EVALUATION OF THE ROLE OF EGFR, SOX2, AND PAX9 IN ORAL CARCINOGENESIS

Timothy John Bates

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

Northern Institute for Cancer Research

Newcastle University

August 2015

Acknowledgements

I would like to acknowledge my three supervisors, Dr M. Robinson, Dr R. Kist, and Professor P. Sloan, without whom this project would not have been possible. Dr Kist and Dr Robinson shared the initial vision for this project and, in collaboration, designed and developed its overall structure. The project therefore arose from a unique collaboration between principle investigators with backgrounds in developmental oral biology and oral pathology. Professor Sloan's experience and insight have proved invaluable to the analysis, interpretation, and presentation of our data.

I would like to thank Miss A. Long and all members of the Research and Development Team, Department of Cellular Pathology, for their essential contribution to the processing, sectioning, and automated immunostaining of both human and mouse tissue during this project.

Dr R. Mahsoub and Miss V. Rook contributed to the manual immunohistochemical staining of sections during this project, work which formed part of their Masters by Research projects. Mr M. Kennedy played a vital role in collecting clinical data for the group of early-stage oral squamous cell carcinoma patients, during a Master of Science project that he undertook at the University of Edinburgh.

I would like to thank Miss M. Elias and all the team in the Dermatology Laboratory, Institute of Cellular Medicine. The generation of the tetracycline-inducible cell lines would not have been possible without the support of Martina and others in the Dermatology team: Carol, Keith, and Emma, to name just a few. It was a privilege to work with this group, whose commitment to and passion for science is an inspiration.

Finally, I would like to thank my parents, Mr and Mrs C.G. Bates, for their patience, wisdom, and support, not only during this project, but throughout my career to date.

Table of Contents

Abbreviations

ACTB	-	Actin, beta
AEG	-	Anophththalmia-esophageal-genital syndrome
Akt	-	Protein kinase B
AP	-	Anterior palate
bp	-	Base pair
CISH	-	Colourimetric *in situ* hybridisation
CNS	-	Central nervous system
CO_2	-	Carbon dioxide
COSHH	-	Control of Substances Hazardhous to Health
CSC	-	Cancer stem cell
Ct	-	Cycle threshold
CT	-	Computerised tomography
cDNA	-	Complementary deoxyribonucleic acid
DAB	-	3-3'-Diaminobenzidine
DAHNO	-	Database for Head and Neck Oncology
DCP	-	Dundee Cell Products
d.f.	-	Degrees of freedom
DMEM	-	Dulbecco's Modified Eagle's Medium
DMSO	-	Dimethylsulphoxide
DNA	-	Deoxyribonucleic acid
DMBA	-	Dimethyl-1,2–benzanthracene
dNTP	-	Deoxyribonucleoside trisphosphate

EGFP	-	Enhanced green fluorescence protein
EGFR	-	Epidermal growth factor receptor
FBS	-	Foetal bovine serum
FGU	-	Functional Genetics Unit
FISH	-	Fluorescence *in situ* hybridisation
4-NQO	-	4 –Nitroquinoline-1-oxide
FP	-	Filiform papilla
G	-	Grams
GA	-	General anaesthesia
GAPDH	-	Glyceraldehyde 3-phosphate dehydrogenase
GCN	-	Gene copy number
H&E	-	Haematoxylin and eosin
HeBS	-	Hepes-buffered saline
HMG	-	High mobility group
HNMDT	-	Head and Neck multi-disciplinary team
HNSCC	-	Head and neck squamous cell carcinoma
HPV	-	Human papilloma virus
HR-HPV	-	High-risk human papillomavirus
IHC	-	Immunohistochemistry
IRES	-	Internal ribosomal entry site
ISH	-	*In situ* hybridisation
Krt	-	Keratin
LC	-	Lung cancer

LOH	-	Loss of heterozygosity
LSCC	-	Lung squamous cell carcinoma
MAPK	-	Mitogen activated protein kinase
µl	-	Microlitres
µm	-	Micrometers
µg	-	Micrograms
mins	-	Minutes
ml	-	Millilitres
µM	-	Micromolar
mM	-	Millimolar
mRNA	-	Messenger ribonucleic acid
MTOR	-	Mammalian target of rapamycin
N	-	Nodal category
ng	-	Nanograms
nm	-	Nanometres
NNK	-	4-methylnitrosamine 1 -3 pyridyl – 1 –buntanone
n.s.	-	Not significant
NUTH	-	Newcastle-upon-Tyne NHS Hospitals
OC	-	Oral cancer
OPMD	-	Oral potentially malignant disorders
OPSCC	-	Oropharyngeal squamous cell carcinoma
OSCC	-	Oral squamous cell carcinoma
PAX	-	Paired box

PBS	-	Phosphate buffered saline
PCR	-	Polymerase chain reaction
PMD	-	Potentially malignant disorder
PPC	-	Percentage of positive cells
PPN	-	Percentage of positive nuclei
pStage	-	Pathological Stage
PVL	-	Proliferative verrucous leukoplakia
qPCR	-	Quantitative real-time polymerase chain reaction
RFP	-	Red fluorescence protein
RIPA	-	Radio-immuno precipitation assay
RNA	-	Ribonucleic acid
ROC	-	Receiver-operator curve
rpm	-	Revolutions per minute
RT-PCR	-	Reverse-transcription polymerase chain reaction
SCC	-	Squamous cell carcinoma
s.d.	-	Standard deviation
secs	-	Seconds
SH2	-	Src Homology 2
SIN	-	Squamous Intra-epithelial Neoplasia
SNOMED	-	Systemised nomenclature for medical terms
SOX	-	SRY-Box
SPT	-	Second primary tumour
SRH	-	Sunderland Royal Hospital

SRY	-	Mammalian testis determing factor
T	-	Tumour category
TBE	-	Tris-borate-EDTA
TBS	-	Tris-buffer saline
TetO	-	Tetracycline-on
TetR	-	Tetracycline-repressor
TGF	-	Transforming growth factor
TIFF	-	Tagged Image Format File
TMA	-	Tissue micro-array
TNM	-	Tumour, Node, Metastasis
3+PC	-	Percentage of strongly-positive cells
3+PN	-	Percentage of strongly-positive nuclei
WHO	-	World Health Organisation
V	-	Volts

★	$p<0.05$
★★	$p<0.01$
★★★	$p<0.001$
★★★★	$p<0.0001$

Abstract

Oral squamous cell carcinoma (OSCC) is a major global healthcare problem. OSCC has devastating consequences for many patients diagnosed with the disease. Outcomes may be improved if the disease is identified in its precursor stages, termed oral potentially malignant disorders (OPMD). Unfortunately, histological assessment of OPMD does not reliably predict which cases will progress to OSCC. Several candidate biomarkers have emerged in recent decades. To date, however, none have been validated for use in clinical practice. This study sought to address the continuing need for biomarkers that stratify OPMD according to their risk of malignant transformation.

Our data show that EGFR gene copy number abnormalities correlate with malignant transformation in OPMD. EGFR genomic gain was also present in a quarter of early-stage OSCC. SOX2 had a heterogeneous expression profile in both OPMD and OSCC, limiting its clinical utility. Nevertheless, the pattern of SOX2 expression suggests it may be a marker of OSCC stem cells and consequently represent a potential chemotherapeutic target. PAX9 is down-regulated in OPMD and early-stage OSCC. Following a course of chemical induction, Pax9-deficient mice were more likely to develop OPMD and OSCC than controls. These findings support the hypothesis that PAX9 has a tumour-suppressor function. In addition to enhanced local sensitivity to chemical induction, Pax9-deficient mice were more susceptible to the toxic systemic effects of treatment. A modified protocol for chemical induction in Pax9-deficient mice is recommended. Paradoxically, our analysis of human tissues showed increased PAX9 expression in OPMD that underwent malignant transformation, suggesting that, in some circumstances, PAX9 may have a tumour-promoting effect. Finally, we summarise the generation of stably transfected cell lines in which PAX9 and SOX2 expression may be manipulated by tetracycline administration. These cell lines will facilitate future studies of the functional role of PAX9 and SOX2 in oral carcinogenesis.

Chapter 1. General Introduction

1.1 Cancer

The term ‘cancer’ encompasses a group of over 200 diseases characterised by the uncontrolled proliferation of abnormal cells (Cancer Research UK, 2010). Each year, there are approximately 13 million new cancer cases and 8 million cancer-related deaths worldwide (Jemal et al., 2011). In the UK, cancer affects more than one-third of the population and accounts for over a quarter of all deaths. Excluding non-melanocytic skin cancers, the four commonest cancers are those involving the lung, breast, gastrointestinal tract, and prostate gland. Lung cancer has the highest mortality rate in the UK (Cancer Research UK, 2010). Metastatic dissemination of tumour cells to distant sites is the principal cause of cancer mortality (Meyer and Hart, 1998).

Cancers vary in their pathogenesis. However, it is now accepted that most cancers develop through a multistep accumulation of genetic alterations (Hanahan and Weinberg, 2011). Observations of the ‘adenoma-carcinoma sequence’ identified in colorectal cancer informed the first step-wise molecular progression model of cancer (Fearon and Vogelstein, 1990). This model has been applied to other cancers, including oral cancer (Califano et al., 1996). Recent advances in cancer research support the view that key genetic events are common to the development of many cancers. These hallmarks include altered expression of oncogenes, suppression of tumour-suppressor genes, and induction of angiogenesis (Hanahan and Weinberg, 2011).

1.2 Oral cancer

1.2.1 Epidemiology

Oral cancer (OC) is the 13th commonest cancer globally. OC has an annual incidence of around 264,000 cases. Each year, OC accounts for up to 130,000 deaths (Jemal et al., 2011). However, the distribution of OC varies markedly around the world. OC is rare in countries such as Japan, Finland, and Greece, but is prevalent in Hungary and parts of northern France (Barnes et al., 2005; Warnakulasuriya, 2009). The highest incidence of OC is in south-east Asia: it is the commonest cancer in Sri Lanka, India, and regions of Pakistan (Indian

Council of Medical Research, 2010; International Agency for Research on Cancer, 2010). Recently, the incidence of OC has been increasing in the developed world. Over the past decade, the incidence in the UK has increased by one-third to ~6500 new cases in 2010 (Cancer Research UK, 2010). Similar increases have been documented in the USA and parts of Europe, particularly among young adults (Schantz and Yu, 2002; Garavello et al., 2010).

1.2.2 Histological classification of oral cancer

A diverse spectrum of epithelial, mesenchymal, and haematolymphoid neoplasms may present in the oral cavity. However, oral squamous cell carcinoma (OSCC) accounts for more than 90% of OC cases. OSCC is a malignant neoplasm of the mucosal squamous epithelium. Histologically, OSCC is characterised by the formation of keratin and/or the presence of intercellular bridges (Barnes et al., 2005). This project will focus specifically on OSCC and its precursor lesions.

OSCC is classified according to its histological differentiation, i.e. the proportion of the OSCC that resembles normal squamous epithelium. This classification was first described by Broders more than 90 years ago (cited by Gnepp (2009)). However, the evidence to date suggests that differentiation does not accurately predict the clinical outcome of OSCC (Roland et al., 1992). Many OSCC comprise a heterogeneous population of malignant cells that show variable differentiation. The majority of OSCC are therefore classified as moderately differentiated (Kearsley and Thomas, 1993).

A multifactorial system for grading OSCC has been described by Jakobsson et al (1973). This system assesses differentiation alongside a range of tumour-host factors, including invasive pattern and peri-tumoral lymphoplasmacytic infiltrate. Of these tumour-host factors, invasive pattern has the greatest prognostic significance. OSCC with a non-cohesive invasive pattern (comprising malignant cells arranged singly or in small nests) behave more aggressively than OSCC with a cohesive invasive pattern (Odell et al., 1994). The adverse prognostic significance of neural invasion, vascular invasion, and positive surgical resection margins is also well documented (Close et al., 1989; Slootweg et al., 2002; Rahima et al., 2004).

1.2.3 Tobacco and alcohol: major environmental risk factors for OSCC

OSCC has a multifactorial aetiology. A range of environmental factors may precipitate the genetic alterations and subsequent aberrations - in cell growth, survival, and migration – that characterise OSCC (Leemans et al., 2011). However, tobacco smoking remains one of the most significant risk factors, accounting for ~70% of OSCC-related deaths in the developed world. The carcinogenic potential of tobacco and alcohol is enhanced through synergistic interactions between constituent molecules and metabolites when the two are consumed together (Danaei et al., 2005; Hashibe et al., 2009). Recently, it has been suggested that all patients diagnosed with either lung cancer or liver cirrhosis – attributable to tobacco and alcohol consumption respectively - should be screened for OSCC and its precursors as an integral part of their care pathway (Salaric *et al.*, 2015).

Tobacco contains more than 60 distinct carcinogens. These include nitrosamines, polycyclic aromatic hydrocarbons, aromatic amines, aldehydes, and phenols. Nitrosamines such as NNK (4-methylnitrosamino-1-(-3-pyridyl)-1-butanone) are among the most significant carcinogens (Hecht, 2003). NNK is metabolised via hydroxylation to methyldiazo-hydroxide, which binds to the guanine residues of DNA. Guanine residues are also bound by aromatic hydrocarbons such as benzopyrene (Hoffman and Wynder, 1986). Chemical modification of guanine residues induces nucleotide transversions in critical genes such as the tumour-suppressor *p53* (Brennan et al., 1995).

In addition to inhalation via smoking, tobacco may be chewed. Tobacco chewing is responsible for the high incidence of OSCC in south-east Asia (Jayalekshmi et al., 2009). Chewed tobacco is often combined with other substances – areca nut/leaf, catechu, and lime - to form ‘paan’. These substances contain molecules that interact to generate further distinct carcinogens, such as arecaidine. Arecaidine is formed by hydrolysis of arecoline - an alkaloid constituent of the areca nut - by calcium hydroxide in lime (Khan et al., 2012).

Alcohol (ethanol) in its pure form is not currently regarded as a carcinogen. Alcoholic drinks, however, may contain carcinogenic by-products, such as

nitrosamines and urethane compounds. Ethanol facilitates the absorption of these molecules in the oral cavity (Hashibe et al., 2009). Ethanol is metabolised to acetaldehyde by dehydrogenases and the cytochrome p450 system. Acetaldehyde induces direct DNA damage. Acetaldehyde also promotes further indirect damage by blocking glutathione, hampering the detoxification of other carcinogens. Induction of the cytochrome p450 enzyme has the potential to activate other pro-carcinogens in both alcohol and tobacco (Hunter et al., 2005).

1.2.4 High-risk human papillomavirus infection in oral cancer and oro-pharyngeal cancer

There is now an established body of evidence to support the role of high-risk human papillomavirus (HR-HPV), notably subtypes 16/18, in the aetiology of oropharyngeal squamous cell carcinoma (OPSCC) (Thariat et al., 2010; Schache et al., 2011; Thavaraj et al., 2011; Robinson et al., 2012). However, convincing evidence that HR-HPV infection is significant in the aetiology of OSCC has yet to emerge (Lopes *et al.*, 2011; Lingen *et al.*, 2013; McCord *et al.*, 2014; Nankivell *et al.*, 2014). HR-HPV infection has been detected in a subset of OSCC precursor lesions and specific histopathological features have been described in these cases (Angiero et al., 2010; Lerman and Woo, 2014). However, the proportion of HPV-positive OSCC remains low (~3%), which suggests that HR-HPV infection is not a major aetiological agent in OSCC formation (Salazar et al., 2014). The role of HR-HPV infection in OSCC formation remains unclear. HR-HPV is not a validated prognostic factor for either OSCC or its precursors (McCord *et al.*, 2014; Nankivell *et al.*, 2014). However, there is emerging evidence that HPV positivity may be a useful prognostic indicator in the small proportion of HPV-positive squamous cell carcinomas presenting at non-oropharyngeal sites, including the oral cavity (Chung et al., 2014).

Given the contrasting aetiological and prognostic significance of HR-HPV in OPSCC and OSCC reported in the literature to date, this project will focus specifically on cases of squamous cell carcinoma and its precursor lesions that are confined to the oral cavity.

1.2.5 The contribution of other infective conditions to OSCC formation

Although HR-HPV infection has yet to be convincingly implicated in the aetiology of OSCC, there are two infective conditions that have a well-documented association. Chronic hyperplastic candidosis is characterised by epithelial hyperplasia and inflammatory changes in both the squamous epithelium and lamina propria. These changes enhance the susceptibility of the epithelium to carcinogenic nitrosamine compounds that are generated by Candida organisms. Tertiary syphilis causes epithelial atrophy, increasing the vulnerability of the epithelium to carcinogens; moreover, medicaments historically used to treat syphilis, notably arsenic compounds, are now known carcinogens (Neville et al., 2009).

1.2.6 Histopathological staging of OSCC

Once a histopathological diagnosis of OSCC has been established, the disease is staged using the 'Tumour, Node, Metastasis' (TNM) classification developed by the International Union Against Cancer (Sobin et al., 2009). A numerical value is assigned for each of the three domains (Table 1-1). These are combined to provide the overall clinical stage (Table 1-2). The clinical stage informs treatment decisions. It may be modified following histopathological examination of the surgical resection specimen.

This project will focus specifically on patients with early-stage OSCC. Early-stage OCCC encompasses patients at both pStage I (pT1 N0 M0) and pStage II (pT2 N0 M0). Tumours in this group may therefore measure up to 4 cm in maximum dimension; however, they have not metastasised either to cervical lymph nodes or to distant sites.

Table 1-1 The 'Tumour, Node, Metastasis' classification of oral squamous cell carcinoma, International Union Against Cancer

Code	Description
T0	No evidence of tumour
T1	Tumour 2 cm or less in greatest dimension
T2	Tumour more than 2 cm but not more 4 cm in greatest dimension
T3	Tumour more than 4 cm in greatest dimension
T4a	Tumour invades through cortical bone, deep/extrinsic muscles of tongue, maxillary sinus or skin
T4b	Tumour invades masticator space, pterygoid plates, skull base, or encases internal carotid artery
N0	No regional lymph node metastasis
N1	Metastasis in a single ipsilateral lymph node, 3 cm or less in greatest dimension
N2a	Metastasis in a single ipsilateral lymph nodes, more than 3 cm but not more than 6 cm in greatest dimension
N2b	Metastases in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension
N2c	Metastasis in bilateral/contralateral lymph nodes, none more than 6 cm in greatest dimension
N3	Metastasis in a lymph node more than 6 cm in greatest dimension
M0	No distant metastasis
M1	Distant metastasis

Table 1-2 Summary of the component TNM categories of each stage of OSCC

Stage	Tumour	Node	Metastasis
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1, T2, T3	N1	M0
Stage IVA	T4a	N0, N1	M0
	T1, T2, T3, T4a	N2	M0
Stage IVB	Any T	N3	M0
	T4b	Any N	M0
StageIVC	Any T	Any N	M1

Adapted from the 'TNM Classification of Malignant Tumours', 7th Edition (Sobin et al., 2009).

1.2.7 Treatment modalities

Surgical resection is currently the preferred treatment modality for early-stage OSCC. Adjuvant radiotherapy may be used for selected cases (Shah and Gil, 2009). Surgical removal of OSCC may be debilitating, leaving patients with dysfunctions in speech, mastication, and swallowing (dysphagia). Radiotherapy is also associated with a range of complications, including mucositis and dysphagia (Campos et al., 2014; Szczesniak et al., 2014). Long-term, radiotherapy may also result in vasculopathy of the carotid artery, thus increasing the risk of cerebrovascular events (Wilbers et al., 2014). These functional impairments may impact negatively on social interactions, and thus contribute to the development of psychological problems (Zwahlen et al., 2008).

Whilst both of the current treatment modalities may cause significant aesthetic and functional impairments, neither guarantees a complete cure. There is, therefore, a pressing need to develop chemotherapeutic agents that either halt or reverse the progress of OSCC and its precursor lesions with minimal long-term impact on patients' function and aesthetics (Braakhuis et al., 2010).

1.2.8 Clinical outcomes for patients with OSCC

Outcomes for patients with OSCC may be improved if the disease is identified in its earliest or precursor stages (Goodson and Thomson, 2010). Due to its anatomical location, OSCC is amenable to early detection and management (Sankaranarayanan et al., 2013). Evidence from high-incidence areas suggests that screening programmes may facilitate early detection of OSCC and improved patient outcomes (Amit et al., 2013). However, despite the increasing standardisation of its management, OSCC is consistently associated with poor survival rates in the developed world (Warnakulasuriya, 2009). Alarmingly, OSCC mortality rates are actually increasing in parts of Europe (Garavello et al., 2010).

Several factors contribute to poor patient outcomes. Despite efforts to raise the profile of OSCC (Mouth Cancer Foundation, 2013), there is generally low public awareness of the disease relative to other cancers (Warnakulasuriya et al., 1999). Lack of awareness is compounded by the clinical course of OSCC. The early stages are often painless. An asymptomatic OSCC may therefore achieve

a significant size and have already seeded metastases prior to the patient seeking help (McGurk et al., 2005). The consequent late-stage presentation may preclude complete surgical clearance. This increases the risk of local recurrence, which is the major cause of death in advanced cases of OSCC (Woolgar et al., 1999).

1.2.9 Second primary tumours and the phenomenon of 'field cancerisation'

Clinical outcomes are further confounded by the high propensity of OSCC to form second primary tumours (SPT). SPT develop in up to 36% of patients with OSCC (Chuang et al., 2008; Atienza and Dasanu, 2012). SPT thus have a higher incidence than primary OSCC, a trend that is consistent across all age groups (Bosetti et al., 2011). SPT portend a poor prognosis. In early-stage (pStage I and pStage II) OSCC, SPT are the commonest cause of treatment failure and death from disease (Day and Blot, 1992; Tsou et al., 2007).

The development of SPT reflects the phenomenon of 'field cancerisation', which was first described by Slaughter et al (1953) more than 60 years ago. Field cancerisation postulates that potentially malignant changes extend into squamous epithelium adjacent to OSCC that appears normal, both clinically and histologically. The concept stemmed from the recognition that common carcinogens, such as alcohol and tobacco smoke, contact multiple epithelial surfaces of the upper aero-digestive tract (Slaughter et al., 1953).

The concept of field cancerisation is supported by recent insights into the genetic and molecular hallmarks of OSCC formation (Braakhuis et al., 2003; Braakhuis et al., 2004; Braakhuis et al., 2010; Leemans et al., 2011). Most synchronous/metachronous primary OSCC are believed to evolve from a single clone of progenitor cells and to have undergone similar genetic alterations (Figure 1-1) (Bedi et al., 1996; Scholes et al., 1998; Braakhuis et al., 2004). The recognition that pre-cancerous genetic changes extend into apparently normal epithelium provides an explanation for the high incidence of SPT (Partridge et al., 2000).

