlapping sequencing with specific primers. As the construct was 3105 bp long, specific primers were needed to sequence precise segments. The quality of sequencing was good and blast analysis indicated that the CRTC1-MAML2 fusion was 100% identical to the human fusion gene.

4.4.3 CRTC1-MAML2 pcDNA+3.1IREsGFP clone Construction and characterisation

To facilitate the transfection experiments and allow us to visualise protein production by fluorescence microscopy, the CRTC1-MAML2 pCR3.1 clone was further sub-cloned into a green fluorescence protein mammalian expression vector (pcDNA3.1+IREs GFP). The clone was subjected to double endonuclease digestion with *NheI* and *NotI* and the fragment ligated into a pre-digested pcDNA3.1+IREs GFP vector. This was then transformed into competent cells and spread on ampicillin resistant agar plates. The positive colonies were detected, selected and cultivated. Minipreps were used to extract plasmids. Restriction enzyme digestion confirmed cloning in the correct direction, 1023 bp bands were detected as predicted (figure 4.6). A standard T7 and R5 sequence reaction was performed, and the sequence further characterised using overlapping sequencing with 6 specific primers (4 forward and 2 reverse) along the whole insert. Blast analysis data indicated that the CRTC1-MAML2 fusion was 100% identical to the human fusion gene.

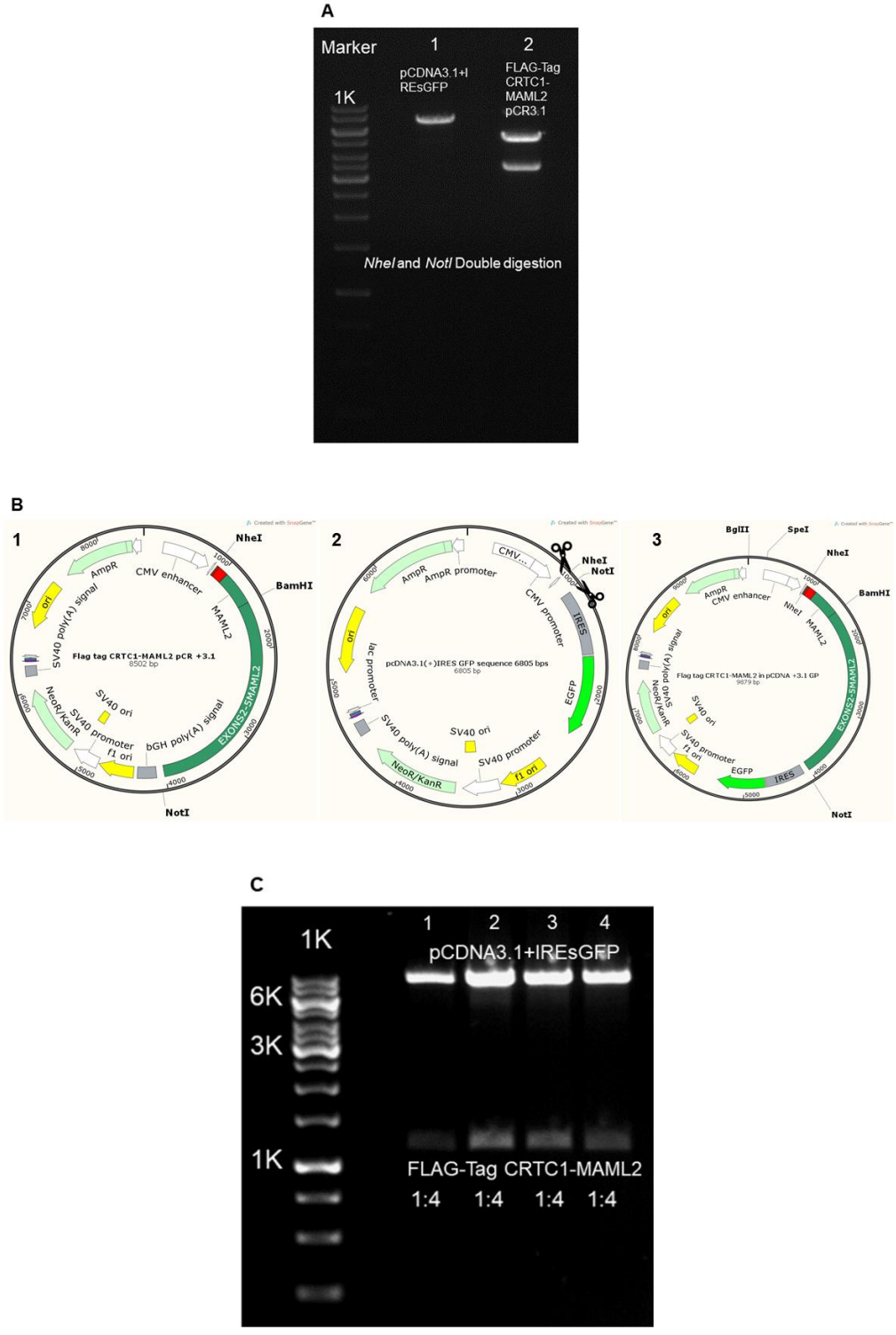


Figure 4.6 synthesis of the CRTC1-MAML2 pCDNA3.1+IREs GFP clone A) An agarose gel of double restriction enzyme digests of the CRTC1-MAML2 pCR3.1 clone and new mammalian expression vector (pCDNA3.1+IREs GFP). Lane 1 indicates *NheI* and *NotI*

digestion of the pCDNA1.3+IREsGFP empty vector resulting in two separated bands: at 6,721 bp (shown) and a small fragment at 84 bp (not shown). Similarly, lane 2 represents *NheI* and *NotI* double cutting causing two separate bands: (pCR3.1) vector at 5,433bp and FLAG-Tag CRTC1- MAML2 fragment at 3,105 bp. **B)** A history map created with SnapGene software illustrates the sub cloning steps. The FLAG-Tag CRTC1-MAML2 pCR3.1 clone was double cleaved at *NheI* and *NotI* sites to release the fusion fragment (1) to subclone into the new pCDNA3.1+IREGsFP empty vector previously double digested at *NheI* and *NotI* sites (2). A representative map of CRTC1-MAML2 pCDNA3.1+IREs GFP clone demonstrating the fusion sequence (Red and Green colour) cloned into a mammalian expression vector (pCDNA3.1+IREs GFP) (3). The map also shows the main restriction enzyme sites ready for diagnostic restriction digest. **C)** Restriction enzyme digestion: a sequence specific diagnostic *BamHI* and *SpeI* digestion of the fusion construct confirmed the expected band sizes (1023 bp) were released from the pCDNA3.1+IREs GFP vector at the top of the gel. Different vector to insert ratios, with fixed insert, ranging from 2:1 to 16:1 were used in sub cloning. The four replicates of the 4:1 ratio were the same size.

4.4.4 Assessing the expression of the cloned constructs

4.4.4.1 Optimisation of transfection conditions

To confirm the production and expression of the recombinant protein, the pCR3.1 and pCDNA+3.1IREsGFP vectors containing the FLAG-Tag CRTC1-MAML2 inserts were transfected into mammalian cell lines and normal primary human salivary gland cells (HuSL). FugeneHD or JetPRIME polyplus reagents were used to transfect 6-well plates seeded with immortalised metastatic pulmonary carcinoma cells NCI-H647, human embryonic kidney cells (HEK 293) and normal human primary sublingual glands (HuSL) with 2 µg of the FLAG-Tag CRTC1-MAML2 pCR3.1 or pCDNA3.1+IREs GFP constructs; where the pCR3.1 vector was used the cells were also co-transfected with 100 ng of a vector encoding green fluorescent protein (pEGFP-N1) to allow assessment of transfection efficiency. To optimise transfection

experiments, in terms of cell viability and transfection efficiency, both FugeneHD and JetPRIME polyplus were used with NCI-H647 and HuSL cells, followed by flow cytometry analysis. The JetPRIME transfection reagent gave better transfection efficiency than FugeneHD, resulting in 35.80% (NCI-H647) and 53.24% (HuSL) cells being transfected by JetPRIME in comparison to 10.73% and 1.54% with FugeneHD (figures 4.7 and 4.8). Neither reagent impacted cell viability with or without the presence of the constructs (figure 4.9).

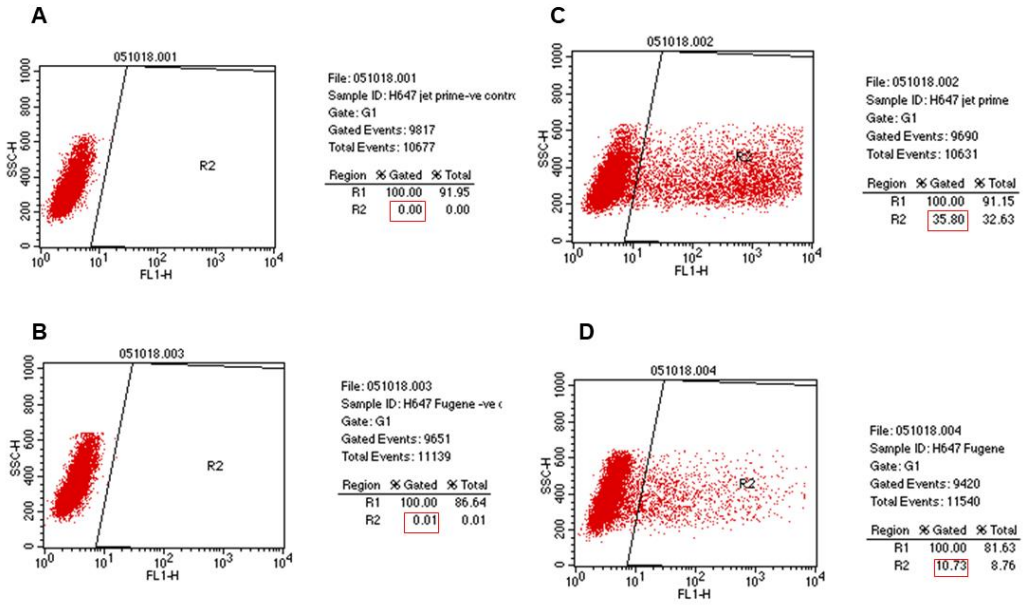


Figure 4.7 Flow cytometry data analysis/ FACS. Representative 2-dimensional dot plots of the flow cytometry data derived from co-transfected NCI-H647 cells with FLAG-Tag CRTC1-MAML2 pCR3.1 and pEGFP-N1 constructs using FugeneHD and JetPRIME transfection reagents; transfection efficiency expressed as percent GFP-positive live cells (red box). **(A and B)** NCI-H647 non-transfected control indicates no positive cells gated at R1and R2 with dot plots. **C)** NCI-H647 48 hours post-transfection using JetPRIME show dot plots are gated at R1 and R2, and 35.80% of the cells as being positive. **D)** NCI-H647 48 hours post-transfection using FugeneHD reagent show dot plots are gated at R1 and R2 and 10.73% of the cells as being positive.

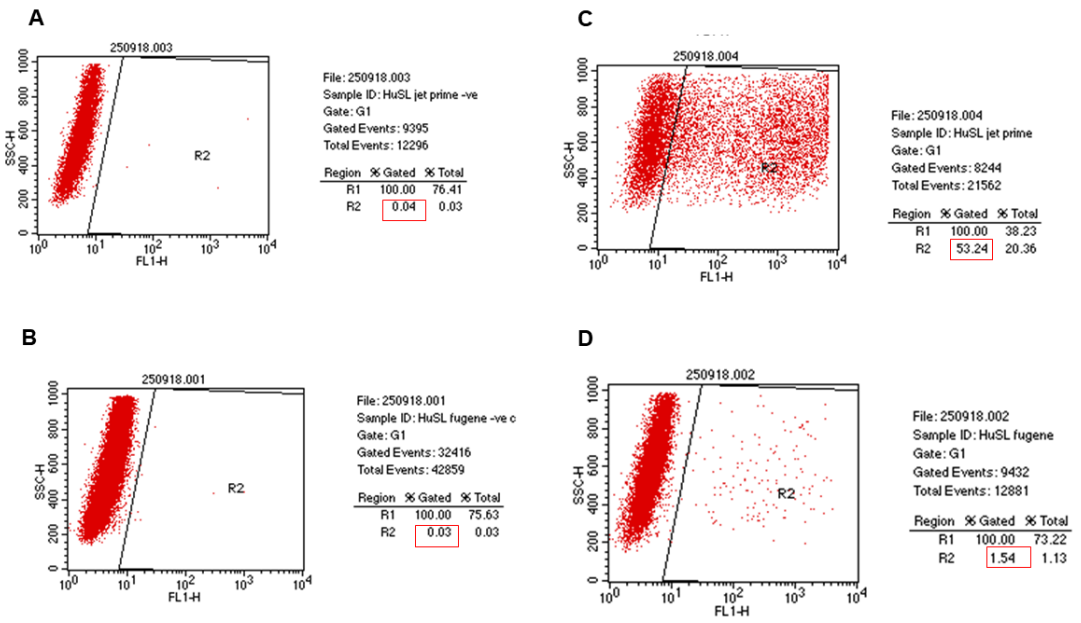


Figure 4.8 Flow cytometry data analysis/ FACS. Representative 2-dimensional dot plots of the flow cytometry data derived from co-transfected normal primary human sublingual gland cells (HuSL) with FLAG-Tag CRTC1-MAML2 pCR3.1 and pEGFP-N1 constructs using FugeneHD and JetPRIME transfection reagents; transfection efficiency expressed as percent GFP-positive live cells (red box). **(A and B)** HuSL non-transfected control indicates no positive cells gated at R1and R2 with dot plots. **C)** HuSL 48 hours post-transfection using JetPRIME show dot plots are gated at R1 and R2, and 53.24% of the cells were positive. **D)** HuSL 48 hours post-transfection using FugeneHD reagent show dot plots are gated at R1 and R2 and 1.54% of the cells were positive.

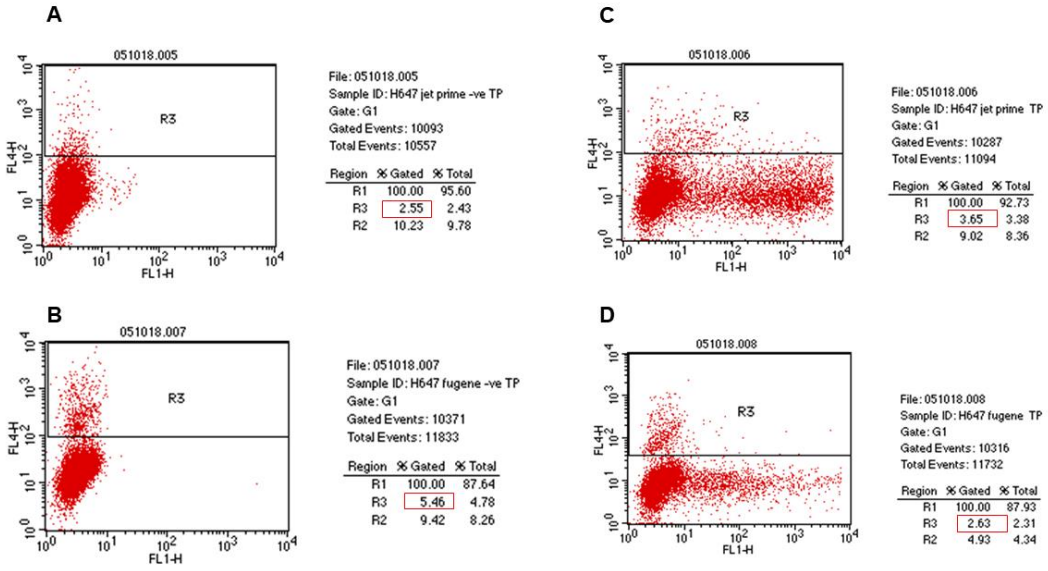


Figure 4.9 Flow cytometry data analysis/ FACS viability Representative 2-dimensional dot plots of the flow cytometry data derived from transfected NCI-H647 cells stained with TO-PRO-3 cell viability stain to determine the level of apoptosis in the cells after transfection with JetPRIME or FugeneHD and analysed with FL4H cell sorter. **(A and B)** Non-transfected control NCI-H647 cells exposed to transfection reagents only indicate a low percentage of dead cells (red box) gated at R3, 2.55% and 5.46% respectively. **(C and D)** NCI-H647 cells transfected with FLAG-Tag CRTC1-MAML2 pCR3.1 construct, 48 hours post-transfection, also demonstrated a low percentage of dead cells (red box) gated at R3 3.65% and 2.63% respectively.

To determine the optimal dilution of JetPRIME reagent (ratio of DNA to transfection reagent), three ratios were assessed: 1:2, 1:4 and 1:5, and transfection efficiency was determined by FACS. This experiment was performed on NCI-H647 and HEK293 cells, as these two cell lines exhibited better transfection rates with JetPRIME than FugeneHD. The best DNA to reagent dilution was 1:2 as this provided the best transfection efficiency in both cell lines, >80% in NCI-H647 cells and > 39% with the HEK293 cells (figure 4.10). These optimal transfection conditions were used for all further experiments.

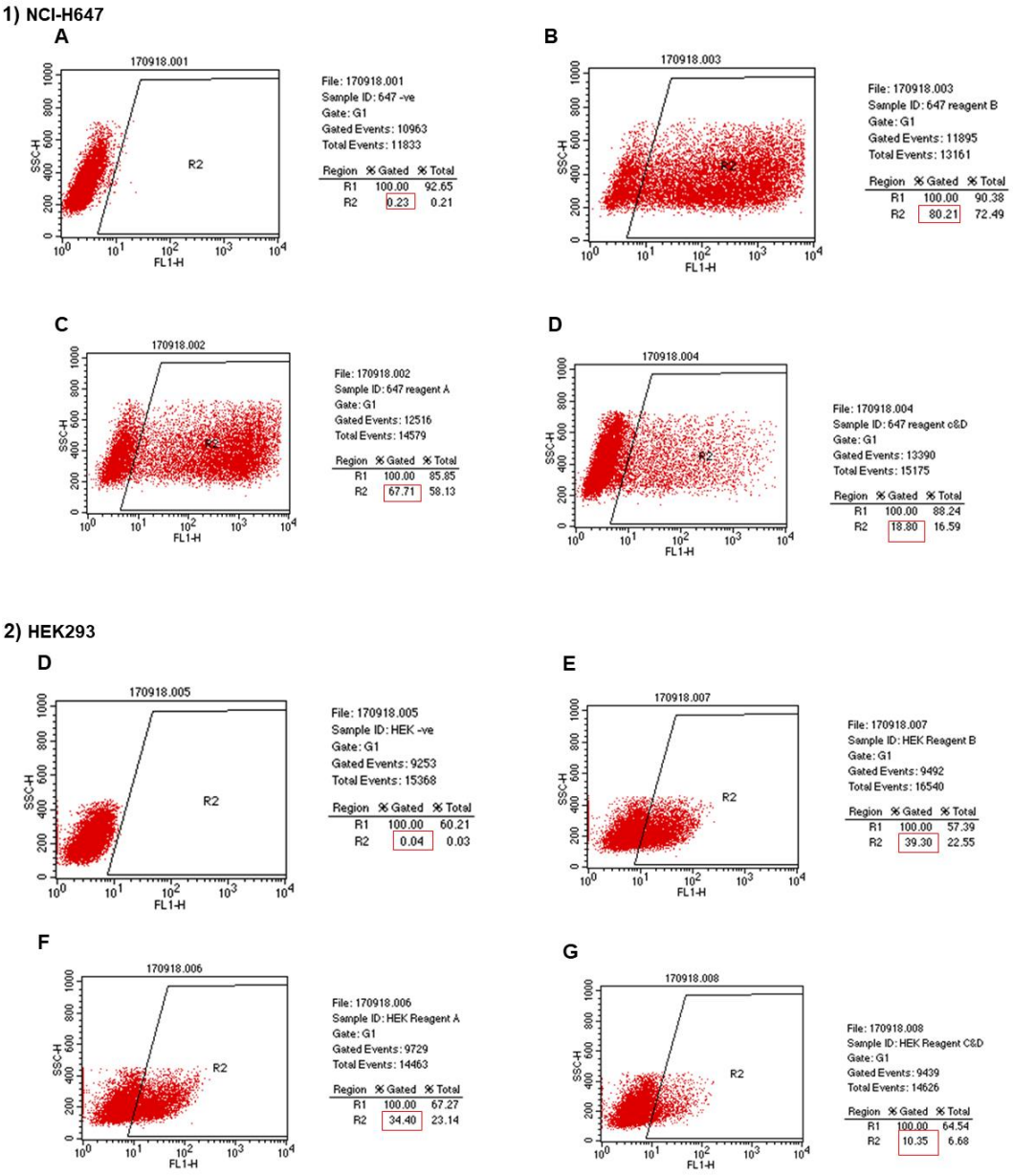


Figure 4.10 Flow cytometry data analysis/ FACS. Representative 2-dimensional dot plots of the flow cytometry data derived from co-transfected NCI-H647 (1) and HEK 293 (2) cells with FLAG-Tag CRTC1-MAML2 pCR3.1 and pEGFP-N1 constructs using three dilutions of JetPRIME transfection reagent (1:2, 1:4 and 1:5); transfection efficiency is expressed as percent GFP-positive live cells (red box). Non-transfected control NCI-H647 and HEK293 cells indicate no positive cells gated at R1 and R2 with dot plots **(images A and D)**. 48 hours post-transfection using 1:2 **(images B and E)**, 1:4 **(images C and F)** and 1:5 **(images D and G)** DNA to transfection reagent shows dot plots gated at R1 and R2.

In addition to determining the optimal transfection reagent and dilution by FACS the NCI-H647, HEK293 and HuSL cells were analysed microscopically (figures 4.11-13), 48 hours after transfection with the JetPRIME reagent, FLAG-Tag CRTC1-MAML2 pCR3.1 and the pEGFP-N1 construct. Cell nuclei were stained with Hoechst 33342 (1:2000 in PBS, Thermo Fisher Scientific, UK) for 3 minutes at room temperature. Cytoplasmic expression of the GFP was clearly noted in the successfully transfected cells which was in line with the FACS analysis data. Microscopic analysis also allowed us to determine that the cells appeared healthy and viable, which again was consistent with our FACS results and indicated that the transfection reagent and the constructs were not cytotoxic.

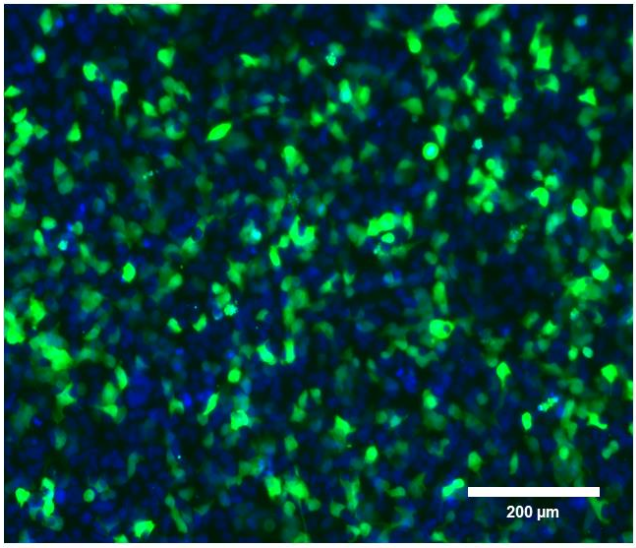


Figure 4.11 Fluorescence microscopy images of transiently transfected NCI-H647 cells with FLAG-Tagged CRTC1-MAML2 and eGFP-N1 using JetPRIME. A representative merged image of eGFP expression (Green) with Hoechst (Blue) staining of the nucleus, 48hrs after transfection, indicated high transfection efficiency rates. Image magnification at 10×.

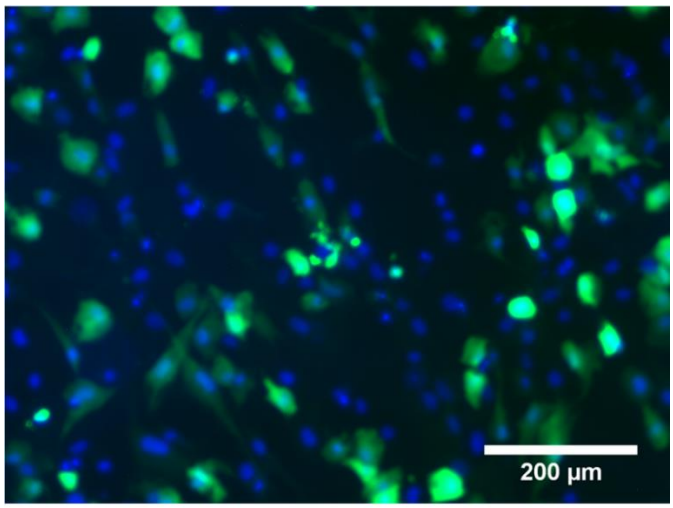


Figure 4.12 Fluorescence microscopy images of transiently transfected normal human primary salivary gland cells (HuSL) with FLAG-Tagged CRTC1-MAML2 and eGFP-N1 using JetPRIME. A representative merged image of eGFP expression (Green) with Hoechst (Blue) staining of the nucleus, 48hrs after transfection, indicated high transfection efficiency rates. Image at 10× magnification.

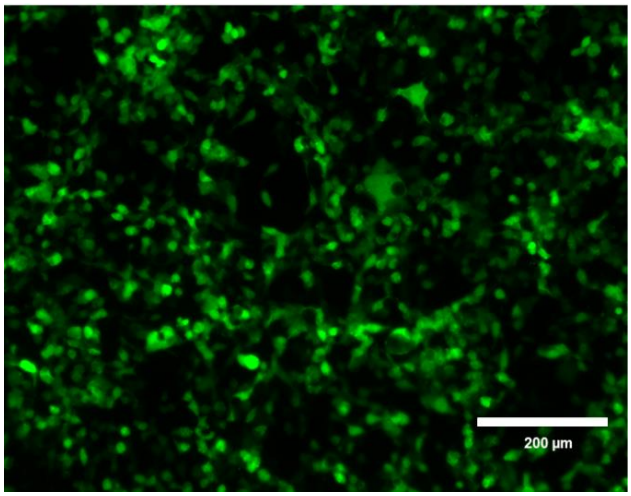


Figure 4.13 Fluorescence microscopy images of transiently transfected HEK293 cells with FLAG-Tagged CRTC1-MAML2 and eGFP-N1 using JetPRIME. A representative image of eGFP expression (Green), 48hrs after transfection, indicated high transfection efficiency rates. Image at 10× magnification.

4.4.4.2 Detection of fusion protein synthesis

Following effective transfection, protein expression was analysed by harvesting cell lysate and western blotting. As a control, NCI-H647 cells were transfected with an alternative FLAG-Tagged construct (F9. hANKS6. pCDNA6, a generous gift from Devon C Smith, Academic Unit of Nephrology, Department of Infection, Immunity & Cardiovascular Disease, The University of Sheffield) to verify the ability of the FLAG antibody to positively bind to expressed protein and to confirm the transfer of high molecular weight protein to a membrane. This construct was a similar expected molecular weight at 100 KDa. The expression of the CRTC1-MAML2 recombinant protein in the pCR3.1 construct was successfully detected using high affinity monoclonal anti-FLAG. A protein of 115 KDa was detected (figure 4. 14).

To validate fusion protein expression results with the MAML2 antibody, normal epithelial cells (HEK293 cells) and NCI-H647 were transfected with pCDNA3.1+IREs GFP vectors with or without FLAG-Tag CRTC1-MAML2 inserts. Whole cell lysates were analysed by western blotting, with known fusion protein expressing MEC (NCI-H292 and UM-HMC-2) being used as positive controls. The CRTC1-MAML2 fusion was detected in transfected HEK293 and NCI-H647 cells and also the NCI-H292 and UM-HMC-2 cells. Two separate bands, at around 125 KDa and 115 KDa, were detected, while only the 125 KDa band was detected in the wild-type HEK293 and NCI-H647 cells (figure 4.15).

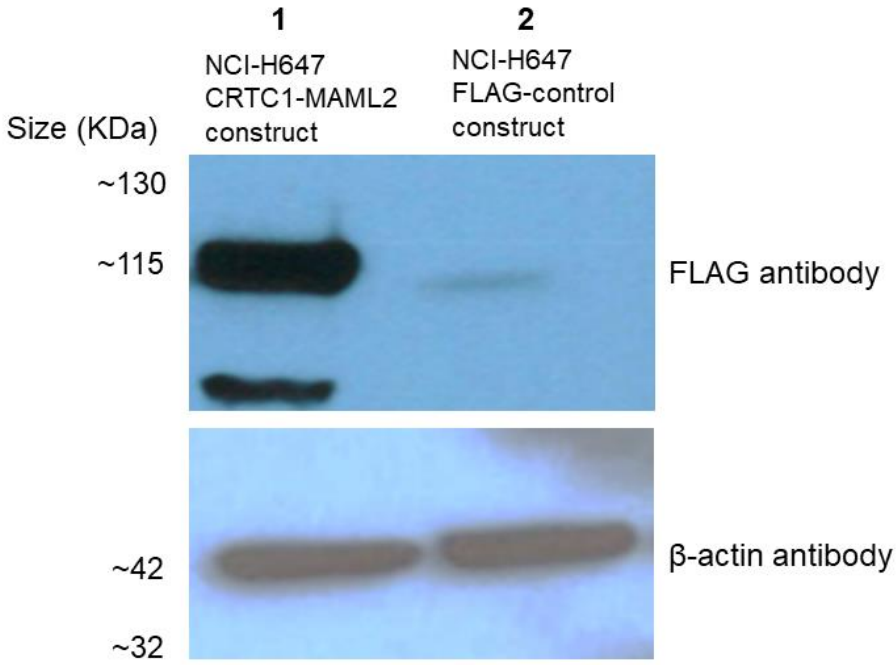


Figure 4.14 CRTC1-MAML2 fusion protein expression in a cell line. NCI-H647 cells were transfected with pCR3.1 vector containing CRTC1-MAML2 or control vector (lane 2) with FLAG-Tag and expression of CRTC1-MAML2 was analysed by Western blotting using an anti-FLAG monoclonal antibody. Lane 1 –NCI-H647 cells transfected with CRTC1-MAML2 construct, 2 – NCI-H647 cells transfected with the control vector. β-actin antibody was used as a protein loading control.

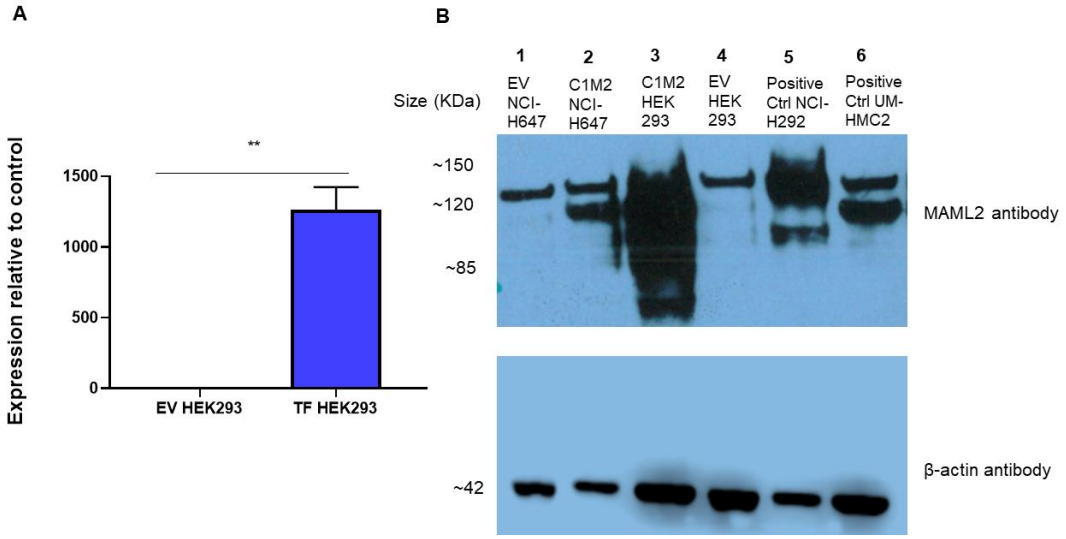


Figure 4.15 CRTC1-MAML2 fusion transcript and protein expression in cell lines. A) CRTC1-MAML2 transcript upregulation in transfected HEK293 cells versus empty vector control. Statistical analysis by two tailed Unpaired t Test (*P* value ** = 0.0054) (N=3, n=3). **B)** Cellular protein was extracted from fusion positive cell lines (NCI-H292 and UM-HMC-2) and transfected NCI-H647 and HEK293 with pCDNA3.1+IREs GFP vector with either the FLAG-Tag CRTC1-MAML2 insert or an empty vector**.** Expression of CRTC1-MAML2 was analysed by Western blotting using an anti-MAML2 polyclonal antibody to compare endogenous MAML2 expression with break apart MAML2. Lane 1- NCI-H647 plus pCDNA3.1+IREs GFP, 2 – NCI-H647 cells plus fusion construct, 3 – HEK293 plus fusion vector, 4 – HEK293 plus pCDNA3.1+IREs GFP vector, 5- NCI-H292. 6- UM-HMC-2. β-actin antibody was used as a protein loading control.

4.4.4.3 Localisation of fusion protein expression

NCI-H647 cells were grown in an 8-chamber slide in order to determine localisation of the CRTC1-MAML2 fusion protein in cells. Four chambers were transfected with the CRTC1-MAML2 construct and four were mock transfected. 48 hours after transfection the cells were stained with an anti-FLAG antibody, in order to avoid

confusion due to endogenous MAML2 expression, and NCI-H647 cells were used in rather than HEK293 cells, as the former are more adherent.

Expression of the CRTC1-MAML2 fusion protein was found in the nucleus of transfected cells (figure 4.16).