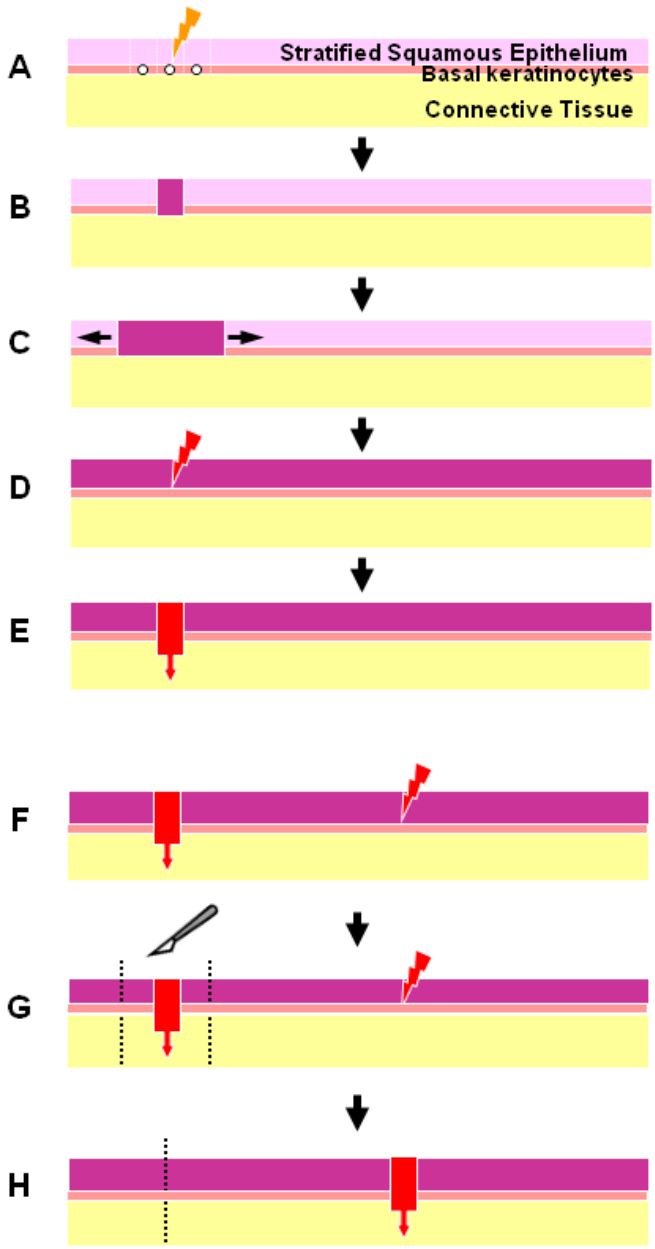


Figure 1-1 Summary of the genetic progression model of field cancerisation

A) Initial Phase: Mutation in basal stem cell. B) "Patch" Phase: Clonal expansion of mutant daughter cells. C) Expanding Field: Lateral expansion of stem cells and descendants having acquired a growth advantage. D) "Second Hit": Precursor lesion develops within field due to secondary mutations and increasing genomic instability. E) Carcinoma Formation: Precursor lesion becomes invasive carcinoma. F) Novel "Second Hit": New precursor lesion independently develops at another site within field. G) Carcinoma Excision: Premalignant field and new precursor lesion remain. H) Tumour recurrence: Second Field Tumour develop from new precursor lesion.

Adapted from Braakhuis et al (2004).

1.2.10 Oral potentially malignant disorders

A range of precancerous lesions and conditions are recognised in the oral cavity. Together, these are termed oral potentially malignant disorders (OPMD) (van der Waal, 2009). Clinically, the majority of OPMD present either as white ('leukoplakic'), red ('erythroplakic') or mixed red and white ('speckled leukoplakic') patches (Barnes et al., 2005). Histopathological examination of the affected squamous epithelium may reveal a spectrum of changes, ranging from squamous cell hyperplasia, through mild, moderate, and severe epithelial dysplasia to carcinoma in-situ (Table 1-3).

Although the 2005 WHO classification is widely recognised and applied by both clinicians and histopathologists, there is continuing debate as to how precancerous epithelial changes should be conceptualised and classified. Two alternative classification systems are currently recognised by the WHO (Barnes et al., 2005). In addition to these, a binary classification system has also been proposed (Kujan et al., 2006). The binary grading system complements the WHO Classification by excluding the intermediate category of 'moderate epithelial dysplasia', which confounds inter-observer agreement. The histopathologist assigns a grade according to the overall perceived 'risk' of the squamous epithelium. The binary system may therefore guide clinicians and facilitate critical management decisions (Kujan et al., 2006).

The debate surrounding the classification of epithelial dysplasia highlights the subjective nature of interpreting the histological features of precancerous lesions and assigning them to precise diagnostic categories. It is recognised that this interpretative process is liable to both intra- and inter-observer variation (Kujan et al., 2007). So far, no histological features have been identified that accurately predict which OPMD will progress to OSCC (Lodi et al., 2006). There is therefore a need to develop biomarkers that enhance prognostication and direct treatment (Mishra, 2012).

Table 1-3 Summary of the current classification systems for epithelial dysplasia

2005 WHO Classification	Squamous Intra-epithelial Neoplasia Classification	Ljubljana Classification of squamous Intra-epithelial lesions	Binary Classification
Squamous cell hyperplasia	-	Squamous cell (simple) hyperplasia	-
Mild dysplasia	SIN 1	Basal/parabasal cell hyperplasia	Low-grade dysplasia
Moderate dysplasia	SIN 2	Atypical hyperplasia	Low-grade OR high-grade dysplasia
Severe dysplasia	SIN 3	Atypical hyperplasia	High-grade dysplasia
Carcinoma in-situ	SIN 3	Carcinoma in-situ	High-grade dysplasia

Adapted from Barnes et al (2005) and Kujan et al (2006).

1.2.11 Genetic and molecular alterations in oral cancer

Analogous to the progression model that is now accepted for colorectal carcinoma (Fearon and Vogelstein, 1990), the genetic alterations characteristic of OSCC are understood to occur in a multistep, sequential pattern (Califano et al., 1996; Leemans et al., 2011). The spectrum of potentially malignant disorders contains common chromosomal, genetic, and molecular aberrations that are also present in OSCC (Mithani et al., 2007). A working model for the formation of OSCC has been widely accepted. First proposed by Califano et al (1996) nearly 20 years ago, this model describes sequential loss of heterozygosity (LOH) in chromosomal regions that contain key cancer-associated genes, and maps these genomic changes to clinical and histological precursor lesions. The model draws attention to alterations in the genes governing growth promotion (proto-oncogenes) and growth inhibition (tumour-suppressor genes) (Califano et al., 1996; Forastiere et al., 2001).

Mutations of proto-oncogenes result in abnormally activated ‘oncogenes’ that confer growth self-sufficiency to cells, thus enabling them to escape normal

growth control mechanisms (Kumar et al., 2005). Early research into proto-oncogenes in OSCC focused largely on c-myc (Eversole and Sapp, 1993), K-ras (Caulin et al., 2004; Vitale-Cross et al., 2004a) and ErbB-1 genes (Wong, 1987). Ki67 is marker of cell proliferation that has been used in the diagnosis and prognostication of breast and prostate cancers, in which the fraction of Ki67 positive cells correlates with the clinical progression of the disease (Scholzen and Gerdes, 2000). Ki67 alterations have also been documented in OPMD. However, Ki67 has yet to be validated as a prognostic marker for detecting the subset of OPMD that are destined to undergo malignant transformation (Kovesi and Szende, 2003). More recent work on OSCC oncogenes has focused on the complex protein-kinase B (Akt) pathway (Amornphimoltham et al., 2004; Amornphimoltham et al., 2005). There is an increasing body of evidence that supports the potential therapeutic significance of this pathway (Czerninski et al., 2009; Wang et al., 2014).

The tumour-suppressor genes of greatest significance in OSCC formation are *p53* and *CDKN2A* (Hunter et al., 2005; Kumar et al., 2005). *CDKN2A* encodes p16INK4a, which controls the cell cycle through inhibition of cyclin-dependent kinases. *CDKN2A* may be inactivated by promoter hyper-methylation in combination with deletions in chromosomal regions 3p and 9p21 (Califano et al., 1996; Forastiere et al., 2001).

p53 encodes a critical tumour-suppressor protein and is mutated in more than 50% of human cancers (Yuen et al., 2001). *p53* mutations are a consistent feature of OSCC (Gasco and Crook, 2003). *p53* is located on chromosome 17p and encodes a transcription factor (p53) that detects DNA damage and other stress signals. p53 acts as a 'gatekeeper' at the G_1 phase of the cell cycle, permitting repair of damaged DNA prior to DNA synthesis and triggering apoptosis if irreparable DNA damage is detected (Liu and Gelmann, 2002). *p53* mutations therefore result in loss of this critical tumour-suppressor function. This confers a significant survival advantage and is frequently an early step in oral carcinogenesis (Leemans et al., 2011). p53 is upregulated in human OPMD and expression of p53 in suprabasal keratinocytes has been associated with malignant transformation (Nylander et al., 2000; Kovesi and Szende, 2003; Varun et al., 2014). The carcinogens present in tobacco smoke, one of the

major risk factors for OSCC, are associated with an increase in both the number and type of *p53* mutations (Brennan et al., 1995).

While Califano's (1996) genetic progression model continues to frame the current molecular understanding of OSCC formation, it remains incomplete. For example, the model does not account for the over-expression of epidermal growth factor receptor (EGFR) that has been identified in a high proportion of OSCC (Grandis and Tweardy, 1993a). The most reliable factors for predicting patients' risk of developing OSCC continue to be their cancer history and the presence/grade of epithelial dysplasia. Markers for LOH, chromosomal aneuploidy, and aberrant p16INK4a/p53 expression remain adjuncts to these clinical and histopathological parameters (Leemans et al., 2011). Advances in the field of ploidy analysis are compromised by the questionable probity of some studies; initial hopes that the detection of aneuploidy might identify epithelium at high-risk of developing OSCC have yet to be widely accepted (Klanrit et al., 2007; Torres-Rendon et al., 2009). Doubts have also emerged as to the reliability of p53 as a standalone prognostic biomarker (Nylander et al., 2000; Takeda et al., 2006). Although it remains the mainstay of diagnosis and prognostication, the reliability of histopathological examination is itself compromised by well-documented intra- and inter-observer variation (Kujan et al., 2007).

There is a continuing need to delineate the genetic alterations and molecular events of OSCC formation in order to inform the development of reliable diagnostic and prognostic biomarkers. An ideal biomarker should be objectively measurable in small biopsy samples, and altered in high-risk tissue in the earliest stages of oral carcinogenesis (Wu et al., 2010).

1.3 Epidermal growth factor receptor

1.3.1 Structure and biological function

Epidermal growth factor receptor (EGFR) is a cell surface tyrosine kinase receptor that binds to members of the epidermal growth factor family of extracellular protein ligands. These ligands include epidermal growth factor and transforming growth factor-α (Herbst, 2004). EGFR is one of four proteins in the ErbB family and is expressed in most epithelial tissues, including the oral squamous epithelium (Citri and Yarden, 2006).

Binding of growth factor ligands to the EGFR extracellular domain induces a conformational change in the EGFR molecule. The conformational change usually involves formation of an active homodimer from two inactive monomers. Alternatively, a heterodimer may also be formed through pairing EGFR with another member of the ErbB receptor family, such as ErbB2/Her2/neu (Yarden and Schlessinger, 1987). Dimerization induces phosphorylation of the tyrosine residues located on the internal domain of EGFR. Phosphotyrosines bind to downstream signalling in proteins that contain the SH2 (Src Homology 2) domain (Molinolo et al., 2009). The activated proteins initiate a range of signal-transduction cascades, including the mitogen-activated protein kinase (MAPK) and protein-kinase B (Akt) pathways. Activation of these pathways stimulate DNA synthesis and contribute to a range of cellular processes, including cell proliferation (Oda et al., 2005).

1.3.2 EGFR as a cancer biomarker

EGFR protein expression and gene copy number are currently used in the prognostication of non-small cell lung carcinoma (Nicholson et al., 2001; Hirsch et al., 2003). They are also used to predict the response of non-small cell carcinoma to EGFR-targeting chemotherapeutic agents (Takano et al., 2005).

EGFR was first heralded as a potential biomarker in OSCC in the early 1990s (Grandis and Tweardy, 1993b). EGFR protein is over-expressed in up to 90% of OSCC. Furthermore, over-expression of EGFR protein in OSCC is associated with poor clinical outcomes (Grandis and Tweardy, 1993a; Grandis, 1998; Kumar et al., 2008). Similarly, EGFR genomic gain is also associated with poor

clinical outcomes (Chung et al., 2006; Temam et al., 2007). To date, the prevalence of EGFR genomic gain that has been reported in OSCC ranges from 9% (in pStage I and pStage II OSCC) to 56% (in mixed pStage I – pStage IV OSCC) (Freier et al., 2003; Rössle et al., 2013; Ryott et al., 2009). These data suggest that EGFR genomic gain is more common in late-stage OSCC and may be a relatively late event in oral carcinogenesis.

In vitro studies have demonstrated that ectopic expression of EGFR is implicated in the transformation of normal oral keratinocytes to a malignant, immortalised phenotype (Goessel et al., 2005). EGFR is critical to many of the cellular processes - such as adhesion, proliferation, and migration – which, when disturbed, encode hallmarks of cancer, including invasion and metastasis (Normanno et al., 2006; Hanahan and Weinberg, 2011). However, the clinical significance of this finding must be interpreted cautiously given the complexity of the EGFR pathway (Forastiere, 2007; Gusterson and Hunter, 2009).

1.3.3 EGFR in oral potentially malignant disorders

The prevalence of EGFR protein over-expression may limit its clinical utility as a biomarker in OSCC. However, it has recently been reported that in OPMD over-expression of EGFR protein correlates with an increased risk of malignant transformation (Ries et al., 2013). There are also data which suggest that OPMD with an abnormal EGFR gene copy number are more likely to progress to OSCC relative to cases with normal EGFR gene copy number (Benchekroun et al., 2010; Poh et al., 2012). These data need to interpreted cautiously: abnormal EGFR gene copy number is not regarded as evidence of genomic gain in the criteria that are currently validated for the interpretation of non-small cell lung carcinoma (Hirsch et al., 2003). However, it is unclear whether criteria validated for the interpretation of EGFR gene copy number in a solid tumour, and involving a different organ, should be applied to OPMD. The evidence from these studies support the contention that EGFR gene copy number abnormalities accumulate in the early stages of oral carcinogenesis. Although EGFR copy number abnormalities do not amount to genomic gain, it is conceivable that they precede genomic gain. Abnormal EGFR gene copy number may therefore be a useful diagnostic adjunct for detecting OPMD at an elevated risk of malignant transformation.

1.4 The *SOX* gene family

The *SOX* (SRY-Box) gene family encode transcription factors that play important roles in regulating the development of both vertebrates and invertebrates (Bowles et al., 2000). To date, ~30 member genes have been identified, each encoding transcription factors which share a 79 amino-acid high-mobility group (HMG) DNA-binding domain that was first identified in SRY, the mammalian testis-determining factor (Chew and Gallo, 2009). The 30 SOX transcription factors are sub-classified into ten different groups according to their functional properties, structural motifs, and homology within the HMG domain. Outside of their conserved domain, SOX transcription factors include variable domains that permit binding to a range of co-regulatory proteins. The resulting 'SOX-partner codes' are stable transcription factor complexes that enable SOX proteins to differentially regulate gene transcription by modulating promoter activity (Chew and Gallo, 2009).

SOX genes have multiple cellular functions, including induction/suppression of proliferation and multipotency, and specifying the terminal differentiation of progenitor cells from a variety of lineages. During embryonic development, *SOX* genes are among the earliest groups to be expressed (Chew and Gallo, 2009). Abnormal *SOX* gene expression and function has been identified in a spectrum of neurological disorders, including tumours. SOX proteins have now been validated as diagnostic markers for paediatric medulloblastoma and ependymoma (de Bont et al., 2008).

1.4.1 SOX2 in development

SOX2 belongs to the *SOX B1* subgroup, along with *SOX1* and *SOX3*. Members of this subgroup are consistently expressed by central nervous system (CNS) progenitor cells but not by mature neuronal cells, indicating their role in maintenance of the progenitor state (Chew and Gallo, 2009). This role has been further specified as the direction of basal cell 'fate' choices, i.e. balancing the rates of renewal and differentiation of progenitor cells. *SOXB1* subgroup genes are therefore important during both tissue morphogenesis and homeostasis. *SOX2* is known to maintain pluripotency of embryonic stem cells (Masui et al., 2007). Over-expression of *SOX2* may actually inhibit terminal differentiation

during neurogenesis (Bani-Yaghoub et al., 2006). In mice, *Sox2* plays a vital role in patterning of the anterior foregut endoderm (Que et al., 2007) and trachea (Que et al., 2009). In the squamous epithelium of the tongue and palate, Sox2 expression is strongest in the basal layer where it promotes proliferation and stabilisation of progenitor cells (Okubo et al., 2009). Sox2 is believed to achieve this through synergistic maintenance of Oct3/4 expression (Masui et al., 2007). In neural stem cells, Sox2 is thought to act in concert with factors such as c-myc, Oct4 and KLF4 – the so-called 'Yamanaka' factors - to reprogramme adult cells to induce pluripotency (Takahashi et al., 2007). In the adult human brain, SOX2 is understood to perform its pro-mitotic regulatory function through the selective up regulation of Notch 1 signalling and expression of cyclin-D1, which together promote entry into the cell cycle (Ellis et al., 2004).

Mutations in *SOX2* cause ocular disturbances and are associated with a range of learning disabilities, motor dysfunctions, and seizures (Ragge et al., 2005). *SOX2* mutations have also been implicated in hypoplasia of the anterior pituitary gland, with subsequent gonadotropin deficiency (Tziaferi et al., 2008). SOX2 deficiency is also implicated in the rare AEG ('Anophthalmia-Esophageal-Genital') syndrome (Williamson et al., 2006). Conversely, SOX2 over-expression has been identified in a range of CNS tumours, including those of astroglial, ependymal, and oligodendroglial lineages (Phi et al., 2008). *SOX2* mutations are also implicated in the tumourigenesis of glioblastoma, an aggressive cancer in which it behaves as an oncogene (Pevny and Nicolis, 2010).

1.4.2 SOX2 in cancer

The oncogenic potential of *SOX2* has been investigated in human oesophageal and lung cancers (Garraway and Sellers, 2006; Bass et al., 2009; Hussenet and du Manoir, 2010). Copy number increases of the 3q26-qter chromosomal region are found in a majority (60-80%) of squamous cell carcinoma (SCC), including those in the oesophagus/head and neck region (Hussenet and du Manoir, 2010). Analysis of the individual contribution of genes mapping to 3q26-qter using a shRNA screen in oesophageal cancer cell lines indicated that *SOX2* was a major contributor to tumorigenesis and may act as a lineage-survival oncogene (Bass et al., 2009).

In lung cancer (LC), SOX2 is over-expressed in both adenocarcinoma and SCC phenotypes (Hussenet et al., 2010; Lu et al., 2010). In mice, up-regulation of Sox2 causes hyperplasia of the bronchiolar epithelium and development of adenocarcinomas. This is attributed to the Sox2-mediated induction of cyclinD1 expression and subsequent cell cycle promotion (Lu et al., 2010). *SOX2* is thus considered a potential chemotherapeutic target for lung cancer (Xiang et al., 2011).

SOX2 may also play a role in tumour differentiation. *In vivo*, up-regulation of SOX2 induces expression of markers of squamous differentiation such as p63 and keratin 6, thus implicating *SOX2* in the differentiation pathway of SCC (Bass et al., 2009). This is in contrast to previous in vitro studies suggesting that increased SOX2 expression contributes to de-differentiation and cellular migration (Hussenet et al., 2010; Hussenet and du Manoir, 2010).

1.4.3 SOX2 in oral cancer

In normal squamous epithelium, *SOX2* promotes epithelial proliferation and the stabilisation of basal progenitor cells (Okubo et al., 2009). Over-expression of SOX2 has been described in SCC arising at a range of sites, including the lung, oesophagus, cervix, and penis (Hussenet et al., 2010; Lu et al., 2010; Maier et al., 2011). Up-regulation of SOX2 has been documented in a subset of OSCC (Freier et al., 2010; Misuno et al., 2013; Huang et al., 2014). Two studies have also reported *SOX2* genomic gain in OSCC (Freier et al., 2010; Kokalj Vokač et al., 2014). Freier et al (2010) reported that SOX2 gene copy number amplification was more common than SOX2 protein over-expression, occurring in >50% of OSCC. SOX2 positivity has also been detected in up to 90% of OPMD in an isolated study by Qiao et al (2013). The authors also identified co-expression of SOX2 and Oct4 - a feature not detected in the normal squamous epithelium - in up to 60% of OPMD (Qiao et al., 2013).

There is also evidence that SOX2 over-expression portends poor clinical outcomes in OSCC, particularly metastasis to cervical lymph nodes (Du et al., 2011; Michifuri et al., 2012; Huang et al., 2014). Michifuri et al (2012) drew attention to the heterogeneous expression of SOX2 in a group of 80 mixed-stage OSCC, categorising the SOX2 pattern as either peripheral or diffuse. The

authors identified a positive correlation between the diffuse pattern of SOX2 staining and risk of metastasis to cervical lymph nodes (Michifuri et al., 2012).

These findings suggest that *SOX2* has an oncogenic function in oral carcinogenesis. This contention is supported by earlier functional studies of oesophageal and lung SCC (Bass et al., 2009; Hussenet et al., 2010; Hussenet and du Manoir, 2010; Lu et al., 2010). There is emerging evidence that suggests *SOX2* is significant in the maintenance and functioning of oral cancer stem cells (CSC) (Boumahdi et al., 2014). Oral CSC show up-regulation of SOX2 and OCT4 *in vitro* (Lim et al., 2011; Bourguignon et al., 2012; Misuno et al., 2013). However, further work is required to delineate the contribution of *SOX2* to the maintenance, survival, and proliferation of CSC (Routray and Mohanty, 2014).

1.5 The *PAX* gene family

The *PAX* ('paired-box') gene family comprises a group of nine genes that encode transcription factors with crucial roles in embryonic development and organ formation (Dahl et al., 1997; Chi and Epstein, 2002; Wang et al., 2008). *PAX* genes have been widely conserved through evolution. They are shared between mammalian species and orthologues have been identified in organisms as diverse as birds, fish, and frogs (Dahl et al., 1997). *PAX* genes are characterised by a common 'paired-box' 128-amino acid DNA-binding domain, which was first identified in Drosophila (Bopp et al., 1986; Wang et al., 2008). The transcriptional activities of *PAX* genes regulate processes such as cellular proliferation, differentiation, migration, and survival (Chi and Epstein, 2002; Robson et al., 2006). Mammalian *PAX* genes are classified into four subgroups according to the presence or absence of two further regions: a homeodomain and octapeptide structural motif (Robson et al., 2006).

Mostly active during embryogenesis and early growth, expression of *PAX* genes gradually declines following completion of development. Residual expression and capacity for re-expression persist into adult life, however, when *PAX* genes are involved in maintaining stem cell pluripotency and directing the regeneration of normal or damaged tissues (Chi and Epstein, 2002). Studies of mice with either spontaneous or induced mutations of *PAX* genes confirm their vital role in

embryogenesis. Mutants are distinguished by overall reduction in growth, and malformation/absence of organs (Dahl et al., 1997; Chi and Epstein, 2002). Mutations of *PAX* genes are known to cause developmental abnormalities in humans. For example, *PAX3* mutations have been identified in patients with Waardenburg syndrome (Wollnik et al., 2003) and deletions of *PAX6* are implicated in the aetiology of aniridia (Davis et al., 2008). Their involvement in stem cell renewal, differentiation, and migration suggests that *PAX* genes have significant potential to contribute to neoplasia - frequently characterised by the recapitulation of developmental processes - should their expression become dysregulated (Robson et al., 2006).

1.5.1 PAX9 in development and disease

In humans, *PAX9* is located on chromosome 14q12 and belongs to group one of the *PAX* gene family. It is highly similar to the other member of this group, *PAX1.* In addition to the paired-box domain, both genes are characterised by the presence of an octapeptide structural motif and the absence of the homeodomain that features in groups II – IV (Peters et al., 1998; Robson et al., 2006). *PAX9* has been identified in a wide range of vertebrates, including chicken, mice and zebra fish.

In mice, Pax1 and Pax9 exhibit similar expression patterns in the sclerotome, i.e. the ventro-medial region of somites, which subsequently forms the vertebral column (Neubuser et al., 1995). They demonstrate overlapping expression patterns in structures derived from the endodermal epithelium of the pharyngeal pouches, including thymus, parathyroid glands, ultimobranchial bodies, Eustachian tubes, and tonsils. There is extensive Pax9 expression in the neural crest-derived mesenchymal tissues that are important in the development of teeth and the craniofacial region (Neubuser et al., 1997; Peters et al., 1998).

The importance of *Pax9* function during development has been emphasised by a study using transgenic mice in which *Pax9* was inactivated. Pax9-deficient mice die shortly after birth, display a cleft secondary palate, and lack all teeth, highlighting the role of Pax9 in neural crest-derived tissues. Furthermore, organs derived from the endodermal pharyngeal pouches - thymus, parathyroid glands and ultimobranchial bodies - are also absent (Peters et al., 1998). Tooth

development in Pax9-deficient mutants is arrested at the ‘bud stage’, and the expression of other inductive mesenchymal genes such as *Bmp4*, *Msx1*, and *Lef1* is strongly decreased (Neubuser et al., 1995).

Heterozygous mutations in *PAX9* cause familial and sporadic cases of hypodontia and oligodontia (Stockton et al., 2000; Mostowska et al., 2003), congenital conditions affecting 5-10% of humans in which variable numbers of teeth are missing (hypodontia defined as < six teeth, and oligodontia as >six teeth, both excluding third molars) (Rose, 1966). The role of Pax9 in facilitating dental field patterning and the ‘minimum dosage’ of Pax9 expression required for normal dental morphogenesis have been defined using novel mouse models for hypodontia and oligodontia (Kist et al., 2005). Reduced levels of wild-type Pax9 mRNA resulted in fewer teeth (ranging from hypodontia to severe oligodontia) and defective formation of enamel with subsequently increased rates of attrition and formation of reparative dentine (Kist et al., 2005).

In addition to its role in structures derived from the endoderm and neural crest, PAX9 has also been implicated in the regional differentiation of the mammalian surface ectoderm. This gives rise to the epidermis of the skin, squamous epithelium lining the oral cavity, and a range of appendages including lingual papillae and salivary glands (Jonker et al., 2004). Although Pax9 expression in mice is normally absent in the epidermis, it is continuously expressed in the lingual epithelium. Here, it is believed to regulate key aspects of epithelial differentiation, notably the morphogenesis of filiform papillae (FP). In mice, FP are small, conical structures on the dorsal surface of the tongue that are believed to enhance retention of food during mastication. They are characterised by a distinctive anterior-posterior polarity that reflects differential expression of specific keratins (Iwasaki, 2002). Pax9-deficiency in mice results in down-regulated expression of so-called ‘hard keratins’ (e.g. Krt1-5, Krt1-24) that form the posterior aspect of FP, and concomitant up-regulation of ‘soft keratins’ e.g. Krt2-1, Krt2-17. These ‘soft keratins’ are commonly expressed in the epidermis but are absent in normal lingual epithelium, indicating mis-differentiation of the tongue epithelium into epidermis (Jonker et al., 2004).

1.5.2 A potential role for PAX9 in cancer

In addition to their normal expression in the epidermis in humans, the 'soft keratins' Krt2-1 and Krt2-17 are over-expressed in dysplastic lesions of the oral squamous epithelium (Bloor et al., 2003). This suggests that the underlying alterations in the differentiation pathway of epithelium in Pax9-deficient mice may be similar to those occurring in epithelial dysplasia (Jonker et al., 2004). Our preliminary observations using biopsy samples of the human tongue have demonstrated progressive down-regulation of PAX9 expression in dysplastic epithelium and the complete absence of PAX9 expression in OSCC [unpublished data]. This is consistent with the study by Gerber et al (2002), who studied biopsies of the human oesophagus. Using a monoclonal antibody to PAX9, which shows nuclear expression in normal oesophageal epithelium, this study revealed PAX9 expression to be lost/significantly reduced in a majority of cancers and pre-cancerous lesions. Furthermore, it identified an inverse relationship between PAX9 expression and clinical course: a decrease in the proportion of PAX9-positive cells correlated with an increase in the malignant behaviour of lesions. This suggests that PAX9 plays a pivotal role in the differentiation of the oesophagus epithelium and provides evidence to support a function for PAX9 in cancer (Gerber et al., 2002).

PAX1 and *PAX9* are generally regarded as having a weaker association with cancer than other *PAX* genes due to the absence of the homeodomain that characterises groups II – IV (Robson et al., 2006). Group II and group III *PAX* genes have an established role in promoting tumourigenesis. Conserved chromosomal translocations have now been identified in several – often aggressive – cancers. These include alveolar rhabdomyosarcoma (translocations in *PAX3/PAX7*), follicular thyroid carcinoma (*PAX8*) and non-Hodgkin lymphomas (*PAX5*) (Robson et al., 2006). The precise mechanisms involved are still unclear. PAX5 protein is thought to regulate p53 transcription (Stuart et al., 1995) and PAX8 to regulate the apoptosis suppressor *Bcl-2* (Hewitt et al., 1997). To date, however, *PAX6* is the only family member directly implicated in tumourigenesis *in vivo*, in the development of a pancreatic cystic adenoma in a transgenic mouse (Yamaoka et al., 2000). This finding has led to

further work assessing the potential of *PAX6* as a chemopreventive target (Robson et al., 2006).

Although the potential role of *PAX9* in squamous epithelial cancers is regarded as that of a tumour-suppressor gene, there is evidence that 'driver gene' candidacy cannot be entirely excluded. Together with the transcription factor genes *TTF-1* and *NKX2-8*, *PAX9* has been identified within the amplified chromosomal region 14q13.3 in lung cancer (LC), leading to over-expression of all three genes. Furthermore, PAX9 expression contributes to the survival of lung cancer cell lines *in vitro*. This has led to the suggestion that *PAX9*, in combination with *TTF-1* and *NKX2-8*, may act as an oncogene in LC (Kendall et al., 2007). There is also evidence that *PAX9* interacts with *c-myb*, a proto-oncogene that enhances the survival of SCC cell lines and is implicated in the pathogenesis of a range of cancers (Lee et al., 2008). These findings are consistent with the functions of other *PAX* family members that have been identified using *in vitro* studies (Muratovska et al., 2003). A follow-up study has shown that PAX9 has only minimal prognostic value in predicting lung cancer outcome; however, the evidence is complicated by the partial retraction of this paper (Hsu et al., 2009). In contrast, other work has identified allelic loss at 14q13.3 in a subset of LC, suggesting that in certain tumours PAX9 and factors such as NKX2-8 may perform tumour-suppressor rather than oncogenic functions (Harris et al., 2011).

1.6 Animal models

The paucity of reliable biomarkers to identify the pre-cancerous lesions at greatest risk of malignant progression has been acknowledged (Kanojia and Vaidya, 2006). In part, this reflects the limited availability of suitable animal models in which to study the mechanisms of OSCC formation and test substances that may halt/reverse these processes (Czerninski et al., 2009). While molecular analysis of human biopsy samples is the ideal, tissue from each of the multiple stages in the evolution of a particular lesion (i.e. hyperplasia, dysplasia, carcinoma in-situ) is usually not available. By contrast, tissue from animal models allows each stage of precancerous development to be reproduced and made readily available for histological and molecular analysis (Herzig and Christofori, 2002; Wong, 2009).

Surrogate animal models that are currently recognised to recapitulate the conditions of the human oral cavity include the rat tongue and Syrian hamster buccal pouch (Eveson and MacDonald, 1978; Kanojia and Vaidya, 2006; Mognetti et al., 2006). Mouse models (tongue, oral cavity) have also been widely used. They are comparatively advantageous because of the potential to easily manipulate the mouse genome and the availability of many mouse mutants (Brudno et al., 2004). Spontaneous cases of OSCC are rare in animals (Thurman et al., 1997). However, development of OSCC has been induced in animal models through several methods: chemical induction, xeno-transplantation, and transgenesis (Mognetti et al., 2006).

1.6.1 Chemical induction of OSCC in animal models

Several chemical carcinogens have been applied to animal models, including arecaidine, coal tar, and polycyclic aromatic hydrocarbons such as DMBA (7, 12-dimethylbenz(a)anthracene and 9,10-dimethyl-1,2–benzanthracene) (Kanojia and Vaidya, 2006). DMBA was one of the first carcinogens successfully used to induce OSCC in animals (Salley, 1954). However, as a potent irritant it caused non-specific inflammation, epithelial necrosis, and proliferation of granulation tissue, which together obscured the interpretation of precancerous lesions. There is also evidence that OSCC induced by DMBA is histologically different to that of human OSCC (Kanojia and Vaidya, 2006; Mognetti et al., 2006). Chemical induction using 4-nitroquinoline 1-oxide (4-NQO) avoids many of these problems and is currently the preferred model for studying OSCC carcinogenesis in mice (Vered et al., 2005).

1.6.2 4-Nitroquinoline 1-oxide (4-NQO)

4-NQO is a synthetic, water-soluble carcinogen that was derived from quinoline in the late 1950s (Nakahara et al., 1957). Initially developed as a possible chemotherapeutic agent, it was found that local application induced cancers of the skin and labial/lingual mucosa of mice (Fujino et al., 1965). A publication in the early 1970s described a comprehensive 4-NQO animal model in which topical application of 4-NQO over seven months generated cancers of the hard palate, tongue base, gingiva, and stomach in rats (Wallenius and Lekholm, 1973). 4-NQO is now known to result in a spectrum of pre-cancerous as well as cancerous lesions, including hyperplasia, squamous cell papilloma (a benign

exophytic neoplasm), various grades of epithelial dysplasia and, most significantly, OSCC (Tang et al., 2004; Kanojia and Vaidya, 2006). Crucially, these changes occur in the absence of non-specific inflammatory changes: 4-NQO is less irritant than other chemical carcinogens such as DMBA, and the histology is similar to that of human OSCC (Eveson and MacDonald, 1978).

Although early studies were confined to topical application, work has since shown that systemic exposure to 4-NQO, which can be readily added to drinking water, causes similar and more reproducible changes (Tang et al., 2004). 4-NQO may induce cancers of the lung following subcutaneous injection (Imaida et al., 1989). However, following oral ingestion it has a recognised fidelity to the oral cavity and oesophagus. Necroscopy studies have failed to demonstrate either cancers or precancerous lesions elsewhere in the digestive tract (stomach, intestine, and bowel) or other viscera (liver, lungs) of mice (Tang et al., 2004). The localisation of 4-NQO induction to the oral cavity and oesophagus is attributed to the high concentration of diaphorase, a reductase that is required to activate 4-NQO, in the mucosa at these sites (Kanojia and Vaidya, 2006).

1.6.3 4-NQO method of action

The carcinogenic action of 4-NQO is potentiated by reduction of its nitro group to 4-hydroxy-aminoquinoline-1-oxide (4-HAQO). 4-HAQO is implicated in the formation of DNA adducts and undergoes further metabolism and acetylation to form the seryl-AMP-enzyme complex that introduces quinoline groups into DNA (Kanojia and Vaidya, 2006). *In vivo*, the 4-HAQO metabolite preferentially reacts with guanine residues (Tada, 1976). The acetylated metabolite, seryl-AMP-enzyme complex, also reacts through its 3rd and 4th positions with guanine at N2 and C8 respectively (Fronza et al., 1992). It is these adducts, which induce a guanine to adenosine substitution, that are implicated in the high mutagenicity of 4-NQO (Kanojia and Vaidya, 2006; Vitale-Cross et al., 2009). As noted previously, nitrosamines such as NNK (4-(methylnitrosamino)-1-(-3-pyridyl)-1-butanone) are among the most significant carcinogens in tobacco and preferentially bind to guanine residues to induce nucleotide transversions. 4-NQO is also associated with the induction of *p53* mutations (Heniford et al., 1993) and exertion of potent oxidative stress by producing free-radical oxygen

species, such as the superoxide radical. These changes are similar to the molecular and genetic alterations induced by tobacco (Hecht, 2003; Kanojia and Vaidya, 2006). Additionally, recent evidence has suggested that 4-NQO induces up-regulation of the Wnt/β-catenin pathway, inappropriate activation of which is implicated in a range of human cancers (Fracalossi et al., 2010).

Together, the site fidelity of and genetic alterations induced by the 4-NQO model simulate key aspects of oral/oesophageal SCC and it is therefore considered a valid model for the study of human OSCC. However, despite its recognition as the best available experimental system to date, the 4-NQO model is associated with certain limitations. In animals, 4-NQO-induced OSCC tend to be well differentiated. By contrast, most human OSCC range from moderately to poorly differentiated (Nauta et al., 1996). There is a markedly reduced tendency for 4-NQO-induced OSCC to metastasize to loco-regional lymph nodes in comparison with human OSCC. Conversely, patterns of local invasion may be more aggressive, with neural invasion reported as a typical finding in 4-NQO models. These suggest underlying differences in the biology of cancerous cells generated by the 4-NQO model. Although careful optimisation of dosage mitigates this complication, 4-NQO-induced OSCC often exhibits a multifocal distribution. Despite the phenomenon of field cancerisation, an initial multifocal presentation of OSCC in humans is rare (Hasina et al., 2009).

1.6.4 Experimental protocols for 4-NQO induction of carcinogenesis in mouse models

Over the past 15 years, workers have attempted to optimise an experimental protocol for the 4-NQO model. Parameters such as the duration, site, and mechanism of carcinogen application, along with the overall duration of treatment and point of sacrifice, have been adjusted (Vitale-Cross et al., 2009). Although lesions in the oral cavity may be induced by brushing the mucosa with highly-concentrated 4-NQO solutions, administration via drinking water is currently preferred due to its ease and reproducibility (Kanojia and Vaidya, 2006; Vitale-Cross et al., 2009).

Early studies administering 4-NQO via drinking water used low doses, ranging from 10–20 µg/mL, for periods of up to 50 weeks. These dosages induced a spectrum of genetic mutations, pre-cancerous lesions, and SCC within the oral

cavity (Ma et al., 1999; Ide et al., 2001). Although one study identified dysplastic epithelial changes following an exposure time as short as two weeks (Nauta et al., 1996), workers have consistently reported a latent period between the start of induction and the detection of epithelial dysplasia/OSCC of up to eight weeks (Dayan et al., 1997; Tang et al., 2004; Vered et al., 2005).

Tang et al (2004) were the first to administer a higher dose of 100 µg/mL to CBA and C57/Bl6 mice. They demonstrated induction of pre-neoplastic and neoplastic lesions in 100% of mice exposed over a 16–week period. This study was the first to compare the outcomes of topical and systemic administration of 4-NQO and to identify the development of oesophageal SCC in addition to lesions of the oral cavity. Tang et al (2004) also found that 100% of mice on a dosage of 50 ug/mL demonstrated premalignant changes after the same 16-week period. A more recent study has confirmed the suitability of administering a 100 µg/mL dosage, which is not reported to result in undue morbidity/mortality. This group found that a sufficient incidence of OSCC was induced after as little as eight weeks to enable assessment of a chemotherapeutic agent (Hasina et al., 2009).

1.6.5 The 4-NQO model and chemotherapy

Chemotherapy may extend the latent period of carcinogenesis and help to maintain patients' quality of life in the terminal stages of cancer. The 4-NQO mouse model has been used in a range of studies assessing the action and efficacy of chemotherapeutic agents. For example, the MTOR (Mammalian Target of Rapamycin) signalling pathway is a recognised therapeutic target. Rapamycin administration has been shown to halt progression of precancerous lesions and promote regression of OSCC using the 4-NQO model (Czerninski et al., 2009). Anti-angiogenic agents, such as ABT-510 and Vandatanib, have also been validated in reducing the incidence of dysplasia and OSCC induced by 4-NQO (Hasina et al., 2009; Zhou et al., 2010). Erlotinib, an inhibitor of the EGFR-STAT3 signalling pathway, has been shown to result in a 69% decrease in epithelial dysplasia/OSCC compared to controls (Leeman-Neill et al., 2011). The chemopreventive potential of dietary xanthophylls has also been explored using the 4-NQO model, in a study that found both astaxanthin and canthaxin consumption to be associated with significantly lower cancer rates than controls

(Tanaka et al., 1995). The potent chemopreventive all-trans retinoic acid, an isomer of 13-cis-retinoic acid that regulates epithelial cell proliferation, has been developed using a 4-NQO model to trial loading microspheres and optimise concomitant celecoxib anti-inflammatory therapy (Park et al., 2005).

1.6.6 Genetic induction of OSCC in mouse models

Studies using transgenic and knockout mice have contributed to our current understanding of tumourigenesis. A range of tumours has been replicated in transgenic models, but genetic modifications resulting in OSCC in mice are still rare (Mognetti et al., 2006). Transgenic models are limited by poor penetrance in mice with heterologous promoters, who may express non-physiological levels of the transgene product. Conversely, homozygous knockout mice may fail to complete embryogenesis or die shortly after birth (Mognetti et al., 2006). Other confounding factors include a tendency to develop multiple primary tumours and, later, failure to develop metastases. It may also be difficult to translate experimental findings regarding isolated mutations in transgenic models to humans, where tumours represent numerous sequential genetic alterations (Fearon and Vogelstein, 1990; Frijhoff et al., 2004).

Notwithstanding these difficulties, transgenic models have been successfully used to assess the contribution of specific genetic alterations to the development of OSCC. These include a study of the role played by the *K-ras* oncogene in promoting OSCC, in a mouse model that used tetracycline as an inducer of transgene expression in basal keratinocytes (Vitale-Cross et al., 2004b). Transgenic mouse models have also helped to delineate the function of *Smad4* and *p53* mutations in OSCC (Opitz et al., 2005; Redman et al., 2005; Bornstein et al., 2009).

It has been suggested that combining the 4-NQO model with a transgenic model in which mice are defective in a certain gene would better reflect the complex contribution of genetic and environmental factors during the formation of OSCC (Wong, 2009). To date, a limited number of studies have combined these models. The role of *p53* in 4-NQO-induced OSCC formation has been studied by a group who applied the 4-NQO model to mice with a dominant-negative *p53* mutation. Using micro-array and gene mapping analysis, this study identified a reduction in p53-dependent apoptosis and cell-cycle arrest

pathways that consequently led to increased cell proliferation (Zhang et al., 2006). A combined 4-NQO/transgenic model has contributed to the evidence implicating HR-HPV in the aetiology of OPSCC (El-Mofty, 2007). Application of 4-NQO to mice transgenic for the HPV oncogenes *E6/E7* demonstrated that these mutants had significantly increased susceptibility to developing OSCC (Strati et al., 2006). Another 'combined model' has also been used to test the hypothesis that the nucleotide excision repair gene *XPA* (xeroderma pigmentosum group A) has a protective effect during oral carcinogenesis (Ide et al., 2001).

1.6.7 Investigating a potential tumour-suppressor role of Pax9 in OSCC

A transgenic mouse line, $Pax9^{flox}$, has been developed that permits conditional, tissue-specific inactivation of *Pax9* (Kist et al., 2007). Conditional inactivation has been achieved by crossing $Pax9^{flox}$ mice with a transgenic strain, K14-Cre, which expresses Cre recombinase from the keratin 14 promoter that is active in the basal keratinocytes of stratified squamous epithelia (Vasioukhin et al., 1999). Tongues of conditional Pax9-deficient mutants exhibit fissuring and a smooth, discoloured surface macroscopically. Histopathological examination and scanning electron microscopy further reveal loss of filiform papillae, reduced thickness of the lamina propria, and focal hyperplastic epithelial growths (Figure 1-2). One aged Pax9-deficient mutant developed an adenoma of the Harderian gland, a specialised lacrimal gland present in mice. Genome wide expression analyses of tongue tissue from these mutants have shown deregulation of numerous genes involved in epithelial differentiation as well as a slight reduction of tumour suppressor gene expression [unpublished data]. We therefore hypothesise that exposing Pax9-deficient mice to 4-NQO may induce more extensive and more rapid pre-cancerous changes and progression to OSCC than in 4-NQO treated control mice.

1.6.8 Normal lingual microanatomy of the Pax9-deficient mouse

The normal mouse tongue shares many of its micro-anatomical features with the normal human tongue. The dorsal surface is covered by specialised mucosa. The ventral surface is covered by non-specialised lining mucosa. Squamous epithelium lining both surfaces is separated from the underlying musculature by a thin lamina propria.

In both humans and mice, the dorsal tongue is lined by specialised squamous epithelium that is characterised by orthokeratosis and the presence of numerous filiform papillae (Figure 1-2). In humans, the ventral tongue is lined by non-specialised squamous epithelium, which is thinner than the dorsal epithelium and has a flat basement membrane. In mice, however, the ventral epithelium shows orthokeratosis, is only slightly thinner than the dorsal epithelium, and has an undulating rete outline. Keratinisation of the ventral surface enables the mouse epithelium to withstand greater frictional forces during mastication than the human tongue. This variation is likely to reflect the mouse diet, which is generally coarser than that of humans.

The tongue of the Pax9-deficient mouse is characterised by a smooth dorsal surface. This is due to failure of development of filiform papillae. Although rudimentary, malformed filiform papillae may be detected, they are short and sparse (Figure 1-2). The dorsal surface is also characterised by fissures and thinning of the lamina propria (Figure 1-2).

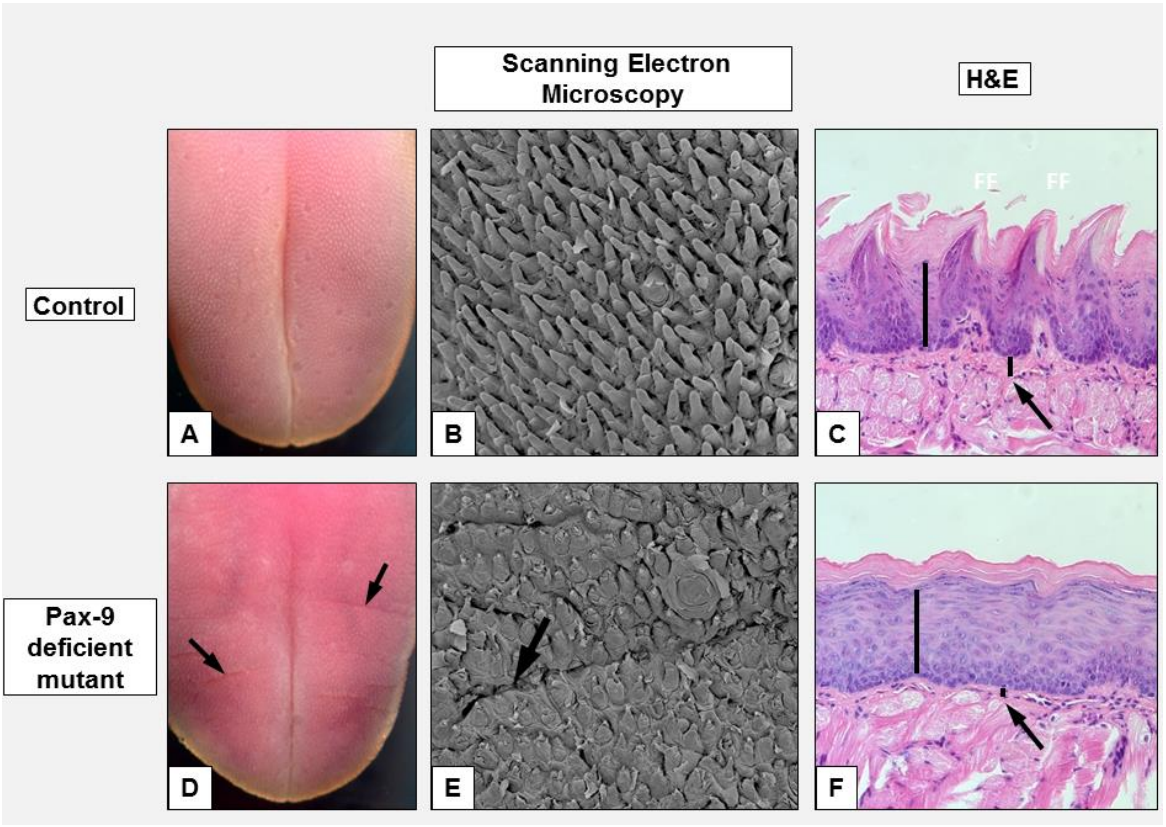


Figure 1-2 Normal micro-anatomy of the tongue of control (wild-type) and Pax9-deficient mice

A) The dorsal surface of the control mouse. B) Scanning electron microscopy shows well-organised filiform papillae. C) Filiform papillae are also seen histologically. D) The dorsal surface of the Pax9-deficient mouse is relatively smooth and shows fissuring (arrows). E) Scanning electron microscopy shows loss of filiform papillae and evidence of fissuring (arrow). F) Histologically, there are sparse filiform papillae that are short and malformed. The epithelium is thinner than the control (bars). There is also thinning of the lamina propria (arrows).

H&E images taken at x100 magnification.