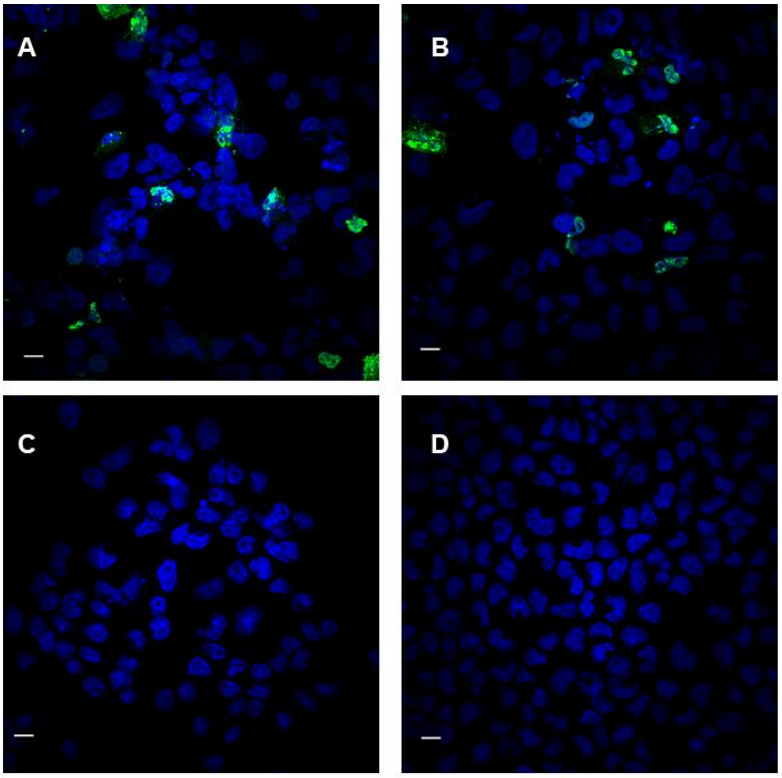


Figure 4.16 Fusion protein localisation in NCI-H647 cells. Representative images **(A & B)** demonstrating positive nuclear expression of the CRTC1-MAML2 fusion protein (green) following transient transfection of NCI-H647 cells. Nuclei (blue) were counterstained with DAPI. Antibody control **(C)** and negative control **(D)**. Images were taken using a confocal microscope (Zeiss 880 AiryScan), scale bar at 10µm.

4.4.5 Analysing the effect of CRTC1-MAML2 on Normal salivary gland cells

4.4.5.1 Expression of the CRTC1-MAML2 transcript in transfected normal salivary gland cells

One of the main aims of this study was to investigate the downstream effect of CRTC1-MAML2 expression on normal salivary gland cells. Initially, normal human primary sublingual gland cells (HuSL), from explanted tissue, were grown in a 6-well plate and transiently transfected with the FLAG-Tag pCR3.1 construct using JetPRIME. The transfected cells were analysed 48 hours post-transfection, the gene expression transcript quantified by qPCR using the F6 R5 fusion primers (as used for construct synthesis). CRTC1-MAML2 transcript expression increased ten-fold (*P***= 0.0051) after normalisation to the housekeeping gene U6. To confirm expression of the fusion transcript, the PCR products were separated on a 1% TAE gel and visualised under UV transillumination. This indicated the product was of the expected size (~720 bp) (figure 4.17 A and B). Expression of the fusion protein was further investigated by western blot analysis with a MAML2 antibody. Two separate bands of the MAML2 protein were detected in transfected cells and the MEC fusion positive cells (NCI-H292 cells) and one band in control HuSL (figure 4.17 C). Western blot analysis with the FLAG antibody, demonstrated a higher, non-specific band (~200 KDa) which did not correspond with the expected protein size. Therefore, after multiple attempts, the FLAG antibody was excluded from the protein analysis of primary cells. Genomic DNA (gDNA) was extracted from transfected HuSL cells and analysed by end point PCR to visualise the genomic integration of the CRTC1-MAML2 translocation using the R6 F5 primers. gDNA was also extracted from HuSL transfected control cells. Our results showed that there was a distinct PCR product at the expected size ~ 700 bp, suggesting genomic integration into the HuSL transfected cells but not the HuSL control cells (figure 4.17 D).

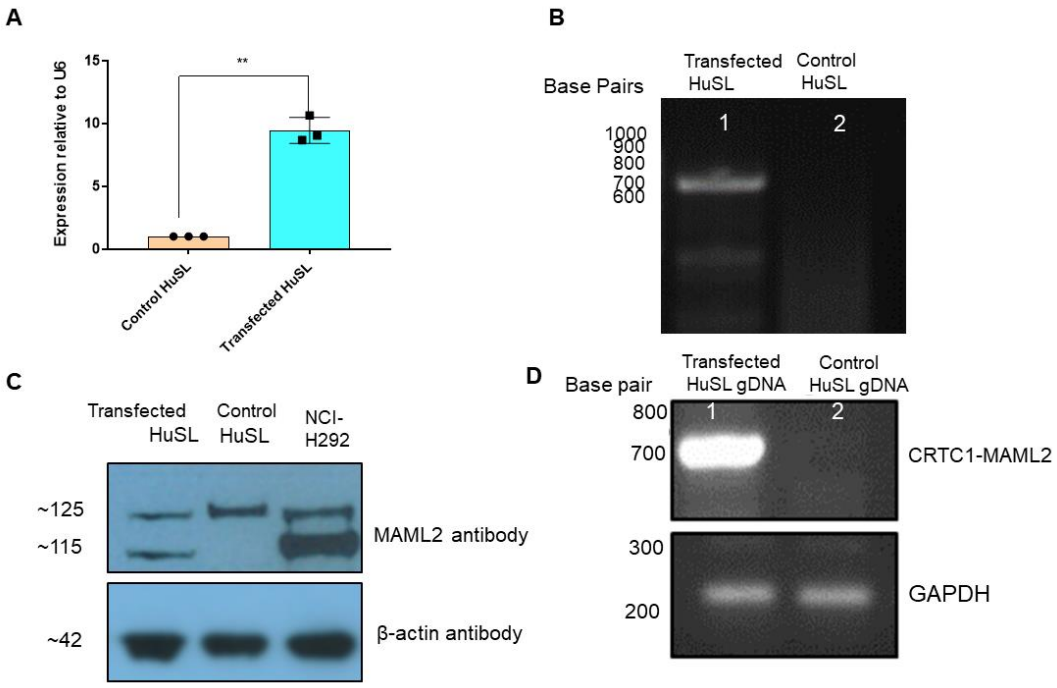


Figure 4.17 CRTC1-MAML2 fusion gene, transcript and protein expression in transfected normal salivary gland cells (HuSL). A) Real-time qPCR analysis of fusion gene expression in normal primary salivary gland cells. Expression was normalised to U6. Statistical analysis by two tailed Unpaired t Test (*P* value ** = 0.0051) (N=3, n=3). **B)** Gel electrophoresis of qPCR products after 40 cycles, using F6 R5 fusion primers confirmed the presence of the CRTC1-MAML2 fusion at the expected product size ~720 bp. **C)** Western blot analysis of transfected HuSL cells showing break apart MAML2 at the expected size compared to controls. β-actin antibody was used as a protein loading control. **D)** An agarose gel of gDNA PCR products of transiently transfected HuSL which indicated the expected size at 720 bp.

To confirm the reproducibility of CRTC1-MAML2 transcript expression in transfected normal primary salivary gland cells, a panel of cells from different donors and sites were studied. Three from sublingual glands (HuSL1, HuSL2 and SLG) and two from submandibular glands (SMG1 and SMG2). A custom-made TaqMan qPCR probe was designed to specifically amplify the fusion transcript. Transfection was carried out in T75 flasks for a greater yield of mRNA. qPCR values were normalised to B2M housekeeping gene. A statistically significant increase in CRTC1-MAML2 gene expression was demonstrated with all of the transfected primary human salivary gland cells when compared to control cells (*P* value **= 0.0024, *** = 0.0002, **** = 0.0001) (figure 4.18). We also assessed endogenous expression of the CRTC1-MAML2 transcript in fusion positive MEC cell lines (NCI-H292 and UM-HMC2) and in a MEC cell line of unknown fusion status (A-253) as further controls. The *de novo* expression of the CRTC1-MAML2 transcript in normal transfected cells was significantly higher than the endogenous expression in those cells which naturally harbour the translocation (77 × higher in the transfected primary cells than the MEC cell lines). This could be due, in part, to the normal transfected cells having a high transfection efficiency rate, resulting in enriched populations of CRTC1-MAML2 fusion positive cells. There was no endogenous expression of the CRTC1-MAML2 transcript in the A-253 cells, nor was there expression in control normal salivary gland cells (figure 4.18-E).

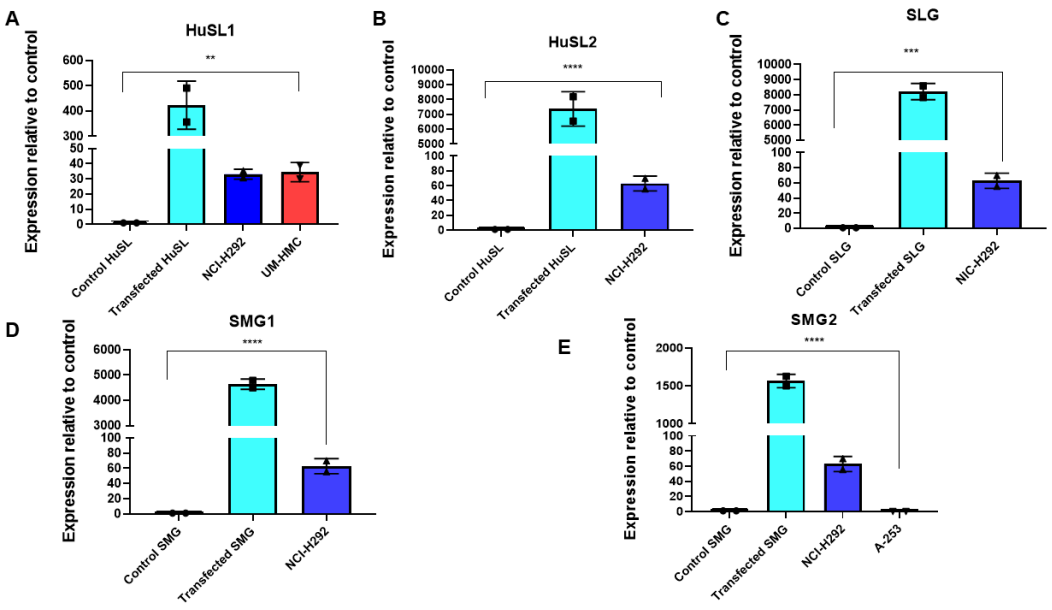


Figure 4.18 CRTC1-MAML2 upregulation in normal transfected primary salivary gland cells versus normal control. (A-E) CRTC1-MAML2 expression in 5 transfected normal salivary gland cells; transcript levels were analysed using qPCR and normalised to B2M endogenous control. Expression was significantly higher than in normal control cells or cells which endogenously express CRTC1-MAML2 (NCI-H292 and UM-HMC-2). No expression was detected in A-253 cells. Statistical analysis was carried out by one-way ANOVA with Bartlett's' correction (*P* value **= 0.0024, *** = 0.0002, **** = 0.0001) (N=2, n=3).

4.4.5.2 Generation of a stable cell line harbouring the CRTC1-MAML2 fusion

Primary human salivary gland cells were used to attempt to generate stable, *de novo* CRTC1-MAML2 expressing cells, which were to be used in functional assays. Antibiotic selection, kill curve experiments, were carried out to determine optimal doses of antibiotics required as both the pCR3.1 and pCDNA3.1+IREs GFP vectors carry the Geneticin, G418 selection gene and the pEGFP-N1 vector carries the Neomycin selection gene which also provides resistance against G418. Initially, NCI-H647, HEK293 and primary human salivary gland cells, at around 70% confluency, were exposed to a range of Geneticin (0 µg/ml to 1 mg/ml) in the absence of penicillin/streptomycin. For HEK293 cells the optimal dose was determined to be 700

µg/ ml as this resulted in 100% cell death after 10 days of treatment (figure 4.19). Similarly, optimal doses for NCI-H647 cells and primary human normal salivary gland cells were determined to be 800 µg/ ml and 600 µg/ ml, respectively (figures 4.20-4.21).

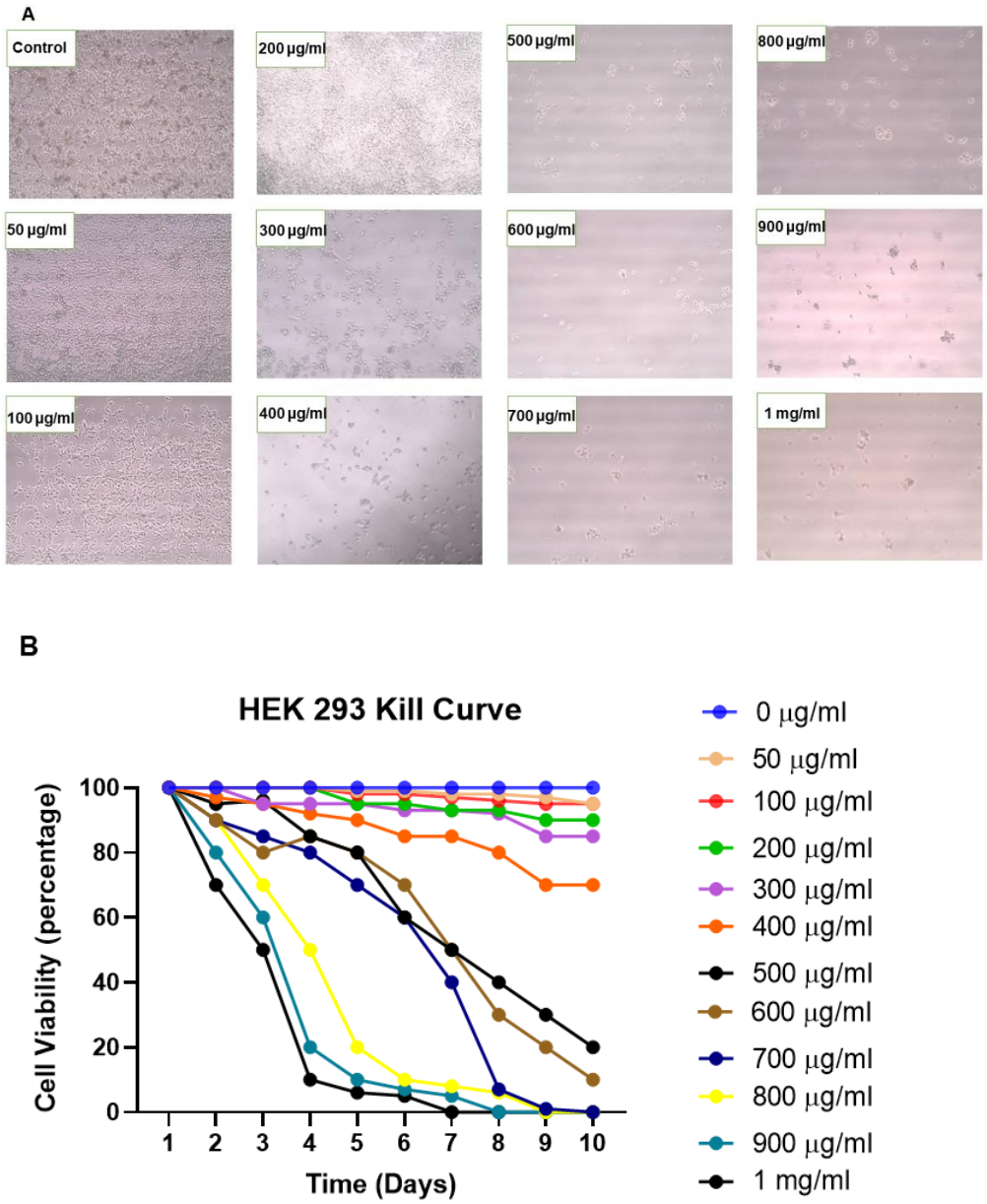


Figure 4. 19 G418 Kill Curve of HEK293 cells. A) Representative light microscopic images of cell response to a range of antibiotic concentrations following 10 days of G418 treatment. 100% of cells treated with concentrations of 700 µg/ ml had died whereas cells treated with lower concentrations remained viable. **B)** Representative graph showing cell viability following G418 treatment (0.05- 1 mg/ ml) for 10 days**.**

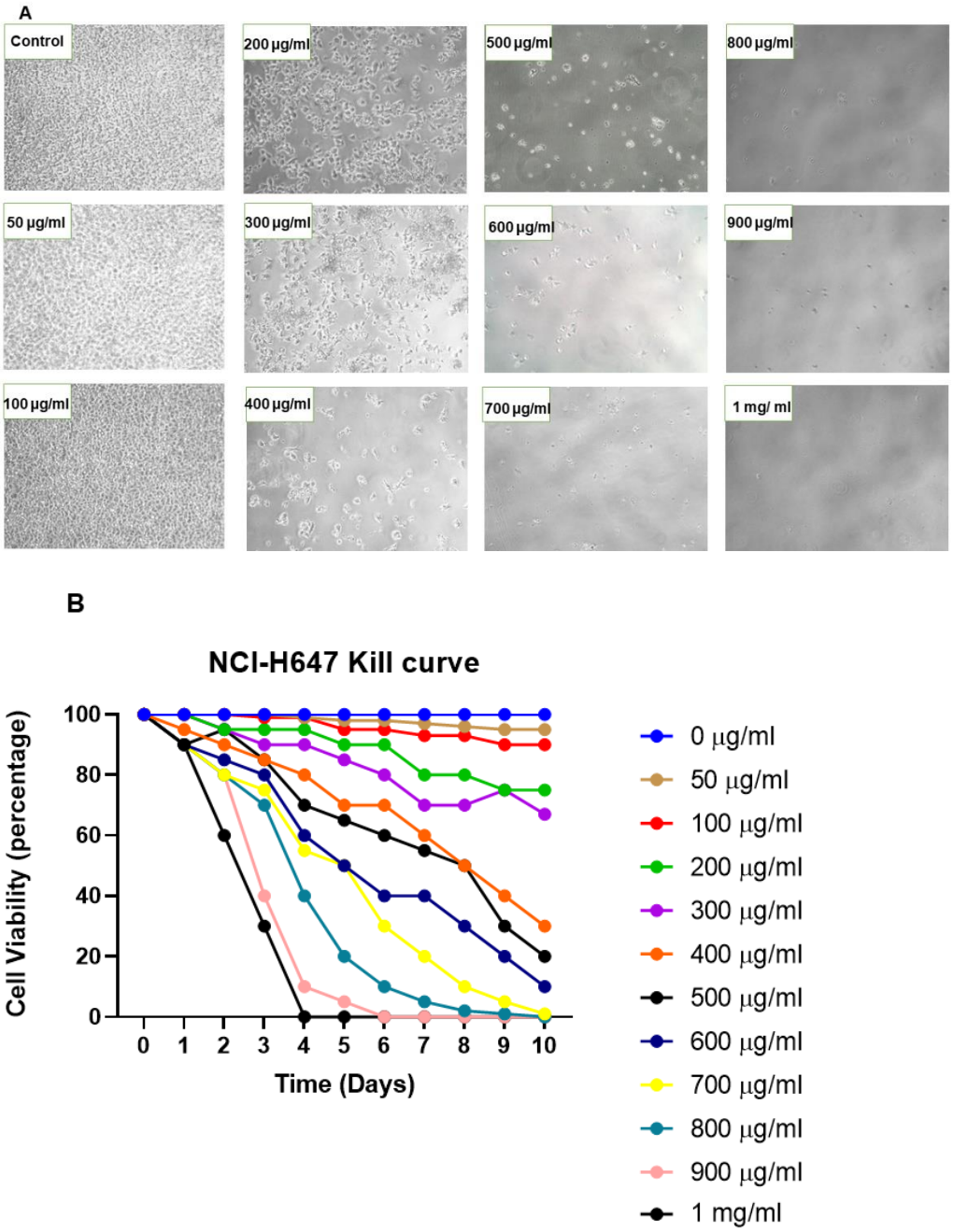


Figure 4.20 G418 Time Kill Curve of NCI-H647 cells. A) Representative light microscopic images of cell response to a range of antibiotic concentrations following 10 days of G418 treatment. 100% of cells treated with concentrations of 800 µg/ ml had died whereas cells treated with lower concentrations remained viable. **B)** Representative graph showing cell viability following G418 treatment (0.05- 1 mg/ ml) for 10 days.

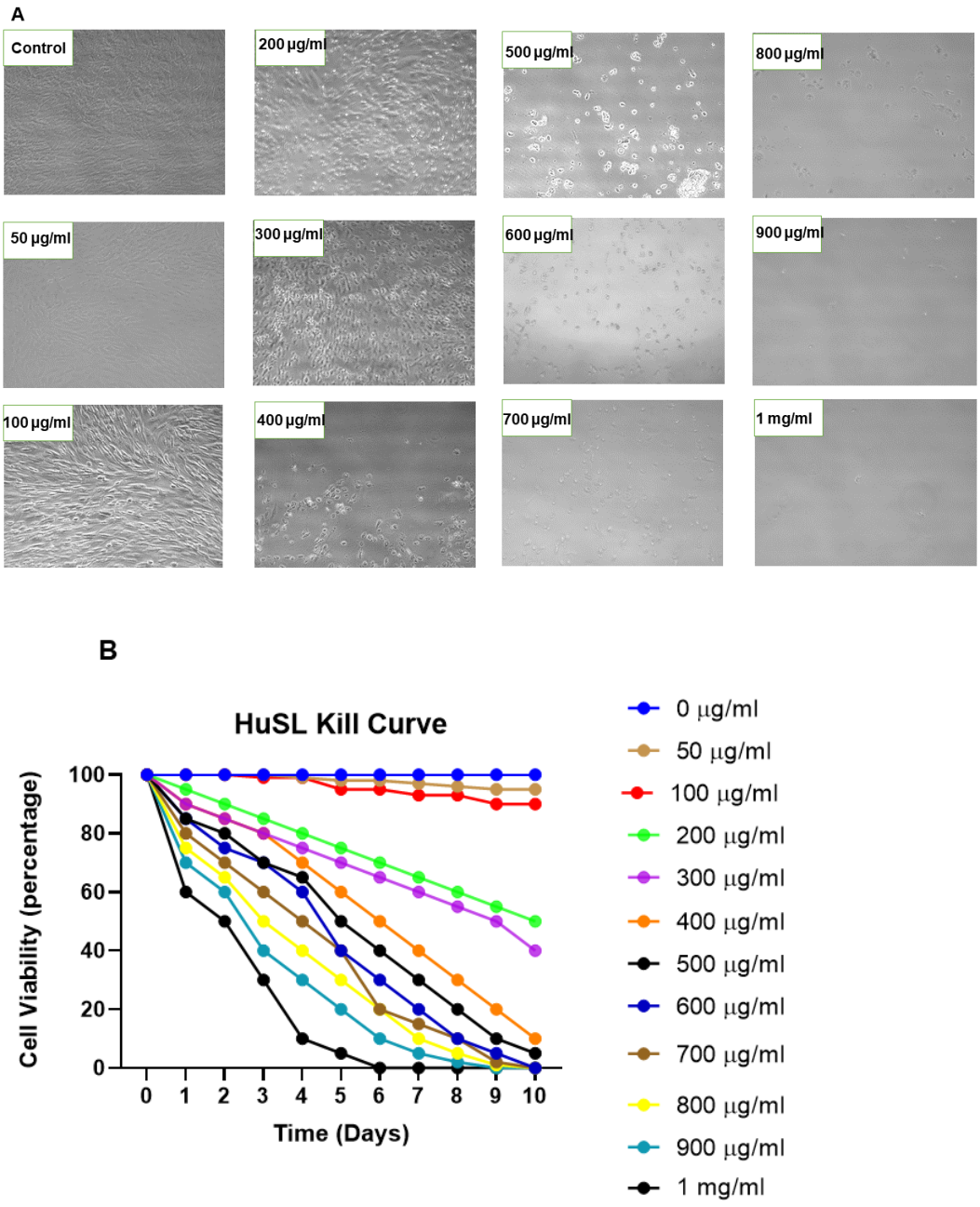


Figure 4.21 G418 Time Kill Curve of HuSL cells. A) Representative light microscopic images of cell response to a range of antibiotic concentrations following 10 days of G418 treatment. 100% of cells treated with concentrations of 600 µg/ ml had died whereas cells treated with lower concentrations remained viable. **B)** Representative graph showing cell viability following G418 treatment (0.05- 1 mg/ ml) for 10 days**.**

NCI-H647, HEK293, HuSL, HPG and SLG cells were transfected with the FLAG-Tag CRTC1-MAML2 pCDNA3.1+IREs GFP construct as previously described in sections 2.4.2-2.4.6. Forty-eight hours post transfection, cells were examined by fluorescence microscope to determine expression of GFP and cells were then either sorted via FACS or selection with G418 was commenced.

For FACS sorting, cells were sorted into GFP positive and GFP negative cells, with the GFP negative cells serving as internal negative controls for all downstream experiments. Sorted cells were expanded for further analysis. Primary cells did not grow well from single cell colonies, despite the addition of conditioned media to the cells, so they were seeded in 48-well plates as poly-clonal populations and expanded in the presence of 10 µg/ml of Rho-Kinase (ROCK-inhibitor) (Y-2763; section 2.4.4). The NCI-H647 and HEK293 colonies were expanded from single cells in 96-well plates. For G418 antibiotic selection, cells were treated with the predetermined optimised dose of geneticin for 10-14 days alongside positive and negative controls. When all non-transfected control cells had died and non-selected cells were growing normally, cells were expanded (2.4.2-2.4.6; (figure 4.22)).

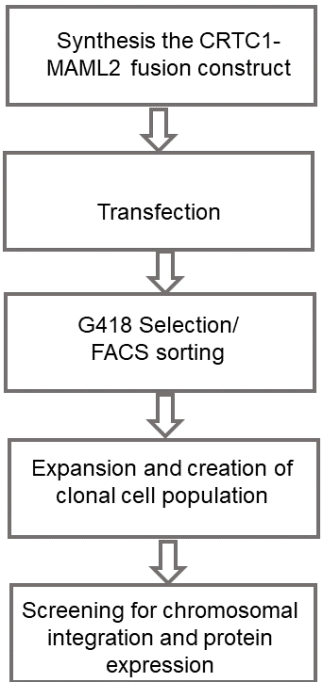


Figure 4.22 Stable cell generation diagram. Summary of stable cell line generation.

gDNA was extracted from the cells and PCR analysis performed to check the integration of the CRTC1-MAML2 translocation gene into the genome. The same pair of primers (F6 R5) previously used for construct synthesis and transient transfection experiments was used to screen the integration. Negative, non-transfected cells were used as internal controls, and the fusion positive MEC cells (NCI-H292) were used as a positive control. GAPDH was used as endogenous control. PCR products were separated on an agarose gel but no successful integration of the translocation genes into the stably transfected cells was noted. The only band observed was that of the endogenous translocated genes of the NCI-H292 cells (figure 4.23).

qPCR analysis was performed on RNA extracted from stably transfected and control cells using TaqMan primers against the fusion transcript junction. The RT-qPCR results showed that, again, there was no significant expression of the CRTC1-MAML2 fusion transcript in the stably transfected normal salivary gland cells (figure 4.23).

Further analysis at the protein level demonstrated CRTC1-MAML2 fusion expression in the transiently transfected HEK293 cells (two independent transfections) and NCI-H292 MEC fusion positive cells (control), as two separate bands at molecular weight of ~125 and ~115 KDa. The fusion protein was also not detected in stably transfected NCI-H647, human normal primary salivary gland cells (HuSL1, HuSL2 and SMG), nor in control HuSL cells. Only the larger molecular weight band at ~125 KDa was detected in the negative samples (figure 4.23).

As stated previously, the generation of stable cell lines to harbour the translocation and produce the CRTC1-MAML2 fusion protein was attempted with two cell lines, a panel of primary human cells, but no cells showed chromosomal integration or fusion protein expression despite the cells having been selected by antibiotic treatment or FACS.

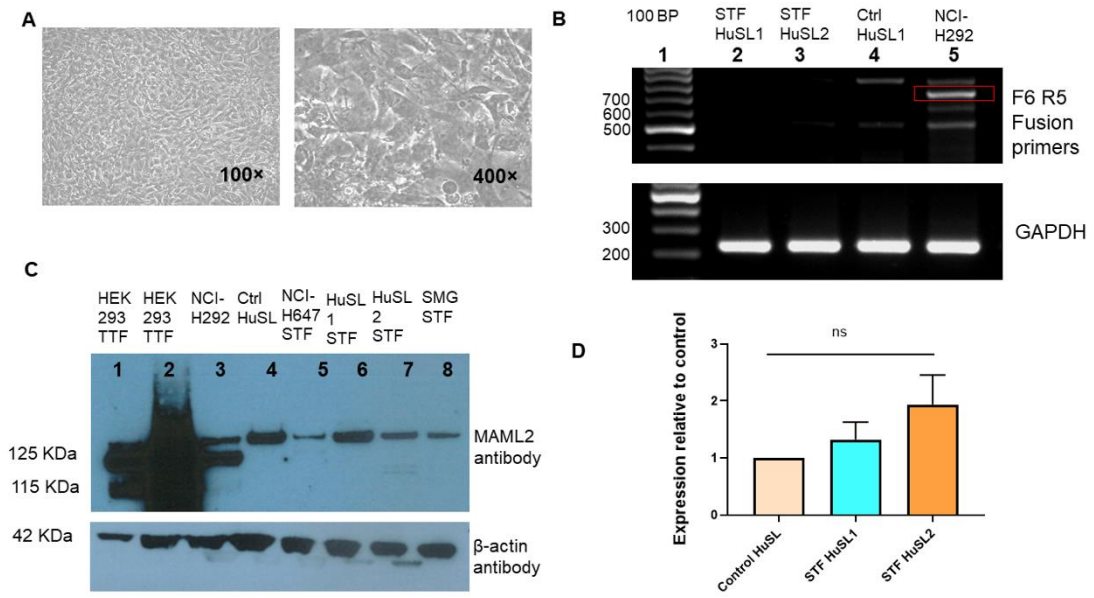


Figure 4.23 Stable transfection analysis. A) Representative phase contrast images of expanded stably transfected HuSL cells (P5). **B)** Resolution of PCR products derived from gDNA of potential stably transfected cells. Primer pairs flanking the fusion target within gDNA of transfected HuSL cell colonies were used for PCR. Products were resolved on a 2% agarose gel. In the presence of the translocated genes in the MEC cells (NCI-H292) expected band sizes were 720 bp for primer pair F6 R5. Lane 1 represents the DNA marker, lanes 2 and 3 represent two different HuSL1 stably transfected cells, lane 3 represents negative internal HuSL control, lane 5 represents positive control fusion band in the red box. (N=1). **C)** Western blot analysis of CRTC1-MAML2 fusion protein expression in the stably transfected primary cells and NCI-H467 cells using an anti-MAML2 polyclonal antibody. Lanes 1 and 2 represent TTF HEK293 cell lysate (exogenous positive control) from two independent previous transfections, 3 – NCI-H292 cell lysate (endogenous positive control), 4 – HuSL internal control (negative control), 5 – Stably transfected NCI-H647 cells, 6 – Stably transfected HuSL1, 7 – Stably transfected HuSL2 cells. 8- Stably transfected SMG cells. (N=3). **D)** RT-qPCR quantification of CRTC1-MAML2 fusion gene expression in stably transfected HuSL cells. Expression is represented as fold change in expression compared to the B2M endogenous control. Data represents the mean ± SD. Experiment carried out in duplicate with two biological repeats and three technical repeats (N=2; n=3); ns $P \leq 0.1529$.

4.4.5.3 Estimating the expression level of the CRTC1-MAML2 fusion in the transient transfected normal salivary gland cells

As we were unable to generate stable cell lines, we reasoned that if the expression of the CRTC1-MAML2 fusion in transiently transfected cells was persistent this would enable us to perform functional assays. An early passage (P1) of normal human primary salivary gland cells (SLG1 and SLG2) was grown in a 6-well plate for 24 hours, cells were then transiently transfected with the pCDNA3.1+IREs GFP vector with or without the FLAG-Tag CRTC1-MAML2 insert. Cells were routinely examined for the expression of GFP, and total RNA was extracted 48, 96 and 144 hours after transfection. qPCR analysis was performed to determine the level of fusion transcript expression over time. The transfected cells maintained expression of GFP until day 6, however, the level of expression decreased with time. This was consistent with the CRTC1-MAML2 fusion transcript expression level which, by day 6, had reduced to almost half of that at 48 hours post-transfection (figure 4.24).

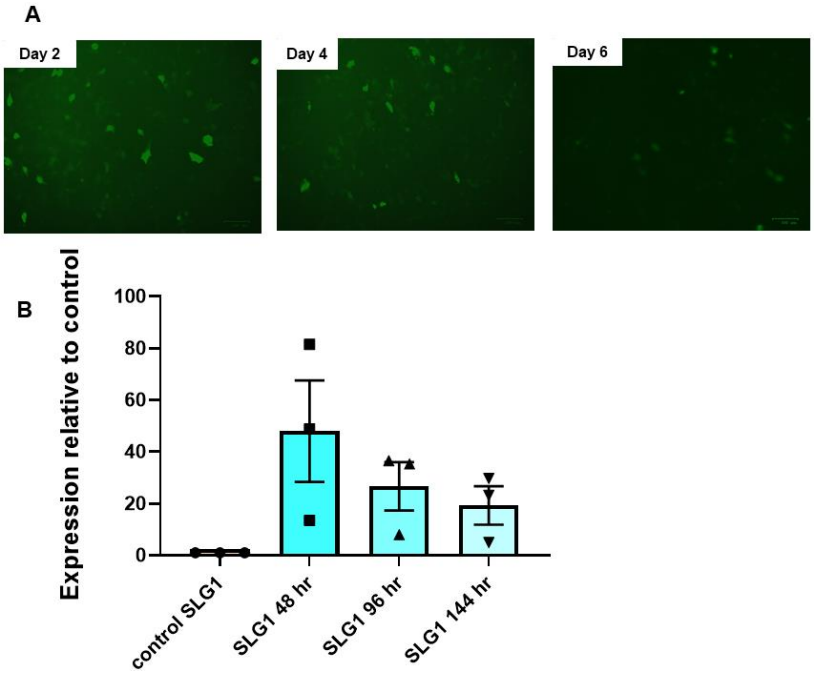


Figure 4.24 The expression pattern of the CRTC1-MAML2 fusion in transiently transfected normal salivary gland cells (SLG1). A) Representative Fluorescent microscopy images of SLG1 cells transiently transfected with the FLAG-Tag CRTC1-MAML2 pCDNA3.1+IREs GFP construct at three time points. GFP expression levels decreased with time. **B)** Quantitative-PCR analysis of transiently transfected SLG1 cells shows decreased transcript expression over time compared to an endogenous control.