1.7 Oral keratinocyte cell lines

1.7.1 Primary culture of normal oral keratinocytes

In vitro, normal human somatic cells exhibit limited replicative capacity. Replicative potential is limited despite optimisation of cells' nutritional and mitogenic requirements (Smith and Pereira-Smith, 1996; Sedivy, 1998). This phenomenon is termed senescence. It reflects the erosion of telomeric DNA with each successive cellular replication. Telomeres are located at the ends of chromosomes and prevent the chromosomes from forming end-to-end fusions. When shortened, however, telomeres trigger the onset of senescence. In primary culture, cells will generally proliferate at first. However, following ~50 to 100 successive replications there is sufficient erosion of telomeres to trigger senescence (Allsopp et al., 1992). Normal oral keratinocytes undergo senescence following an average of three to four passages (Prime et al., 1990). Growth *in vitro* may supported by co-culturing keratinocytes with mesenchymal cells, e.g. mitocyn C-treated 3T3 fibroblasts (Rheinwald and Green, 1975).

Normal, non-neoplastic human fibroblasts may become immortalised and successfully evade senescence *in vitro* (Bodnar et al., 1998). Immortalisation occurs due to ectopic expression of the catalytic subunit of telomerase, hTERT. The hTERT catalytic subunit is usually expressed only by malignant cells (Meyerson et al., 1997). This subunit enables the cells to synthesise telomeres, and thus prevent telomere-induced cellular senescence (Feng et al., 1995).

It has been documented that, *in vitro*, normal oral keratinocytes may spontaneously evade senescence and become immortal (Boukamp et al., 1988). Immortal keratinocytes are also characterised by the ectopic expression of hTERT. However, immortalisation also requires a gain-of-function mutation in the *p16^{INK4A}* gene. There is a complex pattern of p16^{INK4A} protein expression that regulates senescence in a pathway that is independent of telomere-induced shortening. An oral keratinocyte cell line that exhibits both ectopic expression of hTERT and over-expression of p16^{INK4A} has been generated. This OKF6/hTERT cell line evades senescence without exhibiting abnormalities of either growth or differentiation (Dickson et al., 2000).

1.7.2 OSCC cell lines

Two of the oldest cancer cell lines are derived from the oral cavity (Eagle, 1955; Moore et al., 1955). For many years, however, OSCC cell lines have proved difficult to culture and characterise (Krause et al., 1981). *In vitro*, keratinocytes derived from OSCC exhibit diverse phenotypes, and may show characteristics of both normal and malignant cells (Parkinson, 1989). Useful markers of malignancy *in vitro* include resistance to terminal differentiation (Rheinwald and Beckett, 1980), altered expression of TGF-β cell surface receptors (Shipley et al., 1986), and changing patterns of oncogene expression (Field and Spandidos, 1987).

By contrast to their clinical behaviour *in vivo*, cells cultured from OSCC may show little or no growth. Lack of growth may be due to inappropriate cell culture conditions, or insufficient numbers of cancer stem cells (CSC) (Rheinwald and Beckett, 1981). OSCC-derived keratinocytes may be cultured under similar conditions to normal keratinocytes. However, OSCC-derived cells vary in their dependence on mesenchymal support. Mesenchymal support may actually compromise the growth of some OSCC-derived cells. Factors derived from OSCC cultures may be cytotoxic to 3T3 cells, further complicating culture conditions (Rupniak et al., 1985; Prime et al., 1990).

Despite the challenges that the primary culture of malignant oral keratinocytes presents, a range of OSCC-derived cell lines has now been established (Prime et al., 1990; Edington et al., 1995). These cell lines are valuable tools in the study of oral carcinogenesis. They have been characterised with regard to the expression of cytokeratin, vimentin, and involucrin (Sugiyama et al., 1993; Prime et al., 1994a), and tumorigenicity following both subcutaneous and orthotopic transplantation in athymic mice (Prime et al., 1994b; Paterson et al., 2002). They have also been characterised with regard to *H-ras* and *p53* mutations (Yeudall et al., 1993; Yeudall et al., 1995), inhibitory response to TGF-β1 (Prime et al., 1994a), autocrine production of TGF-β isoforms (Fahey et al., 1996) and expression of TGF-β cell surface receptors (Prime et al., 1994a). Several lines have been successfully transfected with plasmid DNA (Paterson et al., 2002). The literature to date suggests that these cell lines will provide excellent models of oral carcinogenesis both *in vitro* and *in vivo*.

1.8 Conclusion

This review has highlighted the global burden of OSCC, the limitations of the treatment modalities currently available, and the consequent need to develop biomarkers that improve the management and prognostication of patients with OPMD and OSCC.

The literature suggests that EGFR gene copy number is a potential biomarker in OPMD and OSCC. However, the nature, prevalence, and biological significance of EGFR gene copy number alterations in the early stages of oral carcinogenesis remain unclear. Similarly, there is emerging evidence that *SOX2* may have an oncogenic function in oral carcinogenesis, is critical to the survival of cancer stem cells, and correlates with clinical outcomes in OSCC. However, there is only a limited number of *SOX2* studies available in the literature, particularly studies that describe SOX2 expression in OPMD. Evidence from an isolated study of oesophageal cancer suggests that *PAX9* has a tumour-suppressor function in the squamous epithelium. To date, however, the role of *PAX9* in oral carcinogenesis is unknown.

1.9 Hypothesis

This study will test the hypothesis that EGFR, SOX2, and PAX9 may be useful prognostic biomarkers in potentially malignant disorders and early-stage squamous cell carcinoma of the oral cavity.

1.10 Aims

1. To determine the expression profiles of EGFR, SOX2, and PAX9 in groups of patients with OPMD and early-stage OSCC
2. To correlate the expression profiles of EGFR, SOX2, and PAX9 with patients' characteristics and clinical outcomes
3. To determine whether conditional Pax9-deficient mice are predisposed to the development of OSCC following exposure to 4-NQO
4. To develop tools that will facilitate exploration of the functional significance of SOX2 and PAX9 in oral carcinogenesis.

Chapter 2. Materials and Methods

2.1 Human tissue samples: oral potentially malignant disorders

Patients with oral potentially malignant disorders (OPMD) were identified retrospectively using two strategies:

1) attendance at an OPMD clinic, Department of Oral and Maxillofacial Surgery, Newcastle-upon-Tyne NHS Hospitals (NUTH)
2) a search of the electronic database, Department of Cellular Pathology, NUTH

2.1.1 Attendance at an OPMD clinic

The characteristics and clinical outcomes of a group of 100 patients attending an OPMD clinic have been published by Diajil et al (2013). Patients in the study presented with an OPMD. Incisional biopsy of the OPMD showed epithelial dysplasia. Dysplastic lesions were managed either by surveillance or laser excision. The study spanned a 13-year period (1997-2009). Patients were followed up for a minimum of 24 months. The study excluded patients who had either a history of oral/upper aero-digestive tract malignancy or a previous OPMD.

2.1.2 Electronic database of the Cellular Pathology Department

A systematic search of the pathology database was carried out using SNOMED codes. The search identified a group of patients with oral squamous cell carcinoma (OSCC) who had previously undergone a biopsy (or multiple biopsies) of an OPMD at the same mucosal subsite. The OPMD was managed either by surveillance or laser excision. The search spanned a 13-year period (1997 - 2009). The search identified a total of 67 cases in which OSCC was preceded by an OPMD. This group was refined by excluding cases with a history of oral/upper aero-digestive tract malignancy; OSCC formation at a separate mucosal subsite to the index OPMD; less than six months between index OPMD biopsy and OSCC diagnosis (the index biopsy may have insufficiently sampled an existing OSCC); cases with non-dysplastic OPMD diagnosed with specific clinico-pathological entities, e.g. chronic hyperplastic candidosis or lichen planus; cases in which the clinical presentation and

histological features were suggestive of the clinical entity proliferative verrucous leukoplakia (PVL).

A separate search of the electronic database identified patients with biopsies that showed normal oral mucosa. The search spanned a six-year period (2007-2012) and identified a total of 24 cases. These cases were used as controls for immunohistochemical analysis (section 2.4).

2.1.3 Histopathological specimens

Haematoxylin and eosin (H&E) stained sections and formalin-fixed paraffin-embedded (FFPE) blocks were retrieved from the Department of Cellular Pathology archive for the OPMD identified by the searches. Cases that satisfied the study's inclusion criteria and had sufficient material for further assays were selected for further analysis. Serial 4 µm sections were taken from each block. Individual sections were mounted on coated slides (Superfrost Plus, Thermo Fisher Scientific, UK). Sectioning was carried out in the Department of Cellular Pathology by Biomedical Scientists.

2.1.4 Histological grading of epithelial dysplasia

H&E sections were reviewed by two pathologists (Dr M. Robinson and Prof. P. Sloan). The grade of epithelial dysplasia was scored independently for each case. Epithelial dysplasia was graded using two classification systems:

1) Squamous Intra-epithelial Neoplasia (SIN) classification (Barnes et al., 2005)
2) Binary (low-grade/high-grade) classification (Kujan et al., 2006).

All discordant cases were reviewed at a meeting between the pathologists and assigned a single grade by consensus.

2.1.5 Clinical outcomes

Clinical outcomes were categorised into four groups: 1) no adverse outcome; 2) local recurrence; 3) new lesion; and 4) malignant transformation, after Diajil et al (2013). The criteria for each of these categories is summarised in Table 2-1.

Table 2-1 Categories of clinical outcome for oral potentially malignant disorders

Clinical outcome category	Criteria
No adverse outcome	No evidence of local recurrence, new lesion formation, or malignant transformation after minimum 24-month follow-up period
Local recurrence	Recurrence of epithelial dysplasia at the same mucosal subsite
New lesion	Development of a new epithelial dysplasia at a separate mucosal subsite
Malignant transformation	Progression from epithelial dysplasia to oral squamous cell carcinoma

Adapted from Diajil et al (2013).

2.1.6 Data collection and management

For each OPMD the following data were collected and entered into an Excel spreadsheet:

- patient's demographic data (sex and age at initial OPMD biopsy)
- oral mucosal subsite of the OPMD
- date of index OPMD biopsy
- patient's risk factors (alcohol and tobacco habits)
- clinical management of the index OPMD (laser excision or surveillance)
- histological grade of epithelial dysplasia (SIN and binary classification)
- clinical outcome (Table 2-1)
- date of any subsequent histological diagnosis (e.g. local recurrence of epithelial dysplasia)
- for the subset of cases that underwent malignant transformation, the histological differentiation and pStage of the subsequent OSCC was also recorded.

2.2 Human tissue samples: early-stage OSCC

Patient data for the early-stage oral squamous cell carcinoma (OSCC) group was collected by Mr M. Kennedy during his Masters' research project.

2.2.1 Patients

Patients were identified retrospectively using a systematic database search. The search identified consecutive cases of early-stage (i.e. pStage I/II) OSCC managed by the Head & Neck Multi-Disciplinary Teams (HNMDT) at NUTH and Sunderland Royal Hospital (SRH). The search spanned an eight-year period (2000-2008). The search involved three stages:

1) A primary search of the DAHNO (DAta for Head and Neck Oncology) national cancer database
2) Data from the DAHNO database was then cross referenced with data recorded by HNMDT at NUTH and SRH
3) Electronic pathology reports archived at the Department of Cellular Pathology were used to verify the diagnosis/histological differentiation and pStage of each OSCC (the Department of Cellular Pathology provides diagnostic pathology services to HNMDT at both NUTH/SRH).

The search excluded cases with: OSCC at pStages III and IV; OSCC of the lip and oropharynx; non-surgical primary management (i.e. radiotherapy/chemotherapy); a previous diagnosis of head and neck cancer; a history of radiotherapy to the head and neck region for non-malignant conditions; patients without follow-up.

The disease status, length of post-surgical follow up, and survival data for cases that satisfied the study's criteria were verified using:

- Hospital patient administration systems, which record:
    - attendance at outpatient clinics
    - most recent survival in patients discharged from the HNMDT
    - date of death (where applicable)
- General medical practitioners' records, to establish survival status of patients with no recent hospital follow-up

- HNMDT clinic correspondence and records, which record:
    - patients' treatment
    - local recurrence and other adverse outcomes
    - last known disease status
- Hospital notes, for a small number of patients for whom survival/disease status could not be determined from the other sources above.

2.2.2 Histopathological specimens

H&E stained sections and FFPE blocks were retrieved from the Department of Cellular Pathology archive. H&E sections were reviewed by one pathologist (Dr M. Robinson) to identify a representative block with sufficient remaining material for further assays. Blocks were serially sectioned by trained laboratory staff in the Department of Cellular Pathology.

2.2.3 Data collection and management

For each OSCC, the following data were collected and entered into an Excel spreadsheet:

- patients' demographic data (sex and age at OSCC diagnosis)
- the oral mucosal subsite of the OSCC
- date of the index OSCC diagnosis
- histological grade of differentiation (Broders' classification)
- clinical outcomes (disease-free survival and overall survival)
- the date of any subsequent histological diagnosis.

2.3 Human tissue samples: OSCC that transformed from OPMD

The pathology database was used to identify the pStage, histological grade of differentiation, and any subsequent diagnoses for patients with OSCC that transformed from a pre-existing OPMD. Details of patients' demographics and risk factors were already available from the OPMD database. Clinical follow-up data (including overall survival and disease status) was obtained from HNMDT meeting records. The data were entered into an Excel spreadsheet.

2.4 Immunohistochemical staining and analysis

2.4.1 Automated immunohistochemistry

For all human OPMD/early-stage OSCC cases, SOX2 and EGFR immunohistochemistry (IHC) was performed on 4 µm sections using an automated platform (Ventana Benchmark Autostainer, Ventana Medical Systems Inc, USA). Staining was carried out in the Department of Cellular Pathology by Biomedical Scientists. SOX2 protein expression was detected using a proprietary rabbit monoclonal antibody (anti-SOX2 SP76 clone, Cell Marque Corporation, USA). EGFR protein expression was detected using a proprietary rabbit monoclonal antibody that binds specifically to the EGFR internal domain (anti-EGFR 5B7 clone, Ventana Medical Systems Inc, USA). Staining was performed using protocols recommended in the manufacturers' instructions. Morphologically normal squamous epithelium provided a positive internal control for each section. A negative control, with primary antibody omitted, was performed for each staining batch.

2.4.2 Manual immunohistochemistry protocol

During this project, manual IHC was performed in order to detect a range of proteins in human, mouse, and cell culture-generated specimens (cell pellets and Surepath™ preparations). Details of each antibody and the specimens on which they were used are summarised in Table 2-2.

Whilst optimising the manual staining protocol for each antibody, IHC was performed on sections of human normal oral mucosa.

Once optimised, manual IHC was performed in batches of between eight and 12 slides. For sections of human OPMD/OSCC and mouse tongues, morphologically normal squamous epithelium provided a positive internal control for each slide. When IHC was performed on cell culture-generated specimens, the batch of slides included one slide with a section of human normal oral mucosa as the positive control.

For each of the antibodies, manual IHC was performed using a commercially-available kit (DAKO, Glostrup, Denmark).

The following description outlines the standard protocol used for manual IHC:

- Sections de-waxed in fresh xylene (4x 5 mins immersions)
- Sections hydrated through graded alcohols:
    - 99% (x2), 95%, and 70% (5 mins each)
- Slides immersed in distilled water (10 mins)
- Slides transferred to citrate buffer, pH 6 (Sigma-Aldrich, UK)
- Slides placed in an automated decloaker (MenaPath, A Menarini Diagnostics, UK) for antigen retrieval:
    - Heated to 125°C and held at 20-25 psi for 30 secs
- (TBS rinse*)
- Periphery of each slide dried carefully and silicone fast well (FastWells, Interchim, France) firmly applied
- Endogenous peroxidase blocked with 300 µl hydrogen peroxide (5 mins)
- (TBS rinse*) Antibody solutions made up (see Table 2-2).
- Slides incubated with 300 µl primary antibody solution (45 mins)
    - Incubated at room temperature in a humid reaction chamber placed on an automated horizontal rocker
- (TBS rinse*)
- PAX9 only: slides incubated with 300 µl secondary antibody (45 mins)
- (TBS rinse*)
- Incubated with 300 µl horseradish peroxidase-labelled polymer (30 mins)
- (TBS rinse*)
- Incubated with 3,3'-Diaminobenzidine (DAB) solution (5 mins)
- Rinsed in tap water to stop DAB reaction (5 mins)
- Sections counterstained in Harris' haematoxylin (5 secs)
- Slides placed in Scott's Bluing solution (30 secs)
- Sections dehydrated through graded alcohols
    - 70%, 95%, 99% (x2) (5 mins each)
- Slides immersed in fresh xylene (2x 5 mins)
- Coverslips placed (Leica CV5030 Coverslipper, Leica, UK).

* N.B. TBS Rinse: tris-buffered saline (TBS) at pH 7.6 (2x 5 mins)

Table 2-2 Summary of antibodies used for each protein assay/specimen by manual immunohistochemistry

Protein assay	Specimens	Primary antibody		Secondary antibody	
		Dilution	Details	Dilution	Details
PAX9	Human OPMD/OSCC Mouse tongues Cell pellets Surepath™ preparations	1/40	Pax-9 (7C2) Rat Anti-Mouse Monoclonal Abcam*	1/100	Immunoglobulins/HRP P0450 Rabbit Anti-Rat Polyclonal DAKO
SOX2	Mouse tongues Cell pellets Surepath™ preparations	1/100	Sox2 (D6D9) Rabbit monoclonal Cell Signalling Technology	N/A	N/A
Ki67	Mouse tongues	1/100	Ki67 (SP6) Rabbit monoclonal Abcam	N/A	N/A
P53	Mouse tongues	1/50	p53 (7F5) Rabbit monoclonal Cell Signalling Technology	N/A	N/A

*Although this PAX9 antibody is now commercially available, we used aliquots of the supernatant generated by hybridoma cells prior to marketing. The precise concentration of the antibody was therefore unknown.

2.4.3 Digital image analysis of immunohistochemical staining

Following IHC, slides were scanned using an Aperio CS2 Scanscope™ platform (Leica Biosystems, UK) at x20 magnification. For human OPMD/OSCC and mouse tongues, the corresponding H&E sections were also scanned. Electronic files were uploaded to the Aperio Spectrum system for digital image analysis (Spectrum Version 11.1.0.751©, Aperio Technologies, Inc, Leica Biosystems, UK). H&E sections were used to guide the annotation of areas on the corresponding IHC-stained sections according to morphology (i.e. areas of histologically normal epithelium, epithelial dysplasia, and OSCC). Areas were annotated using a freehand pen tool (Figure 2-1). Wherever there was sufficient material, a minimum of 1000 cells were selected per morphological area. Annotated areas were then analysed using two standardised Aperio algorithms (nuclear algorithm: PAX9, SOX2, Ki67, and p53; cellular algorithm: EGFR (Cregger et al., 2006)). The algorithms generated a range of parameters, including the total number of nuclei/cells analysed; the percentage of positive nuclei/cells (PPN/PPC); and the percentage of strongly-positive nuclei/cells (3+PN/3+PC). These values were exported into an Excel spreadsheet and collated prior to statistical analysis.

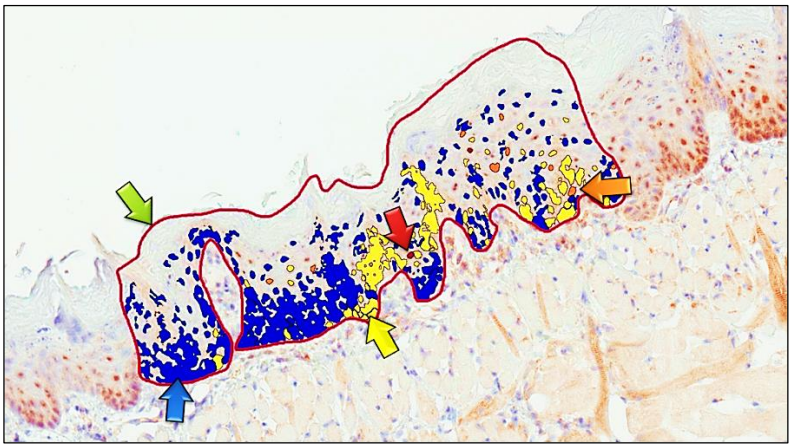


Figure 2-1 Digital image analysis using Aperio software

Areas for digital image analysis were selected using a freehand pen tool (green arrow, red outline). The Aperio nuclear algorithm detected and stratified nuclei according to their signal as either negative (0, blue), weakly positive (1+, yellow), positive (2+, orange), or strongly positive (3+, red). The algorithm used the percentages of 0, 1+, 2+ and 3+ nuclei to calculate the overall percentage of positive nuclei (PPN) and the percentage of strongly positive nuclei (3+PN).

Adapted from a screenshot taken using the Aperio Scanscope software at x100 magnification.

2.5 *In situ* hybridisation for EGFR chromosome 7

2.5.1 Automated staining

The EGFR *in situ* hybridisation (ISH) signal was assessed by a dual colour technique using proprietary reagents (INFORM EGFR-Chromosome 7 dual colour assay, Ventana Medical Systems Inc, USA). This technique detects both the EGFR gene (using silver ISH, seen as black nuclear dots) and chromosome 7 centromeres (using Ultraview Alkaline Phosphatase Red ISH, seen as red nuclear dots) on the same section. Staining was performed by Biomedical Scientists in the Department of Cellular Pathology, using the Ventana Benchmark Autostainer according to the manufacturer's instructions. DNA probes were omitted from negative controls.

2.5.2 Scoring and interpretation of ISH-stained sections

Dual-stained ISH sections were examined independently by two pathologists (Dr M. Robinson, Timothy Bates). Morphologically normal squamous epithelium was used to calibrate the appearance of disomy (seen as two black dots and two red dots). According to the predominant nuclear signal, each case was assigned to one of the descriptive categories described for non-small cell lung carcinoma (Figure 2-2). Actively dividing and overlapping cells were not assessed. All discordant cases were reviewed at a meeting between the pathologists and assigned a single grade by consensus.

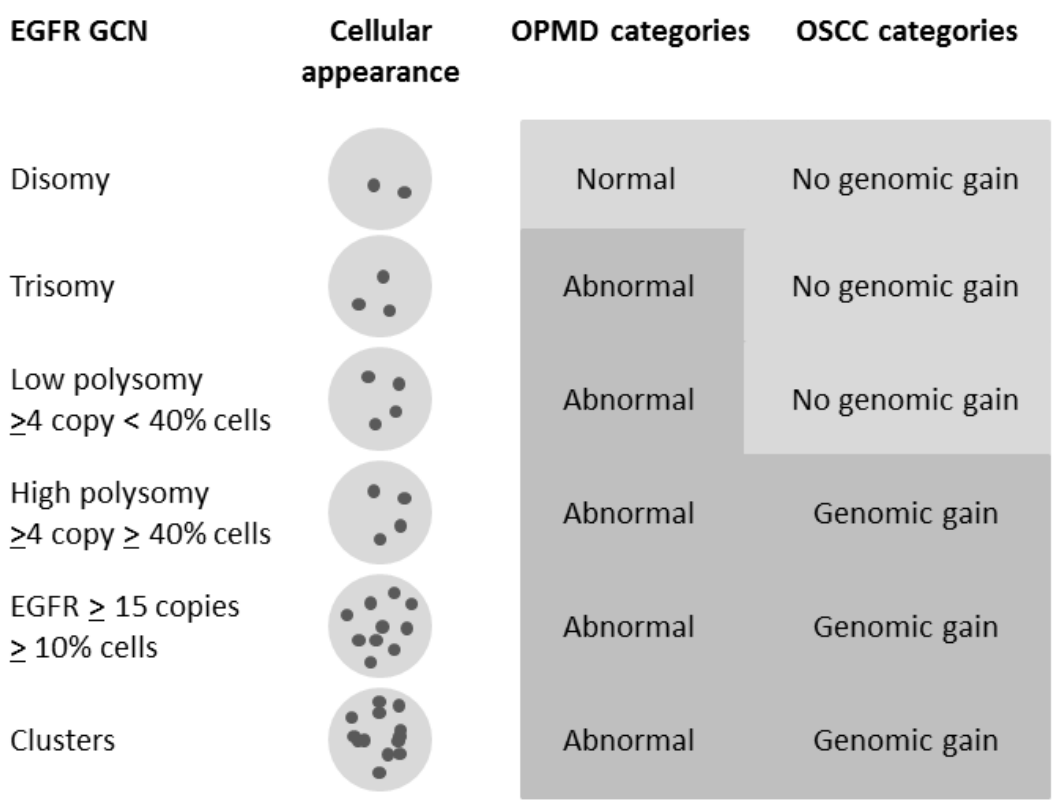


Figure 2-2 Interpretation of dual EGFR gene and chromosome 7 ISH stained sections

EGFR ISH may detect a diverse spectrum of nuclear signals. Normal epithelium is expected to show disomy. The *in situ* hybridisation signals that are interpreted as amounting to genomic gain in lung cancer (high polysomy, ≥15 copies, and clusters) are rare in OPMD. For the analysis of the OPMD group, these signals were therefore grouped alongside cases showing trisomy and low polysomy in a single 'abnormal EGFR gene copy number' category. However, for the analysis of the early-stage OSCC group, only the categories currently validated as indicating EGFR genomic gain were considered.

Adapted from 'Interpretation Guide: Ventana Inform EGFR DNA Probe', Roche.

2.6 Mouse model of oral carcinogenesis

2.6.1 Experimental outline

4-nitroquinolone 1-oxide (4-NQO) is a chemical carcinogen that induces formation of oral squamous cell carcinoma (OSCC) in mice (Kanojia and Vaidya, 2006). This feasibilty study was designed to optimise 4-NQO-induced OSCC formation in Pax9-deficient mutant mice. The purposes of the study were analogous to those of a phase I clinical trial in humans, i.e. to identify the response of Pax9-deficient mutants to 4-NQO treatment, to identify the tolerable 4-NQO dose range for Pax9-deficient mutants, and to identify any potential side effects/systemic toxicity that may be induced by 4-NQO treatment (Cancer Research UK, 2013).

Prior to the 4-NQO experiment, mutant (K14Cre-Pax9$^{flox/flox}$) and control (Pax9$^{flox/flox}$) mice had been crossed using three different genetic backgrounds: C57Bl/6J, FVB/NJ, and a hybrid strain (C57Bl/6J crossed with FVB/NJ). This generated similar numbers of Pax9-deficient mutant and control mice. Mutant mice were characterised by the tissue-specific inactivation of *Pax9* in the oral cavity and upper aerodigestive tract. This is due to the K14-Cre promoter sequence, which induces transcription of Cre-recombinase. Cre-recombinase mediates DNA recombination either side of the floxed *Pax9* sequence, and effectively excises the *Pax9* gene. The mutant phenotype shows loss of filiform papillae from the squamous epithelium of the dorsal tongue. The dorsal tongue may also show fissuring. Control mice lack the activity of the K14-Cre promoter and therefore express normal levels of Pax9. Control mice therefore exhibit a wild-type phenotype with morphologically normal tongues that include filiform papillae.

Groups of mutant and control mice were exposed to 4-NQO administrered via drinking water. Three different concentrations of 4-NQO were used: 10, 20, 50 µg/ml. Exposure to 4-NQO was continued for up to 28 weeks. Mice were sacrificed humanely at planned intervals for analysis. Any mice that unexpectedly became sick as a result of 4-NQO treatment were also sacrificed. The experiment was carried out in the Functional Genetics Unit (FGU) at the Institute of Genetic Medicine, International Centre for Life, Newcastle University.

2.6.2 Preparation of 4-NQO solutions

The concentration of the stock 4-NQO solution was 50 mg/ml. The stock solution was prepared by dissolving 1 g of 4-NQO powder in 20 ml of dimethyl sulphoxide (DMSO) (Sigma-Aldrich, UK). 1.5 ml of the stock solution was aliquoted into 2 ml Eppendorf tubes. Aliquots were stored at -20ºC in a light-proof container in the main laboratory.

The concentration of the experimental 4-NQO solution to be added to the drinking water was 5 mg/ml. The experimental solution was made up on treatment days. One aliquot of stock 4-NQO solution was added to 13.5 ml of propylene glycol (1:10 dilution). The experimental solution was transferred to the FGU and stored at 4ºC in a light-proof container prior to treatment.

2.6.3 4-NQO administration

Each cage was supplied with a 200 ml drinking water bottle. During the active phase of 4-NQO treatment, 4-NQO solution was added to freshly sterilised drinking water at the start of each week. 200 ml of autoclaved drinking water was poured into a sterile water bottle, which was transferred to a designated fume cupboard. Experimental 4-NQO solution was then added using barrier tips. The quantity of experimental solution added depended on the final concentration of 4-NQO planned for each cage (Table 2-3). Water bottles were placed in a light-proof plastic sheath prior to being attached to the cage as light exposure may trigger decomposition of 4-NQO (Tang et al., 2004).

Table 2-3 Pipetting protocol for the preparation of 4-NQO drinking water

Final 4-NQO concentration (µg/ml)	Volume of water (ml)	Volume of experimental 4-NQO solution (µl)	Dilution
50	200	2000	1/100
20	200	800	1/250
10	200	400	1/500

Any excess 4-NQO drinking water from the previous week was decanted into designated 2.5 litre Winchester bottles. These bottles were stored until full. Full bottles were sealed and then removed by waste contractors for disposal as hazardous chemical waste.

2.6.4 Health and safety measures during the handling of 4-NQO

4-NQO is carcinogenic to all mammals, including humans. Personal protective equipment was therefore used at all times when handling 4-NQO. This included a laboratory coat, nitrile gloves, a face mask, and protective glasses.

No 4-NQO spillages occurred and no first aid measures were required during this experiment. However, emergency plans were formulated in conjunction with Newcastle University's Health and Safety Department, as part of the COSHH form risk assessment for the experiment. The emergency measures are summarised in the Appendix.

2.6.5 Cage cleaning and bedding changes

Cages and bedding were replaced by trained FGU technical staff at least once per week. Both were potentially contaminated by 4-NQO. Cage changes were therefore carried out in the designated fume cupboard. Bedding was disposed of in double-layered clinical waste bags. The bags were incinerated according to the FGU's standard waste disposal protocol. Empty cages were collected in autoclave bags. Cages were washed and autoclaved by FGU staff.

2.6.6 Monitoring of mice during exposure to 4-NQO

Mice were checked by trained FGU technical staff every day. Any mice that showed signs of systemic toxicity (such as hypersalivation, a hunched posture) due to 4-NQO treatment were immediately reported to Dr R. Kist. For humane reasons, any sick mice were sacrificed prior to completion of the planned period of 4-NQO treatment.

Mice were weighed each week throughout the 4-NQO experiment until they were sacrificed. Any mouse that weighed <30% of the normal body weight for a mouse of their age was sacrificed for humane reasons. Throughout the experiment, mice were also monitored by a veterinary surgeon on a weekly basis. This was in compliance with Home Office regulations.

2.6.7 Sacrifice for macroscopical and histological analysis

Mice were sacrificed using the FGU's standard CO_2 gas chamber facility. The chamber first was inverted to remove residual CO_2. This was to avoid the distress that sudden exposure to high concentrations of CO_2 may cause the mice. The chamber was lined with a paper towel. Mice were placed in the chamber and the lid sealed. CO_2 was turned on at a low concentration (10-20% CO_2) until mice became drowsy (usually after ~1-2 mins). The CO_2 concentration was gradually increased to ~50-60%. Mice were left for a further 2-3 mins until no signs of life were evident. The CO_2 was turned off. Firm finger pressure was applied to the hind limb of each mouse. If pressure elicited a reflex, the lid was sealed and CO_2 exposure resumed. When no reflexes were elicited, mice were removed for autopsy.

2.6.8 Autopsy procedures

Mice were pinned out supine on a cork board and sprayed with 70% ethanol. Using sharp-ended scissors, a transverse incision was made in the pelvic region. Blunt dissection was then used to separate skin from the peritoneal fascia. A sagittal skin cut was made towards the sternum. Resulting skin flaps were then pinned laterally.

The abdominal cavity was accessed through the peritoneal fascia. Each of the major organs (stomach, liver, spleen, intestines, kidneys) was inspected.

The thorax was accessed by fracturing the sternum in the midline using blunt-ended scissors. The heart, lungs, and oesophagus were inspected. In cases with a possible oesophageal tumour, the oesphagus was blunt dissected off surrounding thoracic viscera and fixed separately.

A midline sagittal skin incision was made from the root of the neck towards the lower border of the mandible. Skin was blunt dissected away from the cervical fascia. The cervical fascia was dissected off the underlying salivary glands. The salivary glands were removed for separate fixation.

A subset of mice showed tumours arising on the buccal mucosa/lip. These were dissected off surrounding structures prior to opening the oral cavity.

The oral cavity was accessed via bilateral sagittal incisions from the labial commissure to the retromolar region. Large, blunt-ended scissors were then used to fracture the temporomandibular joint/ramus of the mandible. The mandible, tongue, and floor of mouth were then dissected away from the surrounding structures and pinned out separately on cork board.

The tongue was dissected away from the floor of mouth and mandible. Forceps were used gently to elevate the tip of the tongue. Sharp-ended scissors were used to make a small transverse incision through the lingual frenum. The tongue was separated by blunt dissection through this incision, working postero-laterally towards the tongue base.

2.6.9 Fixation and cut-up

All tissue separated from the main specimen (tongue, salivary glands, oesophagus, and tumours of the lip/buccal mucosa) was transferred to a 15 ml Falcon tube containing 10% buffered formalin. The tissue was incubated at room temperature on a rocker for 24 hours to facilitate thorough fixation. The main specimen was fixed in 10% buffered formalin and archived.

Following 24 hours of fixation, tongues were subdivided in the coronal plane. Small tongues were bisected into anterior and posterior halves. Larger tongues were trisected into anterior, middle, and posterior thirds. In the majority of cases, the anterior portion was further bisected in the sagittal plane.

2.6.10 Processing, embedding, and sectioning

Each piece of tongue tissue was placed in an individual cassette (except the bisected anterior portions, which were placed in a single cassette). The tissue was processed in the Department of Cellular Pathology by Biomedical Scientists, using a routine overnight processing protool.

Processed tissue was embedded on the most anterior cut surface (medial cut surface for bisected anterior portions). To screen all the tongues, a first-face section was cut at 4 µm and stained with H&E. In selected cases, ribbons were subsquently cut through all remaining tissue. Every tenth section was mounted on a glass slide and stained with H&E. All unstained sections were mounted on

coated slides (Superfrost Plus, Thermo Fisher Scientific, UK) for further assays. Unstained sections were stored at 4ºC to minimise antigen degradation.

2.7 Cell Culture Techniques

For all cell culture techniques, unless otherwise stated, reagents were of molecular biology grade and obtained from either Sigma-Aldrich, UK, or Lonza, UK. Apart from the lentivirus transfections (section 2.9), all cell culture procedures were performed in a level 2 containment tissue culture facility in the School of Dental Sciences, using a Safe Flow 1.2 laminar flow cabinet (BioAir Solutions, UK).

2.7.1 Maintenance of oral squamous cell carcinoma cell lines

Nine commercially-available OSCC-derived cell lines were cultured. The cell lines were maintained in Dulbecco's Modified Eagle's Medium (DMEM/F12 1:1 Mixture, with 15 mM Hepes and 0.6 µg/ml L-glutamine). The final DMEM solution also contained 10% (v/v) foetal bovine serum (FBS), 10 mL penicillin streptomycin, and 0.5 ug/ml hydrocortisone 21-hemisuccinate sodium salt. Cells were cultured in 75 cm^2 flasks and incubated in a dedicated cell-line incubator (InCu Safe, Sanyo Electronics, Japan). The incubator provided a humidified atmosphere of 5% CO_2/95% air and a temperature of 37ºC. Cells were inspected daily for confluence and to screen for contamination using a Leica DM IL inverted-phase contrast microscope (Leica Microsystems, UK). The culture medium was replenished every three to four days.

A control OKF6/hTERT cell line was cultured in keratinocyte serum-free media with 0.6 µg/ml L-glutamine. The final culture medium also contained 0.2 ng/ml human recombinant epidermal growth factor, and 20 µg/ml bovine pituitary extract (Life Technologies, Paisley, UK). The OKF6/hTERT cells were maintained and subcultured in conditions otherwise identical to the OSCC-derived cell lines.

2.7.2 Subculture and cell number determination

Cells were sub-cultured every five to seven days or whenever ~90% confluent. Culture medium was decanted and the cells washed with phosphate-buffered saline (PBS). The cells were enzymatically detached from the floor of the flask

by incubating with 0.025% (v/v) trypsin solution for 5 mins at 37º C. Trypsin was neutralised by the addition of an equal volume of epithelial medium. The cell suspension was centrifuged at 1000 rpm for 5 mins at 4˚C (Jouan CR3i Multifunction Centrifuge, Thermo Scientific, USA). The supernatant was aspirated and discarded. The cell pellet was re-suspended in 5 ml of epithelial medium. The cell number was determined by counting the number of cells present in 20 µl of cell suspension loaded onto a Bright-Line haemocytometer (Hausser Scientific, USA). Cells were counted across sixteen squares in each of four quadrants of the haemocytometer. The mean of all four quadrants was calculated. The mean was multiplied by 1×10^4 to give the number of cells per ml of the cell suspension. The cells were finally re-seeded at a density of 5×10^5 cells per 75 cm^2 flask.

2.7.3 Long-term cell storage

Stocks of OSCC-derived and control OKF6/hTERT cell lines were maintained in liquid nitrogen for long-term storage. The cell number was determined (as above, section 2.7.2). 5×10^5 - 1×10^6 cells were then re-suspended in DMEM with 20% (v/v) FBS and 10% (v/v) DMSO. The cell suspension was transferred to cryotube ampoules (CRYO-S™, Greiner Bio-one, UK). The cryotubes were placed in a -80ºC freezer for 12-24 hours (Ultra Low, Sanyo Electronics Ltd, Japan). Cryotubes were then transferred to liquid nitrogen at a temperature of approximately -196ºC. The level of liquid nitrogen in the storage container was checked on a weekly basis. Fresh liquid nitrogen was added when necessary.

2.7.4 Retrieval of cells for culture

To retrieve cells from long-term liquid nitrogen storage, the cryotube ampoules were thawed rapidly in a water bath (temperature of 37ºC). The cryotube ampoules were quickly sterilised by wiping with 70% alcohol solution. The thawed cell suspension was then added to a 10 ml Falcon tube containing epithelial medium that had been pre-heated to 37ºC. The suspension was then centrifuged at 1000 rpm for 5 mins at 4˚C (Jouan CR3i Multifunction Centrifuge, Thermo Scientific, USA). The supernatant was aspirated and discarded. The cell pellet was re-suspended in 5 ml epithelial medium. The cell suspension was

transferred to a 25 cm^2 tissue culture flask. Cells were maintained in culture as described in section 2.7.1.

2.7.5 Mycoplasma testing

The cell cultures were tested regularly for Mycoplasma infection using a proprietary kit (Mycoplasma Plus PCR Primer Set, Agilent Technologies, USA). Following subculture, 100 µl of supernatant was aspirated and transferred to a 1.5 ml Eppendorf tube. The supernatant was heated to 95˚C on a heat block for 5 mins and then pulsed in a micro-centrifuge for 5 secs at room temperature. 10 µl of StrataClean resin was added to the supernatant and mixed by gentle flicking. The mix was then centrifuged for 5-10 secs at room temperature to pellet the resin component. The supernatant was aspirated and transferred to a fresh 1.5 ml Eppendorf tube. A standard PCR master mix was made up according to the manufacturer's instructions. For each cell line, 5 µl of the treated supernatant was added to 45 µl of the master mix solution. The PCR cycle recommended by the manufacturer was run on a T100™ Thermal Cycler (BioRad, Berkeley, USA). The PCR product was run on a high grade 2% agarose gel with a 100 base-pair ladder.

2.7.6 Preparation of agarose cell pellets

Cells from a >90% confluent 75 cm^2 flask were enzymatically detached using trypsin as for standard subculture (section 2.7.2). Trypsin was neutralised by the addition of an equal volume of epithelial medium. The cell suspension was centrifuged at 1000 rpm for 5 mins at 4˚C. The cell pellet was then re-suspended in 10% buffered formal saline, transferred to a 15 ml Falcon tube, and incubated for 1 hour at room temperature. The suspension was then centrifuged at 10,000 rpm for 10 mins at room temperature. The supernatant was aspirated and discarded. The cell pellet was re-suspended in 200 µl PBS and transferred to a 1.5 ml Eppendorf tube.

A 2% agarose gel was made up in PBS. 1 g of agarose powder (SeaKem® LE Agarose) was added to 50 ml PBS, and microwaved at a low heat. The agarose gel and cell pellet were equilibrated in a water bath heated to 60˚C. Following equilibration, 200 µl agarose gel was added to the cell pellet. The pellet was re-suspended using a glass Pasteur pipette. The agarose cell suspension was

gently vortexed and then centrifuged at 13,000 rpm at room temperature. The pellet was allowed to cool for 1 hour at room temperature. When firm, the pellet was loosened from the sides of the tube by adding a few drops of 10% buffered formalin. The pellet was removed from the tube and longitudinally bisected using a scalpel. Both halves were placed in a histology cassette and immersed in 10% buffered formalin. The cassette was transferred to the Department of Cellular Pathology and processed using a routine overnight programme. The two halves of the pellet were embedded in wax and serially sectioned at 4 µm by Biomedical Scientists.

2.7.7 Preparation of cell monolayers

Following subculture, 250 x10^3 cells were seeded onto glass chamber slides (LAB-TEK, Thermo Fisher Scientific, USA). The chambers slides were incubated for two to three days under standard conditions until the cells were >90% confluent. The slides were transferred to ice and the culture medium was decanted. The cell monolayers were rinsed with 2 x PBS washes. 2 ml of 4% buffered formal saline was added to both chambers and the slides were incubated at 37ºC for 15 mins. The formal saline was removed and the cell monolayers washed with 2x PBS rinses. 2 ml PBS was added to each chamber and the slides were then stored at 4˚C until required for immunohistochemistry.

2.7.8 Preparation of Surepath™ slides for immunohistochemistry

Surepath™ slides were prepared in the Department of Cellular Pathology by Biomedical Scientists. Following subculture (2.7.2) the cell pellet was re-suspended in 10 ml CytoRich® Red alcohol-based fixative (TriPath Imaging, Inc, USA). The cell suspension was incubated for 30 mins. The cell suspension was then centrifuged for 5 mins at 1500 rpm at 4ºC (Hettich Centrifuge, DJB Labcare Ltd, UK). The supernatant was aspirated and discarded. 1 ml of tris-buffered water was added to the pellet. The tube was then vortexed for 15 secs to disperse and mix the cell pellet. Settling chambers were attached to Surepath™ slides. 200 µl of cell suspension was added to each slide. The slides were incubated at room temperature for 10 mins. The settling chambers were removed and the excess fluid was decanted. Slides were transferred to

and stored in a Coplin jar containing 95% alcohol until required for immunohistochemistry.

2.8 Standard transfection of OSCC cell lines using the T-REx™ system

The transfection of H314 and H400 with the PAX9 over-expression/SOX2 knockdown constructs was originally planned using the T-REx™ system (Invitrogen, Life Technologies, Paisley, UK). The transfections were outsourced to a commercial company (Dundee Cell Products Ltd, Whitehall House, 33 Yeaman Shore, Dundee, DD1 4BJ: No. SC303525). Unfortunately, the transfections were unsuccessful and standard transfection of H314 and H400 with T-REx™ plasmids was eventually abandoned (see Chapter 6, section 6.5). Therefore, only a brief overview of the original experimental design and relevant methods are summarised here.

2.8.1 Experimental outline

The T-REx™ system is a tetracycline-regulated mammalian expression system in which the chosen cell line is serially transfected with two separate plasmids. The system uses regulatory elements from the E. coli Tn10-encoded tetracycline resistance operon (Hillen et al., 1983; Hillen and Berens, 1994). The main components include:

- pcDNA6/TR© - this is the regulatory plasmid that constitutively expresses the Tet-repressor (TetR) molecule under the control of the human cytomegalovirus (CMV) promoter (Postle et al., 1984)
- pcDNA4/TO©/lacZ – this is the inducible expression plasmid that contains the gene of interest under the control of the human CMV promoter and two ‘tetracycline-on’ (TetO) sites
    - this plasmid also contains the lacZ gene, which expresses β-galactosidase upon induction with tetracycline as a control
- Tetracycline to induce expression of the gene of interest and the lacZ control.

In the T-REx™ system, the TetR molecule that is constitutively expressed by the regulator plasmid binds to the TetO sequence in the inducible expression plasmid. This blocks transcription of the gene of interest. On addition, tetracycline binds the TetR molecule. Binding induces a conformational change that leads to dissociation from the TetO sequence. Expression of the gene of interest is then promoted by the CMV sequence (Yao et al., 1998) (Figure 2-3).

Generation of the PAX9 pcDNA4/TO©/lacZ plasmid is summarised briefly in Chapter 6. The SOX2 pcDNA4/TO©/lacZ plasmid was designed using the BLOCK-iT™ inducible shRNAi Entry Vector kit (Invitrogen, Life Technologies, Paisley, UK). Addition of tetracycline would lead to expression of SOX2 shRNAi, which would knockdown endogenous SOX2 expression. However, generation of the SOX2 pcDNA4/TO©/lacZ plasmid was abandoned following failure of the initial transfections.

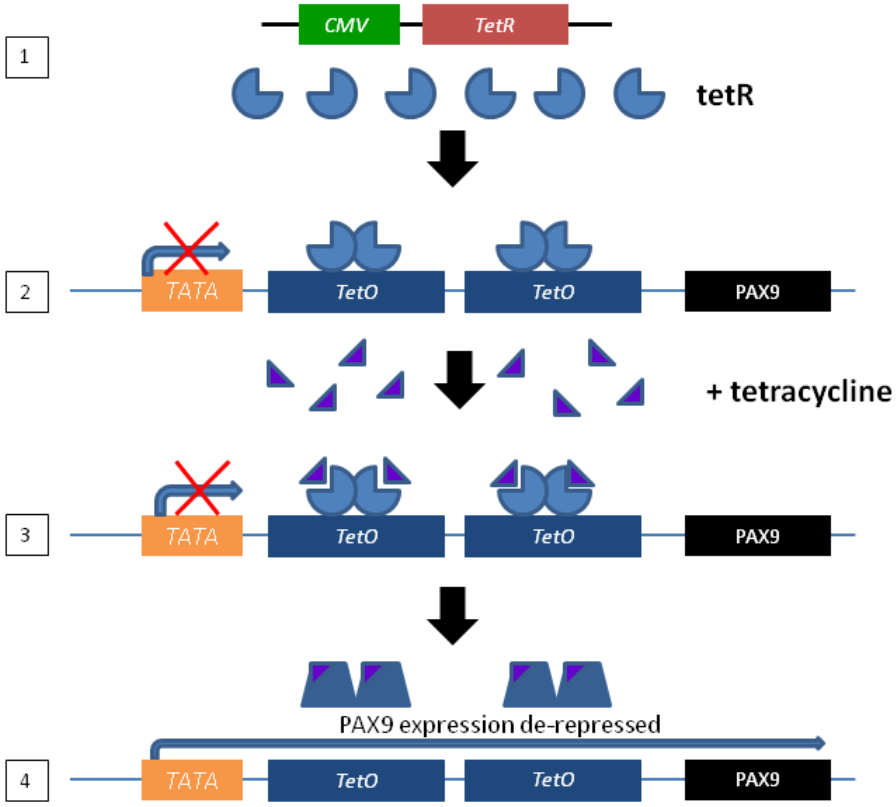


Figure 2-3 The T-Rex™ inducible expression system

1) The regulator plasmid, pcDNA6/TR©, generates high levels of tetracycline-repressor (TetR) protein under the control of a cytomegalovirus (CMV) promoter sequence. 2) The TetR molecule forms a homodimer that binds to the 'Tetracycline-on' (TetO) sequence in the inducible expression plasmid, pcDNA™4/TO/lacz. 3) Tetracycline binds with high affinity to the TetR molecules and induces a conformational change. There is cleavage of the homodimer and dissociation from the TetO sequence in the pcDNA™4/TO/lacz plasmid. 4) The gene of interest (*PAX9* is shown in this example) is de-repressed and transcription is promoted by the TATA box.