4.4.5.4 Downstream effects of the CRTC1-MAML2 fusion gene on transiently transfected normal salivary gland cells

Previously, the CRTC1-MAML2 fusion was implicated in the activation of genes known to be involved in tumour development and progression (Wu *et al*., 2005; Chen *et al*., 2015). However, these experiments were conducted on HeLa cells or fusion positive MEC cells (NCI-H3118) targeting the transcriptional activation domain (TAD) in MAML2 of the translocation. In this study we used optimised transfection conditions for normal HuSL cells, transfected with the FLAG-tag CRTC1-MAML2 construct or empty vector control and analysed cells 48 hours post transfection. To study the effect of the CRTC1-MAML2 fusion gene on transiently transfected normal salivary gland

cells (HuSL), five different target genes known to be regulated by the cAMP signalling pathway were analysed (Wu *et al*., 2005). Exogenous induction of CRTC1-MAML2 fusion gene expression led to an upregulation of four of these genes (DUSP1, STC1, NR4A2, and ATF3) and down regulation of the TFF1 gene. A modest increase in DUSP1, STC1 and ATF3 was demonstrated alongside a much greater upregulation of the NR4A2 gene. A two-tailed t-test revealed a statistically significant difference (figure 4.25).

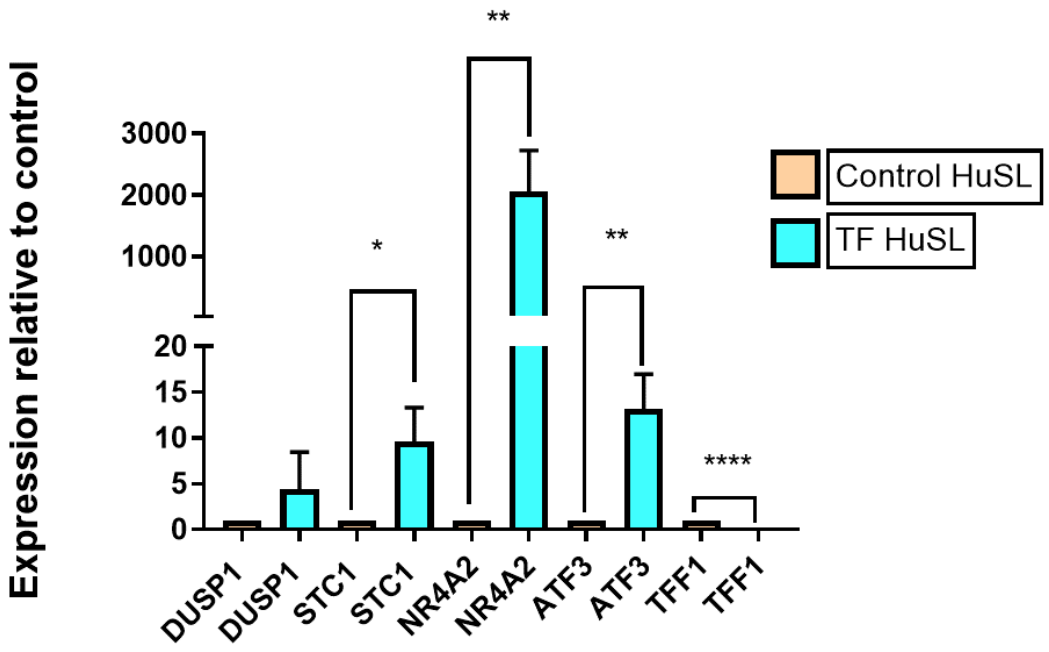


Figure 4.25 The CRTC1-MAML2 fusion gene induces expression of CREB transcriptional target genes. qPCR analysis shows the upregulation of four target genes involved in cAMP signalling pathways and down regulation of one candidate also involved in cAMP signalling in transiently transfected primary salivary gland cells (HuSL) compared to control cells. Cells were analysed 48 hours post transfection. Statistical analysis by two tailed Unpaired t Test (*P* value *=0.0157, **= 0.0061, **=0.0054, **** = 0.0001) (N=3, n=3)

4.5 Discussion

Chromosomal translocations are common in solid human malignancies in general, and in particular in salivary gland tumours, but the molecular mechanism of these translocations is complicated and has yet to be fully elucidated (Zheng, 2013). Synthesis of a recombinant protein via exogenous gene expression is an effective *in vitro* tool that has been extensively used to study the role of fusion proteins in tumourigenesis. In this study our aim was to generate a CRTC1-MAML2 fusion construct, to transfect into normal salivary gland cells, which would enable us to analyse the role of the novel fusion gene in MEC development and progression.

The vectors used in the generation of the fusion gene contained a human cytomegalovirus (CMV) promotor/enhancer, which permitted efficient and high-level expression of the recombinant protein in mammalian cells (Andersson *et al*., 1989, Boshart *et al*., 1985). A T7 promotor and priming site confers *in vitro* transcription in a sense orientation and primary sequencing through the insert. The vector also allowed for efficient transcription termination and polyadenylation of mRNA (Goodwin and Rottman, 1992). Two FLAG-Tag CRTC1-MAML2 constructs were made using two different mammalian expression vectors. Initially we used the pCR3.1 vector, lacking expression of GFP. The pCDNA+3.1IREsGFP vector construct facilitated the cell sorting process and confirmed successful transfection.

We generated a FLAG- tagged CRTC1-MAML2 expression clone by amplifying the fusion gene from MEC fusion positive cell lines, NCI-H292 and UM-HMC2. The FLAG Tag would allow us to detect the expression of the recombinant fusion gene using a specifically designed, high affinity monoclonal antibody. The FLAG-Tag was incorporated into the forward sequence of the primer used to amplify the fusion genes,

just downstream of the start codon. The CRTC1-MAML2 fusion insert contained a Kozak translation initiation sequence (ACC) just upstream of the start codon (ATG) to facilitate correct initiation of translation (Kozak, 1990).

The CRTC1-MAML2 construct was cloned for the first time by Tonon and colleagues (Tonon *et al*., 2003) and then in 2004 by the Enlund group (Enlund *et al*., 2004); these constructs have been used in the field by other scholars and researchers. In the present study we generated our own construct, as the previous constructs were not available to us, both of which were confirmed by nucleotide sequencing. In the database there are only two sequences available for the CRTC1-MAML2 fusion. The sequences, which belong to the previously mentioned groups, are under different accession numbers (AY040324.1 and AY186997.1), however, they are mostly identical (99%). We were able to confirm that our construct was equivalent to these sequences with a consensus of 99.8%. This minute difference (0.2%) between our construct and other scholars' construct was due the presence of a single nucleotide difference at the position 2921 with C to T change. However, this represents a synonymous substitution with both variants being translated to the same amino acid and so there would be no functional consequences.

In the cloning process we used two of the most appropriate MEC cell lines generated from patients. NCI-H292 is derived from a pulmonary MEC, whereas UM-HMC-2 is derived from a salivary gland MEC. To avoid PCR induced mutations, the full length MAML2 pCR4.1 clone was purchased (gene ID 84441).

Following the successful insertion of the full CRTC1-MAML2 fusion into the vectors, they were purified and used in transfection experiments. Initially, the CRTC1-MAML2 fusion was transiently transfected into human embryonic kidney (HEK293) and pulmonary adenocarcinoma (NCI-H647) cells, prior to transfection into primary salivary gland cells. HEK293 cells were chosen as they are regarded as a reliable lab tool, having previously been used extensively for the production of recombinant proteins and the study of protein structure and biological function. These cells are easy to grow, maintain and transfect, besides having high transfection efficiency associated with effective protein translation and processing (Wurm, 2004). NCI-H647 cells are also easy to grow and maintain in culture, and were chosen to optimise the transfection technique and to compare transfection efficiency with the primary cells.

It was necessary to optimise the transfection conditions to ensure efficient transfection and delivery of plasmid to cells whilst maintaining cell viability. Two transfection reagents were compared: FugeneHD (a non-liposomal, lipid-based reagent) and JetPRIME Polypus (based on a polymer formulation). The FACS results and the microscopic analysis of GFP expression indicated that JetPRIME, at a 1:2 (DNA: reagent) ratio, was superior to FugeneHD in effectively delivering the CRTC1-MAML2 plasmid into the cell lines as well as primary cells. A further advantage of using JetPRIME over FugeneHD is that the normal growth media, in the presence of serum, can be used throughout the process, thus ensuring cell viability. JetPRIME was selected as the transfection method of choice and we were able to adapt it for large-scale experiments.

Historically, transfecting primary cells has been challenging and problematic, especially with routine reagent-based techniques. Reported rates of transfection

efficiency are low unless viral-based or electroporation techniques are used (Hamm *et al*., 2002, Han *et al*., 2015, Staedel *et al*., 1994). In this study, however, we were able to successfully transfect human primary cells using JetPRIME with transfection efficiency rates of over 50% and continued cell viability.

The most common expression analysis tools previously used in this field are FLAG and MAML2 antibodies, targeting a FLAG-tagged CRTC1-MAML2 construct (Amelio *et al*., 2014, Chen *et al*., 2015, Chen *et al*., 2014, Chen *et al*., 2018, Komiya *et al*., 2006, Tonon *et al*., 2003, Wu *et al*., 2005). In the current study, protein expression analysis of cell lysate, by western blot, of transiently transfected NCI-H647 cells with a high affinity anti-FLAG monoclonal antibody demonstrated protein expression at the expected molecular weight of the fusion protein and this result was further supported through the use of a human MAML2 antibody identifying exogenous break apart MAML2 (recombinant fusion) from endogenous MAML2. The molecular mass of endogenous MAML2 is 125 KDa (GeneCards® The HUMAN GENE DATABASE, accessed in June 2020) which is consistent with the results of our western blot of wild type control cells (HEK293, NCI-H647). The transiently transfected cells, however, showed two distinct bands at ~125 KDa and around 115 KDa; break apart MAML2 loses 117 amino acids from exon 1. There are discrepancies in the literature in terms of the size of MAML2 protein, as previous findings have either reported the protein being150 KDa or have not indicated the protein size (Chen *et al*., 2018, Komiya *et al*., 2006). One key study, using samples from two MEC cases, confirmed expression of the CRTC1-MAML2 fusion protein to be approximately 110 KDa (Behboudi *et al*., 2006). Our western blot analysis showed *de novo* expression of the CRTC1-MAML2 fusion, in HEK293 and NCI-H647 cells, at the expected molecular weight of 115 KDa with both anti-FLAG and anti-MAML2 antibodies. This demonstrated that we had

successfully transiently transfected the NCI-H647 and HEK293 cells resulting in *de novo* fusion protein production of the predicted size.

CRTC1-MAML2 fusion protein expression in transiently transfected NCI-H647 cells was further studied using immunofluorescence with the anti-FLAG antibody; this allowed us to examine *in situ* expression and to localise the fusion protein to the nucleus. This result broadly supports the work of others (Tonon *et al*., 2003, Wu *et al*., 2005, Behboudi *et al*., 2006). In the study by Behboudi and co-workers (2006), transiently transfected HEK293 cells demonstrated localisation of the GFP-conjugated CRTC1-MAML2 fusion construct in homogeneously distributed fine nuclear granules. In addition, their immunohistochemical study with a polyclonal fusion antibody (LTR12) showed nuclear staining in the majority of the tumour cells (mucous and intermediate) in three fusion positive cases. However, the immunohistochemical study involved only a small minority of cases, images demonstrating expression by all cells within the tumours was not presented and the use of this antibody has not yet been validated in other studies. Similar immunofluorescence studies were carried out by Tonon *et a*l. (2003) and Wu *et al.* (2005), following transient transfection of monkey kidney fibroblast cells (COS7) and human osteosarcoma cells (U2OS), respectively. The fusion construct, detected using GFP-tagged and FLAG-tagged antibodies, indicated distinct nuclear staining.

Primary human salivary gland cells were used in the current study in order to replicate the key findings in the literature, with regard to the functional impact of *de novo* expression of the fusion construct, which had been carried out on immortal cell lines; our study was designed to increase the translational value of our knowledge. Tonon and colleagues (2003) found that transient transfection of the CRTC1-MAML2 fusion

construct into immortalised rat kidney epithelial (RK3E) cells transformed the cells, allowing colonies to form, whereas transfection with an empty vector or full MAML2 construct did not induce any foci formation after 3 weeks. Our study did not show such an effect.

Komiya and colleagues (2006) showed that interfering with endogenous CRTC1-MAML2 translocation via hairpin RNAi in NCI-H292 and NCI-H3118 (fusion positive, parotid origin MEC) cells caused a significant reduction in the growth and survival of MEC fusion positive cells and xenograft growth in *in vivo* models. In the same study, RK3E cells were transfected with either a CRTC1-MAML2 construct, empty vector, full MAML2 construct, or break a part MAML2 (exon 2-5) and the transfected and transformed RK3E cells were then injected into nude mice. The authors detected a palpable tumour mass after 14 days in all mice receiving the cells transfected with the fusion gene but no tumours developed in the 30 control mice. The tumours were excised, total RNA extracted and the CRTC1-MAML2 detected. This data suggested that sustained expression of the CRTC1-MAML2 fusion is essential for tumour growth *in vivo*, however, in the absence of any histological or cytological analysis it is not possible to confirm if these tumours resembled MEC in any way. It is important to remember that the cells transfected and injected into the mice were normal kidney cells from a rat and so there are concerns over whether they are an appropriate tool for this experiment.

Coxon *et al.* (2005) demonstrated that generating two in frame deletions in residues 14 and 20 within the CREB binding domain in exon 1 of the CRTC1-MAML2 translocation resulted in complete loss of foci formation in the RK3E cells. A small in-frame deletion in residues 77-118 within MAML2 of the CRTC1-MAML2 translocation,

however, resulted in no effect on foci formation while a larger in-frame deletion in residues 703-838 of MAML2 did lead to complete loss of foci formation. This suggests that the CREB binding domain within exon 1 of CRTC1 plays an essential role in epithelial transformation and the CRTC1-MAML2 fusion may play a role in activating target genes. Furthermore, microarray analysis conducted on HeLa cells transiently transfected with the CRTC1-MAML2 fusion identified downstream targets which were differentially regulated. Genes regulated by cAMP signalling pathway, such as ATF3, NR4A2 and stanniocalcin, were among the most upregulated genes, with this result being validated via RT-PCR and western blotting. Downstream target genes included those implicated in multiple cellular processes such as cell growth, differentiation and apoptosis, which could be responsible for epithelial cells transformation (Wu *et al*., 2005).

pSuperRetro-GFP/Neo vector-based shRNAs were used by Chen *et al.* (2015) to target the MAML2 transcriptional activating domain (TAD) in the CRTC1-MAML2 fusion gene and MAML2 in fusion positive MEC cells (NCI-H3118) and MAML2 in fusion negative cells (HSY). Gene expression profiles of depleted cells suggested that CRTC1-MAML2 regulated gene expression in both CREB- dependent and independent ways. However, without targeting the CRTC1-MAML2 fusion junction it is difficult to confirm the precise role for the CRTC1-MAML2 fusion in tumourigenesis. In addition, a significant limitation of this study is that the HSY cells, which were used as fusion negative, normal salivary gland cells, were recently disqualified after authentication results confirmed they are in fact HeLa cells (Lin *et al*., 2018).

To date, no previous studies have examined or performed analysis on the effect of the fusion protein on normal salivary gland cells. Initially we performed our

experiments and analysis on sublingual cells and after transient transfection with the fusion gene we harvested DNA, RNA and cell lysate. Our quantitative RT-PCR analysis indicated a statistically significant increase in the expression of the CRTC1-MAML2 fusion transcript compared to control, non-transfected cells. DNA and protein analysis indicated the expected amplicon products and protein of the predicted size. To enhance these findings, we replicated our experiments using a panel of primary human salivary gland cells from 5 donors and we validated our results with custom designed TaqMan probes. All expression patterns of the CRTC1-MAML2 fusion transcript were in line with our previous findings and indicated statistically significant upregulation of the CRTC1-MAML2 transcript (*P* value **= 0.0024, *** = 0.0002, **** = 0.0001).

The use of positive control cells (NCI-H292 and UM-HMC2) provided us with a clear indication of the expression pattern of the endogenous CRTC1-MAML2 fusion, however, our data suggested that an exogenous CRTC1-MAML2 fusion transcript resulted in higher expression levels. This could be due to the very high transfection efficiency of the primary cells which corresponded with significant fold changes in expression.

Generation of stable cell lines permanently expressing the CRTC1-MAML2 fusion protein was unsuccessful despite the use of a number of cell lines (HEK293, NCI-H647) and human primary salivary gland cells from a number of donors (HuSL1 and 2, SLG1 and 2, SMG and HPG). Transfected NCI-H647 and HEK293 cells were cell sorted and colonies generated from single cells, however, no gene or protein expression was detected. The primary cells were also effectively transfected and either FACS sorted or selected in the presence of preoptimised concentrations of

G418. Thus, any surviving cells should have carried and yielded the fusion gene. Primary cells did not produce colonies from single cells, despite several attempts with modifications through the addition of conditioned media. We also attempted to expand cells from polyclonal populations with some success but there was no chromosomal integration of the fusion gene. The reason for this is not clear but could be related to the fact that plasmid integration occurs randomly in the cell genome, a process that is site dependent and has key effects on stable cell generation. Integration into an inactive site of the host chromosome could prevent gene expression and lead to failed transfection, whereas integration into an active site could result in gene integration, but this may not be sufficient to ensure prolonged and permanent gene expression. Another possible explanation is that mammalian cells, and especially primary cells, could be silencing (inactivating) transgene expression rapidly after integration into the host genome.

Primary cells are mortal and so cannot divide or grow indefinitely and thus preventing them from senescing was vital during transfection and expansion. FACS sorted cells, and those undergoing G418 selection, were consistently maintained with 10 µg/ ml of Y-27362, an efficient Rho-Kinase inhibitor (ROCK) that is thought to prolong the life span of normal primary cells such as keratinocytes and other epithelial cells, including adult salivary gland cells, by promoting cell proliferation, inhibiting cell apoptosis and enhancing cell expansion (Coleman *et al*., 2001). The reagent was removed during cell analysis. The exact mechanism of action of ROCK inhibitors is still not clear; however, it has been reported that they upregulate DNA replication and cell division without causing mutations to the cells and the process is completely reversible (Holm *et al*., 2013). In addition, ROCK signalling is considered a key mechanism in salivary gland differentiation and organisation. This claim was supported by a previous study in which salispheres grown from adult mouse submandibular glands progenitor cells

were expanded with Y27632 and had an increased capability for developing proacinar salivary organoids (Daley *et al*., 2012, Daley *et al*., 2009, Daley *et al*., 2011).

There are several factors that could be implicated in the complexity of generating stable cell populations. Some have been previously mentioned, such as the integration site or silencing by the host cells following successful vector integration. One strategy that could help to overcome such obstacles would be linearisation of the plasmids as this could increase the likelihood of vector integration into the cell chromosome and accelerate the stable integration process (Wurm, 2004). Linearisation of the vector can be achieved with restriction enzyme digestion upstream of the promoter and gene of interest and absence in the insert sequencing. In this study we did linearise our constructs using *BgI II* endonuclease, however, time restrictions meant we have not yet been able to test the linearised product.

Strategies which could allow us to fully determine whether the plasmid had been successfully integrated into the host genome include those designed to overcome the host cell genome preventing gene expression or causing protein degradation. For example, butyrate could be used to inhibit silencing of the gene by blocking the deacetylation of histones (HDACs) while the proteasome inhibitor (MG 132) reduces protein degradation by mammalian cells in a reversible and potent way (Steliou *et al*., 2012, Wang *et al*., 2019).

Transient gene expression is widely used in academic research and preclinical studies. Although DNA expression and protein yield appear finite, scale-up approaches could lead to greater expression and yields. In this study we continued

our transient transfection approach, studying in detail the efficacy and longevity of *de novo* gene expression, and were able to show that gene expression gradually decreased over time such that 6 days after transfection expression levels were reduced by almost two thirds. As a result, functional assays were carried out on cells which had been transfected no longer than 48 hours prior. This experimental model allowed us to answer the third question of this study and assess the downstream effects of CRTC1-MAML2 transcript expression on normal primary salivary gland cells. In particular, quantitative PCR analysis showed that the CRTCT1-MAML2 transcript caused an upregulation of four genes implicated in cAMP signalling. These results are mostly in agreement with a previous study by Wu *et al.* (2005) in which the N-terminal of the CREB binding domain of the fusion protein was shown to regulate and mediate transforming activity via upregulation of the expression of genes such as NR4A2, ATF3, DUSP1, TFF1, NDRG1 and STC1, which are involved in the normal cell cycling (Wu *et al*., 2005). Our study showed an increase in the expression of NR4A2, ATF3, DUSP1 and STC1 transcripts.

Previously published studies have also shown that the Notch-target genes, HES1 and HES5, were implicated in cell fate and were upregulated in MEC cell lines and cells having *de novo* expression of CRTC1-MAML2 fusion (Tonon *et al*., 2003 and Behboudi *et al*., 2006). The downstream target genes of the cAMP (FLT1, NR4A2 and CLDN7) and Notch (HES1 and HES5) pathways were analysed by RT-PCR in a series of *ex vivo* MEC tissue samples. The FLT1, NR4A2 and HES1 genes were not expressed or were expressed at a very low level in fusion positive MEC; in contrast, the HES5 was highly expressed in fusion positive MEC samples (Behboudi *et al*., 2006). The low or lack of expression of CREB target genes in the previous study is contrary with our findings and those from others, however, in our study we did not analyse the Notch target genes, which makes comparison difficult to Behboudi's data.

The CRTC1-MAML2 fusion has also been shown to induce CREB-independent activities through interactions with AP-1 and MYC nuclear factors (Amelio *et al*., 2014, Canettieri *et al*., 2009). The CRTC1-MAML2 fusion interacted with MYC and this complex was shown to be essential for cell transformation and expansion of RK3E epithelial cells in a colony focus assay. Microarray analysis of CRTC1-MAML2 transfected HEK293 cells also showed that MYC target genes were regulated by the CRTC1-MAML2 fusion (Amelio *et al*., 2014). Moreover, the AP-1 complex contains numerous transcriptional factors which are controlled by mitogenic stimuli and are implicated in key biological and pathological activities like proliferation and differentiation. AP-1 has a synergistic effect with CRTC1 in controlling cell transformation, and the depletion of AP-1 in MEC cells (NCI-H292) led to disturbance in cell proliferation and transformation. Furthermore, interference of AP-1 in RE3K cells transfected with the CRTC1-MAML2 fusion caused a reduction in the number of foci formed (Canettieri *et al*., 2009). Thus, the CRTC1-MAML2 fusion could be implicated in AP-1 and MYC dependent cellular transformation and regulating signalling driving carcinoma development.

Four of the five candidate genes with altered expression following transient transfection of CRTC1-MAML2 fusion gene into normal sublingual salivary gland cells (HuSL) were upregulated. DUSP1, dual-specificity-phosphatase1, has previously been shown to be involved in cell proliferation, differentiation, transformation, inflammation, cell stress, cell arrest and apoptosis (Owens and Keyse, 2007). Increased expression of DUSP1 transcript has been described in a range of epithelial malignancies, including prostate, pancreatic, gastric, breast, lung, colon and bladder cancers (Shen *et al*., 2016). *In vivo* work conducted on a nude mouse model showed

that knocking down DUSP1 in pancreatic cancer cells reduced tumorigenesis by delaying tumour formation; additionally, a lower number of colonies was formed in clonogenic assays (Liao *et al*., 2003). However, the expression of DUSP1 has also been shown to be inversely related to cancer progression, with a loss of DUSP1 in higher histological grade tumours (Shen *et al*., 2016).

Stanniocalcin-1 (STC1) is an autocrine/paracrine factor secreted by wide range of tissues, with the STC1 protein able to control the transport of intestinal and renal calcium and phosphate, cell metabolism, or cellular calcium/phosphate homeostasis. Recent evidence supports the association of STC1 expression and carcinogenesis in a range of cancers (Leung and Wong, 2020). The association with tumour prognosis, however, is contradictory. Our study showed a statistically significant increase in expression of STC1 in the primary cells transfected with the CRTC1-MAML2 fusion construct. A study by Chen *et al.* (2015) in which the expression of the CRTC1-MAML2 fusion in a MEC fusion positive cell line (NCI-H3118) was knocked down resulted in STC1 downregulation (Chen *et al*., 2015), corroborating our results.

Importantly, our results also indicated an upregulation in the nuclear receptor 4 NR4A2 (Nurr1), a transcriptional factor nuclear receptor and member of a steroid nuclear hormone superfamily previously shown to be over expressed in a range of diseases including gastrointestinal cancer (Han and Cao, 2012). The CRTC1-MAML2 fusion has previously been reported to be a nuclear protein, functioning as a transcriptional co-activator (Tonon *et al*., 2003), which could explain the significant upregulation of the NR4A2 transcript seen in our experiments. Our results are in agreement with the results from Wu *et al.* (2005) where the microarray, RT-PCR and protein analysis showed upregulation of the NR4A2 transcript and protein in the transiently transfected HeLa with the CRTC1-MAML2 (Wu *et al*., 2005). However, this

result contradicts other results from Behboudi *et al.* (2006) who found that the fusion positive MEC samples showed lower expression of the NR4A2 transcript.

ATF3, activating transcriptional factor 3, is a member of the cAMP family; expression has been correlated with breast cancer development and metastasis, via enhancing cell proliferation, invasion and protecting the cells against apoptosis. However, conflicting data has suggested that ATF3 can act as a cancer suppressor by promoting the apoptosis and reducing invasion and metastasis (Thompson *et al*., 2009). Our results showed upregulation of ATF3 transcription in the HuSL cells transfected with the CRTC1-MAML2 fusion.

Surprisingly, our study showed slightly downregulated expression of the TFF1 transcript. This contradicts results from a previous study that used HeLa cells, where the cells showed a 10× fold increase in the expression level of the TFF1, when transfected with the CRTC1-MAML2 fusion (Wu *et al*., 2005). In addition, Pelden *et al* (2013) studied breast cancer cell lines and reported that TFF1 enhanced the inhibition of apoptosis and promoted tumourigenesis. TFF1 was the first described member of the Trefoil factor family after expression was reported in breast cancer cell line (MCF7) (Tomasetto and Rio, 2005). TFF1 interacts with MUC5A mucin to form a polymer to protect the epithelial component of various parts of the body including the conjunctiva, gastric mucosa and tracheal mucosa (Ruchaud-Sparagano et al., 2004). It is still not clear why TFF1 expression was decreased in normal primary salivary gland cells after introduction of the CRTC1-MAML2 fusion gene.

4.5.1 Conclusion

In summary, CRTC1-MAML2 fusion constructs were successfully generated in mammalian expression vectors. The CRTC1-MAML2 fusion was effectively expressed in human primary salivary gland cells and mammalian cell lines. Attempts were made to generate stable cell lines expressing the CRTC1-MAML2 fusion; however, neither gene nor protein expression could be detected. Transient transfection of normal human primary salivary gland cells with the CRTC1-MAML2 fusion led to significant upregulation of four target genes which are involved in many critical cellular processes such as cell growth and differentiation. This suggests that the CRTC1-MAML2 fusion may interfere with normal cell functioning, leading to cell transformation and subsequently driving tumourigenesis. As it was not possible for us to establish a stable cell line, with which we could have further investigate the role of the fusion gene, we continued our functional studies with transiently transfected cells. These experiments are described in detail in chapter 5.

4.5.2 Future work

- Perform stable transfection experiments with linearised plasmids.
- Optimise cell treatment with butyrate and MG132 to enhance the stable transfection experiments, promote gene integration and protein expression.
- Conduct RNA sequence analysis experiments on transfected primary human normal salivary gland cells with *de novo* CRTC1-MAML2 expression to validate key studies conducted on mammalian cancer cell lines.

Chapter 5

The Effect of the CRTC1-MAML2 Fusion Protein on Cell Behaviour

5. Chapter 5

5.1 Introduction

In order to successfully determine the role of a fusion gene in tumourigenesis it is necessary to either completely silence the gene in expressing cells or to introduce *de novo* expression in non-expressing cells prior to carrying out functional assays. Our strategy was to introduce the gene into normal human salivary gland cells and then compare function with non-transfected normal cells. As outlined in chapter 4 we were unsuccessful in our attempts to create a stable cell line following transfection of the fusion gene into normal salivary gland cells or into non-expressing cell lines. An alternative to the production of stable cell lines would be to carry out large-scale transient transfections and to use these cells for functional assays.

It has been reported, and previously mentioned in this thesis, that the CRTC1-MAML2 fusion protein is implicated in tumour initiation and progression via 1) an initiation of clusters of epithelial in RK3E cells *in vitro* (Tonton *et al*., 2003), 2) the development of subcutaneous masses following the introduction of fusion transformed RK3E to immunocompromised mice (Chen *et al*., 2014); and 3) a significant decrease in cell growth and survival following silencing, *in vitro* and *in vivo* (Chen *et al*., 2014).

Based on our results discussed in chapter 3, not all cells constituting a MEC harbour the translocation, therefore, performing a functional analysis on transiently transfected cells may more closely replicate the *in vivo* status of a MEC tumour *in situ*. In chapter 4, we successfully demonstrated exogenous expression of transiently

expressed CRTC1-MAML2 fusion protein in human immortalised cell lines and human primary salivary gland cells (HuSL). In this chapter we demonstrate the utility of optimised, large-scale, transient transfection to study the effect of the CRTC1-MAML2 on basic cell functions associated with tumourigenicity, survival and metastasis. Two cell lines, HEK293 and NCI-H647, were selected as the data we presented in chapter 4, demonstrates that effective transfection of these cells can be accomplished reproducibly.

5.2 Aim and objectives

The aims of this chapter, therefore, were to elucidate the function of the fusion protein, assess the effect of ectopic expression on cell behaviour and how this is related to tumour survival and progression. Transient transfection experiments were used to induce the ectopic expression of the fusion protein and the downstream effect of the CRTC1-MAML2 fusion protein was monitored using standard functional assays.

5.3 Materials and Methods

The relevant materials and methods utilised in this chapter are detailed in chapter 2, under headings 2.2, 2.4, 2.5, 2.6 and 2.7 as indicated below:

- HEK293 and NCI-H647 cells were cultured as outlined in section 2.2.3.
- Transient transfection of HEK293 and NCI-H647 cells were performed as outlined in section 2.4.2.
- Successful transfection was confirmed by gene expression (RT-qPCR) as outlined in sections 2.5.2, 2.5.3 and 2.5.6.

- Successful transfection was also confirmed by protein expression, (western blotting) which was performed as outlined in sections 2.6.1, 2.6.3, 2.6.4, 2.6.5, 2.6.6 and 2.6.7.
- The transiently transfected cell lines were subjected to conventional cell functional assays including proliferation, migration, clonogenic and invasion, assays, which were conducted as outlined in sections 2.7.1-2.7.4.

5.4 Results

5.4.1 Determine the expression of the CRTC1-MAML2 fusion in transiently transfected cells

Prior to conducting the functional assays on HEK293 and NCI-H647 cells, it was crucial to confirm exogenous CRTC1-MAML2 transcript and protein expression. HEK293 and NCI-H647 cells were transiently transfected with pCDNA3.1+IREs GFP, with either the FLAG-Tag CRTC1-MAML2 insert or an empty vector, using the JetPRIME transfection method described in section 2.4.2. Forty-eight hours post-transfection the expression level was determined, with functional tests being initiated simultaneously. Our microscopic analysis indicated that both cell lines were efficiently transfected and after successfully detecting mRNA and protein (figure 5.1 and 5.2) the functional assays were continued. Our protein analysis was validated by two antibodies, anti-FLAG and anti-MAML2, with bands of the expected protein size, 115 KDa with the FLAG antibody and 115 KDa with the MAML2 antibody (a break apart in the lower band) being detected (figure 5.1 C and 5.2 C).

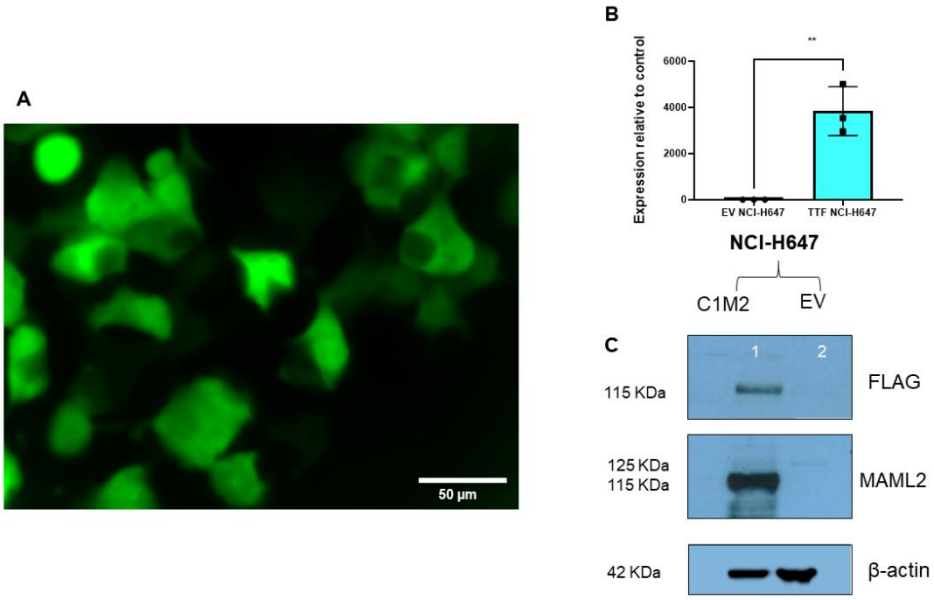


Figure 5.1 CRTC1-MAML2 fusion transcript and protein expression in NCI-H647. A) A representative image of GFP expression of transiently transfected NCI-H647. **B)** CRTC1-MAML2 transcript upregulation in transiently transfected NCI-H647 cells versus empty vector control. Statistical analysis was carried out using two tailed Unpaired t Test (*P* value ** = 0.0033) (N=3, n=3). **C)** Cellular protein was extracted from transfected NCI-H647 and expression of recombinant CRTC1-MAML2 in the cell lysate was analysed by Western blotting using anti-FLAG and anti-MAML2 antibodies. Lane numbers represent the following: 1- NCI-H647 cells transfected with - FLAG-tag CRTC1-MAML2 pCDNA3.1+IREs GFP vector (NCI-H647 C1M2), 2- NCI-H647 cells transfected with the pCDN3.13.1+IREsGFP empty vector (NCI-H647 EV). β-actin antibody was used as a protein loading control.