2.8.2 Cell lines and attempted transfections

In order to minimise the risk of using contaminated cell lines for the transfections, fresh stocks of H314 and H400 cells were sent directly to Dundee Cell Products (DCP) from the European Collection of Cell Cultures, Public Health England. DCP received the cell lines at passage numbers 15 (H314 - #06092003) and 18 (H400 - #06092006). The cell lines were thawed down, maintained, and subcultured by DCP Biomedical scientists according to standard protocols specified by the ECCC. These are the same protocols as those used during the selection of suitable cell lines at Newcastle University (see section 2.7). DCP initially used Lipofectamine® 2000 Transfection Reagent (Life Technologies, Paisley, UK). Calcium phosphate was also used in later attempts at the transfections. Approximately 150-200 µg of plasmid pcDNA6/TR© DNA was used to transfect each cell line. Further information regarding the experimental difficulties encountered by DCP are described in Chapter 6.

2.9 Lentivirus transfection of cell lines

The lentivirus transfections were performed in a level 3 containment tissue culture facility in the Dermatology Research Laboratory, School of Medicine, Newcastle University. The experimental plan for the lentivirus transfections is summarised in the Appendix and further discussed in Chapter 6. The following sections summarise the standard protocol used for each of the lentivirus transfections.

2.9.1 Transfection of lentivirus-producing 293T cells with target DNA sequence

- Day 1: 293T cells were seeded into 100 mm tissue culture dishes. Cells were seeded at a density of 2-3 x 10^5 cells/ml in 10 ml of complete growth medium
- Day 2: The lentivirus transfection solution was prepared (components detailed in Table 2-4). The plasmid solution was added to the calcium chloride ($CaCl_2$) solution. The combined plasmid/$CaCl_2$ solution was then added drop wise (one drop per second) to the 2x HeBS solution. Air was continuously passed through the HeBS solution while adding each drop in

order to prevent precipitation. The transfectant solution was incubated for 30 mins at room temperature before being added drop-wise to the 293T cells.

- Day 3: The 293T culture medium was aspirated and discarded in a sealed container as hazardous chemical waste. 293T cells were gently washed with 10 ml PBS to eliminate residual precipitate. 15 ml of fresh complete growth medium was added gently to avoid disrupting the cells. The plate was incubated for three days under standard conditions
- Day 6: When the 293T cells were >90% confluent, the lentivirus particles were harvested from the supernatant. The supernatant was gently aspirated from the culture dish, transferred to a sterile 50 ml Falcon tube, and centrifuged (2000 rpm, 4˚C, for 20 mins) to pellet cell debris. The supernatant was then filtered into fresh 15 ml Falcon tubes through an Acrodisc® syringe 0.45 µm filter (Pall Corporation, USA) to remove any residual debris. Recombinant lentiviral solution was stored at -80˚C until required.

Table 2-4 Pipetting protocol for day 2 lentiviral transfectant solution

Component	Quantity
DNA mixture	
pMD2.G envelope plasmid	5 µg
pCMVdeltaR8.91 packaging plasmid	15 µg
Lentivirus transfer vector	20 µg
Special water	(make up to 250 µl)
Calcium chloride solution	250 µl
2x HeBS	500 µl

2.9.2 Transduction of OSCC cell line with recombinant lentivirus solution

- Day 1: The OSCC cell line selected for transfection was sub-cultured under standard conditions (section 2.7.2) and the pellet re-suspended. 2 x 10^4 cells were seeded into each well of a six-well plate. Two six-well plates were seeded for each cell line
- Day 4: The OSCC cell line was transduced when cells were 50-70% confluent. The transduction solution comprised 3 ml of recombinant lentiviral solution; 9 ml of DMEM without penicillin/streptomycin; and 6 µl polybrene. The culture media was decanted from the six-well plates and cells were rinsed in PBS (x2). 2 ml of lentiviral/DMEM mix was added to each well. The plates were centrifuged at 1300 rpm for 1.5 hours. The supernatant was decanted and cells washed with 2x PBS. 2 ml culture media was added to each of the wells. The transduced cells were incubated under standard conditions until >90% confluent
- Day 7: The transduced OSCC cells were subcultured. 0.5 ml of trypsin was used per well. The cell suspension was seeded into a 25cm^2 flask and incubated under standard conditions. After 24 hours, the culture media was replaced with fresh media that contained antibiotic to select transduced clones (either 0.5 µg/ml puromycin or 5 µg/ml blasticidin).

2.9.3 Doxycycline induction and fluorescence imaging

Following one week of culture in the selection antibiotic (either puromycin or blasticidin) the transduced cells were expanded until there were sufficient cells for maintenance, characterisation, and long-term storage (2.7.3).

Cells for initial characterisation were seeded into six-well plates at a density of 1 x 10^5 cells per well. At 24 hours, the culture media was replaced and the cells treated with doxycycline at six different concentrations: 0, 25, 50, 100, and 500 ng/ml, and 1 µg/ml. When the majority of wells were confluent, the cells were inspected using an inverse fluorescence microscope (Leica DMI4000B, Leica Microsystems, UK) and TRITC/CY3.5 (tetramethylrhodamine Isothiocyanate/Cyanine) filter (excitation wavelength -532 nm; excitation wavelength - 570 nm). Fluorescent images were taken for cases with positive signals and to also confirm the absence of a signal in negative cases. Light

microscopic images were taken to confirm the relative confluence of the induced cells. All images were taken at x100 magnification.

2.10 Molecular biology and biochemical techniques

2.10.1 RNA extraction from cultured cells

Following subculture, cells selected for RNA isolation were seeded into a six-well plate at a density of 2 x 10^4 cells per well. The cells were incubated under standard conditions. When the cells were >90% confluent, the plates were transferred to ice. The culture medium was decanted and cells washed with PBS (2x). 1 ml of TRIzol® reagent (Life Technologies, Paisley, UK) was added to each well and pipetted vigorously to lyse the cells thoroughly. The lysate was transferred to a 1.5 ml Eppendorf tube and stored at -80°C until required.

2.10.2 RNA purification from the TRIzol® lysate

A 2 ml phase-lock gel tube (5 Prime, UK) was centrifuged at 11,000 g for 30 secs. The lysate was added to the tube and incubated at room temperature (5 mins). 200 µl chloroform was added and the tube was shaken vigorously (15 secs). Following incubation at room temperature (3 mins) the tubes were centrifuged at 11,000 g at 4°C for 15 mins (Centrifuge 5417B, Eppendorf, Germany). The aqueous phase was transferred to a non-stick RNAse-free tube (Ambion® Life Technologies, UK). 500 µl isopropanol was added and mixed by inversion to precipitate the RNA. Following incubation at room temperature (10 mins) the samples were centrifuged at 11,000 g at 4°C for 20 mins.

The supernatant was decanted and the RNA pellet washed by adding 1 ml of 70% ethanol. Samples were vortexed and centrifuged at 7,000 g at 4°C for 5 mins. The supernatant was decanted. The RNA pellet was air-dried for ~5 mins and dissolved in 30 µl RNAse-free water. The solution was incubated at 55°C for 5 mins with an open lid to facilitate RNA dissolution and ethanol evaporation. 1 µl RNAse inhibitor was added to each sample. The samples were transferred to ice. The RNA concentration was determined (NanoDrop 2000 UV-Vis Spectophotometer, Thermo Scientific, UK). Samples were stored at -80ºC until required.

2.10.3 Reverse transcription of RNA to generate cDNA samples

The RNA concentration of each sample was adjusted to 400 ng/µl. RNA integrity was checked by running 1 µl of RNA solution on a 1% TBE gel. 5 µl of RNA solution was subsequently used for reverse transcription (see first reaction, below). The remaining RNA solution was stored at -80ºC.
Reverse transcription was carried out using a commercially available kit (Precision NanoScript™ Reverse Transcription Kit, Primer Design, Southampton, UK). Barrier tips were used throughout. The protocol involved two separate reactions:

1. In the first reaction, 5 µl RNA (400 ng/µl) was added to 3.5 µl nuclease-free water and 5.5 µl oligodT primer (180 ng/µl), using 0.5 ml RNA free Eppendorf tubes. The solution was heated to 70ºC for 5 mins and placed on ice
2. In the second reaction, the product from the first reaction was added to 11 µl of mix two (see Table 2-5 for pipetting protocol) and mixed by gentle flicking. The mix was heated to 37ºC for 2 hours, then 70ºC for 15 mins to stop the reaction, then cooled and placed on ice.

5 µl of the PCR product was added to 20 µl nuclease-free water to make a 1:5 diluted experimental solution. The remaining PCR product was retained as a stock solution. Both were stored at -20ºC until required.

Table 2-5 Pipetting protocol for RNA reverse transcription mix two

Reagent	Quantity (µl)
Nuclease-free water	3
10mM dNTP	1.5
RNAase inhibitor	0.5
5x M-MLV reaction buffer	5
M-MLV RT enzyme	1
Total	**11**

2.10.4 Semi quantitative reverse transcription polymerase chain reaction

The primers for PAX9, SOX2, and ACTB reverse transcription polymerase chain reaction (RT-PCR) were designed using software from and supplied by eurofins mwg operon (Eurofins Genomics, Ebersberg, Germany). For each sample, 1 µl experimental cDNA solution (2.10.3) was added to a standard master mix (see Table 2-6).

Table 2-6 Pipetting protocol for PAX9, SOX2, and ACTB RT-PCR

Reagent	Quantity (µl)
cDNA (1:5 dilution)	1
5 x Green GoTaq buffer (15mM $MgCl_2$)	5
10mM dNTP	0.5
Forward primer*	0.5
Reverse primer**	0.5
GoTaq-Polymerase (HotStart)	0.15
Nuclease free water	17.35
Total	**25**

*Forward primers: hPAX9-RT1-F; hSOX2-Endo-F2; hACTB-RT1-F

**Reverser primers: hPAX9-RT1-R; hSOX2-Endo-R2; hACTB-RT1-R

The PCR products were run on a 2% agarose gel (SeaKem® LE Agarose, Lonza, UK). The gel was imaged using ultraviolet light (GelDoc-It®^{2} Imager, UVP, Germany). The gel was photographed and saved as a tagged image format file (TIFF). The TIFF image was opened and the bands were analysed using the densitometry tool. Rectangles of equal size were placed over each band (numbers 1 - 30) and the resulting measurements were saved and exported into an Excel spreadsheet. The total density was divided by the total background and the result was calculated for each gene. Normalisation was achieved by dividing the score for PAX9 or SOX2 by the score for ACTB. Values were then ranked.

2.10.5 Quantitative real-time polymerase chain reaction

Quantitative real-time reverse transcription PCR (qPCR) was performed to analyse DNA amplification during the exponential phase and thus detect quantitative differences between the expression of PAX9/SOX2 in each sample.

Commercially-available TaqMan® probes for PAX9, SOX2, and ACTB were added to a TaqMan® Universal Master Mix (both from Applied Biosystems, Life Technologies, UK) were used to perform qPCR on cDNA samples from the H314 and H400 pTRIPZ™ cell lines. The samples, probes, and master mix were pipetted according to the manufacturers' instructions. Triplicate samples were then loaded onto 96-well plates. The plates were run on an Applied Biosystems® 7500 Real-Time PCR machine (Life Technologies, UK).

The fluorescent probes containing each gene of interest (PAX9, SOX2) included a reporter dye (FAM-6) at the 5' end, and a non-fluorescent quencher dye at the 3' end. The quencher dye normally suppresses fluorescence of the reporter dye. However, during PCR amplification the probes are cleaved by the 5'-3' nuclease activity of DNA polymerase. Reporter and quencher dyes are separated, increasing the fluorescence signal. PCR products accumulate with each replication cycle. A threshold fluorescence value is determined for each cycle - this is the cycle threshold (Ct) value. The Ct value detects the point at which fluorescence from accumulated PCR product exceeds the predetermined background value. A sample with a low Ct value therefore has a high level of the gene of interest as the threshold is achieved early.

A reference gene, ACTB, was amplified at the same time as PAX9/SOX2. ACTB is theoretically expressed at a constant level across all the samples. Relative differences in the PAX9/SOX2 levels of each sample were then calculated using the comparative Ct method described by Livak and Schmittgen (2001).

1. $\Delta Ct = Ct_{(\text{gene of interest})} - Ct_{(\text{housekeeping gene})}$
2. $\Delta\Delta Ct = \Delta Ct_{(\text{expression})} - \Delta Ct_{(\text{control})}$
3. $2^{-\Delta\Delta Ct}$

These calculations were performed using an Excel spreadsheet. The mean fold values were then used to rank samples according to the level of PAX9/SOX2.

2.10.6 Protein isolation from cultured cells

Following subculture, the cell line selected for protein isolation was seeded into a six-well plate (2 x 10^4 cells per well). The cells were incubated under standard conditions until >90% confluent.

The six-well plates were placed on ice. The culture medium was aspirated and discarded. Cells were washed with PBS (2x). 120 µl of RIPA buffer (Sigma-Aldrich, UK) containing protease inhibitor and sodium deoxycholate were added to each well. The cells were scraped off the bottom of the plate using a cell scraper (Corning®, Sigma-Aldrich, UK). The cell lysate was aspirated at transferred to a 1.5 ml Eppendorf tube. The tube was vortexed for 30 secs. The cell lysate was stored at -80°C until required.

2.10.7 Protein quantification

Protein samples were thawed on ice and sonicated briefly (3x 3 secs). The samples were centrifuged at 13,000 rpm at 4ºC for 10 mins and returned to ice.

Protein was quantified using the Pierce BCA Protein Assay Kit (Thermo Scientific, UK). A 96-well plate was set up with 2 x 25 µl of each template sample in rows one and two. 5 µl of protein lysate was added to 70 µl of distilled water in row three and mixed thoroughly. 2 x 25 µl aliquots were pipetted into rows four and five as triplicate repeats. 200 µl of Pierce solution was added to each well. The plate was incubated at 37ºC for 30 mins. The plate was read using a SoftMax Pro® Microplate reader (Miratech, UK) at a wavelength of 562 nm. Data was exported to an Excel spreadsheet to the mean protein quantity for each sample.

Protein samples were aliquoted and diluted to produce a uniform quantity of protein. The precise quantity varied according to the range of concentrations determined for each cell line. Each aliquot was adjusted by adding RIPA buffer (Sigma-Aldrich, UK) to equilibrate the protein suspensions. 4x LDS sample buffer and 10x sample reducing agent (NuPAGE®, Life Technologies) were then added to each aliquot. Quantities were varied for each sample so that every 10 µl contained: 1 µl of 10 x reducing agent; 2.5 µl LDS; and 6.5 µl protein sample.

2.10.8 Western blotting

Protein samples were thawed and heated to 70˚C on a heat block (10 mins). Tubes were pulsed briefly to remove condensation and briefly vortexed. Pre-cast polyacrylamide gels (NuPAGE® Novex® 4-12% Bis-Tris Gels, Life Technologies) were placed in a tank filled with running buffer (NuPAGE® MES SDS, Life Technologies). 500 µl of antioxidant was added to the central chamber of the tank. Protein samples were loaded on the gel with a 10-170 kDa protein ladder (PageRuler Prestained Protein Ladder, Pierce Protein Products, Thermo Scientific, UK). The gel was run at 200 V for 30 mins.

The gel was removed and placed on PDVF membrane (Sigma-Aldrich, UK) in a transfer cassette lined by sponge and filter paper. The cassette was placed in a tank containing transfer buffer (NuPAGE® Transfer Buffer, Life Technologies) with 200 ml methanol and 2 ml of antioxidant (NuPAGE® Antioxidant, Life Technologies). The tank was placed in an ice bucket. The transfer was run at 100V for 1 hour.

The primary antibody solution was made up in a 50 ml Falcon tube. 1 g of bovine serum albumin powder and 40 µl of sodium azide were added to 20 ml of TBS Tween. The solution was mixed thoroughly. The primary antibody was added at a concentration of 1/1000. The PDVF membrane was transferred to the antibody solution and incubated overnight at 4˚C on a mixer.

The secondary antibody solution was made up in a 50 ml Falcon tube. 1 g of milk powder was added to 20 ml TBS Tween and mixed thoroughly. The secondary antibody was added at a concentration 1/1000.

After washing with TBS Tween, the membrane was transferred to the secondary antibody solution and incubated at room temperature on a mixer (1 hour). The membrane was washed and placed on cling-film. 2 ml of low sensitivity enhanced chemiluminescent substrate (Themo Scientific, UK) was added to the membrane drop-wise and incubated for 5-10 mins.

The membrane was placed between acetate sheets in a transfer cassette. The cassette was taken into the dark room. An X-ray film (Kodak, USA) was removed and placed on top of the acetate sheet. The cassette was closed an

incubated for 1 min. The film was removed and placed in developer for ~30 secs. A light viewer was used to assess the presence of bands. If bands were feint then the film was returned to the developer solution. The film was then rinsed in water and placed in fixing solution. The film was rinsed then allowed to air dry.

2.11 Statistical Analysis

Statistical analysis was performed using SPSS for Windows (version 21, SPSS Inc., Chicago, Illinois, USA). The selection and interpretation of specific tests was informed by discussion with Dr Simon Kometa, a statistician at Newcastle University.

Continuous data was first explored using a descriptive analysis. This analysis generated mean average values and standard deviations for each subgroup. The analysis included a Shapiro-Wilk and Kolmogorov–Smirnov tests of normality. The Shapiro-Wilk test result was used due to the relatively small numbers of cases in each subgroup (<50 cases).

A Shapiro-Wilk test result of $p>0.05$ indicated that the data had a normal distribution. Normally distributed data were analysed using parametric tests. For comparison of multiple subgroups analysis was performed using a one-way analysis of variance (ANOVA) test with the Bonferroni correction. Binary comparisons were performed using the independent sample T-test. Results were considered significant at $p<0.05$.

A Shapiro-Wilk test result of $p<0.05$ indicated that the data significantly deviated from a normal distribution. These data were analysed using non-parametric tests. For comparison of multiple subgroups analysis was performed using the Kruskal–Wallis test with the Bonferroni correction. Binary comparisons were performed using the Mann-Whitney U-test. Results considered as significant at $p<0.05$.

Kaplan-Meier time-to-event analysis was performed using Log Rank (Mantel-Cox) calculations. For OPMD, adverse outcome or malignant transformation were designated as the defined event (=1). For early-stage OSCC, adverse outcome or death from disease was designated as the defined event (=1).

Cases were grouped according to a range of criteria (e.g. grade of epithelial dysplasia; T-category).

Receiver-operator curves (ROC) were generated by plotting the cumulative frequency distribution of a given characteristic (e.g. high grade epithelial dysplasia) for detecting a positive result (e.g. malignant transformation). The curve plotted the true positive rate against the false positive rate (1-specificity). Cases were classified using a binary score (0, 1) according to the characteristic under analysis (e.g. low grade epithelial dysplasia = 0; high-grade epithelial dysplasia = 1). The ROC indicated the sensitivity of the characteristic in detecting the chosen positive result as a function of fall out. Fall out is equal to 1 – specificity and indicates the probability of a case that is negative for the given characteristic having a score of 1.

Ordinal data generated by the scoring system was analysed by cross tabulation. The Kappa test was used to analyse agreement between two independent scores (e.g. grade of epithelial dysplasia and EGFR ISH score). The Pearson's Chi-squared test was used to analyse the significance of the distribution of cases according to row and column classifiers (Field, 2013).

2.12 Photomicrography

All photomicrographs were taken with an Axiocem HRc camera on an Axioplan 2 microscope using AxioVision software v.4.3 (Carl Zeiss, Germany) and processed usng Adobe Photoshop v7.0 (Adobe Systems, Inc.).

2.13 Ethical Approval

The study had a favourable ethical opinion from the National Research Ethics Service (NRES) Committee North East, Sunderland (REC reference: 11/NE/0118).

Chapter 3. EGFR Protein Expression and Gene Copy Number in Potentially Malignant Disorders and Early-Stage Squamous Cell Carcinoma of the Oral Cavity

3.1 Introduction

3.1.1 Oral squamous cell carcinoma

Oral squamous cell carcinoma (OSCC) is a major global healthcare problem (Jemal et al., 2011). The incidence of OSCC is increasing in the developed world, particularly among young adults (Schantz and Yu, 2002; Garavello et al., 2010). OSCC is associated with poor outcomes. Up to 50% of patients diagnosed with OSCC will die of the disease within five years (Barnes et al., 2005; Warnakulasuriya, 2009). Outcomes for patients with OSCC may be improved if the disease is identified in its earliest stages (Goodson and Thomson, 2010). Patients diagnosed with OSCC are staged using the 'Tumour, Node, Metastasis' (TNM) classification that has been developed by the International Union Against Cancer, Switzerland (Sobin et al., 2009). This study focused on patients with pStage I and II ('early-stage') OSCC. An early-stage OSCC may measure up to 4 cm in maximum dimension; however, it has not metastasised either to regional lymph nodes or distant sites (Sobin et al., 2009).

3.1.2 Oral potentially malignant disorders

OSCC may be preceded by clinically-recognisable lesions, termed oral potentially malignant disorders (OPMD) (van der Waal, 2009). OPMD encompass a broad spectrum of entities, including proliferative verrucous leukoplakia, chronic hyperplastic candidosis, and oral submucous fibrosis (Barnes et al., 2005). This study focused on lesions that presented clinically as either white, red, or mixed white/red patches, and subsequently showed epithelial dysplasia on histological examination. There is continuing debate as to how the spectrum of changes that characterise epithelial dysplasia should be classified. Therefore, this study applied both the 'Squamous Intra-epithelial Neoplasia' (SIN) classification (Barnes et al., 2005) and the binary 'high-grade' versus 'low-grade' classification described by Kujan et al (2006).

3.1.3 Epidermal growth factor receptor: a prognostic lung cancer biomarker

Epidermal growth factor receptor (EGFR) is a cell surface tyrosine kinase receptor that is expressed in most epithelial tissues (Citri and Yarden, 2006). EGFR protein expression and gene copy number are currently used in the prognostication of non-small cell lung carcinoma (Nicholson et al., 2001; Hirsch et al., 2003) and the prediction of its response to EGFR-targeted chemotherapeutic agents (Takano et al., 2005).

3.1.4 EGFR is a candidate biomarker in oral carcinogenesis

EGFR is considered a candidate biomarker in oral carcinogenesis. Over-expression of EGFR protein has been reported in up to 90% of OSCC and is associated with poor clinical outcomes (Grandis and Tweardy, 1993a; Grandis, 1998; Kumar et al., 2008). In OPMD, over-expression of EGFR protein is associated with an increased risk of malignant transformation (Ries et al., 2013).

EGFR genomic gain is associated with poor clinical outcomes in OSCC (Freier et al., 2003; Rössle et al., 2013; Ryott et al., 2009). There is evidence that OPMD with abnormal EGFR gene copy number are at an increased risk of malignant transformation compared to OPMD with normal EGFR gene copy number (Benchekroun et al., 2010; Poh et al., 2012).

3.2 Aims

The aims of this chapter are:

1. To summarise the patient characteristics, clinical outcomes, and histopathological features of a group of OPMD and a group of early-stage OSCC cases. These data also underpin the analysis of two novel biomarkers, SOX2 and PAX9, which is described in Chapter 4
2. To describe the profiles of EGFR protein expression and EGFR gene copy number in the two groups
3. To correlate the EGFR protein expression and gene copy number profiles with the patient characteristics and clinico-pathological features of each group.

3.3 Results

3.3.1 OPMD group: patient characteristics, mucosal subsite, and histological grade of epithelial dysplasia

A total of seventy-eight OPMD satisfied the study's inclusion criteria and had sufficient remaining material for further assays (*n* = 78). The cases were identified using two distinct strategies: 1) attendance at an oral epithelial dysplasia clinic; 2) a search of the electronic database at the Department of Cellular Pathology. The majority (72%) of cases were identified using the first strategy. Both strategies are detailed further in Chapter 2 (section 2.1).

Overall, the group of OPMD had a male predominance (62.8% males; male:female ratio = 1.7:1). The mean age was 58.6 years (range: 30 - 94). The majority of patients either smoked tobacco during the study period (i.e. from 1997 to 2009) or were former tobacco smokers. Similarly, the majority of patients either consumed alcohol during the study period or had a history of alcohol consumption (Table 3-1). The majority of OPMD presented on the floor of mouth or ventro-lateral surface of the tongue. Smaller proportions involved subsites such as the buccal mucosa or palate (Table 3-1).

The epithelial dysplasia present in each OPMD was assigned to a single histological grade according to the worst area. Using the SIN classification, there was moderate agreement between the two independent scoring pathologists (Kappa value - 0.678; 95% confidence interval 0.544 - 0.812). SIN 1 was the single commonest histological grade. The remaining cases were evenly distributed between SIN 2 and SIN 3 (Table 3-1).

There was also moderate agreement between the two scoring pathologists when epithelial dysplasia was graded according to the binary classification (Kappa value - 0.697; 95% confidence interval 0.541 - 0.853). By contrast to the SIN classification, however, high-grade epithelial dysplasia was the commonest histological grade (Table 3-1).

Table 3-1 Characteristics of the group of cases with oral potentially malignant disorders (*n* = 78)

Characteristic	Number (%)
Alcohol consumption:	
During study period	56 (71.8)
Prior to but not during study period	3 (3.8)
None	7 (9.0)
Not known	12 (15.4)
Tobacco consumption:	
During study period	45 (57.8)
Prior to but not during study period	13 (16.6)
None	13 (16.6)
Not known	7 (9.0)
Mucosal subsite:	
Tongue	36 (46.1)
Floor of mouth	29 (37.2)
Buccal mucosa	5 (6.4)
Masticatory mucosa	8 (10.3)
Histological grade of dysplasia:	
SIN classification:	
SIN 1	31 (39.7)
SIN 2	23 (29.5)
SIN 3	24 (30.8)
Binary classification:	
High-grade epithelial dysplasia	44 (56.4)
Low-grade epithelial dysplasia	34 (43.6)
Clinical outcome (end of study period):	
No adverse outcome	30 (38.5)
Adverse outcome:	48 (61.5)
Local recurrence	15 (19.2)
New lesion	11 (14.1)
Malignant transformation to OSCC	22 (28.2)

3.3.2 Clinical management of and outcomes for patients with OPMD

Each of the 56 cases identified using strategy 1) were managed by laser excision. The majority (54%) of these cases had no adverse outcome. The remainder had adverse outcomes, proceeding to develop either a local recurrence (26%) or a new dysplastic lesion (20%). Consistent with its design and purpose, each of the 22 cases identified by strategy 2) underwent malignant transformation. The majority (73%) of these cases had been managed by surveillance.

When OPMD identified using both strategies were combined, the overall majority of cases had an adverse outcome (either local recurrence, new lesion formation, or malignant transformation). Malignant transformation was the single most common adverse outcome (Table 3-1). Kaplan-Meier time-to-event analysis did not detect a significant correlation between clinical outcome and either patient demographics (age, sex) or risk factors (tobacco/alcohol consumption). Similarly, there was no correlation between clinical outcome and the mucosal subsite of the OPMD [data not shown].

3.3.3 Grade of epithelial dysplasia and clinical outcome of OPMD

Kaplan-Meier time-to-event analysis did not identify a significant correlation between the risk of malignant transformation and grade of epithelial dysplasia when stratified according to the SIN classification (Figure 3-1A). By contrast, there was a significant correlation when epithelial dysplasia was graded using the binary classification. OPMD with high-grade epithelial dysplasia were significantly more likely to undergo malignant transformation than cases with low-grade epithelial dysplasia (Figure 3-1B). Used in isolation, however, high-grade epithelial dysplasia showed low sensitivity (72.7%; positive predictive value – 36.4%) and low specificity (50%; negative predictive value - 82.4%) for detecting the subset of OPMD that were destined to undergo malignant transformation. Table 3-4 summarises the sensitivity/specificity values of high-grade epithelial dysplasia for detecting cases that underwent malignant transformation, alongside the corresponding values for EGFR protein expression and EGFR gene copy number.

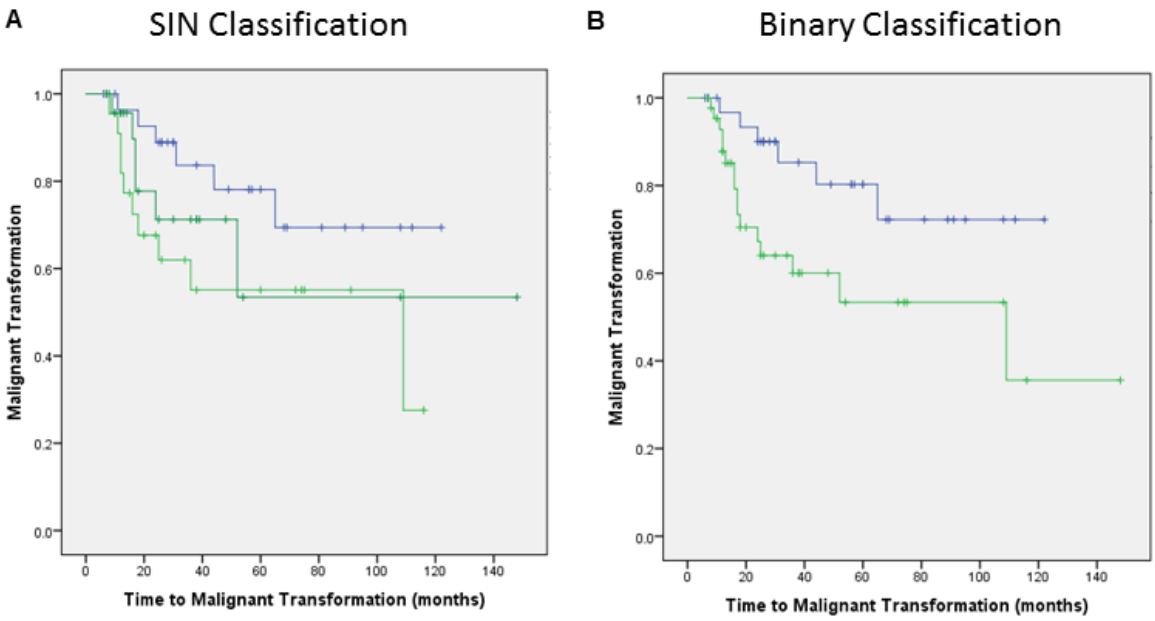


Figure 3-1 Kaplan Meier time-to-event analysis showing malignant transformation in OPMD stratified according to grade of epithelial dysplasia

Colour index: Blue line – SIN 1/low-grade epithelial dysplasia; green line – high grade epithelial dysplasia (SIN 2 – pale green; SIN 3 – dark green).

A) There was no correlation between the risk of malignant transformation and epithelial dysplasia when cases were graded using the SIN classification ($p>0.05$, x^2 value - 3.79, 2 d.f.). B) By contrast, there was a significant correlation between the risk of malignant transformation and epithelial dysplasia when cases were graded according to the binary classification ($p<0.05$, x^2 value - 4.97, 1 d.f.).

3.3.4 Clinico-pathological features of OSCC arising from OPMD

Twenty-two of the 78 OPMD in the study underwent malignant transformation to OSCC (Table 3-1). The mean transformation time (i.e. the interval from the index OPMD biopsy to the first OSCC diagnosis) was 27.2 months (s.d. - 5.4 months). The OSCC was managed by surgical excision in 20 cases and by chemoradiotherapy in two cases.

Three of the subsequent OSCC were among the group of early-stage OSCC described in the following section (see 3.3.5). The majority of the remaining cases were also classified as early-stage OSCC. However, there were three cases which involved the alveolar mucosa and showed evidence of bone invasion. These cases were therefore classified as category T4, pStage IV. The clinico-pathological features of the cases are summarised in Table 3-2. The cases had a mean disease-free survival of 43.7 months (s.d. – 8.1 months) and mean overall survival of 50.2 months (s.d. - 8.8 months).

Table 3-2 Clinico-pathological features of OSCC arising from OPMD (*n* = 22)

Characteristic	Number (%)
Mucosal subsite:	
Tongue	11 (50)
Floor of mouth	7 (32)
Buccal mucosa	1 (4.5)
Masticatory mucosa	3 (13.5)
Histological grade of differentiation:	
Well differentiated	6 (27)
Moderately differentiated	13 (59)
Poorly differentiated	3 (14)
pStage:	
pStage I	14 (64)
pStage II	3 (13.5)
pStage IV	3 (13.5)
pStage unavailable	2 (9)
Overall survival:	
Alive:	
Free from disease	9 (41)
Deceased:	
Free from disease	8 (36.5)
With disease	3 (13.5)
Lost to follow-up	2 (9)

3.3.5 Early-stage OSCC group: patient characteristics, mucosal subsite, pStage, and histological differentiation

A total of ninety-two early-stage OSCC satisfied the study's inclusion criteria and had sufficient remaining material for further assays (*n* = 92). The group had a male predominance (58.7% males; male:female ratio = 1.4:1). The mean age was 61.8 years (range: 33 - 93). The majority of cases presented on the ventro-lateral tongue and floor of mouth. Smaller proportions involved subsites such as the buccal mucosa, gingiva/alveolar ridge, or palate (Table 3-3). Moderately differentiated OSCC was the single commonest group when cases were grouped according to tumour morphology. The majority of cases were at pStage I (i.e. the tumour had a maximum dimension of $\leq$2 cm).

3.3.6 Clinical outcomes of patients diagnosed with early-stage OSCC

The overall survival and disease status of patients with early-stage OSCC are summarised in Table 3-3. The mean overall survival time was 55.9 months (s.d. - 24.4 months). The mean disease-free survival time was 46.8 months (s.d. - 25.2 months).

Sixty-seven patients were alive at the end of the study period:

- Sixty-six of these patients were alive and free from disease. However, several patients in this group had been treated for adverse outcomes during the study period. These included:
    - Local recurrence (seven cases)
    - Second primary tumour formation (two cases)
    - Regional cervical lymph node metastasis (one case)
- One patient was alive with disease. This patient had developed a local recurrence. The index tumour was a well differentiated squamous cell carcinoma arising on the buccal mucosa, pStage I.

The remaining 25 cases died during the study period.

- Fifteen of these patients died free from disease. However, three of these patients had been successfully treated for adverse outcomes prior to death. These included:
    - Local recurrence (two cases)
    - Regional cervical lymph node metastasis (one case)
- Ten patients died from disease:
    - One patient succumbed to the index tumour, a moderately differentiated squamous cell carcinoma arising on the ventro-lateral tongue, pStage II
    - Each of the nine remaining cases experienced adverse outcomes subsequent to apparently successful management of the index tumour. These included:
        - Local recurrence (two cases)
        - Regional cervical lymph node metastasis (four cases)
        - Loco-regional recurrence (one case)
        - Second primary tumour formation (one case)
        - Distant metastasis (pulmonary metastasis, one case).

The clinical outcomes of early-stage OSCC did not correlate with patient characteristics (age, sex); mucosal subsite; or pStage [data not shown]. Kaplan-Meier time-to-event analysis did not detect a correlation between clinical outcome and histological differentiation (Figure 3-2).

Table 3-3 Characteristics of the group of cases with early-stage oral squamous cell carcinoma (*n* = 92)

Characteristic	Number (%)
Mucosal subsite:	
Tongue	49 (53.3)
Floor of mouth	20 (21.7)
Buccal mucosa	9 (9.8)
Masticatory mucosa	14 (15.2)
Histological grade of differentiation:	
Well differentiated	20 (21.6)
Moderately differentiated	63 (68.5)
Poorly differentiated	9 (9.9)
pStage:	
pStage I	75 (81.5)
pStage II	17 (18.5)
Overall survival:	
Alive:	67 (72.8)
Free from disease	66 (71.7)
With disease	1 (1.1)
Deceased:	25 (27.2)
Free from disease	15 (16.3)
With disease	10 (10.9)

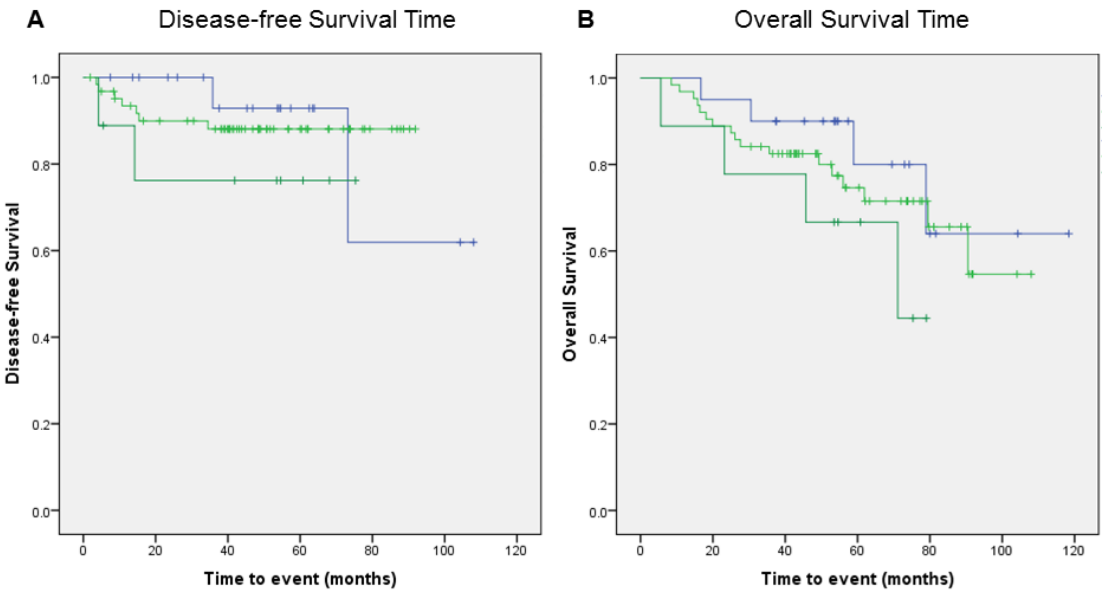


Figure 3-2 Kaplan-Meier time-to-event analysis comparing overall and disease free survival in the group of early-stage OSCC stratified according to histopathological differentiation

Colour index: Blue line – well differentiated; pale green line – moderately differentiated; dark green line – poorly differentiated.

Kaplan-Meier time-to-event analysis showed that histological grade of differentiation in early-stage OSCC did not correlate with clinical outcome measured in terms of either A) disease-free survival ($p>0.05$, x^2 value - 1.3, 2 d.f.) or B) overall survival ($p>0.05$, x^2 value – 2.2, 2 d.f.).

3.3.7 Normal epithelium: EGFR protein expression and gene copy number in normal tissue used as an internal control

Normal epithelium was annotated for image analysis in a subset of 60 cases (30 OPMD and 30 OSCC). The mean number of cells analysed per case was 3341 (range: 661 - 10594). The mean percentage of positive cells (PPC) was 85.3% (range: 58.1% - 98.9%). The mean percentage of strongly-positive cells (3+PC) was 20.4% (range: 0.8% - 56.1%). Nuclei in the normal epithelium consistently showed disomy, the gene copy number signal that is expected for normal, healthy epithelium (Figure 3-4C).

3.3.8 OPMD: EGFR protein expression

Dysplastic epithelium was annotated for image analysis in each of the 78 OMPD. The mean number of cells analysed per case was 2455 (range: 1175 - 9638). The mean PPC was 85.4% (range: 54.8% - 99.7%). The mean 3+PC was 33.8% (range: 0% - 68.7%).

Comparison of image analysis data showed that EGFR protein was over-expressed in dysplastic epithelium relative to normal epithelium (Figure 3-3 and Figure 3-4). The majority (66.6%) of dysplastic areas had PPC values that were higher than the mean PPC value of the normal epithelium. However, a pairwise comparison showed that the mean PPC values of dysplastic and normal epithelium did not differ significantly ($p>0.05$, Independent T-test). The majority (78.2%) of dysplastic areas also had 3+PC values that were higher than the mean 3+PC value of normal epithelium. A pairwise comparison showed that dysplastic epithelium had a significantly higher mean 3+PC value than normal epithelium ($p<0.01$, Independent T-test, Figure 3-3).

There was a positive correlation between grade of epithelial dysplasia and EGFR protein expression (Figure 3-4). High-grade epithelial dysplasia had a significantly higher mean 3+PC value relative to low-grade epithelial dysplasia. Similarly, cases with SIN 2 and SIN 3 had a significantly higher mean 3+PC value than cases with SIN 1 (One-way ANOVA; Bonferroni correction, Figure 3-3).

Mean 3+PC values did not differ significantly when OPMD were grouped according to the four individual clinical outcomes [data not shown]. OPMD with adverse outcomes (i.e. local recurrence, new lesion formation, or malignant transformation) had a significantly higher mean 3+PC value than cases with no adverse outcome (Figure 3-3). However, the difference between the mean 3+PC values of OPMD that underwent malignant transformation and those which did not (i.e. OPMD with no adverse outcome, local recurrence, or new lesion formation) was not statistically significant (Figure 3-3). This was confirmed by Kaplan-Meier time-to-event analysis (Figure 3-5B). High EGFR protein expression (defined as a 3+PC value greater than the mean 3+PC value of the normal epithelium) had low sensitivity and specificity for detecting cases that were destined to undergo malignant transformation (Figure 3-6 and Table 3-4). No significant differences in PPC/3+PC values were detected when cases were analysed according to patient demographics (age, sex), risk factors (alcohol/tobacco habits), or OPMD mucosal subsite [data not shown].

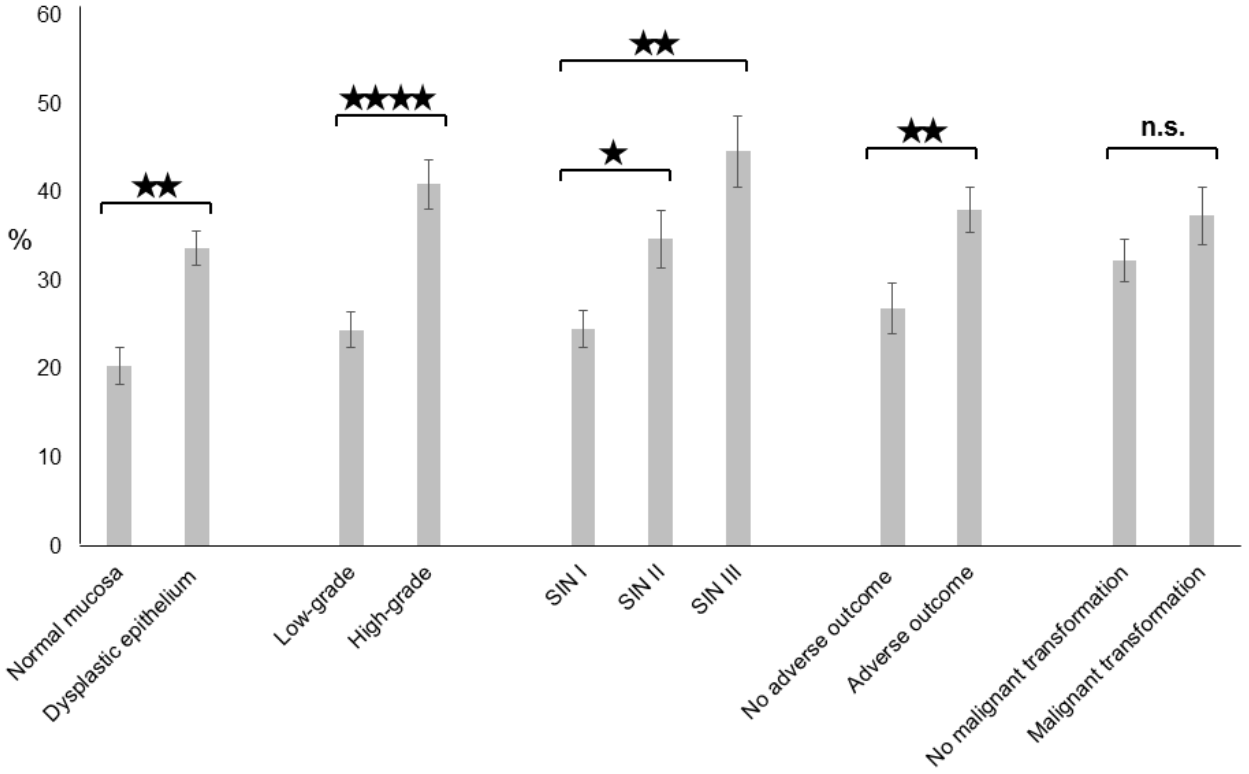


Figure 3-3 Bar chart comparing the mean EGFR 3+PC values of normal epithelium and OPMD stratified according to grade of epithelial dysplasia and clinical outcome

The mean 3+PC of dysplastic epithelium was significantly higher than the mean for normal epithelium ($p<0.01$). The mean 3+PC of high-grade epithelial dysplasia was significantly higher than the mean for low-grade epithelial dysplasia ($p<0.0001$). Cases with SIN 2 and SIN 3 had a significantly higher mean 3+PC than cases with SIN 1 ($p<0.05$ and $p<0.01$ respectively). OPMD with adverse outcomes had a significantly higher mean 3+PC than cases with no adverse outcome ($p<0.01$). However, there was no significant difference between the mean 3+PC values of OPMD that underwent malignant transformation and those which did not (n.s. – not significant).

3.3.9 OPMD: EGFR gene copy number

Each OPMD was assigned to either a ‘normal’ or ‘abnormal’ category according to its predominant EGFR gene copy number signal. This approach to the interpretation of EGFR gene copy number in OPMD has been applied in a previous study (Benchekroun et al., 2010). It is further detailed in Chapter 2, Figure 2.2. An abnormal EGFR gene copy number signal was detected in a total of 16 OPMD (20.5% of cases).

Fifteen of these cases showed either trisomy or low polysomy (Figure 3-4I):

- Eight subsequently underwent malignant transformation. Each showed high-grade epithelial dysplasia (six cases graded as SIN 2; two cases as SIN 3)
- Seven cases did not progress to OSCC:
    - Two cases had no adverse outcome
    - However, five of these cases had other adverse outcomes:
        - Local recurrence (four cases)
        - New lesion formation (one case).

One case showed clusters, a signal that is considered evidence of genomic gain according to the manufacturer’s scoring system (Figure 3-4L). The patient, a 58 year-old male, had a history of both alcohol and tobacco consumption. The OPMD biopsy showed high-grade dysplasia (also graded as SIN 3). Malignant transformation was diagnosed after a 17-month interval.

Kaplan-Meier time-to-event analysis detected a significant correlation between abnormal EGFR gene copy number and the risk of malignant transformation (Figure 3-5). Receiver-operator curve analysis was used to compare the reliability of abnormal EGFR gene copy number in detecting cases destined to undergo malignant transformation relative to high-grade epithelial dysplasia and high EGFR protein expression (Figure 3-6). This analysis showed that abnormal EGFR gene copy number was the most reliable marker of malignant transformation. Reliability was further enhanced when abnormal EGFR gene copy number and high-grade epithelial dysplasia were combined to form a single marker (i.e. positive cases had both high-grade epithelial dysplasia and an abnormal EGFR gene copy number (Figure 3-5C and Figure 3-6). Overall, 60% of cases with both high-grade epithelial dysplasia and an abnormal EGFR

gene copy number signal underwent malignant transformation (Figure 3-7). The superior detection of this combined marker was confirmed by comparison of sensitivity/specificity and positive/negative predictive values of all four markers (i.e. high-grade epithelial dysplasia, high EGFR protein expression, abnormal EGFR gene copy number, and high-grade epithelial dysplasia/abnormal EGFR gene copy number combined). These values are summarised in Table 3-4.