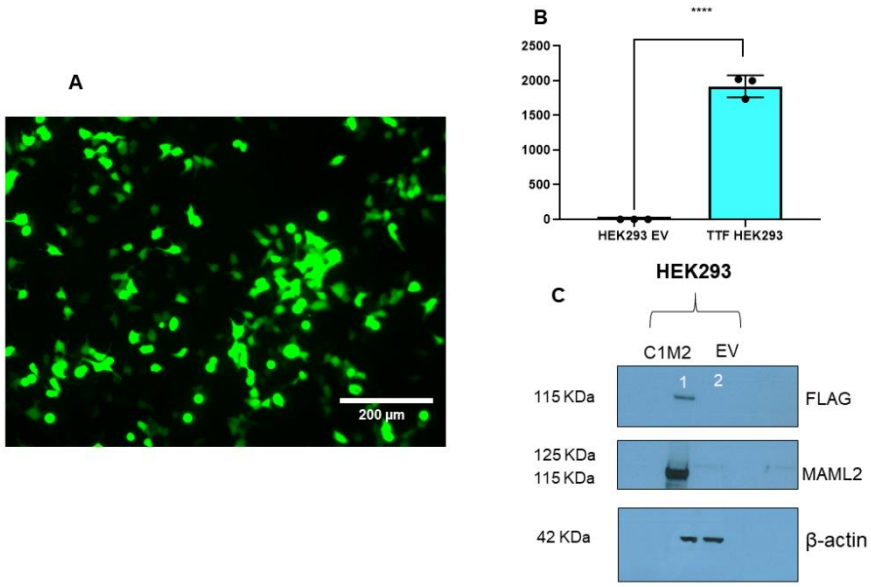


Figure 5.2 CRTC1-MAML2 fusion transcript and protein expression in HEK293 cells. A) A representative image of GFP expression of transiently transfected HEK293. **B)** CRTC1-MAML2 transcript upregulation in transfected HEK293 cells versus empty vector control. Statistical analysis was carried out using two tailed Unpaired t Test (*P* value **** = <0.0001) (N=3, n=3). **C)** Cellular protein was extracted from transfected HEK293 and expression of recombinant CRTC1-MAML2 in the cell lysate was analysed by Western blotting using anti-FLAG and anti-MAML2 antibodies. Lane numbers represent the following: 1- HEK293 cells transfected with - FLAG-tag CRTC1-MAML2 pCDNA3.1+IREs GFP vector (HEK293 C1M2), 2 – HEK293 cells transfected with the pCDN3.13.1+IREsGFP empty vector (HEK293 EV). β-actin antibody was used as a protein loading control.

5.4.2 Effect of CRTC1-MAML2 fusion protein on cell migration and invasion

After confirming the exogenous expression of the CRTC1-MAML2 fusion transcript and protein we proceeded to study the effect of the fusion portion on cell behaviour. The ability of the cells to migrate was assessed by performing a scratch assay. To ensure that the cells were only migrating and not proliferating to fill the gap, a low dose of Mitomycin C, a cell cycle arrest antagonist (Grada *et al*., 2016), was added

to the cells 2 hours before commencing the assay, and after 2 hours the media was replaced with fresh media.

HEK293 cells were excluded from the assay after three unsuccessful attempts at gap creation. In order to initiate the scratch assay, the cells needed to be at 100% confluent, but as HEK293 cells are always loosely adherent, when they reach 70-80% confluency, they routinely begin to detach from the bottom of the flask and float in the media. It is therefore not possible to create a scratch wound with these cells and so they had to be excluded from the study.

Transiently transfected NCI-H647 cells took 32 hours to close the gap, whereas with the control cells 26% of the wound remained open at this same time point (figure 5.3 A and B). The exogenous expression of the CRTC1-MAML2 fusion protein thus had a statistically significant effect on cell migration. (*** P= 0.0004). Treatment with Mitomycin C prior to the test confirmed that the cells had truly migrated and not proliferated to fill the gaps.

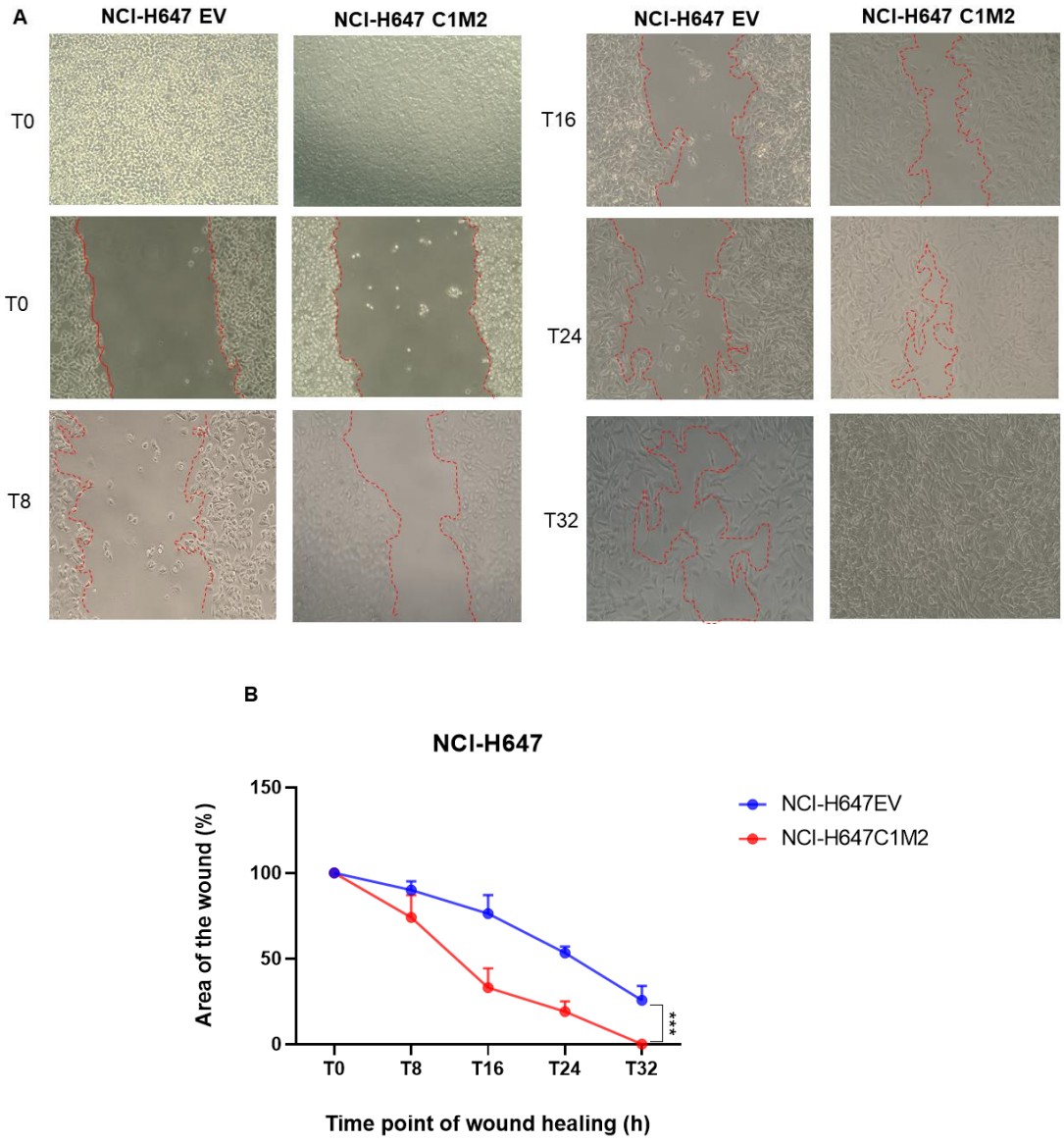


Figure 5.3 Wound healing assay. A) Representative bright field images of a scratch assay with NCI-H647 transiently transfected with the pCDNA3.1+IREs GFP vector with FLAG-Tag CRTC1-MAML2 insert (C1M2) or empty vector (EV). Exogenous expression of the CRTC1-MAML2 fusion protein accelerated wound closure. After 32 hours the transiently transfected NCI-H647 cells had achieved complete wound closure but a gap remained in the control cells. **B)** Representative graph showing quantification of the migration assay and indicating that the exogenous expression of the CRTC1-MAML2 fusion protein had a statistically significant effect on the cell migration compared to the control group (*** *P*= 0.0004). All data are representative of experiments performed in triplicate and are presented as mean ±SD; data was normalised to T0 and mean area for each group represented in percentage.

To investigate the effect of the CRTC1-MAML2 fusion protein on the invasive behaviour of cells, an 8 μm pore transwell invasion assay was performed. Two coating materials, Matrigel and Fibronectin, were used separately as a basement membrane matrix, in initial experiments with HEK293 cells. Cells were cultured in 1% serum (to prevent cell proliferation) and allowed to invade the matrix, moving towards media containing 10% serum, for 48 hours. The non-invading cells were swabbed away, and the invading cells fixed and stained with H&E. The analysis was carried out using the cell counter function on Image J software. 7 replicates were performed, images of each membrane were captured at 400× magnification using a light microscope (OLYMPUS BX51), the invaded cells counted, average score calculated for each membrane and the values used for statistical analysis.

The Matrigel coated inserts allowed more accurate analysis than the fibronectin coated inserts, due to the better image quality and our ability to differentiate between the pores and invaded cells in the quantification process. Therefore, the invasion assay for the NCI-H647 cells was only performed on pre-coated Matrigel inserts. The results indicated that the CRTC1-MAML2 fusion protein significantly increased the invasion capacity of both the HEK293 and NCI-H647 cells after 48 hours (P= <0.0017 and 0.0012) (figure 5.4 and 5.5).

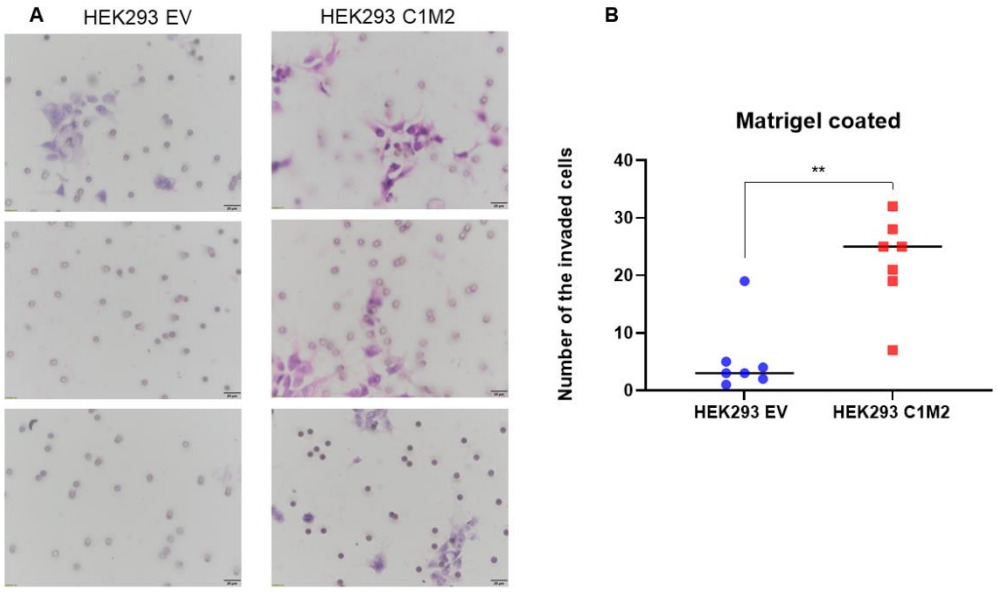


Figure 5.4 Invasion assay of the HEK293 cells. The invasion of CRTC1-MAML2 (C1M2) fusion transfected cells compared to empty vector (EV) negative control cells was analysed using Matrigel-coated Transwells, whereby cells were left to invade for 48 hours. A) Representative images of invading HEK293 cells following transfection with CRTC1-MAML2 (right panel) or empty vector control (left panel). Scale bar at 20 µm. B) Quantitative analysis revealed that the transfected cells were effective in increasing the cells' invasive capacity (*P*= 0.0017).

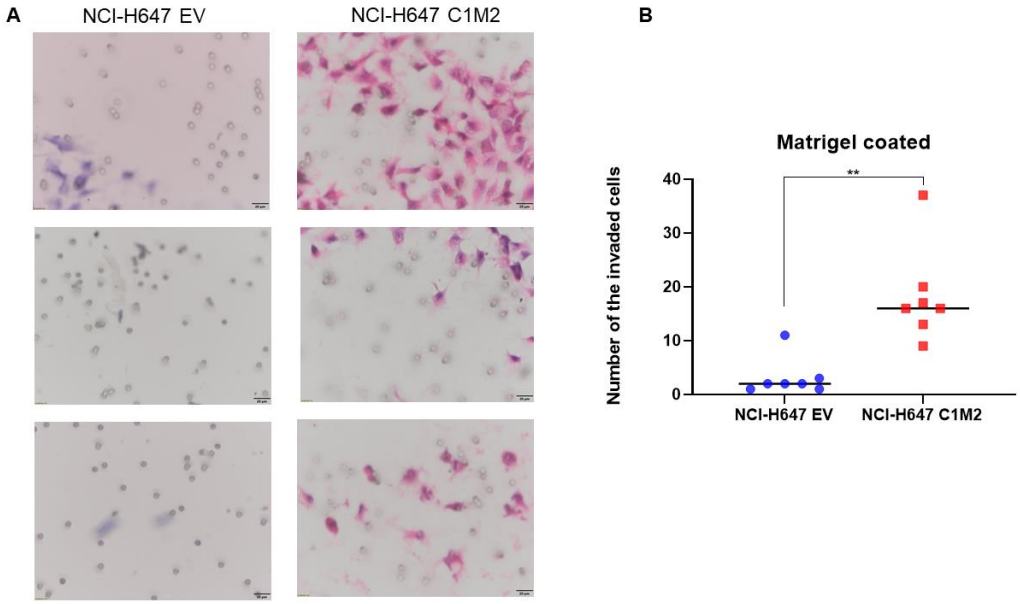


Figure 5.5 Invasion assay of the NCI-H647 cells. The invasion of CRTC1-MAML2 fusion transfected cells (C1M2) compared to empty vector (EV) negative control cells was analysed using Matrigel-coated Transwell, whereby cells were left to invade for 48 hours. A) Representative images of invading NCI-H647 cells following transfection with the CRTC1-MAML2 (right panel) or empty vector control (left panel). Scale bar at 20 µm. B) Quantitative analysis revealed that the transfected cells were effective in increasing the cells' invasive capacity (*P*= 0.0012).

5.4.3 Effect of CRTC1-MAML2 fusion protein on cell proliferation and survival

Previous data has shown that the CRTC1-MAML2 fusion protein, when it is exogenously expressed in RK3E cells, impacts cell proliferation as measured by the ability to develop colonies (Tonon *et al*., 2003). In this study two assays were used to study the effect of the CRTC1-MAML2 fusion protein on cell proliferation; colony formation and viability assessed through a clonogenic assay and an EdU cell proliferation assay.

The EdU proliferation assay was analysed using two approaches: fluorescence microscopy and flow cytometry. No differences in proliferation were observed between the CRTC1-MAML2 transiently transfected cells and their counterpart controls in either of the analyses (figures 5.6-5.9).

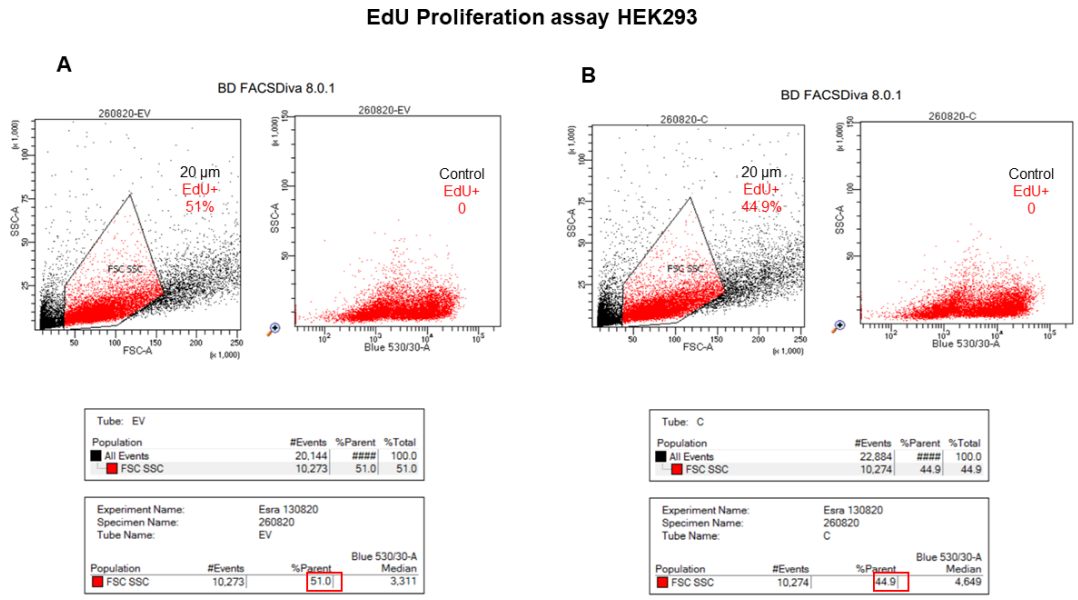


Figure 5.6 Dot plot of EdU-488 proliferation assay. HEK293 cells transiently transfected with the pCDNA3.1+IREs GFP vector with FLAG-Tag CRTC1-MAML2 insert (C) or empty vector (EV) were incubated with 20 µm EdU for 4 hours (right panels A and B) or 0 µm EdU control (left panels A and B). Images were acquired on a BD™ LSR II cytometer with cells excited using a 488 nm laser and data analysed using FACSDiva 8.0.1. The percentage of gated cells is highlighted.

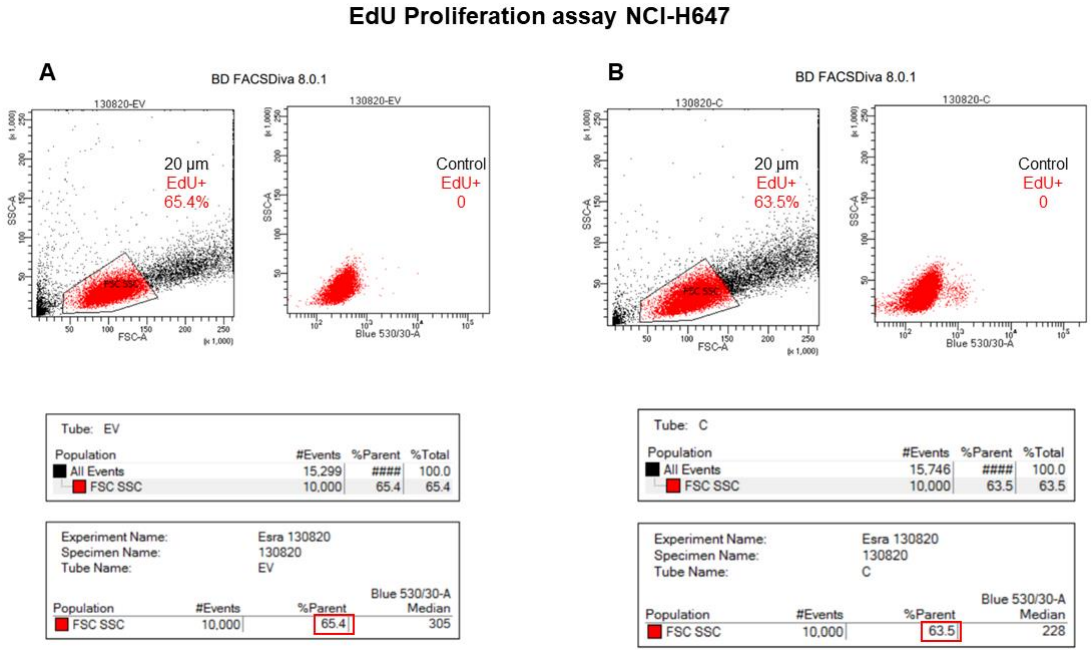


Figure 5.7 Dot plot of EdU-488 proliferation assay. NCI-H647 cells transiently transfected with the pCDNA3.1+IREs GFP vector with FLAG-Tag CRTC1-MAML2 insert (C) or empty vector (EV) were incubated with 20 µm EdU for 4 hours (right panels A and B) or 0 µm EdU control (left panels A and B). Images were acquired on a BD™ LSR II cytometer with cells excited using a 488 nm laser and data analysed using FACSDiva 8.0.1. The percentage of gated cells is highlighted.

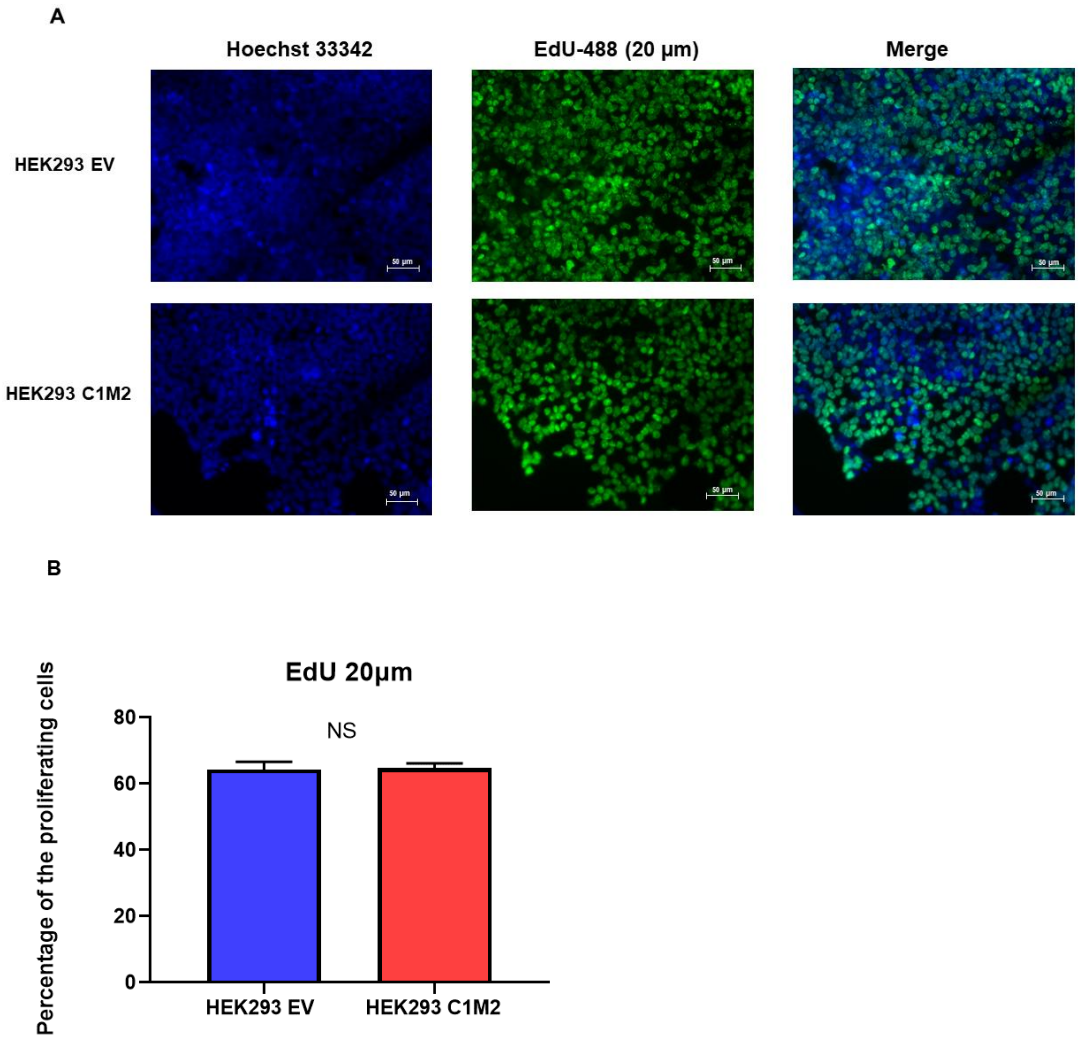


Figure 5.8 EdU staining of proliferating cells. A) Representative fluorescence microscopy images of HEK293 cells transiently transfected with CRTC1-MAML2 (C1M2) and pCDNA3.1+IREsGFP empty vector (EV) (1.5 ×10^5 cells/ well in 24 plate) and incubated with 20 µm EdU for 4 hours. Cells were analysed using a ZEISS AxioCam MRm2 fluorescent light microscope (Carl Zeiss, UK) and imaged using Image AxioVision Rel.4.8 image software at 20× magnification power. DNA (Blue) was stained with Hoechst 33342 (ab145597). Green cells are EdU positive. B) The proliferation of HEK293 cells transiently transfected with CRTC1-MAML2 and empty vector negative control cells was analysed by counting the number of EdU positive cells and dividing with total number of cell nuclei stained with Hoechst 33342 (ab 145597) using the cell counter plugin in image J. Results represent the mean ± SD (N=3) NS P = 0.8166. Scale bar at 50 µm.

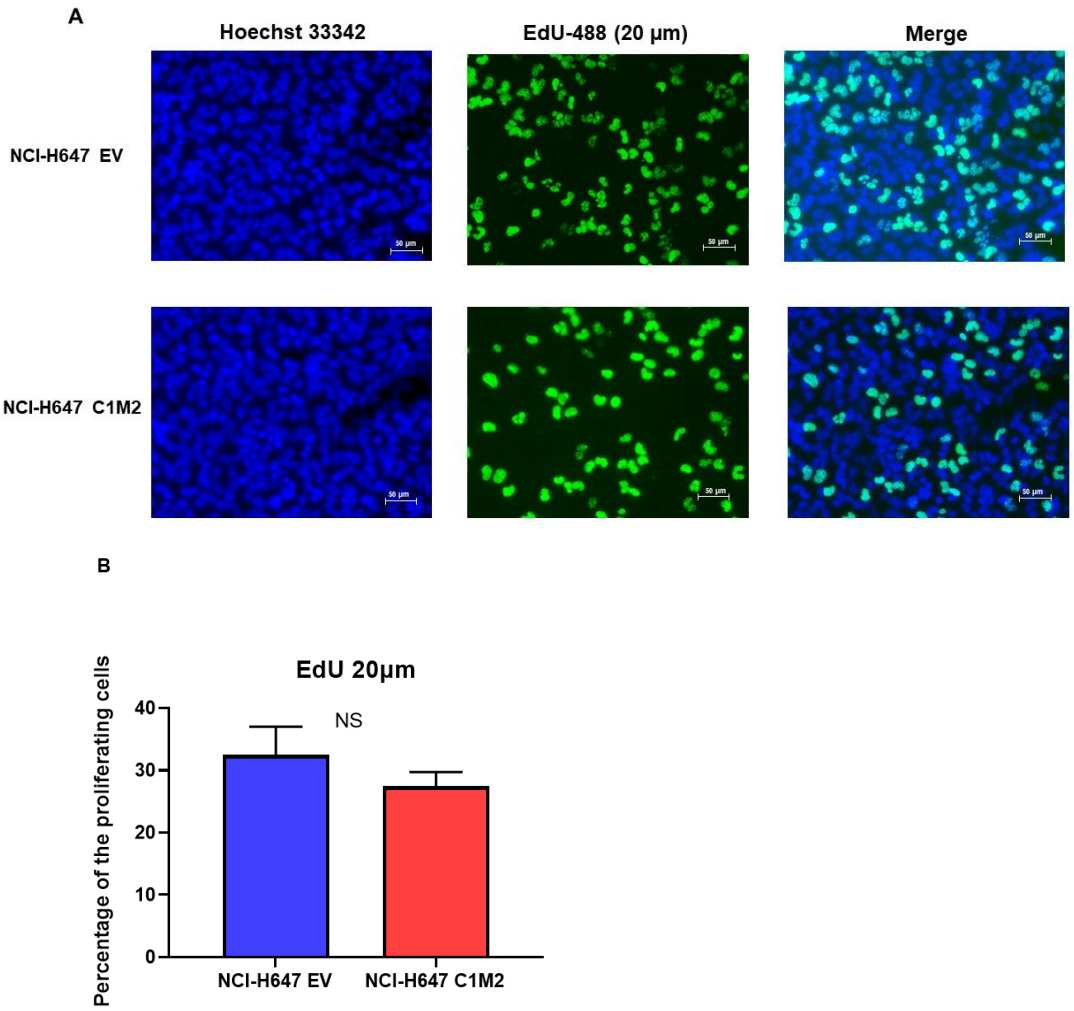


Figure 5.9 EdU staining of proliferating cells. A) Representative fluorescence microscopy images of NCI-H647 cells transiently transfected with CRTC1-MAML2 (C1M2) and pCDNA3.1+IREsGFP empty vector (EV) (1.5 ×10^5 cells/ well in 24 plate) and incubated with 20 µm EdU for 4 hours. Cells were analysed using a ZEISS AxioCam MRm2 fluorescent light microscope (Carl Zeiss, UK) and imaged using Image AxioVision Rel.4.8 image software at 20× magnification power. DNA (Blue) was stained with Hoechst 33342 (ab 145597). Green cells are EdU positive. B) The proliferation of NCI-H647 cells transiently transfected with CRTC1-MAML2 and empty vector negative control cells was analysed by counting the number of EdU positive cells and dividing by the total number of cell nuclei stained with Hoechst 33342 (ab145597) using the cell counter plugin in image J. Results represent the mean ± SD (N=3) NS P = 0.1613. Scale bar at 50 µm.

The second approach used to investigate the effect of the CRTC1-MAML2 fusion protein on cell viability was the clonogenic assay. Clonogenicity is a means of assessing a cells' ability to undergo clonal expansion from a single cell while ensuring their "unlimited" proliferation and differentiation capacity" (Rajendran and Jain, 2017). Our data indicated that after 10 days incubation in normal growth media, CRTC1-MAML2 transfected HEK293 and NCI-H647 cells showed an increase in the number of colonies formed and in the survival of the cells (figure 5.10 A and 5.11 A). The number of colonies formed was assessed by scanning the plates and quantifying the number of colonies which, as suggested by Franken *et al*. (2006), would only consist of ≥ 50 cells. Our analysis revealed a statistically significant increase in the number of colonies in the CRTC1-MAML2 transfected HEK293 and NCI-H647 cells (P= 0.0444 and 0.0008, respectively) (figure 5.10 B and 5.11 B). Surprisingly, the differences between the NCI-H647 transfected and control groups were more significant than between the HEK293 cell groups, despite the fact that the NCI-H647 is a pulmonary carcinoma cell line, whereas the HEK293 are derived from normal kidney epithelial cells. Our results indicated that the CRTC1-MAML2 fusion induced cell survival and expansion.

Among the colonies formed in each cell line, there was no difference in terms of the size of the colonies between the two groups (transfected vs EV controls). However, the colonies formed by the NCI-H647 cells were larger than those formed by the HEK293 cells. Notably, the colonies formed by both cell lines and groups (C1M2 construct and EV control) were mostly regular in shape with defined edges and without any morphological distinction among the groups (figure 5.12 A and B).

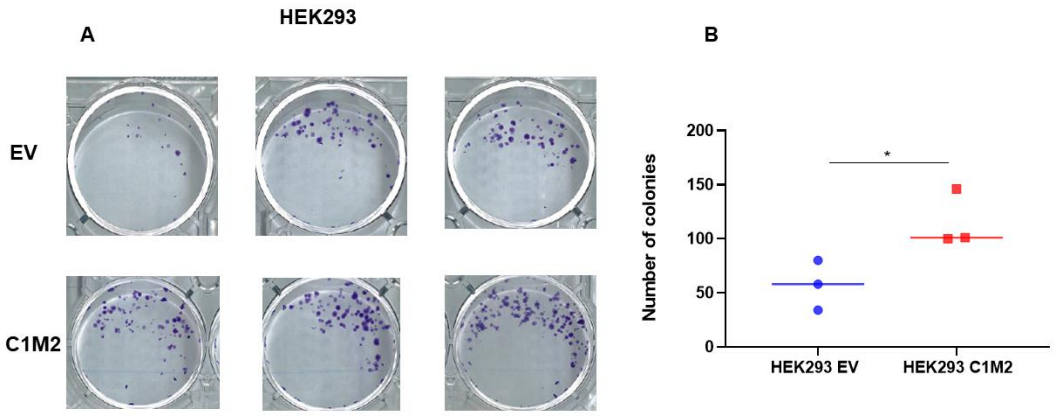


Figure 5.10 HEK293 Clonogenic assay. A) Representative images of single-cell colonies of HEK293 cells transfected with CRTC1-MAML2 (C1M2) and empty vector (EV) (control) and cultured in normal growth media for 10 days. Cells then were fixed and stained with 0.2% Crystal Violet solution. **B)** Quantification of the results revealed a statistically significant (*P*= 0.0444) increase in colony formation and survival of CRTC1-MAML2 transfected HEK293 cells compared to controls.

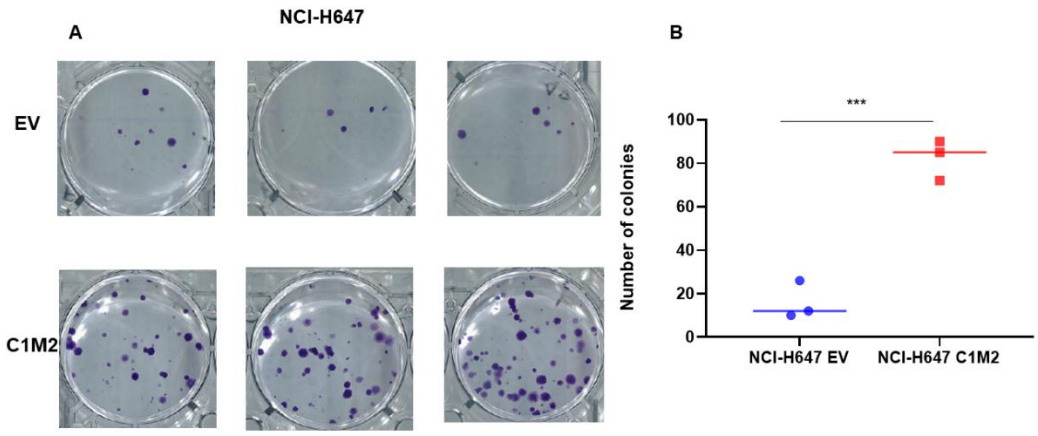
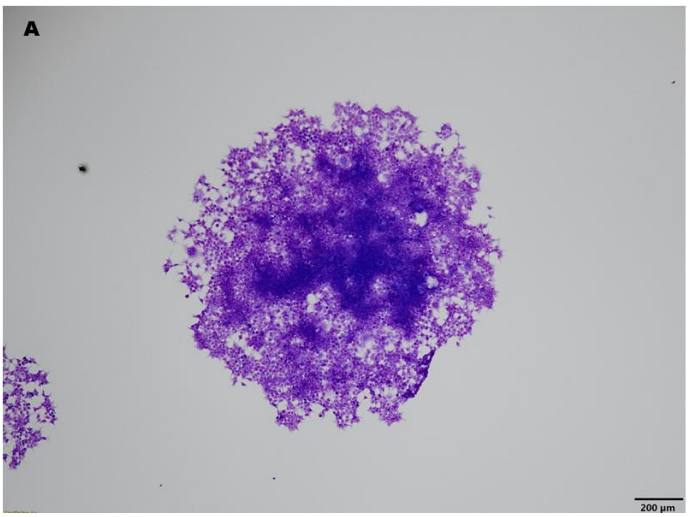


Figure 5.11 NCI-H647 Clonogenic assay. A) Representative images of single-cell colonies cells transfected with CRTC1-MAML2 (C1M2) and empty vector (EV) (control) and cultured in normal growth media for 10 days. Cells then were fixed and stained with 0.2% Crystal Violet solution. **B)** Quantification of the results revealed a statistically significant (*P*= 0.0008) increase in colony formation and survival of transfected CRTC1-MAML2 NCI-H647 cells compared to controls.



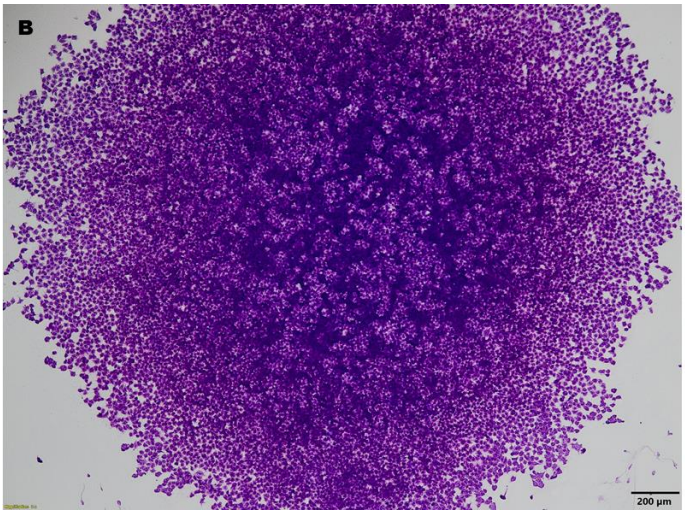


Figure 5.12 Clonogenic assay. Representative light microscopic images of transiently transfected HEK293 **(A)** and NCI-H647 **(B)** cells indicating differences in colony size. Scale bar at 200 µm.

5.5 Discussion

Assessing cell behaviour is crucial for the study of protein function. It is important to map the role of the protein in a specific cellular pathway or biological process. The main aim of the work described in this chapter was to identify the role of the CRTC1-MAML2 fusion protein in influencing cell behaviour. To achieve this, two cell lines that lack endogenous expression of the fusion protein, HEK293 and NCI-H647, were selected for gene transfection experiments and then used for downstream functional assays. Our initial *in vitro* data suggests that the CRTC1-MAML2 fusion protein has a significant effect on cell migration, survival and invasion, but it does not play a role in cell proliferation.

Previously (chapter 4) I have shown that traditional end-point *in vitro* analysis provides limited information on cell physiology. To our knowledge, no previous work has investigated the role of the CRTC1-MAML2 fusion protein on cell migration or invasion in normal or cancerous human cell lines.

Our previous work (chapter 4) was aimed at developing stably transfected cell lines expressing the novel fusion gene. Whilst we were not able to achieve this aim, we were able to effectively carry out reliable and reproduceable transient transfection, with significant transfection efficiency, for use in functional analyses. In this chapter we performed the optimised transfection methods on the selected cells and confirmed their success with western blotting and qPCR analyses. Transient transfection further allowed us to obtain a heterogeneous cell population which more closely replicates the *in vivo* nature of the disease.

Cell migration and invasion are important investigations in cancer research as they provide a better understanding of the molecular mechanisms of tumourigenesis, particularly the ability of a tumour to metastasise. Specifically, this study provides information on the effect of the CRTC1-MAML2 fusion protein on the rate of migration by NCI-H647 cells and also on their ability, and that of HEK293 cells, to invade.

A migration scratch assay is a basic and easy assay to perform but in order to ensure that we were only detecting cell migration, it was necessary to add Mitomycin C to inhibit cell proliferation and prevent any change in cell number. Although this assay is straightforward and efficient there is one associated limitation, namely variation of the gap width during wound creation. To avoid such an issue, we used three (N=3) biological repeats and during data analysis normalised our measurements to the original gap area created at time zero (T0).

As stated previously (section 5.4.2) it was not possible to perform the scratch assay on HEK293 cells due to cell detachment immediately after the wounds were created. This issue was previously reported in the literature as a limitation of the migration assay with some types of cells, including the HEK293 cells (Justus *et al*., 2014). The CRTC1-MAML2 transfected NCI-H647 cells demonstrated a statistically significant increase in cell migration.

The scratch assay assesses the ability of cells to migrate horizontally, whereas the trans-well invasion assay investigates the movement of cells vertically. Moreover, using coated membranes with ECM substitutes, such as Fibronectin or Matrigel, the trans-well invasion assay measures the ability of the cells to dissolve the extracellular

matrix (ECM) and invade through the pores in the trans-well membrane. The presence of a chemoattractant promotes the directed migration of cells towards the physical barrier. To mimic the *in vivo* process of cancer cell invasion through the basement membrane, we coated the top of the trans-well membranes with fibronectin or used commercially available trans-wells pre-coated with Matrigel. There are advantages in using pre-coated trans-wells as these will be standardised and have a homogeneous matrix layer, in addition, during the analysis, it was easier to differentiate between the pores in the membranes and invaded cells.

Matrigel is a basement membrane derivative, comprised of a number of extracellular matrix proteins (56% laminin, 31% collagen IV and 8% entactin) and numerous growth factors such as bFGF, EGF, IGF-1, TGF-β (Kleinman *et al*., 1982). Fibronectin is one of the component of the extracellular matrix secreted by fibroblasts, plays a major role in cell adhesion, differentiation, growth and migration and it is important in the wound healing process. Dysregulation in fibronectin function is associated with various pathological conditions including cancer (Pankov and Yamada, 2002).

In our study we chose an applicable chemoattractant for our cells (10% FBS) and serum-starved the seeded cells in 1% FBS. We allowed 48 hours for the cells to invade; it was not possible to monitor the rate of invasion. Before serum-starving, the cells were left for 4 hours to attach and form boundaries with the basement membrane via integrin-fibronectin or integrin-Matrigel interactions. Our results suggest that CRTC1-MAML2 could be implicated in the invasion process by promoting the enzymatic disassociation of cells from the coating membrane and this facilitated movement towards the chemoattractant.

It was crucial to study the effect of the CRTC1-MAML2 fusion protein on the ability of cells to migrate and invade, as these two cell functions are important for cancer progression and metastasis and they are responsible for morbidity and mortality due to surrounding tissue destruction (Xu *et al*., 2018). Study of cell migration and invasion can assist in determining the metastatic characteristics of cancer cells. Metastasis is responsible for 90% of cancer deaths from solid tumours (Kramer *et al*., 2013). Quantification and analysis of the migration speed and time will provide more accurate information about the *in vivo* event which could subsequently be used for treatment strategies.

Measuring the DNA synthesis of living cells, such as in the EdU assay, is an optimal method of assessing the rate of cell proliferation, while a clonogenic assay is considered an alternative approach to measure a cell populations' ability to proliferate sufficiently to form a colony. The EdU assay is thought to give the most robust, sensitive and direct indication of cell proliferation (Menyhárt *et al*., 2016). First developed by Salic and Mitchison in 2008, the EdU assay is based on the incorporation of 5-ethynyl-2'-deoxyuridine into cellular DNA during replication. This reaction is detected via fluorescent azides (iFluor-488), which are small enough to diffuse freely through native tissue and DNA, by covalent crosslinking that uses "click chemistry"; this technology gives confidence that the EdU assay is more sensitive, quick and easier to perform than the 5-bromo-2'-deoxyuridine (BrdU) assay, which requires DNA denaturation and addition of a detection antibody. The EdU assay can be used to quantify proliferating cells by flow cytometry or fluorescence microscopy.

Surprisingly, our proliferation assay results indicated that the CRTC1-MAML2 fusion protein has no effect on the proliferation rate of either the HEK293 or NCI-H647, and

this was consistent across the two analytical approaches used, flow cytometry and fluorescence microscopy. Although there were no significant differences between the transfected cells and control cells, the flow cytometry data did show a trend towards a lower proliferation rate in the CRTC1-MAML2 transfected cell lines. Nonetheless, the mitotic rate of our MEC cohort (chapter 3, section 3.4.1) was generally low, which is support that the MEC is a slow growing tumour. Further experiments using a larger panel of cell lines could validate this result. Currently, there is no published data demonstrating any effect of the fusion protein on cell proliferation for comparison or validation.

This result is quite interesting and might also be related to the clinical course of the disease. MECs are usually slow growing tumours and so the presence of the fusion gene, as our results suggest, does not affect the proliferation of the tumour cells. However, if we reflect on all of my findings, the fusion protein does significantly impact many cell behaviours that are associated with aggressiveness (migration, invasion, survival) which are not related to an increase in proliferation. Therefore, the presence of the translocation is probably interfering directly with mechanisms associated with cell motility, ECM degradation and pro-survival or anti-apoptotic.

Our results are consistent those of Tonon *et al.* (2003) who assessed the effect of the CRTC1-MAML2 fusion protein on cell survival and the ability of the cells to undergo sufficient proliferation to form colonies; their results suggested that the CRTC1-MAML2 fusion enabled transiently transfected RK3E cells to form colonies as in the absence of the fusion gene the cells did not form colonies.

Our study, and that of Tonon *et al.* (2003), suggest that transient exogenous expression of the CRTC1-MAML2 fusion protein alone is not sufficient to promote cell proliferation, other metabolic changes and the presence of other cell promoters may be required. Amelio *et al*. (2014) demonstrated that following transient transfection of RK3E cells with CRTC1-MAML2 fusion the addition of a MYC oncoprotein, and Notch and CREB transcriptional co-activators, was required for an increase in the number of irregularly-shaped foci. These findings suggest that the CRTC1-MAML2 fusion may initiate cell transformation, but other promoters and genetic alterations are required to induce the progression of the carcinoma. Anzick *et al*. (2010), using high-resolution oligonucleotide array comparative genomic hybridisation (CGH) identified deletions within the CDKN2A (p16) gene in 5 CRTC1-MAML2 fusion positive, fatal MEC cases and metastatic MEC fusion positive cell lines (NCI-H292, NCI-H3118, 10H2, 181A2) and a lack of such deletions in 9 CRTC1-MAML2 fusion positive MEC cases with good prognosis; patients alive, free of disease, or other cause of death was documented without recurrence of MEC.

These results further support our findings in chapter 3, regarding the expression pattern of the CRTC1-MAML2 translocation in MEC clinical cases. Expression of the fusion transcript was mild in most of our cohort, and intense in only three cases, two of which showed fusion positive cells with distinct histological architectures such as direct contact with stroma and neural invasion (figures 3.10 and 3.11).

NCI-H647 is a cancer cell line and as such will grow as colonies, however, our data showed that the CRTC1-MAML2 fusion protein significantly affected the cells' ability to form colonies compared to control cells (P= 0.0008). The impact of the fusion gene

on the colony forming ability of HEK293 cells, immortalised normal epithelial cells, was not so significant (*P*= 0.0444).

NCI-H647 cells were established from metastatic non-small cell adenocarcinoma, from a patient who received radiotherapy prior to the generation of the cell lines (www.ATCC.com, NCI-H647 [H647] (ATCC® CRL-5834™, accessed in September 2020). HEK293 cells are originally derived from human embryonic kidney cells, which have been immortalised by transfection with adenovirus 5 DNA, and express adrenal, progenitor and neural cell markers; they are, therefore, thought to originate from a transformed embryonic adrenal precursor cell. Despite the HEK293 cells being used by many researchers as normal cells, their tumourigenicity is still debated, mainly due to changes in the expression of cancer-associated genes such as, AKT2, CUX1, MAD2B and their chromosomal instability, which was mainly due to the immortalised reagent (adenoviral E1) (Stepanenko and Dmitrenko, 2015). Importantly, in this study an accurate control (mock transfection with an empty vector) was used to investigate the effect of the CRTC1-MAML2 fusion protein on cell behaviour.

Our results suggest that the exogenous expression of the CRTC1-MAML2 fusion protein could lead to the development of a more aggressive *in vitro* phenotype, which may clinically lead to the development of potentially metastatic tumours.

Further research is needed to more resolutely elucidate the role of the CRTC1-MAML2 fusion in malignant cell growth and behaviour. The data discussed in this chapter suggest that the CRTC1-MAML2 fusion influences cell migration, invasion and survival. To further investigate this hypothesis, it would be of interest to perform

these assays on CRISPR gene edited MEC fusion positive cell lines, by silencing the endogenous translocation gene. This would enable assessment of the effect of lack of endogenous expression of the CRTC1-MAML2 on MEC cell populations and provide further insight into the function of the fusion protein.

5.5.1 Conclusion

The CRTC1-MAML2 fusion protein appears to be implicated in tumourigenic cell behaviour in HEK293 and NCI-H647 cells *in vitro*. The CRTC1-MAML2 fusion induced migration in the NCI-H647 cells and increased survival and invasion in the HEK293 and NCI-H647 cells. Interestingly, no effect on HEK293 and NCI-H647 cell proliferation was observed.

5.5.2 Future work

Future work would include scanning electron microscopy to identify any morphological differences between the CRTC1-MAML2 transfected and control cells, in terms of cell size, nuclear condensation or highlighting any phenotypic changes in the cells.

This assessment was performed on cell lines that were chosen due to their ability to express the exogenous CRTC1-MAML2 fusion protein after transient transfection. To develop the full picture of the role of the fusion protein, additional studies using normal primary human salivary gland cells will be needed. In chapter 4, we demonstrated the *de novo* expression of the CRTC1-MAML2 translocation transcript in transiently transfected primary salivary gland cells and that this was maintained for six days after transfection with some reduction in expression over time. It would be valuable to carry

out large scale transient transfection experiments and snap freeze the cells to perform functional assays after confirming the ectopic expression of the CRTC1-MAML2 fusion at a gene and protein level.

Chapter 6

Development of *In vitro* Organoid Three-Dimensional Models from Normal Primary Salivary Gland Cells, Transfected Normal Primary Salivary Gland Cells and Mucoepidermoid Carcinoma Cell Lines

6. Chapter 6

6.1 Introduction

Salivary glands are distinctive exocrine structures whose function is to produce saliva. As outlined previously in chapter 1, the mature salivary gland structure consists of organised acini, branching ducts and terminated ducts in connective tissue stroma. The salivary unit is composed of two main categories of cells: luminal, where the saliva is produced and secreted, and abluminal cells, including a basal cell layer and myoepithelial cells which contract to help excrete the saliva. Under normal circumstances, saliva production and excretion occur in an organised manner with controlled interactions and signals between epithelial cells and the surrounding microenvironment. Interactions between cells, adhesion molecules, junctional proteins and receptors allow mono-directional flow of the solutions and the exchange of minerals and electrolytes.

During diseases such as neoplasms, the organised structure of the salivary gland unit is lost as the tumour progresses, leading to the destruction of most of the gland and overgrowth of one or more neoplastic cell type. Salivary gland tumours are a heterogenous and clinically diverse group. Study of the initiation and tumourigenesis of these tumours in standard 2D cultures remains limited and does not recapitulate

the exact *in vivo* situation. My study demonstrates the development and validation of novel, three-dimensional structures from normal salivary gland cells.

6.2 Aims and objectives

The primary aim of this chapter was to establish three-dimensional, organoid, *in vitro* models using normal salivary gland cells, those that have been transfected with the CRTC1-MAML2 fusion gene and from established MEC cell lines.

Our objectives were to use an early passage of human primary normal salivary gland cells from explanted tissue to develop the organoid structures and, using pre-optimised conditions, transfect early passage human primary salivary gland cells and utilise the transfected cells in developing further 3D models. A further objective was to establish 3D MEC models from MEC cell lines and to validate all models using microscopic and histological techniques.

6.3 Materials and Methods

The relevant materials and methods utilised in this chapter are detailed in chapter 2, under headings 2.2, 2.5 and 2.8 as indicated below:

- Growth and culture of the normal human primary salivary gland cells and cell lines were carried out as outlined in section 2.2.2, 2.2.3.
- β-galactosidase (SA-β-Gal) assay was performed as indicated in section 2.2.11.
- Immunofluorescence double staining of human primary salivary gland cells was conducted as described in section 2.2.12.
- RT-PCR analysis of mRNA from normal human primary salivary gland cells was carried out as described in sections 2.5.2, 2.5.3, 2.5.4 and 2.5.5.

- Development of 3D Organoid models of salivary glands was performed as described in section 2.8.1.
- Development of 3D Organoid models of CRTC1-MAML2 transfected salivary glands cells was performed as described in section 2.8.2.
- Development of 3D models of MEC cell lines was performed as described in the section 2.8.3.
- Fixation, processing and sectioning of 3-D models was performed as described in section 2.8.4.
- Periodic Acid Schiff staining of three-dimensional organoid models was performed as described in section 2.8.5.
- Immunofluorescence staining of 3D organoid models was performed as described in section 2.8.6.

6.4 Results

6.4.1 Characterisation of human primary normal salivary gland cells

It was crucial to re-confirm the origin and the nature of the cells we were working on; this work had initially been performed by Dr Zulaiha A Rahman and Miss Rachel Furmidge (two previous members of our laboratory). A range of analyses were performed on the normal human primary salivary gland cells, including senescence activity, transcriptional and protein analysis.

A β-galactosidase (SA-β-Gal) assay allowed us to confirm that the normal primary cells are senescence with age, and they will stop proliferating after certain passaging. HuSL cells at passage 1-4 continued to proliferate and stained negatively for SA- β-Gal activity (0%) but later passage HuSL cells (P7 and above) stained positively for SA- β-Gal activity (~90%) and also showed changes in characteristics and phenotype; the cells flattened, were enlarged and vacuoles were present inside the cells (figure 6.1).

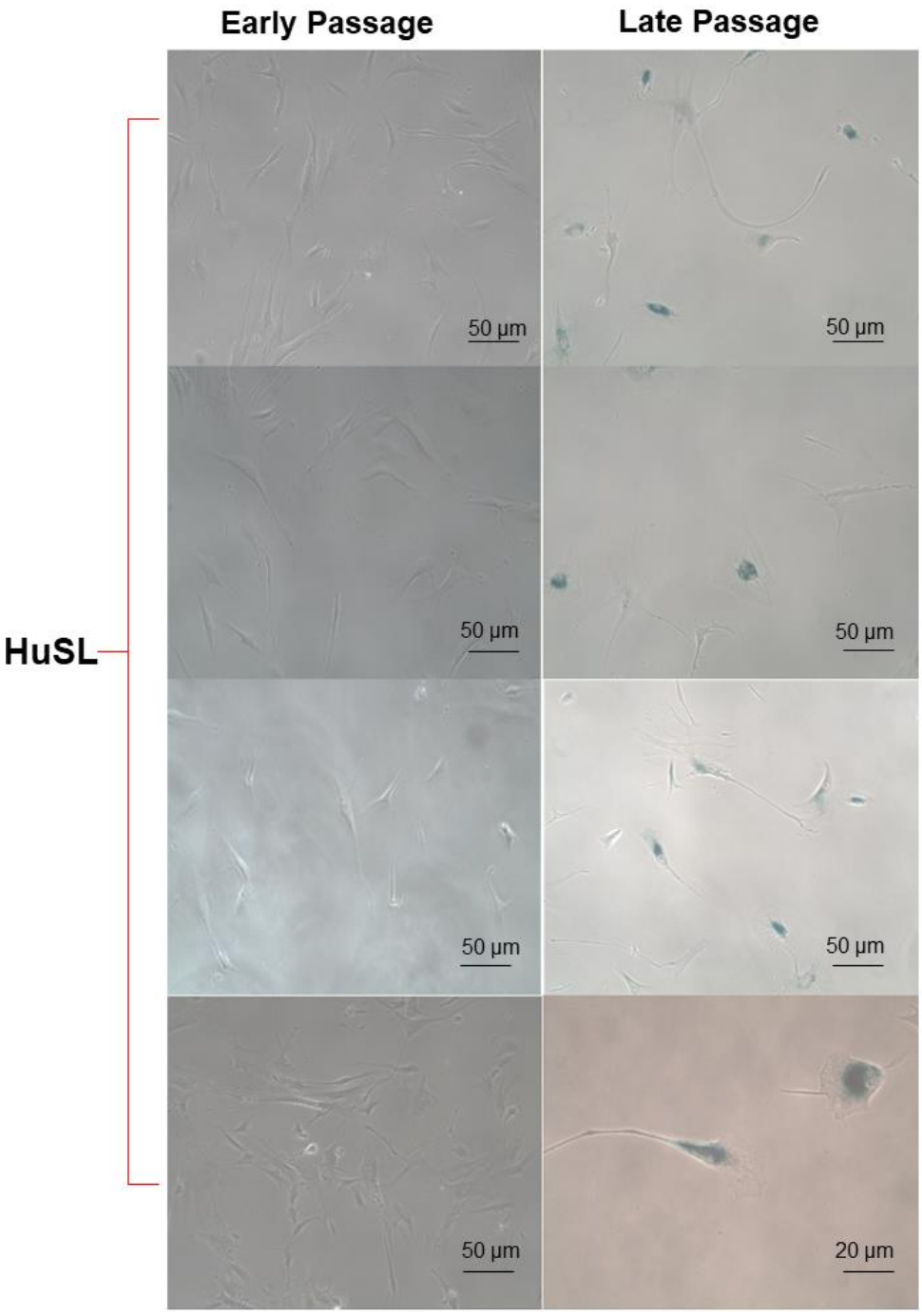


Figure 6.1 β-galactosidase (SA-β-Gal) assay. Early (P2) and late passage (P7+) human primary normal salivary gland cells (HuSL) were seeded per well in a 6-well plate with normal growth media and left to adhere overnight. Senescence associated β Gal activity was

measured using a Senescence Detection Kit. The β-galactosidase activity was increased in late passage cells. Scale bars represent 50 and 20 μm.

RT-PCR analysis was used to detect the expression of a selected panel of salivary gland genes with RNA extracted from normal human primary salivary gland cells (figure 6.2). The gene profiling analysis showed that for most of the genes studied there was expression, of varying intensity: Nanog (a progenitor cell marker), E-cadherin (tight junction marker), C-Kit (stem cell marker), amylase (acinar cell marker), CK5 (basal cell marker) and mucin genes. Although RT-PCR is semi-quantitative at best, my results suggest greater expression of amylase in parotid cells, which is as expected as the parotid gland is composed mainly of serous cells while the submandibular glands and sublingual glands are composed of a mixture of mucous and serous cell types. Expression of CK5 was weak, particularly in the parotid and submandibular gland cells. In addition, expression of MUC5B (mucin secretion of normal cells) was weak in the parotid and sublingual gland cells and moderate in the submandibular gland cells, whereas there was no expression of MUC5AC and MUC2 (mucin secretion of tumour cells) in any normal cells and the multiple bands detected were not consistent with the expected product size. Human Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and human Ornithine Decarboxylase Antizyme (hOAZ1) were used as internal controls.

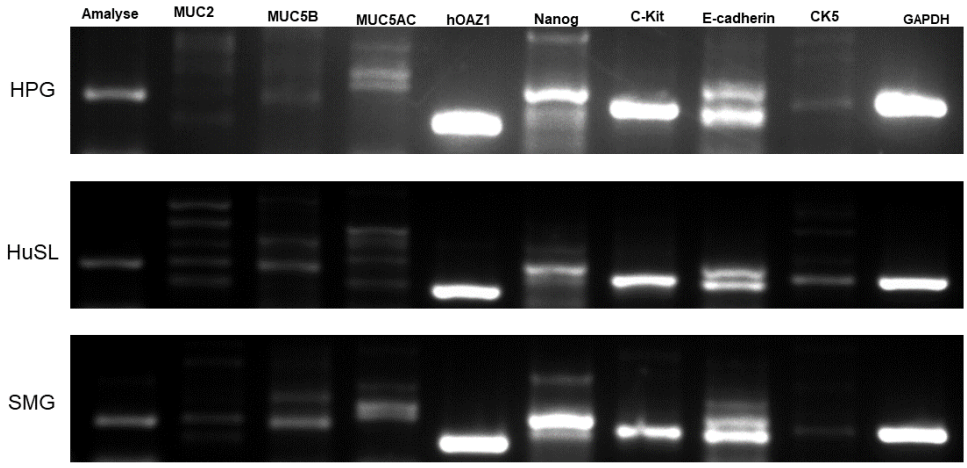


Figure 6.2 Reverse-Transcriptase PCR analysis of mRNA of human primary normal salivary gland cells (HPG, HuSL and SMG). A 2% agarose gel for end-point RT-PCR analysis demonstrated mRNA expression of epithelial markers, stem cell markers and EMT markers. GAPDH and hOAZ1 were used as internal controls, DNA ladder is 100 bp.

Double immunofluorescence protein analysis and confocal microscopy were conducted on monolayers of salivary gland cells to validate the origin of the cells. Positive staining of the basal epithelial cell marker, CK5, and stronger expression of the epithelial-to-mesenchymal transition marker, vimentin, clearly indicated that these cells were a mixed population of epithelial and mesenchymal cells. (figure 6.3).

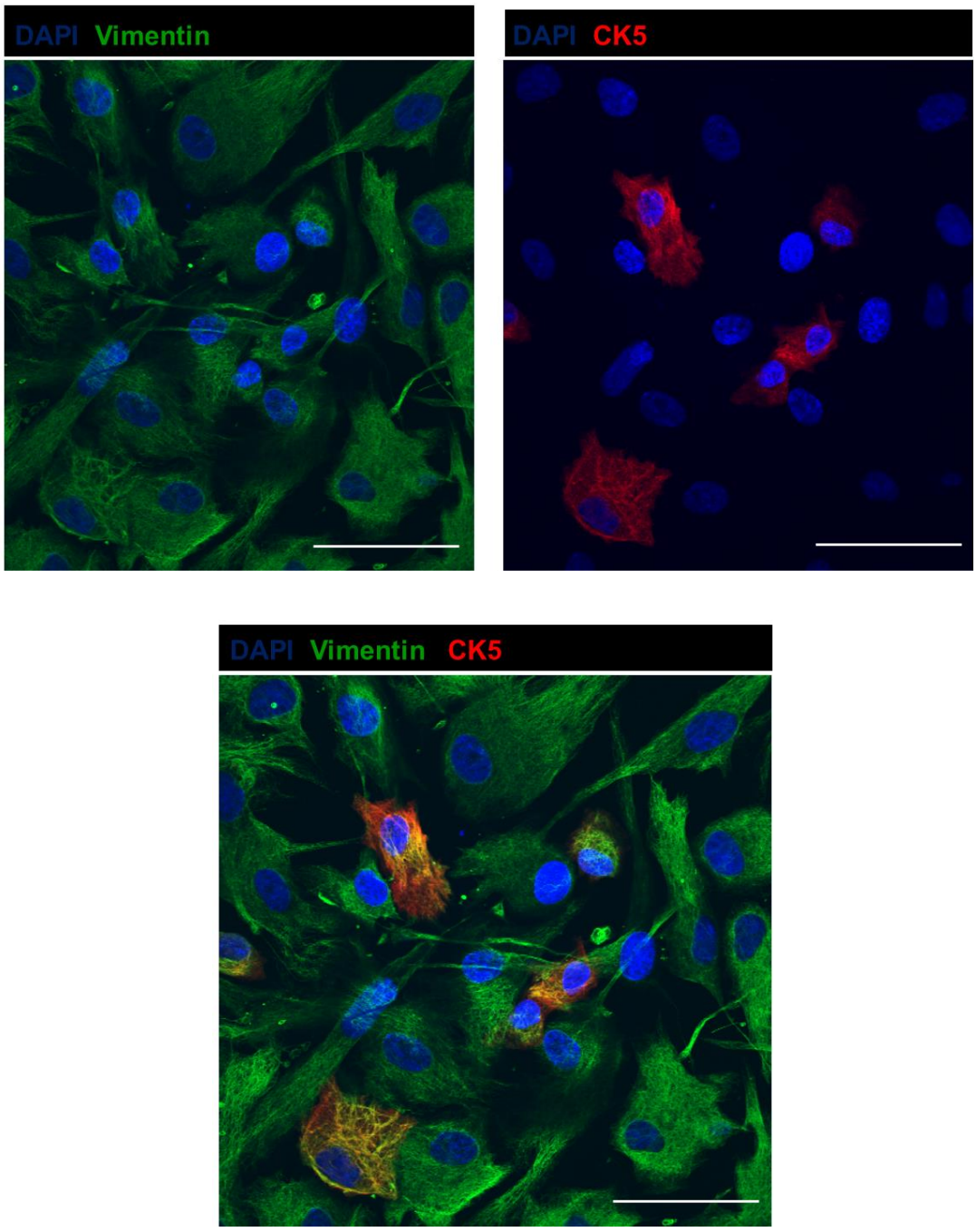


Figure 6.3 Immunofluorescence analysis of monolayers of human primary normal salivary gland cells. Representative immunofluorescence confocal microscopic images show the expression of vimentin (green) and CK5 (red) in normal explanted cells. Scale bar represents 50 µm.

6.4.2 Development of three-dimensional salivary gland organoid model

A salivary gland organoid model was primarily established and optimised in our laboratory by a previous doctoral colleague (Dr Zulaiha Rahman) and methods established were repeated in this study. An early passage of expanded sublingual cells (HuSL) was mixed with Matrigel, seeded onto trans-well inserts and submerged in media containing Wnt3-Rsp-1 and Rho-Kinase (ROCK-inhibitor, Y-2763). After three days, an Air-Liquid interface (ALI) was established through the removal of culture media from the apical surface of the cells. Growth media from this point did contain Wnt3-Rsp-1, but not Y-2763, and this time point was considered day zero of model growth. The model was allowed to continue to develop for a further 14 days.

It was evident by day 3 that the cells had started to aggregate and proliferate to form small regular spherical-like structures, and when the ALI culture was established the spherical structures continued to grow, differentiate and increase in size. On day 14 the organoids had proliferated and differentiated into mini-spherical structures (figure 6.4 A).

Primary analysis was carried out by Haematoxylin and Eosin staining of formalin-fixed, paraffin embedded sections and this identified epithelial cells growing in a circular uniform pattern, adherent at the periphery and decreasing in size towards the centre of the structure. We also noted the presence of smaller budding-like structures close to larger structures in the connective tissue matrix and the presence of elongated connective tissue cells with a fibroblast-like morphology (figure 6.4 B).

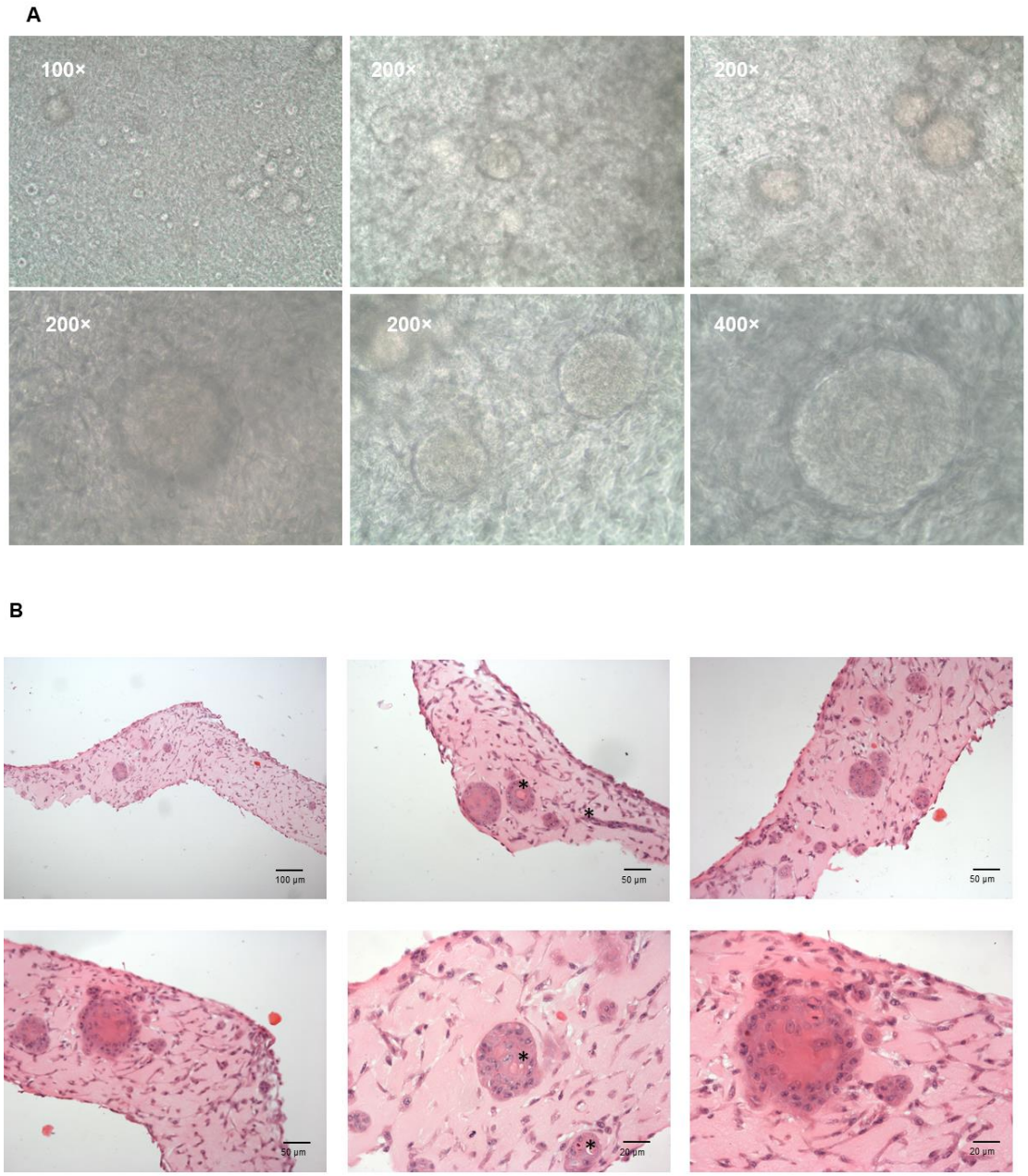


Figure 6.4 ALI culture of human salivary gland organoid models. Primary salivary gland cells at low passage (1-4) were used to develop 3D organoid cell cultures on inserts at an ALI. Cells were mixed with Matrigel (ECM), and grown in media containing R-Spondin1 and Wnt3aA. A) Representative phase-contrast images of 3D organoids on day 14 showing proliferation and differentiation of the cells into small, round structures resembling salivary gland acini. Organoids and small cell clusters were observed dispersed throughout the ECM substitute. The morphology of the organoids was stable and consisted of round aggregates

with a smooth periphery. No lumen or buddings were noted with phase contrast microscopic analysis. B) H&E staining of formalin fixed paraffin embedded sections confirmed the presence of organoids with round morphology and smooth periphery (folded-like structure), small cell clusters and single cells dispersed within the ECM. Central cystic formations suggestive of a lumen were noted in some organoids (asterisks).

6.4.3 Development of three-dimensional salivary gland organoid model of CRTC1-MAML2 transfected cells

To investigate whether the presence of the CRTC1-MAML2 translocation alters the morphology and organisation of 3D organoids, early passage HuSL cells were transfected with the CRTC1-MAML2 construct and the model established using the same culture conditions as previously determined for the growth of organoids from normal primary cells.

As with the normal cells, by day 3 the transfected cells had started to aggregate and develop organoid structures, however, the resultant organoid structures were not completely spherical, they were more irregular in shape and of varying size (figure 6.5).

H&E analysis revealed that the histology mainly lacked the regular, well-defined organoid structure and the cells were mostly mixed with the ECM. In some sections there were some organoid structures that were very small in size but, unlike the normal cells, the cells were arbitrarily distributed (figure 6.6).

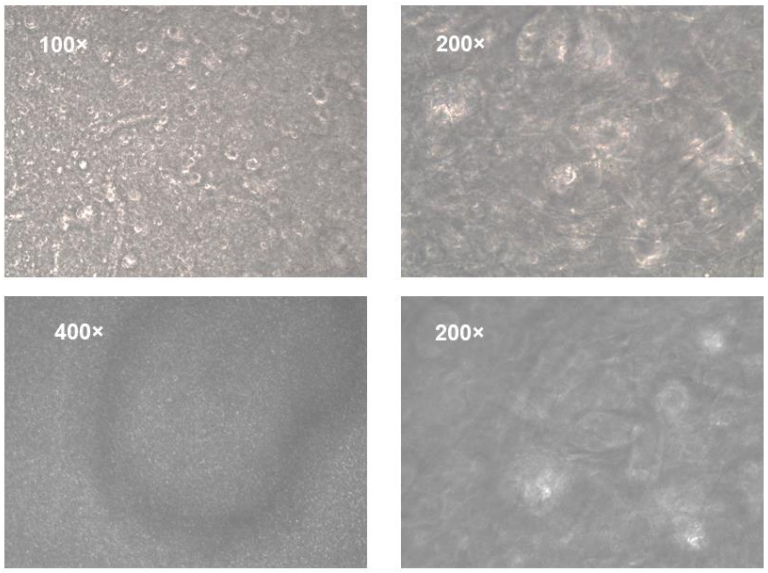


Figure 6.5 ALI culture of human salivary gland organoid model transfected with the CRTC1-MAML2 construct. Primary salivary gland cells with low passage number were initially transfected with the CRTC1-MAML2 construct and after 48 hrs were harvested and used to develop a 3D organoid cell culture at an ALI. The cells were mixed with Matrigel and grown in media containing R-Spondin1 and Wnt3aA**.** On day 14, representative phase-contrast images showed proliferation and differentiation of the cells into varying sizes and slightly irregular organoid structures.

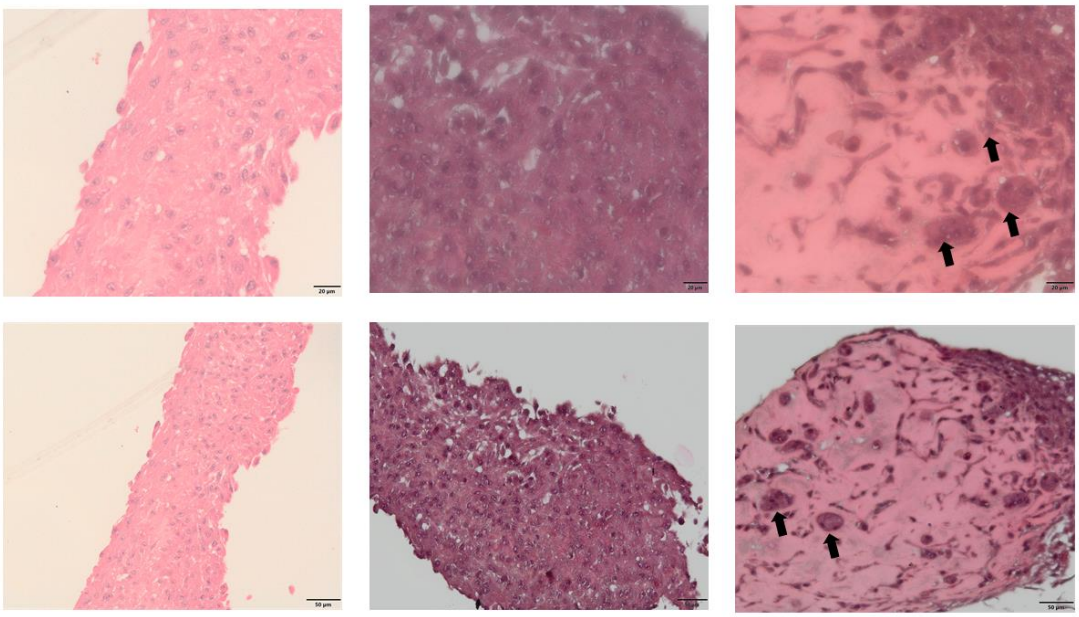


Figure 6.6 Representative H&E images of 3D organoid structure of transfected normal salivary gland cells. The histological analysis revealed irregular and atypical cell structures mixed in the ECM. One section exhibited the presence of small organoids (arrows).

6.4.4 Establishment of three-dimensional mucoepidermoid carcinoma organoid models

Three MEC cell lines were also used to develop *in vitro* MEC organoid models. Two of these were of salivary gland origin, UM-HMC-2 cells derived from a lymph node metastasis and A-253 cells derived from a submandibular gland primary tumour, with a third cell line, NCI-H292, being derived from a lymph node metastasis of a pulmonary MEC; 2D cultures of these cells are shown in figure 6.7. Unlike the organoid models established from primary cells, no growth factors or specific conditioned media were added to these models, only the standard growth media with essential components for each cell line, as outlined in section 2.2.3 was used.

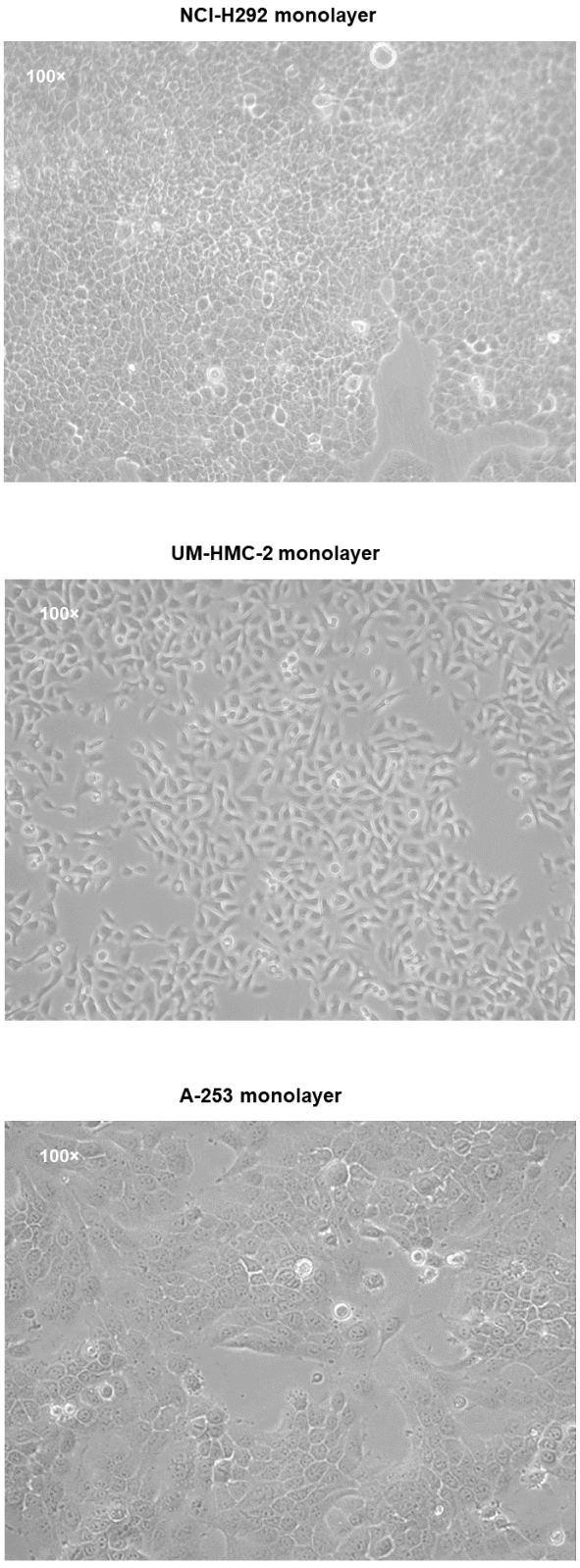


Figure 6.7 Mucoepidermoid carcinoma cell lines. Phase contrast images of MEC cell lines growing in monolayer and showing homogeneous growth of polygonal epithelial islands.

The culture of MEC 3D models was carried out in the same way as that for the 3D models of normal salivary glands. By day 14 the three MEC had proliferated and differentiated to form 3D organoids (figure 6.8).

Haematoxylin and Eosin staining was used to assess the histology and morphology of the model. Additionally, Periodic acid-Schiff (PAS) staining was used to confirm the differentiation status and highlight mucin production by cells in the models. We further confirmed the organisation and development of these models at a protein level using immunofluorescence staining with confocal image analysis.

All MEC cell lines formed round 3D clusters suggestive of tumour organoids and 3D aggregates with irregular morphology were dispersed within single cells in the ECM substitute (figure 6.8, A-C).

Histological analysis revealed that the tumour cells in ECM retained the histopathology of the cells from which they were derived with the epithelial cells being organised as small islands resembling a glandular pattern. Interestingly, in comparison to the models established from normal cells, the epithelial cells grew in clusters of irregular, round aggregations of varying size and the number of clusters was greater (figure 6. 8, D-F).

All 3D models derived from MEC cell lines showed clear positive PAS staining suggesting that they were secreting a high level of mucopolysaccharide and thus

producing mucin (figure 6.8, G-I,). The 3D models derived from A-253 cells appeared to contain more PAS in the cells. Interestingly, most of the PAS positive cells had morphological features similar to mucous cells present in MEC, such as abundant foamy cytoplasm with mucin vacuoles (figure 6.9).

Immunofluorescence analysis, with confocal microscopy, indicated successful development and differentiation of 3D MEC organoid models with clear expression of actin filaments (figure 6.10-11).

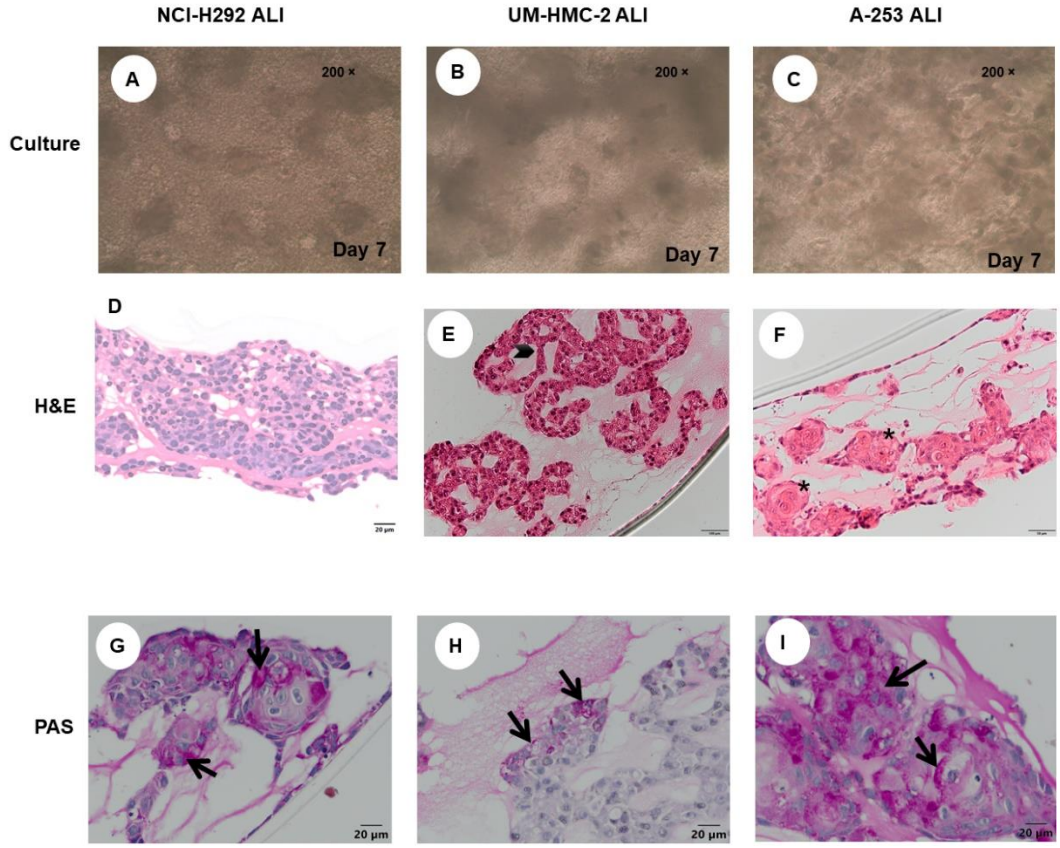


Figure 6.8 MEC Organoid 3D model. Representative images of MEC cell lines NCI-H292, UM-HMC-2 and A-253 3D cell culture under **(A-C)** phase contrast, **(D-F)**, H&E histological analysis and **(G-I)** PAS staining. **(A-C)** Phase-contrast inverted microscopic images of stages of development of organoid models, from MEC cell lines, showing morphology following 7 days of growth in Matrigel at an ALI culture. All MEC cell lines formed round 3D clusters suggestive of tumour organoids and 3D aggregates, with irregular morphology and single cell dispersing within the ECM substitute. **(D-F)** Histological analysis revealed more cellular structures compared to the 3D matrix derived from normal salivary gland tissue and apparent growth of epithelial cells in clusters. Cells formed irregular round aggregations (asterisks) or (arrowhead). **(G-I)** All MEC cell lines in the 3D cultures showed clear positive staining for Periodic acid-Schiff (PAS) in mucous-secreting cells as indicated by the arrows, suggesting mucin production. The 3D models derived from the A-253 cells had the highest percentage of PAS positive cells.

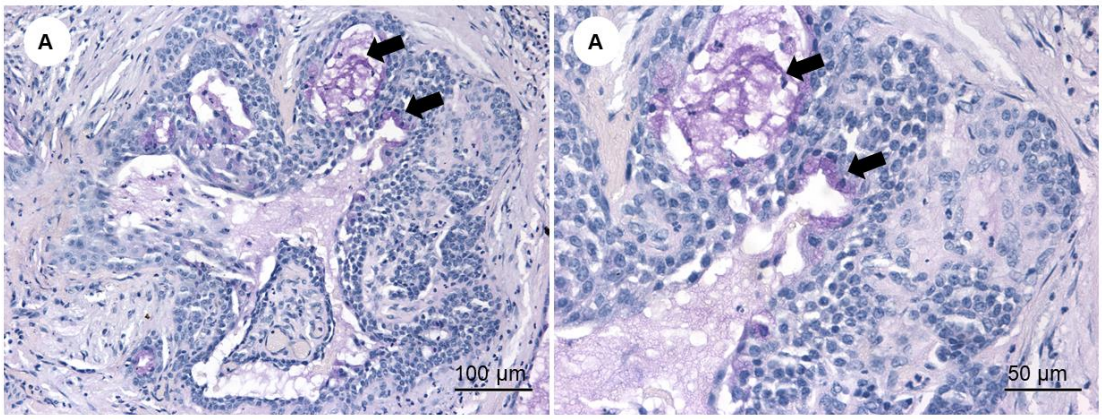


Figure 6.9 Mucoepidermoid carcinoma sections. Representative microscopic images of Periodic acid-Schiff (PAS) staining showing positive tumour cells secreting mucin (indicated with arrows). Scale bars at 100 and 50 µm.

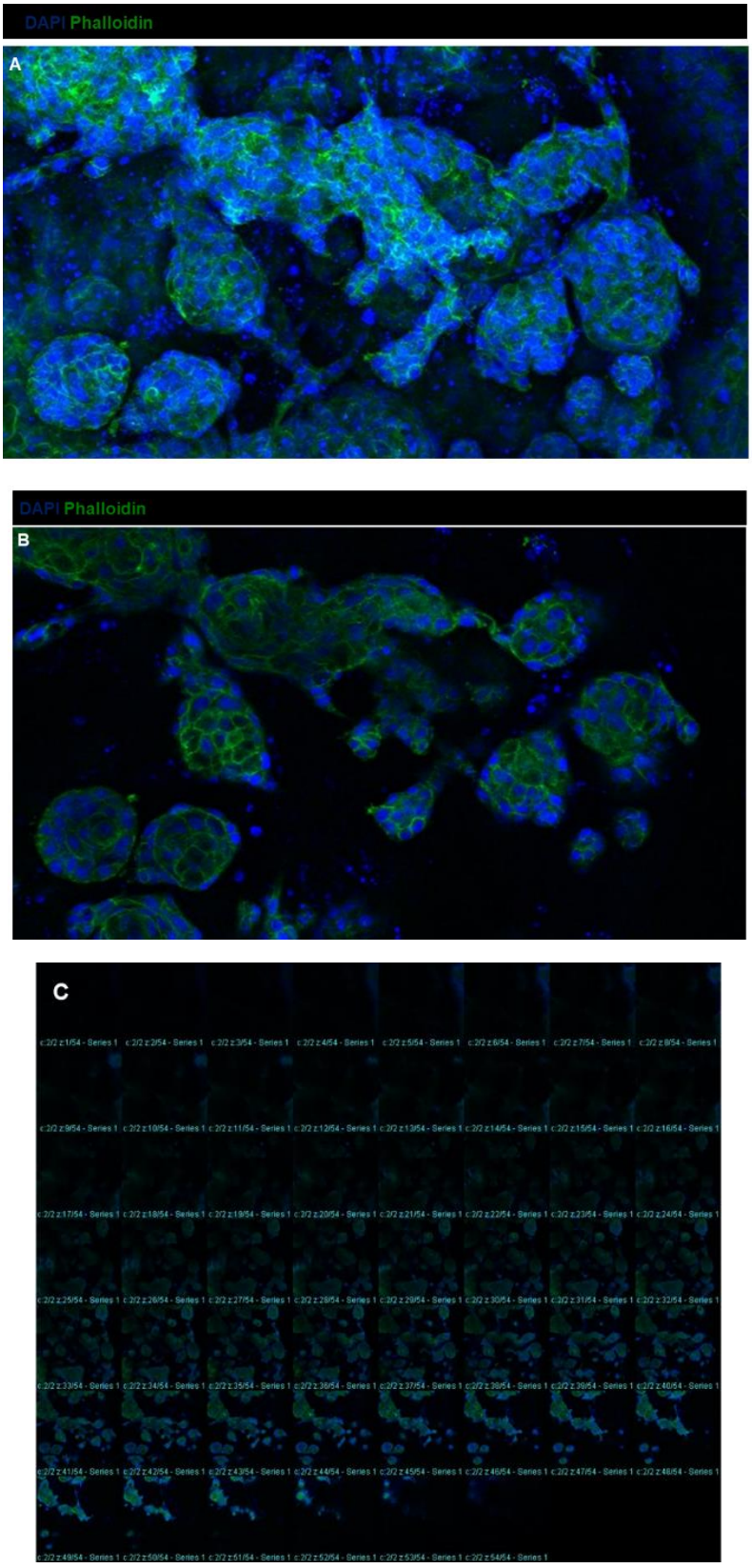


Figure 6.10 Morphology of the A-253 3D organoid model. Representative confocal fluorescence microscopic images of MEC cell line organoids showing Phalloidin staining

(green) of key filamentous actin junctions, counter stained with DAPI (blue) for nuclei DNA. **A)** a compact Z project processed with Fiji ImageJ software demonstrating the development of organoids in different dimensions. **B)** One processed layer of the stalk and **C)** montage of the whole stalk layer (108).

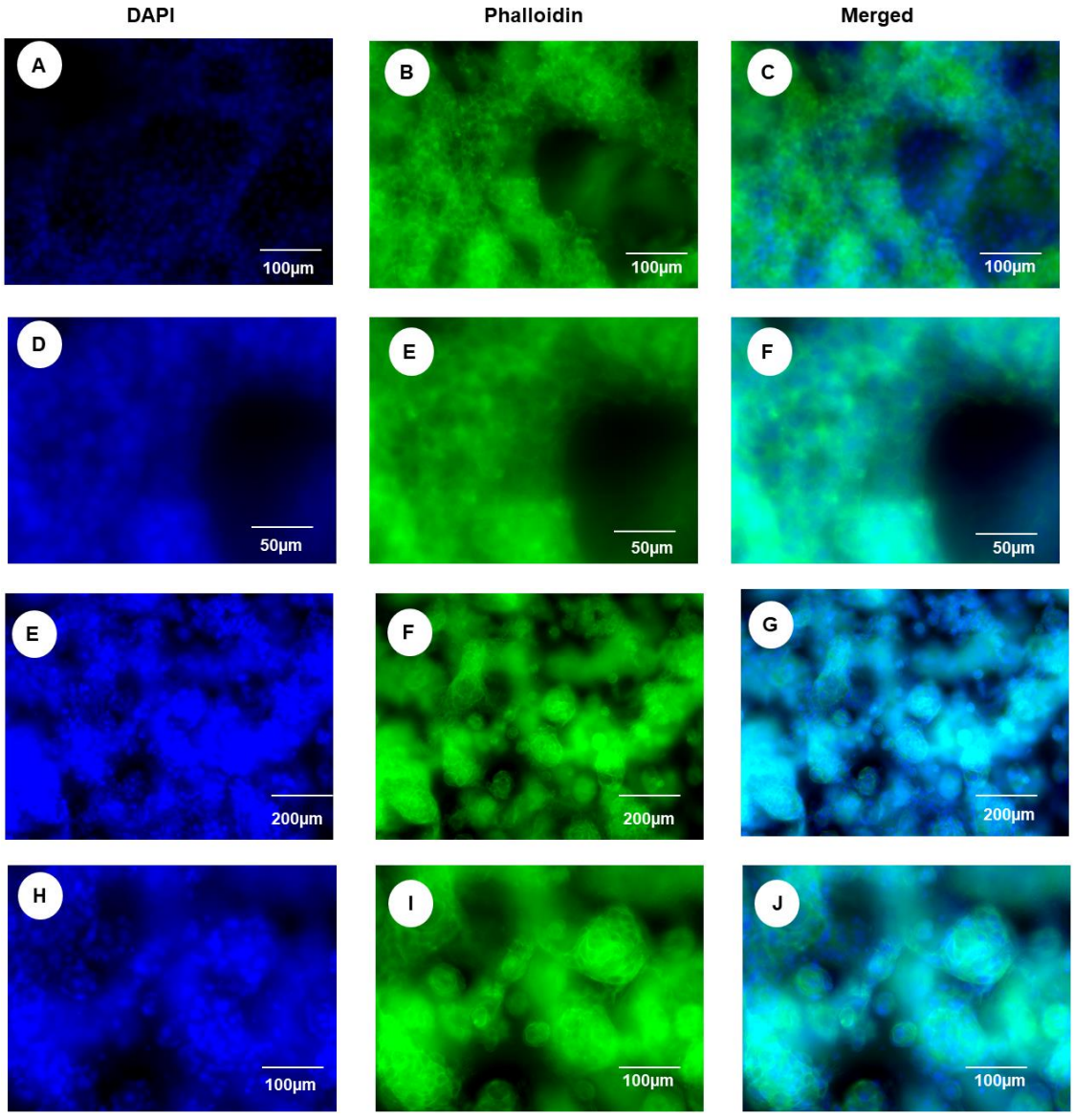


Figure 6.11 Immunofluorescence staining of MEC 3D structure. Representative fluorescence microscopic images of MEC 3D organoid structures: (A-F) UM-HMC-2 cells, (G-J) A-253 cells stained with phalloidin (Green) to identify actin filaments and counter stained with DAPI (Blue) nuclear marker.

6.5 Discussion

The concept of developing 3D *in vitro* cell cultures has gathered significant interest and has been applied to many mammalian cell types (Ruedinger *et al*., 2015). Areas of potential use for such cultures include models of tissue morphogenesis and organogenesis, models of diseases such as tumours and infections, drug and toxicity testing, personalised and regenerative medicine.

The three-dimensional model has certain advantages over two-dimensional cultures and *in vivo* models. 2D cultures lack the cellular heterogeneity, true cell to cell contact and cell/extracellular matrix interactions, which occur in a 3D model and more closely mimics the *in vivo* situation. *In vivo* animal models must be ethically considered, are expensive, can take significant time to establish, do not always replicate the human *in vivo* situation and it is often not possible to study the cellular processes, particularly the initiation of disease. In contrast, our knowledge surrounding the growth of cells in 3D models has expanded greatly, along with the availability of the necessary tools, such that it is now possible, as mentioned above, to develop and manipulate models for multiple functional assays whilst retaining a structure similar to that found *in vivo*.

A recent publication demonstrated generation of 3D organoids from different parts of the body in a micro physiological system using patient-derived induced pluripotent stem cells, from the same or different donors and a microfluid device (Ramme *et al*., 2019) further illustrating the potential for such model culture systems.

In this study, we re-established and validated 3D salivary gland organoid models with normal cells (Rahman *et al*., 2019, unpublished data), (Maimets *et al*., 2016, Nanduri

et al., 2014) and were able to further use these techniques to develop 3D models of MEC cell lines and also normal cells transfected with the CRTC1-MAML2 fusion gene.

The primary cells were subjected to analysis and characterisation to re-confirm that they were normal primary cells. Our data suggest that the expression of epithelial, myoepithelial, tight junction proteins and stem cell markers confirms heterogeneous populations of epithelial and mesenchymal origin, similar in nature to the salivary gland cell components (Martinez-Madrigal and Micheau, 1989).

The presence of Wnt3 conditioned media is essential for organoid growth and expansion (Clevers *et al*., 2014). Wnt3 (Wingless-related integration site) is a member of the Wnt protein super family and thus implicated in multi-cell functions such as assisting progenitor adult stem cell proliferation, migration, cell-cell communication and self-renewal (Pinto *et al*., 2003). Depletion of Wnt3 in cultures can cause changes in cell morphology, mutation and subsequently cell death. Similarly, Roof plate-specific Spondin-1 (Rsp-1) is important for the growth of organoids acting as a regulator for Wnt signalling, stimulating epithelial cell proliferation and supporting epithelial stem cell healing and renewal (Clevers, 2016).

Growth supplements were essential for cell growth and differentiation, including epidermal growth factor (EGF) and fibroblast growth factor (FGF) (Dignass and Sturm, 2001). EGF is potent stimulator of stem cell genesis and epithelial cell growth and proliferation, while FGF is important for salivary gland organisation and survival (Mariz *et al*., 2018, Patel *et al*., 2011); the exact mechanism of action for these functions remains unclear. Similarly, as discussed in section 4.5, the ROCK inhibitor

(Y-27632) is essential during the early stage of culture for progenitor cells to proliferate and survive without changing phenotypically (Holm *et al*., 2013).

The use of Matrigel as an ECM substitute not only provides a physical support for the cells to grow and differentiate into a 3D structure, but also contains a number of ECM components including Laminin, Collagen IV, Heparin Sulfate Proteoglycans, entactin/nidogen, and also angiogenic and adipogenic growth factors (FGF2 and VEGF-A) (Hughes *et al*., 2010).

Histological analysis allowed us to confirm that, after seeding into Matrigel, the normal salivary gland cells proliferated and differentiated into mini gland-like structures. Previous data from our laboratory demonstrated gene expression of detectable levels of amylase (serous acinus secretion), Aquaporin 5 (acinus cell marker) and MUC7 (mucous acinus secretion) (Rahman *et al*., 2018, unpublished data). These markers of normal human salivary gland development and differentiation indicate that the models would likely maintain their physiological functions. Positive transcript signals for CK5, αSMA and epithelial cell markers were also detected suggesting that the 3D structures are more physiologically relevant than 2D cultures, show a higher degree of structural complexity and, due to modelling of the microenvironment, retain the homeostasis of the organ of origin.

Growing 3D models at ALI allowed us to more closely mimic the *in vivo* situation; the models were placed on a porous membrane that allowed the basal exchange of nutrients and apically, cells embedded in Matrigel, were exposed to the air. The ALI

culture condition was previously described by Su *et al*. (2016) in the development of a 3D slice organotypic culture of human submandibular gland.

We demonstrated, in some organoid structures, the presence or initiation of a lumen towards the centre (figure 6.4, B), not at the periphery where cells were often compact. Detailed analysis using IHC for CK7 (ductal cell marker) is, however, needed to confirm whether these structures were truly the beginnings of a duct. Further analysis would also allow us to investigate the functionality of 3D models through secretion of amylase.

Creating the translocation genetic event by transiently transfecting normal salivary gland cells could allow us to model, in 3D, the initiation and pathogenesis of MEC in normal cells. Our preliminary results indicated that the 3D organoid produced from the transfected cells was distinctive from its normal counterpart. Transfection of normal sublingual gland cells prior to development of an organoid, could affect cell behaviour and organisation, and so the cells were randomly mixed in ECM. Small organoid structures within these models (figure 6.6), suggested that after transfection the cell population was a mixture of transfected and non-transfected cells; the normal (non-transfected) cells maintaining organisation in small organoids. Further work is needed to confirm the identity of the cells in the 3D structures and to detect expression of the fusion protein within these models.

Matrigel is derived from the mouse Engelbreth-Holm-Swarm sarcoma (EHS) whilst Myogel, an alternative ECM, is derived from a benign human leiomyoma (smooth

muscle) (Salo *et al*., 2015). Data from our previous study showed that Matrigel was superior to Myogel, as a culture matrix, in that it provided the necessary physical support for organoid development and also that it resulted in more ECM being produced. Matrigel allowed us to develop 3D organoids with cells embedded, whereas cells tended to "fall through" the Myogel and thus only form a monolayer. Our data were consistent with observations from similar studies indicating that Matrigel encourages the growth and differentiation of human sublingual and parotid cells, either used separately or when mixed with collagen (Joraku *et al*., 2007, Maria *et al*., 2011). Despite our reservations over the use of an animal product in the development of a human cell model we thus adopted Matrigel as an ECM for our 3D *in vitro* models but continue to investigate alternative matrices which are either derived from human tissues or are synthetic. Other challenges arising from the use of a bio-active scaffold such as Matrigel include the batch-to-batch variations or interference with cell behaviour. Matrigel is synthesised from mouse sarcoma, which severely limits clinical applications in humans. As an alternative, modifications could be made to the Myogel to alter its rigidity and physical properties and thereby allow this ECM to truly support cell growth as 3D structures. A previous study suggested that this could be achieved by adding collagen and other substitutes (Salo *et al*., 2015) but, due to time constraints, we have not yet been able to replicate this in our laboratory.

We adopted the same techniques developed for growth of normal salivary gland cells as 3D organoids for the development of organoids from established MEC cell lines (table 2.21, section 2.8.3,). This is the first time such models have been described. A prior study using fresh, poorly differentiated MEC samples from seven patients, grown

in 2D or encapsulated in alginate gel microcapsules as 3D structures, demonstrated distinctive gene expression profiles for tumour angiogenesis and modelling of the tumour hypoxic microenvironment (Yang and Guo, 2017). In 3D, cells proliferated more actively and showed significantly higher potential for the promotion of tumour angiogenesis, via the secretion of VEGF-A a key angiogenesis factor, and a significant decrease of TSP-1 a key anti-angiogenic factor. Furthermore, the 3D model exhibited significantly greater expression of the hypoxia marker, HIF1α. (Yang and Guo, 2017). These results indicate that poorly differentiated MEC cells grown in 3D express key molecules important in cancer cell tumourigenesis and survival.

H&E staining of our 3D organoids revealed structural similarity to MEC tumours, with the presence of large numbers of cells organised in varying sizes of cluster; this was contrary to the H&E staining of normal salivary gland organoids. PAS staining demonstrated that the models not only maintained histological type, but also differentiated to produce mucin thereby mimicking histological sections of MEC. PAS and mucicarmine staining are considered diagnostic aids in the identification of mucous cells in MECs, especially in high-grade cases, where it is often difficult to identify mucous cells using routine H&E staining.

Immunofluorescence staining and confocal microscopy analysis showed organisation and formation of actin filaments within cells, however, more markers would enable the identification of other proteins such as mucins, epithelial markers and tumour suppression markers. Differential expression of a number of classes of membrane-bound mucins (MUC1, MUC2, MUC4 and MUC5AC) has been demonstrated in MECs, where MUC1 and MUC4 were correlated with specific differentiation grades

of MEC, MUC5AC was expressed by all MEC grades and MUC2 was not expressed in MECs (Robinson *et al*., 2020). Cytokeratin 14 and p63 are expressed by epidermoid and basal cells in MEC, while strong expression of the proliferating cell nuclear antigen (PCNA) and p53 are related to high-grade MEC (Namboodiripad, 2014). Also, intense staining of some inflammatory cytokines such as tumour necrosis factor alpha (TNF α) and deletion in malignant brain tumour one (DMBT1) occurs in MEC, where they are thought to act as tumour suppressor genes (Namboodiripad, 2014).

Two of our 3D MEC models were derived from metastatic tumours, and thus could provide a useful tool for testing chemotherapy and targeted therapy for advanced cases. Previous data has demonstrated that cells grown in 3D structures had a slower proliferation rate, maintained their phenotypic characteristics, and their histology bore closer resemblance to the primary tumours than cells grown in 2D (Lee *et al*., 2013). 3D models from 31 epithelial ovarian cancer cell lines were established and characterised and shown to be more representative of biological, histological and molecular features of the primary tumours than the 2D cultures. In addition, differences in sensitivity to chemotherapeutic reagents were demonstrated.

In future investigations, it might be possible to use the MEC 3D models to create a mutation in the CRTC1-MAML2 fusion using gene editing CRISPR Cas-9, to study the effect of the fusion interfering with neoplastic cells. Complete silencing of the endogenous translocation gene (the junction between exon1 in the CRTC1 gene and exon 2 in the MAML2 gene), would allow a better understanding of the loss of function of the fusion protein in 3D models and could provide downstream target candidates.

An example of this is a study done by Firth and colleagues (2016), who conducted functional gene editing for cystic fibrosis in lung epithelial cells generated from patient iPSC. This study demonstrated the use of the CRISPR Cas-9 technology to correct the mutation (deletion in F508) in the cystic fibrosis transmembrane regulator gene (CFTR) which naturally happened in cystic fibrosis patients. This correction was significantly precise and efficient and led to the differentiation and maturation of the epithelial airway cells, with the expression of normal CFTR (Firth *et al*., 2015).

The ultra-low passage approach used in this study is more representative of the patient situation and opens opportunities for testing patient-specific chemotherapy, targeted therapy, genetic signatures and mutations. However, further optimisation is needed before the ideal organoid that truly replicates the *in vivo* situation can be developed. Factors which need to be considered include geometry, mechanics and chemistry while the inclusion of a blood supply and innervation also need to be considered. This would allow a truer representation of physiological function.

The 3D model we have described, however, can still be considered an invaluable tool that could accelerate translation of research in cancer biology. We strongly believe that the low-passage human primary cells, or those taken directly from patients, can be used to further our knowledge of tumorigenesis and the pathogenesis of salivary gland tumours but could also be applied to many other disciplines.

6.5.1 Conclusion

In this study we have developed 3D organoid models of normal primary salivary gland cells in ALI- culture. The organoids grow in a more spherical and regular pattern and showed signs of the initiation of lumen formation. Furthermore, we were able to develop a 3D organoid model of CRTC1-MAML2 transiently transfected normal primary salivary gland cells in ALI-culture, and this enabled us to compare the histological organisation of normal and transfected 3D models. We were also able to develop 3D organoid models from established MEC cell lines.

6.5.2 Future Directions

- Perform immunohistochemical analysis of 3D models using specific markers of normal salivary glands such as, amylase, MUC7 (secretion markers) to confirm the differentiation status and function of the cells within the models.
- Perform a BaseScope assay to detect the spatial expression of the CRTC1-MAML2 transcript in 3D models of transiently transfected HuSL cells.
- Perform protein expression analysis using western blotting to investigate differences between 3D models of normal and transfected salivary gland cells (HuSL).
- Scanning electron microscopic (EM) analysis of 3D models will identify ultra-structures and fully investigate the cellular differences between normal and transfected HuSL. This will help us to further understand the true structure of our cultures.
- Matrigel is a murine ECM and an alternative animal-free ECM would allow further investigation in relation to human studies and has the potential to provide a

regenerative tool for use in the clinic. Future studies should, therefore, investigate ECM derived from human tissue or synthetic products which are biodegradable.

Chapter 7: General Discussion

7. Chapter 7

7.1 Discussion

Cancer is a major cause of death and the most significant obstacle to increasing life expectancy in the 21st century (WHO statistics, 2018). Despite their rarity, salivary gland cancers are a significant cause of death in the population, with 40% and 30% five year survival rates for patients with stage 4 cancer in major and minor salivary glands, respectively (https://www.cancerresearchuk.org/about-cancer/salivary-gland-cancer/survival, accessed in October 2020). In addition, according to the predicted figures from the International Agency of Cancer Research (WHO, 2018), the estimated number of new cases and deaths of salivary gland cancer will increase in the next twenty years (figure 7.1). Amongst these cancers, MEC is considered the most common salivary gland cancer among children, adolescents and in the adult population (Jones *et al*., 2008).

The presence of the unique t(11;19)(21q-p13) translocation in a high proportion of MEC, which leads to the formation of the CRTC1-MAML2 fusion protein, has attracted the attention of scientists around the world to investigate the role of this oncoprotein in the tumourigenesis of MEC. Understanding the function of this fusion protein could help with the development of novel targeted therapeutics to treat advanced fusion positive cases with poor clinical outcome. Moreover, detection of the CRTC1-MAML2 translocation could help with accurate diagnosis in identifying those cases that harbour the CRTC1-MAML2 fusion.

The aim of the work outlined in this thesis was to develop a novel *in situ* technique to detect the expression of the CRTC1-MAML2 transcript in MEC tissue samples and to perform an *in vitro* investigation into the role of the fusion protein in the tumourigenesis of MEC. Multiple approaches were utilised in this study to assist in elucidating its function.

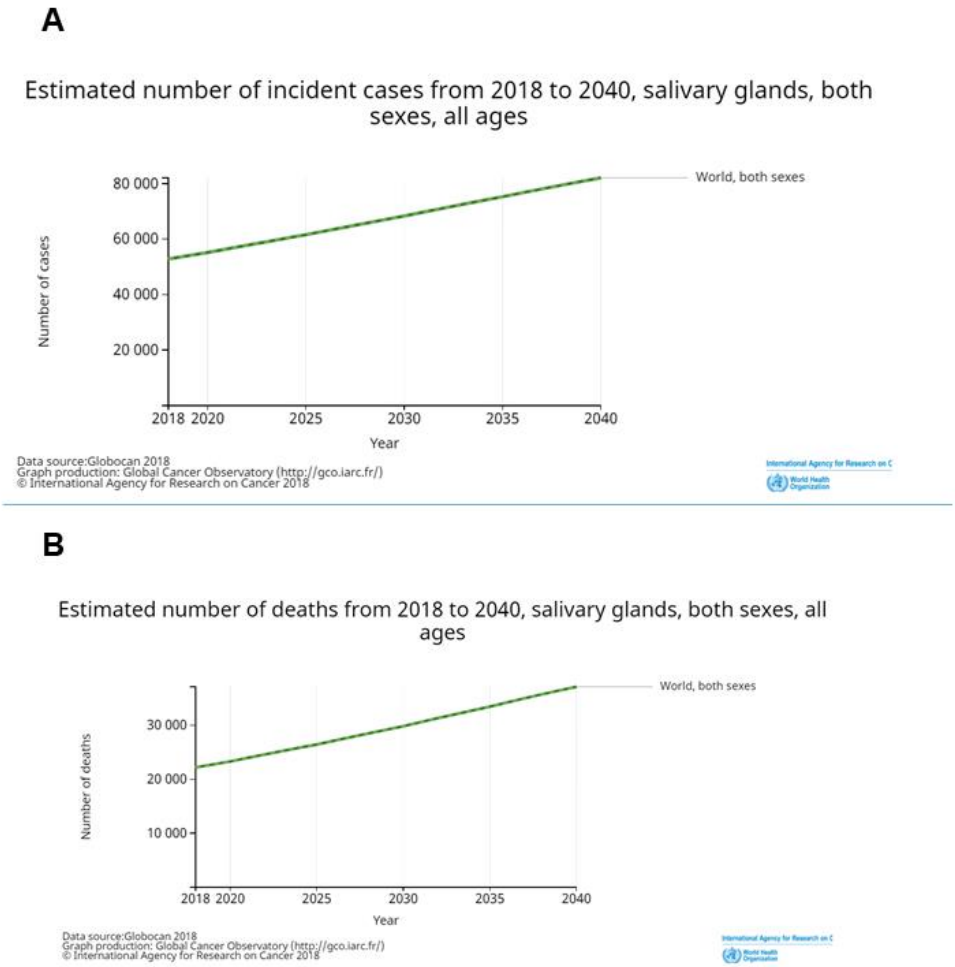


Figure 7.1 Line graph of the estimated number of incidence cases of salivary gland cancers from 2018 to 2040, in males and females and across all ages **(A)**. The estimated number of deaths of salivary gland cancers from 2018 to 2040, both sexes, all ages **(B)**. Data from Globocan, cancer tomorrow, accessed in October 2020.

In chapter 3, we showed for the first time the development and application of the BaseScope BA-Hs-CRTC1-MAML2-FJ probe, which allowed us to visualise and quantify the CRTC1-MAML2 transcript in MEC histological samples. Although the expression of the CRTC1-MAML2 translocation is typically detected via other techniques (RT-PCR and/or FISH), our novel approach could prove to be superior and more efficient, however, further study is necessary to confirm this. Our data indicated that the expression pattern of the CRTC1-MAML2 fusion in tumours varied among the cases, and most of the cases scored low in terms of the percentage of cells carrying the translocation.

Our chapter 3 results inspired us to further investigate whether this translocation gene is an innocent bystander or a tumour promoter. Although prior data from others suggested that the translocation *per se* is responsible for the generation of the tumour, irrespective of the exact cell type in which the fusion protein is expressed, it was necessary to develop an *in vitro* tool to help improve our understanding of the role of the CRTC1-MAML2 fusion protein. The first step towards this model was to establish an efficient, cost-effective and reproducible method for transfecting mammalian cell lines and primary salivary gland cells to enable *de novo* expression of the fusion gene.

In chapter 4, the CRTC1-MAML2 fusion construct was established, transfected, transcribed and translated in relevant cells. One of the more significant findings to emerge from this study relates to the upregulation of the DUSP1, STC1, NR4A2, and ATF3 genes when CRTC1-MAML2 was expressed in the HuSL cells, as these genes have been implicated in a variety of cellular processes affecting cell growth, such as proliferation, survival, differentiation and apoptosis. We were unsuccessful in our

attempts to generate a stable cell line expressing the fusion gene using primary salivary gland cells. As we believe this would be an extremely useful tool in our future analysis we aim to continue these studies beyond this thesis. As a substitute for stably transfected cells, we transiently transfected cells to further investigate the role of the CRTC1-MAML2 on cell behaviour with the results being presented in chapter 5.

Growing cells in monolayer or suspension is not a true representation of cellular environment. Therefore, we felt it was essential to further examine our *in vitro* tool through the development of 3D models as presented in chapter 6.

7.2 Study Strengths

The present study provided a comprehensive investigation into the novel CRTC1-MAML2 fusion protein incorporating clinical data, histological tumour samples, cellular and molecular studies. The use of multiple techniques ranging from gene to protein to functional experiments and 3D models, which led to the validation of our results through independent means, represents one of the main strengths of this study (The project story diagram, figure 7.2). Additionally, the use of primary human salivary gland cells lines for the first time allowed comparisons to be made between previous published studies using transformed cell lines and normal cells.

In this study we developed an *in vitro* tool to improve our understanding of the role of CRTC1-MAML2 fusion protein and identify significant downstream targets contributing to the development of MEC. Designing a specific CRTC1-MAML2 fusion TaqMan probe allowed us to accurately detect the expression of the CRTC1-MAML2 transcript in our transfected primary cells, mammalian cell lines and control cells.

Moreover, the use of fusion positive MEC cells in this study has enabled us to investigate the expression pattern of the fusion *in situ* and in solution, clone the fusion construct, and validate our experiments.

7.3 Study Limitations

Our clinical cohort was limited to a small sample size (N=29). This was restricted by the number of cases accessible to us. It would be beneficial to examine a larger cohort and, as the majority of the cases in our cohort were low grade (N= 21), include more high and intermediate grade cases. Notwithstanding the relatively limited sample, this work offers valuable insights into visualising the CRTC1-MAML2 transcript in MEC histological samples and could provide an alternative diagnostic technique to detect the translocation.

Our study has shown for the first time the application of an RNA probe that can specifically detect the CRTC1-MAML2 fusion and our data suggests that the probe is very specific and is able to detect the translocation in FFPE histological samples. Nonetheless, further research is needed to validate and determine the effectiveness of our novel technique against the currently available gold standard techniques (RT-PCR and FISH).

Our results with HuSL cells indicated that the exogenous expression of the CRTC1-MAML2 fusion significantly altered the expression of CREB candidate genes, however, the study could be further supported by validation of candidate gene experiments on a panel of primary cells, especially parotid and submandibular gland cells, in which MEC is more frequently observed than in sublingual glands. A further

issue that was not addressed in our study was the downstream targets of NOTCH signalling; despite this limitation the study certainly adds to our understanding of the effect of the CRTC1-MAML2 on normal primary human sublingual gland cells.

Whilst this study involved only two mammalian cell lines with which to study the effect of CRTC1-MAML2 on cell behaviour, our findings suggest that the fusion protein significantly influences tumourigenesis through effects on cell survival, invasion and migration.

The project story

Development of a BaseScope RNA probe for CRTC1-MAML2 fusion detection in MEC cases and development of an *in vitro* tool for investigating into the role of the CRTC1-MAML2 fusion protein.

Chapter 3 **Designing BA-Hs-CRTC1-MAML2-FJ probe**

1. Validating the probe on MEC fusion positive cell lines

   Probe was specific to the translocation transcript

2. Applying the novel technique on MEC tissue cases.

   Probe specifically detected the translocation in 34% MEC cases. Only a low number of the cells were harbouring the fusion.

Chapter 4 **Development CRTC1-MAML2 *In vitro* tool**

1. Cloning the CRTC1-MAML2 fusion from MEC cell lines.
2. Optimising an efficient transient transfection technique for mammalian cells.
3. Confirming the protein expression and colocalising in the cells.
4. Designing a fusion specific probe for qPCR transcript analysis.
5. Studying the downstream effect of the CRTC1-MAML2 transcript on primary salivary gland cells.

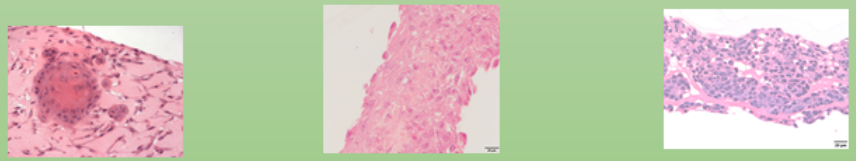
Chapter 5 **The effect of the CRTC1-MAML2 fusion protein on cell behaviour**

Using the in vitro tool and strategy developed in chapter 4 to perform cell functional assays

Migration assay | **Invasion assay** | **Proliferation assay** | **Clonogenic assay**

Chapter 6 **Development of *in vitro* organoid three-dimensional models**

Using the in vitro tool and strategy developed in chapter 4 to perform 3D organoids of normal, transfected and cancer cell lines



Normal HuSL cells | C1M2 TF HuSL cells | MEC cells

Figure 7.2 The project story.

7.3 Clinical Implications

The BaseScope assay described in this thesis is a method that could easily be adopted and applied in clinical laboratories to help diagnose and classify fusion positive cases for any potential future therapeutics. With standardisation in tissue preparation and fixation, by following the guidelines and avoiding the over fixation of tissues, (between 6 hours and 72 hours) this assay can easily be performed.

The work described in this thesis suggests that the CRTC1-MAML2 fusion is not simply an innocent bystander. Our preliminarily results support a role for CRTC1-MAML2 fusion in tumourigenesis via inducing the expression of downstream genes in HuSL cells and influencing cell behaviour in mammalian cell lines. Following our successful transfection experiments, we now aim to use these strategies and primary cells in future studies to identify and target pathways.

The MEC 3D models established and analysed in this thesis will be involved in ongoing research and will assist the testing of new therapeutics for patients with advanced disease.

7.4 Final Conclusion

Through this thesis, I was able to develop a novel alternative diagnostic technique that will assist the detection of the CRTC1-MAML2 transcript in MEC clinical cases. The CRTC1-MAML2 transcript appears to have an effect on the expression of the CREB target genes in cultured monolayer normal human primary sublingual gland cells. The 3D culture used in this study allowed the growth of cells after transfection with the fusion gene, which enabled the initial stages of the disease process to be

investigated not only in monolayer cultures but also in the more physiologically relevant 3D models where cell-cell and cell-matrix interactions and effects of the fusion gene can be studied. This 3D system could provide an effective tool for the study not only of the pathogenesis of disease but also the application of new therapeutics.

Overall, the data of this thesis indicate that although CRTC1-MAML2 expression was low in most of the cases studied, it could be implicated in driving tumourigenicity via upregulation of CREB targets genes in primary cells and influencing cell migration, invasion and survival, without influencing cell proliferation. Further investigation is therefore required to explore the exact mechanistic role of the CRTC1-MAML2 in the maintenance of MECs.

8. References

ALTAREJOS, J. Y. & MONTMINY, M. 2011. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. *Nat Rev Mol Cell Biol,* 12**,** 141-51.

AMAN, P. Fusion genes in solid tumors. Seminars in cancer biology, 1999. Elsevier, 303-318.

AMELIO, A. L., FALLAHI, M., SCHAUB, F. X., ZHANG, M., LAWANI, M. B., ALPERSTEIN, A. S., SOUTHERN, M. R., YOUNG, B. M., WU, L. & ZAJAC-KAYE, M. 2014. CRTC1/MAML2 gain-of-function interactions with MYC create a gene signature predictive of cancers with CREB–MYC involvement. *Proceedings of the National Academy of Sciences,* 111**,** E3260-E3268.

ANDERSSON, S., DAVIS, D., DAHLBÄCK, H., JÖRNVALL, H., RUSSELL & DW 1989. Cloning, structure, and expression of the mitochondrial cytochrome P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme. *Journal of Biological Chemistry,* 264**,** 8222-8229.

ANZICK, S. L., CHEN, W. D., PARK, Y., MELTZER, P., BELL, D., EL-NAGGAR, A. K. & KAYE, F. J. 2010. Unfavorable prognosis of CRTC1-MAML2 positive mucoepidermoid tumors with CDKN2A deletions. *Genes Chromosomes Cancer,* 49**,** 59-69.

ARGANI, P. & LADANYI, M. 2003. Recent advances in pediatric renal neoplasia. *Advances in anatomic pathology,* 10**,** 243-260.

AUCLAIR, P. L., GOODE, R. K. & ELLIS, G. L. 1992. Mucoepidermoid carcinoma of intraoral salivary glands evaluation and application of grading criteria in 143 cases. *Cancer,* 69**,** 2021-2030.

BAKER, A. M., HUANG, W., WANG, X. M., JANSEN, M., MA, X. J., KIM, J., ANDERSON, C. M., WU, X., PAN, L., SU, N., LUO, Y., DOMINGO, E., HEIDE, T., SOTTORIVA, A., LEWIS, A., BEGGS, A. D., WRIGHT, N. A., RODRIGUEZ-JUSTO, M., PARK, E., TOMLINSON, I. & GRAHAM, T. A. 2017. Robust RNA-based in situ mutation detection delineates colorectal cancer subclonal evolution. *Nat Commun,* 8**,** 1998.

BAO, B., AHMAD, A., AZMI, A. S., ALI, S. & SARKAR, F. H. 2013. Overview of cancer stem cells (CSCs) and mechanisms of their regulation: implications for cancer therapy. *Curr Protoc Pharmacol,* Chapter 14**,** Unit 14.25.

BARNES, L. 2005. *Pathology and genetics of head and neck tumours*, IARC.

BARNES, L., EVESON, J. W., SIDRANSKY, D. & REICHART, P. 2005. *Pathology and genetics of head and neck tumours*, IARC.

BEAN, G. R., KRINGS, G., OTIS, C. N., SOLOMON, D. A., GARCÍA, J. J., VAN ZANTE, A., CAMELO-PIRAGUA, S., VAN ZIFFLE, J. & CHEN, Y. Y. 2019. CRTC1-MAML2 fusion in mucoepidermoid carcinoma of the breast. *Histopathology,* 74**,** 463-473.

BEHBOUDI, A., ENLUND, F., WINNES, M., ANDRÉN, Y., NORDKVIST, A., LEIVO, I., FLABERG, E., SZEKELY, L., MÄKITIE, A. & GRENMAN, R. 2006. Molecular classification of mucoepidermoid carcinomas—prognostic significance of the MECT1–MAML2 fusion oncogene. *Genes, Chromosomes and Cancer,* 45**,** 470-481.

BELL, D. & EL-NAGGAR, A. K. 2013. Molecular heterogeneity in mucoepidermoid carcinoma: conceptual and practical implications. *Head Neck Pathol,* 7**,** 23-7.

BELL, D. & HANNA, E. Y. 2012. Salivary gland cancers: biology and molecular targets for therapy. *Curr Oncol Rep,* 14**,** 166-74.

BIRKELAND, A. C., FOLTIN, S. K., MICHMERHUIZEN, N. L., HOESLI, R. C., ROSKO, A. J., BYRD, S., YANIK, M., NOR, J. E., BRADFORD, C. R., PRINCE, M. E., CAREY, T. E., MCHUGH, J. B., SPECTOR, M. E. & BRENNER, J. C. 2017. Correlation of Crtc1/3-Maml2 fusion status, grade and survival in mucoepidermoid carcinoma. *Oral Oncol,* 68**,** 5-8.

BOSHART, M., WEBER, F., JAHN, G., DORSCH-H, K., FLECKENSTEIN, B. & SCHAFFNER, W. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *cell,* 41**,** 521-530.
BRENNER, J. C. & CHINNAIYAN, A. M. 2009. Translocations in epithelial cancers. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer,* 1796**,** 201-215.
Cancer Research UK, (2020), Cancer Research UK. [online] Available at: http://www.cancerresearchuk.org/ Accessed 3 October 2020.
Cancer Tommorw IARC Globocan (2018), [online] Available at: https://gco.iarc.fr/tomorrow/graphic-line? Accessed 3 October 2020.
CANETTIERI, G., CONI, S., DELLA GUARDIA, M., NOCERINO, V., ANTONUCCI, L., DI MAGNO, L., SCREATON, R., SCREPANTI, I., GIANNINI, G. & GULINO, A. 2009. The coactivator CRTC1 promotes cell proliferation and transformation via AP-1. *Proc Natl Acad Sci U S A,* 106**,** 1445-50.
CHAPMAN, S., LIU, X., MEYERS, C., SCHLEGEL, R. & MCBRIDE, A. A. 2010. Human keratinocytes are efficiently immortalized by a Rho kinase inhibitor. *The Journal of clinical investigation,* 120**,** 2619-2626.
CHAPMAN, S., MCDERMOTT, D. H., SHEN, K., JANG, M. K. & MCBRIDE, A. A. 2014. The effect of Rho kinase inhibition on long-term keratinocyte proliferation is rapid and conditional. *Stem cell research & therapy,* 5**,** 60.
CHEN, J., LI, J.-L., CHEN, Z., GRIFFIN, J. D. & WU, L. 2015. Gene expression profiling analysis of CRTC1-MAML2 fusion oncogene-induced transcriptional program in human mucoepidermoid carcinoma cells. *BMC cancer,* 15**,** 803.
CHEN, Z., CHEN, J., GU, Y., HU, C., LI, J., LIN, S., SHEN, H., CAO, C., GAO, R. & LI, J. 2014. Aberrantly activated AREG–EGFR signaling is required for the growth and survival of CRTC1–MAML2 fusion-positive mucoepidermoid carcinoma cells. *Oncogene,* 33**,** 3869-3877.
CHEN, Z., LIN, S., LI, J. L., NI, W., GUO, R., LU, J., KAYE, F. J. & WU, L. 2018. CRTC1-MAML2 fusion-induced lncRNA LINC00473 expression maintains the growth and survival of human mucoepidermoid carcinoma cells. *Oncogene,* 37**,** 1885-1895.
CLEVERS, H. 2016. Modeling development and disease with organoids. *Cell,* 165**,** 1586-1597.
CLEVERS, H., LOH, K. M. & NUSSE, R. 2014. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *science,* 346.
COLEMAN, M. L., SAHAI, E. A., YEO, M., BOSCH, M., DEWAR, A. & OLSON, M. F. 2001. Membrane blebbing during apoptosis results from caspase-mediated activation of ROCK I. *Nature cell biology,* 3**,** 339-345.
COSTA, M. R., SUGITA, D. M., VILELA, M. H., DA SILVA MENDONÇA, R. P., DE MORAIS, D. T., JÚNIOR, P. C., COSTA, T. R. & BARREIRA, B. M. 2015. Mucoepidermoid carcinoma of the penis: Case report and literature review. *Can Urol Assoc J,* 9**,** E27-9.
COXON, A., ROZENBLUM, E., PARK, Y.-S., JOSHI, N., TSURUTANI, J., DENNIS, P. A., KIRSCH, I. R. & KAYE, F. J. 2005. Mect1-Maml2 fusion oncogene linked to the aberrant activation of cyclic AMP/CREB regulated genes. *Cancer research,* 65**,** 7137-7144.
DALEY, W. P., GERVAIS, E. M., CENTANNI, S. W., GULFO, K. M., NELSON, D. A. & LARSEN, M. 2012. ROCK1-directed basement membrane positioning coordinates epithelial tissue polarity. *Development,* 139**,** 411-422.
DALEY, W. P., GULFO, K. M., SEQUEIRA, S. J. & LARSEN, M. 2009. Identification of a mechanochemical checkpoint and negative feedback loop regulating branching morphogenesis. *Developmental biology,* 336**,** 169-182.
DALEY, W. P., KOHN, J. M. & LARSEN, M. 2011. A focal adhesion protein-based mechanochemical checkpoint regulates cleft progression during branching morphogenesis. *Developmental Dynamics,* 240**,** 2069-2083.

DAVIES, K. D., LE, A. T., SHEREN, J., NIJMEH, H., GOWAN, K., JONES, K. L., VARELLA-GARCIA, M., AISNER, D. L. & DOEBELE, R. C. 2018. Comparison of molecular testing modalities for detection of ROS1 rearrangements in a cohort of positive patient samples. *Journal of Thoracic Oncology,* 13**,** 1474-1482.

DIGNASS, A. U. & STURM, A. 2001. Peptide growth factors in the intestine. *Eur J Gastroenterol Hepatol,* 13**,** 763-70.

EL-NAGGAR, A. K. 2017. What is new in the World Health Organization 2017 histopathology classification? *Current treatment options in oncology,* 18**,** 43.

ELLIS, G. & AUCLAIR, P. 2008. Tumors of the salivary glands: AFIP Atlas of tumor pathology, series 4. *Washington: Armed Forces Institute of Pathology (AFIP)*.

ENLUND, F., BEHBOUDI, A., ANDRÉN, Y., OBERG, C., LENDAHL, U., MARK, J. & STENMAN, G. 2004. Altered Notch signaling resulting from expression of a WAMTP1-MAML2 gene fusion in mucoepidermoid carcinomas and benign Warthin's tumors. *Exp Cell Res,* 292**,** 21-8.

EVIATAR, J. A. & HORNBLASS, A. 1993. Mucoepidermoid carcinoma of the lacrimal gland: 25 cases and a review and update of the literature. *Ophthalmic plastic and reconstructive surgery,* 9**,** 170-181.

FEHR, A., RÖSER, K., HEIDORN, K., HALLAS, C., LÖNING, T. & BULLERDIEK, J. 2008. A new type of MAML2 fusion in mucoepidermoid carcinoma. *Genes Chromosomes Cancer,* 47**,** 203-6.

FENG, Y., WANG, Y., WANG, Z., FANG, Z., LI, F., GAO, Y., LIU, H., XIAO, T., ZHOU, Y., ZHAI, Q., LIU, X., SUN, Y., BARDEESY, N., WONG, K. K., CHEN, H., XIONG, Z. Q. & JI, H. 2012. The CRTC1-NEDD9 signaling axis mediates lung cancer progression caused by LKB1 loss. *Cancer Res,* 72**,** 6502-11.

FIRTH, A. L., MENON, T., PARKER, G. S., QUALLS, S. J., LEWIS, B. M., KE, E., DARGITZ, C. T., WRIGHT, R., KHANNA, A., GAGE, F. H. & VERMA, I. M. 2015. Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. *Cell Rep,* 12**,** 1385-90.

FOOTE, F. W. & FRAZELL, E. L. 1954. *Tumors of the major salivary glands*, National Academies.

FRANKEN, N. A., RODERMOND, H. M., STAP, J., HAVEMAN, J. & VAN BREE, C. 2006. Clonogenic assay of cells in vitro. *Nat Protoc,* 1**,** 2315-9.

FREDRIKSSON, S. & BÜLOW, L. 2001. Chimeric Genes, Proteins.

GARCÍA, J. J., HUNT, J. L., WEINREB, I., MCHUGH, J. B., BARNES, E. L., CIEPLY, K., DACIC, S. & SEETHALA, R. R. 2011. Fluorescence in situ hybridization for detection of MAML2 rearrangements in oncocytic mucoepidermoid carcinomas: utility as a diagnostic test. *Hum Pathol,* 42**,** 2001-9.

GARTRELL, R., PAULI, J. & ZONTA, M. 2015. Primary cutaneous mucoepidermoid carcinoma: a case study with a review of the literature. *Int J Surg Pathol,* 23**,** 161-4.

GHOSH-LASKAR, S., MURTHY, V., WADASADAWALA, T., AGARWAL, J., BUDRUKKAR, A., PATIL, N., KANE, S., CHAUKAR, D., PAI, P., CHATURVEDI, P. & D'CRUZ, A. 2011. Mucoepidermoid carcinoma of the parotid gland: factors affecting outcome. *Head Neck,* 33**,** 497-503.

GOH, G. H., LIM, C. M., VANACEK, T., MICHAL, M. & PETERSSON, F. 2017. Spindle Cell Mucoepidermoid Carcinoma of the Palatine Tonsil With CRTC1-MAML2 Fusion Transcript: Report of a Rare Case in a 17-Year-Old Boy and a Review of the Literature. *Int J Surg Pathol,* 25**,** 705-710.

GOODWIN, E. C. & ROTTMAN, F. 1992. The 3'-flanking sequence of the bovine growth hormone gene contains novel elements required for efficient and accurate polyadenylation. *Journal of Biological Chemistry,* 267**,** 16330-16334.

GRADA, A., OTERO-VINAS, M., PRIETO-CASTRILLO, F., OBAGI, Z. & FALANGA, V. 2017. Research Techniques Made Simple: Analysis of Collective Cell Migration Using the Wound Healing Assay. *J Invest Dermatol,* 137**,** e11-e16.

GU, Y., LIN, S., LI, J., NAKAGAWA, H., CHEN, Z., JIN, B., TIAN, L., UCAR, D., SHEN, H. & LU, J. 2012. Altered LKB1/CREB-regulated transcription co-activator (CRTC) signaling axis promotes esophageal cancer cell migration and invasion. *Oncogene,* 31**,** 469-479.

GUPTA, R., BALASUBRAMANIAN, D. & CLARK, J. R. 2015. Salivary gland lesions: recent advances and evolving concepts. *Oral surgery, oral medicine, oral pathology and oral radiology,* 119**,** 661-674.

GUZZO, M., LOCATI, L. D., PROTT, F. J., GATTA, G., MCGURK, M. & LICITRA, L. 2010. Major and minor salivary gland tumors. *Crit Rev Oncol Hematol,* 74**,** 134-48.

HAMM, A., KROTT, N., BREIBACH, I., BLINDT, R. & BOSSERHOFF, A. K. 2002. Efficient transfection method for primary cells. *Tissue Eng,* 8**,** 235-45.

HAN, N. R., LEE, H., BAEK, S., YUN, J. I., PARK, K. H. & LEE, S. T. 2015. Delivery of episomal vectors into primary cells by means of commercial transfection reagents. *Biochem Biophys Res Commun,* 461**,** 348-53.

HAN, Y. F. & CAO, G. W. 2012. Role of nuclear receptor NR4A2 in gastrointestinal inflammation and cancers. *World J Gastroenterol,* 18**,** 6865-73.

HOLM, F., NIKDIN, H., KJARTANSDOTTIR, K. R., GAUDENZI, G., FRIED, K., ASPENSTRÖM, P., HERMANSON, O. & BERGSTRÖM-TENGZELIUS, R. 2013. Passaging techniques and ROCK inhibitor exert reversible effects on morphology and pluripotency marker gene expression of human embryonic stem cell lines. *Stem Cells and Development, 2*2**,** 1883-1892.

HU, H. J., ZHOU, R. X., LIU, F., WANG, J. K. & LI, F. Y. 2018. You cannot miss it: Pancreatic mucoepidermoid carcinoma: A case report and literature review. *Medicine (Baltimore),* 97**,** e9990.

HUGHES, C. S., POSTOVIT, L. M. & LAJOIE, G. A. 2010. Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics,* 10**,** 1886-90.

IORGULESCU, G. 2009. Saliva between normal and pathological. Important factors in determining systemic and oral health. *J Med Life,* 2**,** 303-7.

JANJUA, O. S., QURESHI, S. M., KHAN, T. S. & ALAMGIR, W. 2012. Bcl-2 protein expression in mucoepidermoid carcinoma of salivary glands: a single institution experience. *Hematology/Oncology and Stem Cell Therapy,* 5**,** 96-100.

JASKOLL, T., HTET, K., ABICHAKER, G., KAYE, F. J. & MELNICK, M. 2011. CRTC1 expression during normal and abnormal salivary gland development supports a precursor cell origin for mucoepidermoid cancer. *Gene Expr Patterns,* 11**,** 57-63.

JOHNSON, N. W., JAYASEKARA, P. & AMARASINGHE, A. 2011. Squamous cell carcinoma and precursor lesions of the oral cavity: epidemiology and aetiology. *Periodontology 2000,* 57**,** 19-37.

JOLLY, S., LANG, V., KOELZER, V. H., FRIGERIO, C. S., MAGNO, L., SALINAS, P. C., WHITING, P. & PALOMER, E. 2019. Single-Cell quantification of mRNA expression in the human brain. *Scientific reports,* 9**,** 1-9.

JONES, A., CRAIG, G., SPEIGHT, P. & FRANKLIN, C. 2008. The range and demographics of salivary gland tumours diagnosed in a UK population. *Oral oncology,* 44**,** 407-417.

JORAKU, A., SULLIVAN, C. A., YOO, J. & ATALA, A. 2007. In-vitro reconstitution of three-dimensional human salivary gland tissue structures. *Differentiation,* 75**,** 318-324.

JUST, P. A., CAZES, A., AUDEBOURG, A., CESSOT, A., PALLIER, K., DANEL, C., VACHER-LAVENU, M. C., LAURENT-PUIG, P., TERRIS, B. & BLONS, H.

2012. Histologic subtypes, immunohistochemistry, FISH or molecular screening for the accurate diagnosis of ALK-rearrangement in lung cancer: a comprehensive study of Caucasian non-smokers. *Lung Cancer,* 76**,** 309-15.
JUSTUS, C. R., LEFFLER, N., RUIZ-ECHEVARRIA, M. & YANG, L. V. 2014. In vitro cell migration and invasion assays. *J Vis Exp*.
KANG, H., TAN, M., BISHOP, J. A., JONES, S., SAUSEN, M., HA, P. K. & AGRAWAL, N. 2017. Whole-Exome Sequencing of Salivary Gland Mucoepidermoid Carcinoma. *Clin Cancer Res,* 23**,** 283-288.
KLEINMAN, H. K., MCGARVEY, M. L., LIOTTA, L. A., ROBEY, P. G., TRYGGVASON, K. & MARTIN, G. R. 1982. Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan from the EHS sarcoma. *Biochemistry,* 21**,** 6188-93.
KOMIYA, T., PARK, Y., MODI, S., COXON, A., OH, H. & KAYE, F. 2006. Sustained expression of Mect1-Maml2 is essential for tumor cell growth in salivary gland cancers carrying the t (11; 19) translocation. *Oncogene,* 25**,** 6128-6132.
KOSLOW, M., O'KEEFE, K. J., HOSSEINI, Z. F., NELSON, D. A. & LARSEN, M. 2019. ROCK inhibitor increases proacinar cells in adult salivary gland organoids. *Stem Cell Res,* 41**,** 101608.
KOZAK, M. 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proceedings of the National Academy of Sciences,* 87**,** 8301-8305.
KRAMER, N., WALZL, A., UNGER, C., ROSNER, M., KRUPITZA, G., HENGSTSCHLÄGER, M. & DOLZNIG, H. 2013. In vitro cell migration and invasion assays. *Mutat Res,* 752**,** 10-24.
LAURIE, S. A., HO, A. L., FURY, M. G., SHERMAN, E. & PFISTER, D. G. 2011. Systemic therapy in the management of metastatic or locally recurrent adenoid cystic carcinoma of the salivary glands: a systematic review. *The lancet oncology,* 12**,** 815-824.
LEE, J. M., MHAWECH-FAUCEGLIA, P., LEE, N., PARSANIAN, L. C., LIN, Y. G., GAYTHER, S. A. & LAWRENSON, K. 2013. A three-dimensional microenvironment alters protein expression and chemosensitivity of epithelial ovarian cancer cells in vitro. *Lab Invest,* 93**,** 528-42.
LEUNG, C. C. T. & WONG, C. K. C. 2020. Characterization of stanniocalcin-1 expression in macrophage differentiation. *Transl Oncol,* 14**,** 100881.
LIAO, Q., GUO, J., KLEEFF, J., ZIMMERMANN, A., BÜCHLER, M. W., KORC, M. & FRIESS, H. 2003. Down-regulation of the dual-specificity phosphatase MKP-1 suppresses tumorigenicity of pancreatic cancer cells. *Gastroenterology,* 124**,** 1830-1845.
LIN, L. C., ELKASHTY, O., RAMAMOORTHI, M., TRINH, N., LIU, Y., SUNAVALA-DOSSABHOY, G., PRANZATELLI, T., MICHAEL, D. G., CHIVASSO, C. & PERRET, J. 2018. Cross-contamination of the human salivary gland HSG cell line with HeLa cells: A STR analysis study. *Oral diseases,* 24**,** 1477-1483.
LIU, C., ZHAO, Y., CHU, W., ZHANG, F. & ZHANG, Z. 2015. Mucoepidermoid carcinoma of esophagus combined with squamous carcinoma of lung: A case report and literature review. *Journal of cancer research and therapeutics,* 11**,** 658.
LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods,* 25**,** 402-408.
MAIER, T., GÜELL, M. & SERRANO, L. 2009. Correlation of mRNA and protein in complex biological samples. *FEBS Lett,* 583**,** 3966-73.
MAIMETS, M., ROCCHI, C., BRON, R., PRINGLE, S., KUIPERS, J., GIEPMANS, B. N., VRIES, R. G., CLEVERS, H., DE HAAN, G. & VAN OS, R. 2016. Long-

term in vitro expansion of salivary gland stem cells driven by Wnt signals. *Stem cell reports,* 6**,** 150-162.
MARIA, O. M., MARIA, O., LIU, Y., KOMAROVA, S. V. & TRAN, S. D. 2011. Matrigel improves functional properties of human submandibular salivary gland cell line. *The international journal of biochemistry & cell biology,* 43**,** 622-631.
MARIZ, B. A. L. A., SOARES, C. D., MORAIS, T. M. L., FONSECA, F. P., DE CARVALHO, M. G. F. & JORGE, J. 2018. Expression of FGF-2/FGFR-1 in normal mucosa, salivary gland, preneoplastic, and neoplastic lesions of the oral cavity. *J Oral Pathol Med,* 47**,** 816-822.
MARK, J., DAHLENFORS, R., EKEDAHL, C. & STENMAN, G. 1980. The mixed salivary gland tumor—a normally benign human neoplasm frequently showing specific chromosomal abnormalities. *Cancer Genetics and Cytogenetics,* 2**,** 231-241.
MARTINEZ-MADRIGAL, F. & MICHEAU, C. 1989. Histology of the major salivary glands. *Am J Surg Pathol,* 13**,** 879-99.
MATTA, A. & RALHAN, R. 2009. Overview of current and future biologically based targeted therapies in head and neck squamous cell carcinoma. *Head & neck oncology,* 1**,** 6.
MENYHÁRT, O., HARAMI-PAPP, H., SUKUMAR, S., SCHÄFER, R., MAGNANI, L., DE BARRIOS, O. & GYŐRFFY, B. 2016. Guidelines for the selection of functional assays to evaluate the hallmarks of cancer. *Biochim Biophys Acta,* 1866**,** 300-319.
MIYABE, S., OKABE, M., NAGATSUKA, H., HASEGAWA, Y., INAGAKI, A., IJICHI, K., NAGAI, N., EIMOTO, T., YOKOI, M. & SHIMOZATO, K. 2009. Prognostic significance of p27Kip1, Ki-67, and CRTC1-MAML2 fusion transcript in mucoepidermoid carcinoma: a molecular and clinicopathologic study of 101 cases. *Journal of oral and maxillofacial surgery,* 67**,** 1432-1441.
MORAN, C. A. & SUSTER, S. 2003. Primary mucoepidermoid carcinoma of the pleura. A clinicopathologic study of two cases. *Am J Clin Pathol,* 120**,** 381-5.
NAKAYAMA, T., MIYABE, S., OKABE, M., SAKUMA, H., IJICHI, K., HASEGAWA, Y., NAGATSUKA, H., SHIMOZATO, K. & INAGAKI, H. 2009. Clinicopathological significance of the CRTC3–MAML2 fusion transcript in mucoepidermoid carcinoma. *Modern Pathology,* 22**,** 1575-1581.
NAMBOODIRIPAD, P. A. 2014. A review: Immunological markers for malignant salivary gland tumors. *Journal of oral biology and craniofacial research,* 4**,** 127-134.
NANDURI, L. S., BAANSTRA, M., FABER, H., ROCCHI, C., ZWART, E., DE HAAN, G., VAN OS, R. & COPPES, R. P. 2014. Purification and ex vivo expansion of fully functional salivary gland stem cells. *Stem Cell Reports,* 3**,** 957-64.
NORDKVIST, A., GUSTAFSSON, H., JUBERG-ODE, M. & STENMAN, G. 1994. Recurrent rearrangements of 11q14–22 in mucoepidermoid carcinoma. *Cancer genetics and cytogenetics,* 74**,** 77-83.
O'NEILL, I. D. 2009. t (11; 19) translocation and CRTC1-MAML2 fusion oncogene in mucoepidermoid carcinoma. *Oral oncology,* 45**,** 2-9.
OKABE, M., MIYABE, S., NAGATSUKA, H., TERADA, A., HANAI, N., YOKOI, M., SHIMOZATO, K., EIMOTO, T., NAKAMURA, S. & NAGAI, N. 2006. MECT1-MAML2 fusion transcript defines a favorable subset of mucoepidermoid carcinoma. *Clinical cancer research,* 12**,** 3902-3907.
OKUMURA, Y., MIYABE, S., NAKAYAMA, T., FUJIYOSHI, Y., HATTORI, H., SHIMOZATO, K. & INAGAKI, H. 2011. Impact of CRTC1/3–MAML2 fusions on histological classification and prognosis of mucoepidermoid carcinoma. *Histopathology,* 59**,** 90-97.
OKUMURA, Y., NAKANO, S., MURASE, T., UEDA, K., KAWAKITA, D., NAGAO, T., KUSAFUKA, K., URANO, M., YAMAMOTO, H. & KANO, S. 2020. Prognostic

impact of CRTC1/3-MAML2 fusions in salivary gland mucoepidermoid carcinoma: A multiinstitutional retrospective study. *Cancer science*.
OWENS, D. & KEYSE, S. 2007. Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases. *Oncogene,* 26**,** 3203-3213.
PANKOV, R. & YAMADA, K. M. 2002. Fibronectin at a glance. *J Cell Sci,* 115**,** 3861-3.
PATEL, N., SHARPE, P. T. & MILETICH, I. 2011. Coordination of epithelial branching and salivary gland lumen formation by Wnt and FGF signals. *Developmental biology,* 358**,** 156-167.
PELDEN, S., INSAWANG, T., THUWAJIT, C. & THUWAJIT, P. 2013. The trefoil factor 1 (TFF1) protein involved in doxorubicin-induced apoptosis resistance is upregulated by estrogen in breast cancer cells. *Oncology reports,* 30**,** 1518-1526.
PÉREZ-DE-OLIVEIRA, M. E., WAGNER, V. P., ARAÚJO, A. L. D., MARTINS, M. D., SANTOS-SILVA, A. R., BINGLE, L. & VARGAS, P. A. 2020. Prognostic value of CRTC1-MAML2 translocation in salivary mucoepidermoid carcinoma: Systematic review and meta-analysis. *J Oral Pathol Med,* 49**,** 386-394.
PINTO, D., GREGORIEFF, A., BEGTHEL, H. & CLEVERS, H. 2003. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & development,* 17**,** 1709-1713.
RAMME, A. P., KOENIG, L., HASENBERG, T., SCHWENK, C., MAGAUER, C., FAUST, D., LORENZ, A. K., KREBS, A. C., DREWELL, C., SCHIRRMANN, K., VLADETIC, A., LIN, G. C., PABINGER, S., NEUHAUS, W., BOIS, F., LAUSTER, R., MARX, U. & DEHNE, E. M. 2019. Autologous induced pluripotent stem cell-derived four-organ-chip. *Future Sci OA,* 5**,** FSO413.
ROBINSON, L., VAN HEERDEN, M. B., KER-FOX, J. G., HUNTER, K. D. & VAN HEERDEN, W. F. P. 2020. Expression of Mucins in Salivary Gland Mucoepidermoid Carcinoma. *Head Neck Pathol*.
RODEN, A. C., GARCÍA, J. J., WEHRS, R. N., COLBY, T. V., KHOOR, A., LESLIE, K. O. & CHEN, L. 2014. Histopathologic, immunophenotypic and cytogenetic features of pulmonary mucoepidermoid carcinoma. *Mod Pathol,* 27**,** 1479-88.
RUCHAUD-SPARAGANO, M.-H., WESTLEY, B. & MAY, F. 2004. The trefoil protein TFF1 is bound to MUC5AC in humangastric mucosa. *Cellular and molecular life sciences,* 61**,** 1946-1954.
RUEDINGER, F., LAVRENTIEVA, A., BLUME, C., PEPELANOVA, I. & SCHEPER, T. 2015. Hydrogels for 3D mammalian cell culture: a starting guide for laboratory practice. *Appl Microbiol Biotechnol,* 99**,** 623-36.
SALDANA, M. J. 1994. *Pathology of pulmonary disease*, Lippincott Williams & Wilkins.
SALO, T., SUTINEN, M., HOQUE APU, E., SUNDQUIST, E., CERVIGNE, N. K., DE OLIVEIRA, C. E., AKRAM, S. U., OHLMEIER, S., SUOMI, F., EKLUND, L., JUUSELA, P., ÅSTRÖM, P., BITU, C. C., SANTALA, M., SAVOLAINEN, K., KORVALA, J., PAES LEME, A. F. & COLETTA, R. D. 2015. A novel human leiomyoma tissue derived matrix for cell culture studies. *BMC Cancer,* 15**,** 981.
SASAHIRA, T., KIRITA, T. & KUNIYASU, H. 2014. Update of molecular pathobiology in oral cancer: a review. *International journal of clinical oncology,* 19**,** 431-436.
SCHUMACHER, Y., APARICIO, T., OURABAH, S., BARAILLE, F., MARTIN, A., WIND, P., DENTIN, R., POSTIC, C. & GUILMEAU, S. 2016. Dysregulated CRTC1 activity is a novel component of PGE2 signaling that contributes to colon cancer growth. *Oncogene,* 35**,** 2602-14.
SCHWARZ, S., STIEGLER, C., MÜLLER, M., ETTL, T., BROCKHOFF, G., ZENK, J. & AGAIMY, A. 2011. Salivary gland mucoepidermoid carcinoma is a clinically, morphologically and genetically heterogeneous entity: a clinicopathological

study of 40 cases with emphasis on grading, histological variants and presence of the t(11;19) translocation. *Histopathology,* 58**,** 557-70.
SEETHALA, R., HOSCHAR, A., BENNETT, A., ARROSSI, A., DAVISON, J., KRASINSKA, A. & HUNT, J. Reproducibility of grading in salivary gland mucoepidermoid carcinoma and correlation with outcome: does system really matter? LABORATORY INVESTIGATION, 2008. NATURE PUBLISHING GROUP 75 VARICK ST, 9TH FLR, NEW YORK, NY 10013-1917 USA, 241A-241A.
SEETHALA, R. R. 2009. An update on grading of salivary gland carcinomas. *Head Neck Pathol,* 3**,** 69-77.
SEETHALA, R. R., DACIC, S., CIEPLY, K., KELLY, L. M. & NIKIFOROVA, M. N. 2010. A reappraisal of the MECT1/MAML2 translocation in salivary mucoepidermoid carcinomas. *Am J Surg Pathol,* 34**,** 1106-21.
SHAROVA, L. V., SHAROV, A. A., NEDOREZOV, T., PIAO, Y., SHAIK, N. & KO, M. S. 2009. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res,* 16**,** 45-58.
SHEN, J., ZHANG, Y., YU, H., SHEN, B., LIANG, Y., JIN, R., LIU, X., SHI, L. & CAI, X. 2016. Role of DUSP1/MKP1 in tumorigenesis, tumor progression and therapy. *Cancer Med,* 5**,** 2061-8.
SHINOMIYA, H., ITO, Y., KUBO, M., YONEZAWA, K., OTSUKI, N., IWAE, S., INAGAKI, H. & NIBU, K.-I. 2016a. Expression of amphiregulin in mucoepidermoid carcinoma of the major salivary glands: a molecular and clinicopathological study. *Human Pathology,* 57**,** 37-44.
SHINOMIYA, H., ITO, Y., KUBO, M., YONEZAWA, K., OTSUKI, N., IWAE, S., INAGAKI, H. & NIBU, K. I. 2016b. Expression of amphiregulin in mucoepidermoid carcinoma of the major salivary glands: a molecular and clinicopathological study. *Hum Pathol,* 57**,** 37-44.
SOOD, S., MCGURK, M. & VAZ, F. 2016. Management of Salivary Gland Tumours: United Kingdom National Multidisciplinary Guidelines. *J Laryngol Otol,* 130**,** S142-S149.
SPEIGHT, P. M. & BARRETT, A. W. 2009. Diagnostic difficulties in lesions of the minor salivary glands. *Diagnostic histopathology,* 15**,** 311-317.
SPEIGHT, P. M. & BARRETT, A. W. 2020. Salivary gland tumours: diagnostic challenges and an update on the latest WHO classification. *Diagnostic Histopathology*.
STAEDEL, C., HUA, Z., BROKER, T. R., CHOW, L. T., REMY, J.-S. & BEHR, J.-P. 1994. High-efficiency transfection of primary human keratinocytes with positively charged lipopolyamine: DNA complexes. *Journal of investigative dermatology,* 102**,** 768-772.
STEFANOU, D., GOUSSIA, A. C., ARKOUMANI, E., METAFRATZI, Z. M., SYMINELAKIS, S. & AGNANTIS, N. J. 2004. Mucoepidermoid carcinoma of the thymus: a case presentation and a literature review. *Pathol Res Pract,* 200**,** 567-73.
STELIOU, K., BOOSALIS, M. S., PERRINE, S. P., SANGERMAN, J. & FALLER, D. V. 2012. Butyrate histone deacetylase inhibitors. *Biores Open Access,* 1**,** 192-8.
STENMAN, G. Fusion oncogenes and tumor type specificity—insights from salivary gland tumors. Seminars in cancer biology, 2005. Elsevier, 224-235.
STENMAN, G. 2013. Fusion oncogenes in salivary gland tumors: molecular and clinical consequences. *Head and neck pathology,* 7**,** 12-19.
STENMAN, G., PERSSON, F. & ANDERSSON, M. K. 2014. Diagnostic and therapeutic implications of new molecular biomarkers in salivary gland cancers. *Oral Oncol,* 50**,** 683-90.

STEPANENKO, A. A. & DMITRENKO, V. V. 2015. HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene,* 569**,** 182-90.
STEWART, F. W., FOOTE, F. W. & BECKER, W. F. 1945. Muco-epidermoid tumors of salivary glands. *Annals of Surgery,* 122**,** 820.
SU, X., FANG, D., LIU, Y., RAMAMOORTHI, M., ZEITOUNI, A., CHEN, W. & TRAN, S. D. 2016. Three-dimensional organotypic culture of human salivary glands: the slice culture model. *Oral Dis,* 22**,** 639-48.
THOMPSON, M. R., XU, D. & WILLIAMS, B. R. 2009. ATF3 transcription factor and its emerging roles in immunity and cancer. *J Mol Med (Berl),* 87**,** 1053-60.
TIRADO, Y., WILLIAMS, M. D., HANNA, E. Y., KAYE, F. J., BATSAKIS, J. G. & EL-NAGGAR, A. K. 2007. CRTC1/MAML2 fusion transcript in high grade mucoepidermoid carcinomas of salivary and thyroid glands and Warthin's tumors: implications for histogenesis and biologic behavior. *Genes Chromosomes Cancer,* 46**,** 708-15.
TOMASETTO, C. & RIO, M. C. 2005. Pleiotropic effects of Trefoil Factor 1 deficiency. *Cell Mol Life Sci,* 62**,** 2916-20.
TONON, G., MODI, S., WU, L., KUBO, A., COXON, A. B., KOMIYA, T., O'NEIL, K., STOVER, K., EL-NAGGAR, A. & GRIFFIN, J. D. 2003. t (11; 19)(q21; p13) translocation in mucoepidermoid carcinoma creates a novel fusion product that disrupts a Notch signaling pathway. *Nature genetics,* 33**,** 208-213.
VOGEL, C. & MARCOTTE, E. M. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet,* 13**,** 227-32.
WANG, D., XU, Q., YUAN, Q., JIA, M., NIU, H., LIU, X., ZHANG, J., YOUNG, C. Y. & YUAN, H. 2019. Proteasome inhibition boosts autophagic degradation of ubiquitinated-AGR2 and enhances the antitumor efficiency of bevacizumab. *Oncogene,* 38**,** 3458-3474.
WHO, World Health Organisation, (2020) [online], Available at: https://www.who.int/data/gho/data/themes/noncommunicable-diseasesAccessed October 2020.
WU, L., LIU, J., GAO, P., NAKAMURA, M., CAO, Y., SHEN, H. & GRIFFIN, J. D. 2005. Transforming activity of MECT1-MAML2 fusion oncoprotein is mediated by constitutive CREB activation. *EMBO J,* 24**,** 2391-402.
WURM, F. M. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol,* 22**,** 1393-8.
XU, L., GORDON, R., FARMER, R., PATTANAYAK, A., BINKOWSKI, A., HUANG, X., AVRAM, M., KRISHNA, S., VOLL, E., PAVESE, J., CHAVEZ, J., BRUCE, J., MAZAR, A., NIBBS, A., ANDERSON, W., LI, L., JOVANOVIC, B., PRUELL, S., VALSECCHI, M., FRANCIA, G., BETORI, R., SCHEIDT, K. & BERGAN, R. 2018. Precision therapeutic targeting of human cancer cell motility. *Nat Commun,* 9**,** 2454.
YANG, S. & GUO, L.-J. 2017. Microencapsulation of low-passage poorly-differentiated human mucoepidermoid carcinoma cells by alginate microcapsules: in vitro profiling of angiogenesis-related molecules. *Cancer cell international,* 17**,** 1-10.
ZHENG, J. 2013. Oncogenic chromosomal translocations and human cancer. *Oncology reports,* 30**,** 2011-2019.
ZHU, Y., SHARP, A., ANDERSON, C. M., SILBERSTEIN, J. L., TAYLOR, M., LU, C., ZHAO, P., DE MARZO, A. M., ANTONARAKIS, E. S., WANG, M., WU, X., LUO, Y., SU, N., NAVA RODRIGUES, D., FIGUEIREDO, I., WELTI, J., PARK, E., MA, X. J., COLEMAN, I., MORRISSEY, C., PLYMATE, S. R., NELSON, P. S., DE BONO, J. S. & LUO, J. 2018. Novel Junction-specific and

Quantifiable In Situ Detection of AR-V7 and its Clinical Correlates in Metastatic Castration-resistant Prostate Cancer. *Eur Urol,* 73**,** 727-735.

9. Appendix

4% Paraformaldehyde preparation

40 ml of distilled water was heated up to 60°C. 4 g of Paraformaldehyde (Sigma) added to the water under the fume hood with constant stirring. A few drops of 1M NaOH was added to the solution with continuing stirring until all the powder was dissolved and the solution became clear. The solution was removed from the heat and 50 ml of 2× PBS was added. The PH was adjusted to 7.2 with HCL. The solution allowed to cool at room temperature and distilled water was topped up to 100 ml. The solution was stored at - 20°C for long periods and at 4°C for short usage.

Sequencing Data

GAC TAC AAA GAC GAT GAC GAC AAG flag sequence

GCTAGC NheI sequence

ACC Kozak sequence

ATG start codon sequence

>FLAG-Tag CRTC1-MAML2

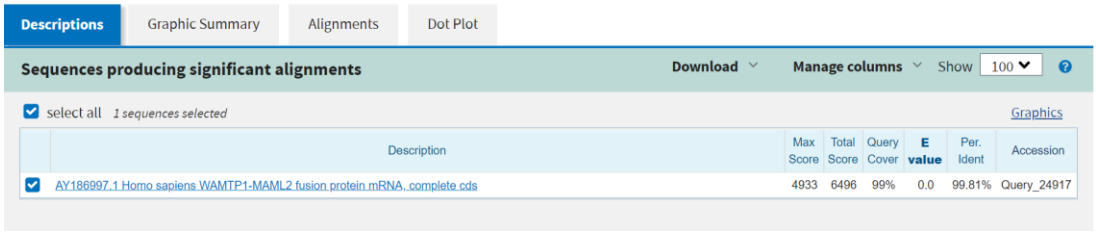
>FLAG-Tag CRTC1-MAML2

GCTAGCACCATGGACTACAAAGACGATGACGACAAGGCGACTTCGAACAATCCGCGGAAATTCAGCGA
GAAGATCGCGCTGCACAATCAGAAGCAGGCGGAGGAGACGGCGGCCTTCGAGGAGGTCATGAAGGACC
TGAGCCTGACGCGGGCCGCGCGGCTCCAGGGTTCCTTGAAAAGAAAACAGGTAGTTAACCTATCTCCT
GCCAACAGCAAGCGACCCAATGGCTTTGTGGACAACTCATTTCTTGATATCAAAAGAATTCGTGTTGG
GGAGAATCTCTCTGCAGGACAAGGTGGCCTCCAAATAAACAATGGACAAAGTCAGATTATGTCAGGGA
CCTTGCCTATGAGCCAAGCACCCCTGCGAAAGACTAACACTCTGCCATCCCATACACATTCTCCTGGC
AATGGCCTGTTTAACAATGGGCTTAAAGGAGGTAAAGAAGGAGCCAGGAGAGACTCTGTCTTGCAGTAA
GCACATGGATGGCCAAATGACCCAAGAGAATATTTTTCCTAATAGGTACGGAGACGACCCTGGAGAAC
AACTGATGGATCCTGAGCTGCAGGAACTGTTCAATGAACTGACCAACATATCTGTGCCTCCCATGAGT
GACCTTGAACTGGAGAACATGATCAATGCCACCATAAAGCAGGATGACCCATTTAACATTGACTTGGG
TCAGCAAAGCCAGAGGAGCACACCTAGGCCCTCCTTACCCATGGAGAAAATAGTGATCAAAAGTGAAT
ACTCACCGGGCTTGACTCAGGGCCCCTCAGGCTCTCCTCAGCTGAGGCCCCCATCAGCTGGCCCCGCA
TTCTCCATGGCCAACTCTGCCCTCTCCACTTCGTCTCCAATCCCTTCAGTCCCTCAGAGCCAGGCTCA
GCCTCAGACAGGCTCCGGAGCAAGCCGGGCCTTGCCAAGCTGGCAGGAAGTATCCCATGCCCAGCAGC
TCAAACAGATAGCTGCTAATCGTCAGCAGCATGCCCGGATGCAGCAGCACCAGCAGCAGCACCAGCCT
ACCAACTGGTCAGCCTTGCCCTCCTCTGCTGGACCATCACCAGGTCCATTTGGGCAGGAGAAAATCCC
CAGCCCTTCTTTTGGTCAGCAGACATTCAGCCCACAGAGCTCCCCCATGCCTGGGGTAGCTGGCGGCA
GCGGCCAGTCGAAAGTAATGGCTAACTACATGTACAAGGCCGGCCCCTCAGCCCAGGGTGGGCACCTA
GATGTCCTCATGCAGCAAAAGCCTCAGGATCTCAGTCGAAGTTTTATTAACAACCCGCACCCAGCCAT
GGAGCCCCGTCAGGGCAACACCAAGCCTTTGTTTCATTTTAACTCAGATCAAGCGAACCAGCAGATGC
CTTCTGTTTTGCCTTCCCAGAACAAGCCTTCTCTCCTACACTACACCCAACAGCAACAGCAGCAACAG
CAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAACAGCAACAGCAACAGCA
GAGTTCAATTTCAGCTCAACAACAGCAACAGCAGCAGAGCTCAATTTCAGCCCAACAGCAGCAGCAGC

AGCAACAACAGCAGCAGCAGCAGCAACAACAACAGCAACAACAGCAGCAGCAGCAGCAGCAACAACCA
TCTTCTCAGCCTGCCCAATCTCTACCAAGCCAGCCTTTGCTAAGGTCACCTTTGCCACTTCAGCAAAA
GCTCCTACTTCAGCAAATGCAGAATCAGCCCATTGCAGGAATGGGATACCAAGTCTCCCAACAACAGA
GACAGGATCAACACTCTGTGGTAGGCCAGAACACAGGCCCCAGTCCAAGTCCTAACCCCTGCTCAAAT
CCAAACACTGGAAGTGGTTACATGAACTCCCAGCAATCACTGTTGAATCAGCAATTGATGGGAAAGAA
GCAGACTCTACAGAGGCAGATCATGGAGCAGAAACAGCAACTTCTTCTCCAGCAGCAGATGCTGGCTG
ACGCGGAGAAAATTGCTCCACAAGATCAGATAAACCGACATTTGTCAAGGCCACCTCCAGATTATAAA
GACCAAAGAAGAAATGTGGGCAATATGCAACCAACTGCTCAGTATTCTGGTGGCTCATCCACAATAAG
CTTAAACTCTAACCAGGCTTTGGCAAACCCAGTTTCAACACACACCATTTTAACTCCCAATTCCAGCC
TCCTGTCTACTTCTCACGGGACAAGAATGCCATCATTATCTACAGCAGTTCAGAATATGGGGATGTAT
GGAAATCTGCCTTGTAATCAACCTAACACATACAGTGTCACTTCAGGAATGAATCAATTGACCCAACA
GAGAAACCCAAAGCAATTGTTAGCAAATCAAAACAACCCTATGATGCCACGGCCACCTACCTTAGGGC
CAAGTAATAATAACAATGTAGCCACTTTTGGAGCTGGATCTGTTGGTAATTCACAACAATTGAGACCA
AATTTAACCCATAGTATGGCAAGCATGCCACCACAGAGAACATCAAACGTAATGATCACATCCAACAC
AACTGCACCAAACTGGGCCTCTCAAGAAGGAACAAGCAAACAGCAAGAAGCCCTGACGTCTGCAGGAG
TCCGCTTCCCCACAGGTACACCTGCAGCCTATACCCCAAATCAGTCACTGCAACAGGCAGTAGGTAGC
CAGCAATTTTCCCAGAGGGCAGTGGCTCCTCCTAACCAGTTAACACCAGCAGTGCAAATGAGACCCAT
GAACCAAATGAGCCAAACACTAAATGGGCAAACCATGGGTCCCCTCAGGGGTCTGAATCTCAGACCCA
ATCAGCTAAGCACACAGATTTTGCCTAATTTGAATCAGTCAGGAACAGGGTTGAATCAGTCGAGGACG
GGCATCAACCAGCCACCATCCCTGACGCCCAGCAATTTTCCTTCACCCAACCAAAGTTCCAGGGTTTT
TCAAGGAACTGACCACAGCAGTGACTTAGCTTTTGACTTCCTCAGCCAACAAAATGATAACATGGGCC
CTGCCCTAAACAGTGATGCTGATTTCATTGATTCTTTATTGAAGACAGAGCCTGGTAATGATGACTGG
ATGAAAGACATCAATCTTGATGAAATCTTGGGGAACAATTCCTAA

Overlapping sequence

Alignment between the FLAG Tag CRTC1 MAML2 constructs and the Reference sequence.


Descriptions
Graphic Summary
Alignments
Dot Plot
Sequences producing significant alignments
Download
Manage columns
Show
100
select all
1 sequences selected
Graphics
Description
Max Score
Total Score
Query Cover
E value
Per. Ident
Accession
AY186997.1 Homo sapiens WAMTP1-MAML2 fusion protein mRNA, complete cds
4933
6496
99%
0.0
99.81%
Query_24917