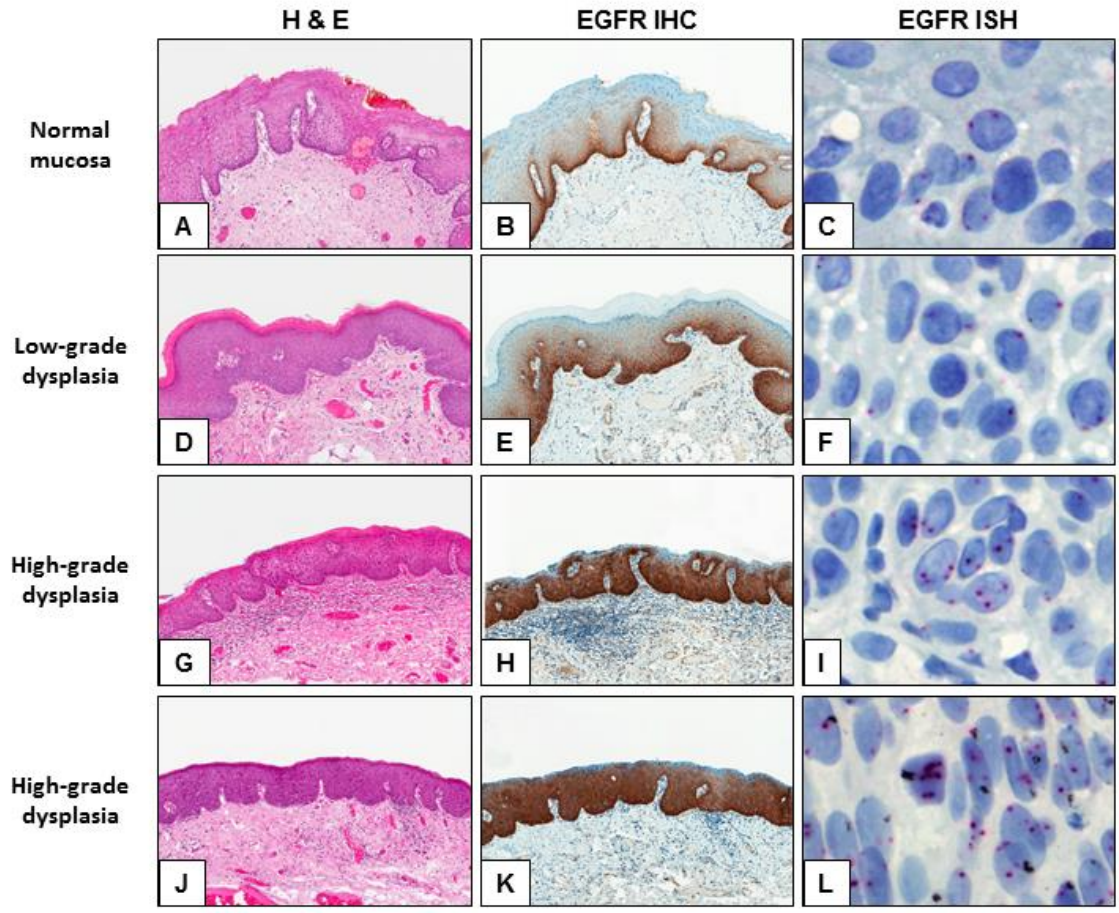


Figure 3-4 Comparison of EGFR protein expression and EGFR gene copy number in normal mucosa and OPMD

A) Normal mucosa at the edge of an OPMD biopsy. B) In normal epithelium, EGFR protein expression is strongest in the basal/parabasal layers. C) Nuclei of normal keratinocytes show disomy. D) Low-grade epithelial dysplasia. E) There is increased EGFR protein expression relative to the normal epithelium. F) However, nuclei continue to show disomy. G) and J) High-grade epithelial dysplasia. H) and K) EGFR protein expression is increased relative to both normal epithelium and low-grade epithelial dysplasia. I) The first example of high-grade epithelial dysplasia shows low polysomy (3-4 nuclear signals in ≤40% of cells). L) The second example shows clustered EGFR gene copy number signals, consistent with genomic gain.

H&E & EGFR IHC x100 magnification; EGFR ISH x400 original magnification.

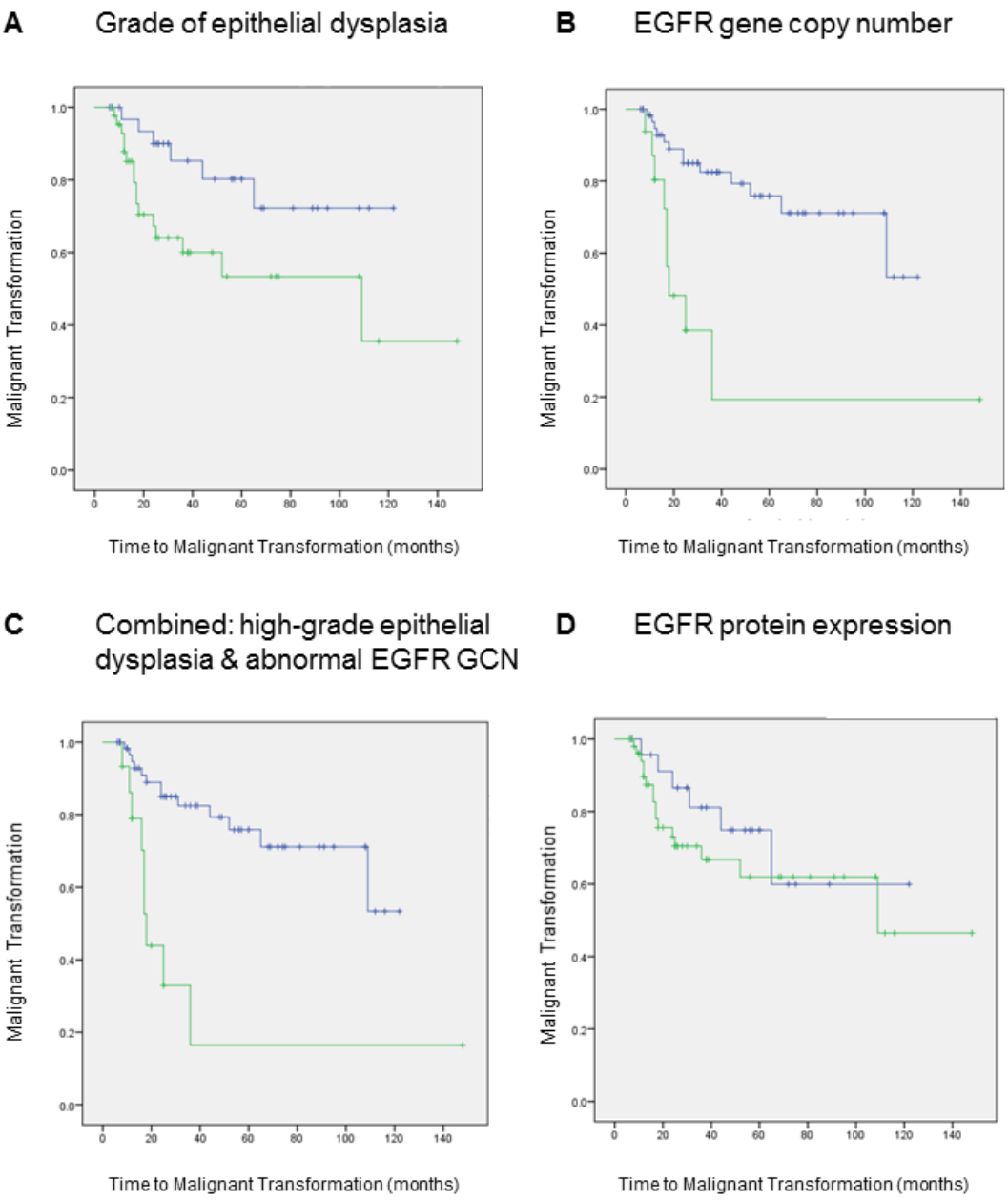
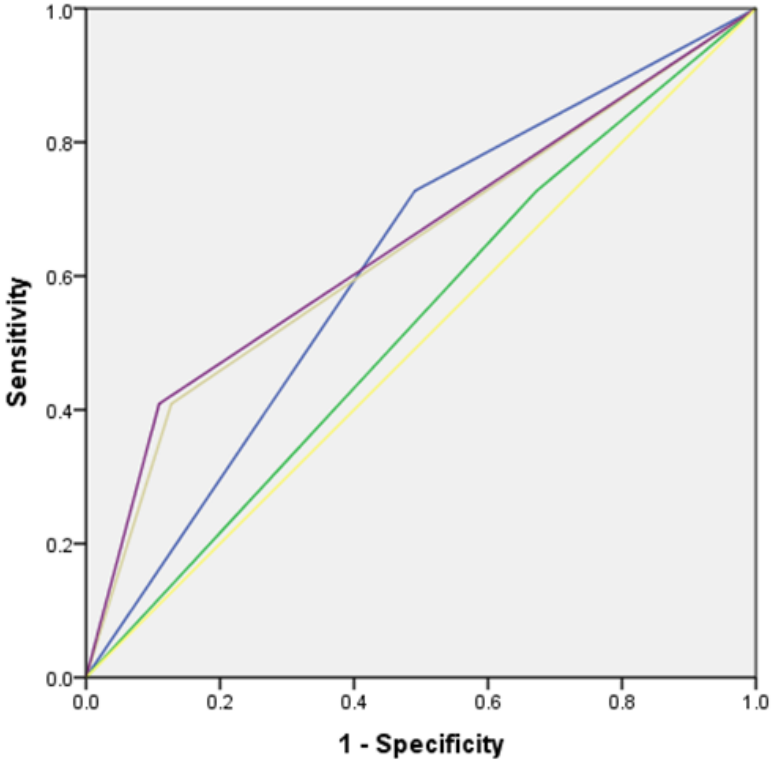


Figure 3-5 Kaplan-Meier time-to-event analysis showing malignant transformation in OPMD stratified according epithelial dysplasia, EGFR protein expression, and EGFR gene copy number

Colour index: Blue line - low grade epithelial dysplasia, low EGFR protein expression, normal EGFR gene copy number; negative combined score. Green line - high-grade epithelial dysplasia; high EGFR protein expression; abnormal EGFR gene copy number; positive combined score.

A) There was a significant correlation between malignant transformation and high-grade epithelial dysplasia ($p<0.05$, x^2 value - 4.97, 1 d.f.). B) There was a significant correlation between malignant transformation and abnormal EGFR gene copy number ($p<0.0001$; x^2 value - 13.9, 1d.f.). C) A significant correlation was also identified when abnormal EGFR gene copy number and high-grade epithelial dysplasia were combined ($p<0.0001$; x^2 value - 16.1, 1d.f.). D) By contrast, there was no correlation between malignant transformation and high EGFR protein expression ($p>0.05$).



Test Variable	Area beneath curve	Standard error	Asymptotic Significance	Asymptotic 95% Confidence Interval	
				Lower	Upper
High-grade epithelial dysplasia	0.61	0.069	p>0.10	0.48	0.75
High EGFR protein expression	0.53	0.074	p>0.649	0.496	0.788
Abnormal EGFR Gene copy number	0.64	0.074	p<0.05*	0.50	0.79
Combined high-grade epithelial dysplasiaand abnormal EGFR gene copy number	0.65	0.074	p<0.05*	0.50	0.79

Figure 3-6 Receiver-operator curve analysis comparing the detection of OPMD destined to undergo malignant transformation by epithelial dysplasia, EGFR protein expression, and EGFR gene copy number

Colour index: Blue – high grade epithelial dysplasia; green – high EGFR protein expression; cream – abnormal EGFR gene copy number; purple – combined high grade epithelial dysplasia and abnormal EGFR gene copy number; yellow – reference line.

The combined marker (high-grade epithelial dysplasia and abnormal EGFR gene copy number) has the greatest area beneath the curve. This suggests it is the most reliable detector of OPMD destined to undergo malignant transformation. Both abnormal EGFR gene copy number and the combined category show an asymptotic significance of $p<0.05$. By contrast, high-grade epithelial dysplasia and high EGFR protein expression are not significant at $p<0.05$, suggesting they do not reliably detect OPMD that undergo malignant transformation.

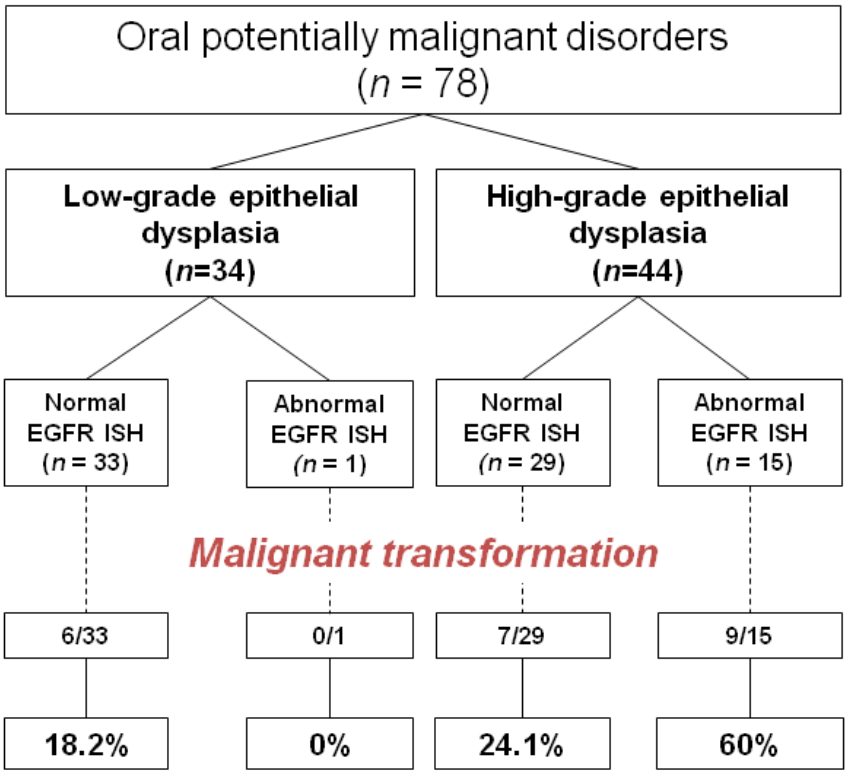


Figure 3-7 Diagnostic algorithm showing distribution of OPMD that underwent malignant transformation according to grade of epithelial dysplasia and EGFR gene copy number signal

The majority (60%) of OPMD that transformed to OSCC showed both high-grade epithelial dysplasia and an abnormal EGFR signal. This was more than double the proportion of cases with high-grade epithelial dysplasia and a normal EGFR signal, and more than triple the proportion of cases that were negative for both markers (i.e. cases with low-grade epithelial dysplasia and a normal EGFR signal).

Table 3-4 Detection of OPMD that underwent malignant transformation using high-grade epithelial dysplasia, high EGFR protein expression, and abnormal EGFR gene copy number

	High-grade epithelial dysplasia	High EGFR protein expression	Abnormal EGFR gene copy number	Combined dysplasia/gene copy number
Sensitivity	72.7	72.7	40.9	40.9
Specificity	50	14.0	87.5	89.1
Positive predictive value	36.4	30.2	56.3	60.0
Negative predictive value	82.4	50.0	79.0	79.0

3.3.10 OSCC arising from OPMD cases: EGFR protein expression

Malignant transformation occurred in 22 OPMD (section 3.3.4). Biopsy material was available for analysis in 21 of these cases. The mean number of cells analysed per case was 10,240 (range: 2442 – 30,730). The mean PPC was 88.5% (range: 66.8 – 99.7%). The mean 3+PC was 37.9% (range: 13.9 – 68.7%).

The mean 3+PC for the group of transformed OSCC was significantly higher than the mean 3+PC for the normal epithelium ($p<0.0001$, Independent T-test; Figure 3-8). It was also significantly higher than the mean 3+PC for OPMD. The difference was significant when comparing both the entire group of OPMD and the subset that underwent malignant transformation ($p<0.0001$, Independent T-test). However, the mean PPC value for transformed OSCC did not differ significantly from the mean PPC of normal epithelium ($p>0.05$, Independent T-test).

The mean PPC and 3+PC values for transformed OSCC were lower than the corresponding values for early-stage OSCC (Figure 3-8). However, neither of these trends was statistically significant ($p>0.05$, Independent T-test (PPN) and Mann-Whitney U-test (3+PN)).

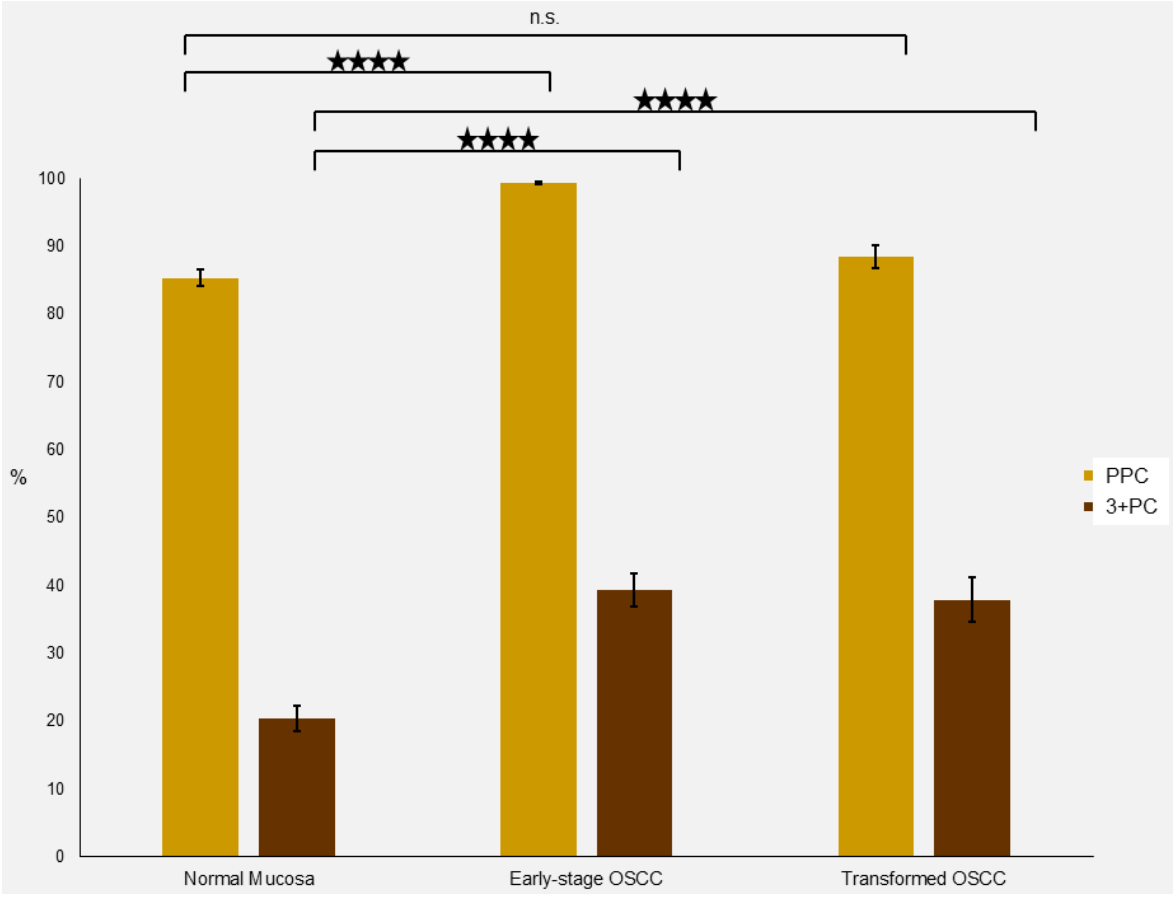


Figure 3-8 Bar chart comparing EGFR protein expression of normal mucosa with early-stage OSCC and the group of OSCC that transformed from OPMD

EGFR protein was over-expressed in both early-stage OSCC and the group of transformed OSCC relative to normal mucosa. For early-stage OSCC, both the mean PPC and 3+PC values were significantly higher than the corresponding values of the normal mucosa ($p<0.0001$, Independent T-tests). The mean 3+PC of the transformed OSCC group was also significantly higher ($p<0.0001$, Mann-Whitney U-test). Differences between the mean PPC and 3+PC values of early-stage and transformed OSCC were not statistically significant.

3.3.11 OSCC arising from OPMD cases: EGFR gene copy number

The majority of the OSCC cases in this subset showed an aberrant EGFR gene copy number, either an abnormal signal (seven cases) or EGFR genomic gain (five cases). Just less than half of the cases showed a normal EGFR signal (nine cases).

The differences between the EGFR gene copy number category of the index OPMD and the subsequent OSCC are summarised in Table 3-5. In one-third of cases (seven cases, 33%) the transition from the index OPMD to OSCC was associated with a move to a higher EGFR gene copy number category:

- In five cases, the index OPMD showed a normal EGFR signal which progressed in the subsequent OSCC to either an abnormal EGFR signal (three cases) or EGFR genomic gain (two cases)
- In two cases, the index OPMD showed an abnormal EGFR signal which progressed in the subsequent OSCC to EGFR genomic gain.

However, in the majority of cases (12 cases, 57.1%) the index OPMD and subsequent OSCC showed the same EGFR gene copy number signal:

- In the single case of OPMD that showed clusters (i.e. EGFR genomic gain) clusters were also detected in the subsequent OSCC
- An abnormal EGFR gene copy number signal was detected in both OPMD and OSCC in four cases
- A normal EGFR signal was detected in seven cases.

In two cases, the index OPMD showed an abnormal EGFR gene copy number but the subsequent OSCC showed a normal signal.

Table 3-5 Summary of the EGFR gene copy number categories assigned to the OPMD that underwent malignant transformation and the subsequent OSCC

EGFR gene copy number signal		Number (%)
OPMD	**OSCC**	
Normal	Normal	7 (33)
Normal	Abnormal	3 (14)
Normal	Genomic gain	2 (10)
Abnormal	Normal	2 (10)
Abnormal	Abnormal	4 (19)
Abnormal	Genomic gain	2 (10)
Genomic gain	Genomic gain	1 (5)
Total		**21 (100)**

3.3.12 Early-stage OSCC: EGFR protein expression

Areas of OSCC were annotated for image analysis in each of the 92 cases. The mean number of cells analysed per case was 10219 (range: 578 - 38509). The mean PPC was 99.4% (range: 82.7% - 100%). The mean 3+PC was 39.4% (range: 0.1% - 79.2%).

Comparison of image analysis data showed that EGFR protein was over-expressed in early-stage OSCC relative to the normal epithelium (Figure 3-8, Figure 3-9). OSCC had higher PPC and 3+PC values than the corresponding mean values for normal epithelium in 98.9% and 73.9% of cases, respectively. Pairwise comparisons showed that these differences were both statistically significant ($p<0.0001$, Independent T-test; Figure 3-8). Small differences between the PPC and 3+PC values of OSCC stratified according to differentiation were not significant ($p>0.05$, One-way ANOVA with Bonferroni correction, Figure 3-9E, H, and K). There was no correlation between the mean PPC/3+PC values when analysed according to patient demographics (age, sex), OSCC mucosal subsite, or pStage. There was no correlation between PPC/3+PC values and either overall or disease-free survival [data not shown].

3.3.13 Early-stage OSCC: EGFR gene copy number

Of the 92 cases in the study, one case was excluded from EGFR gene copy number analysis due to technical difficulties that hampered interpretation. Satisfactory staining in the remaining 91 cases enabled the scoring pathologists to reach a consensus.

EGFR genomic gain was identified in 23 (24.7%) early-stage OSCC. EGFR genomic gain was visualised as either high polysomy (11 cases) or clusters (12 cases) (Figure 3-9). All of the cases that showed EGFR genomic gain were at pStage I. EGFR genomic gain was associated with a reduction in mean overall survival time (50.2 months compared to 57.7 months for cases with normal gene copy number) and mean disease-free survival (45.6 months compared to 47.7 months for cases with normal gene copy number). However, Kaplan-Meier time-to-event analysis showed that neither of these trends was statistically significant (both $p>0.05$).

3.3.14 Correlation between EGFR expression and EGFR gene copy number

There was a positive correlation between EGFR gene copy number and EGFR protein expression. EGFR protein expression was significantly higher in OPMD with abnormal EGFR gene copy number than in cases with normal EGFR gene copy number ($p<0.0001$, mean values of 49.9% (s.d. - 12.1) and 29.3% (s.d. - 15.7) respectively). EGFR protein expression was significantly higher in OSCC with EGFR genomic gain relative to cases with no genomic gain ($p<0.01$, mean values of 51.2% (s.d. - 21.9) and 35.9% (s.d. - 22.5) respectively).

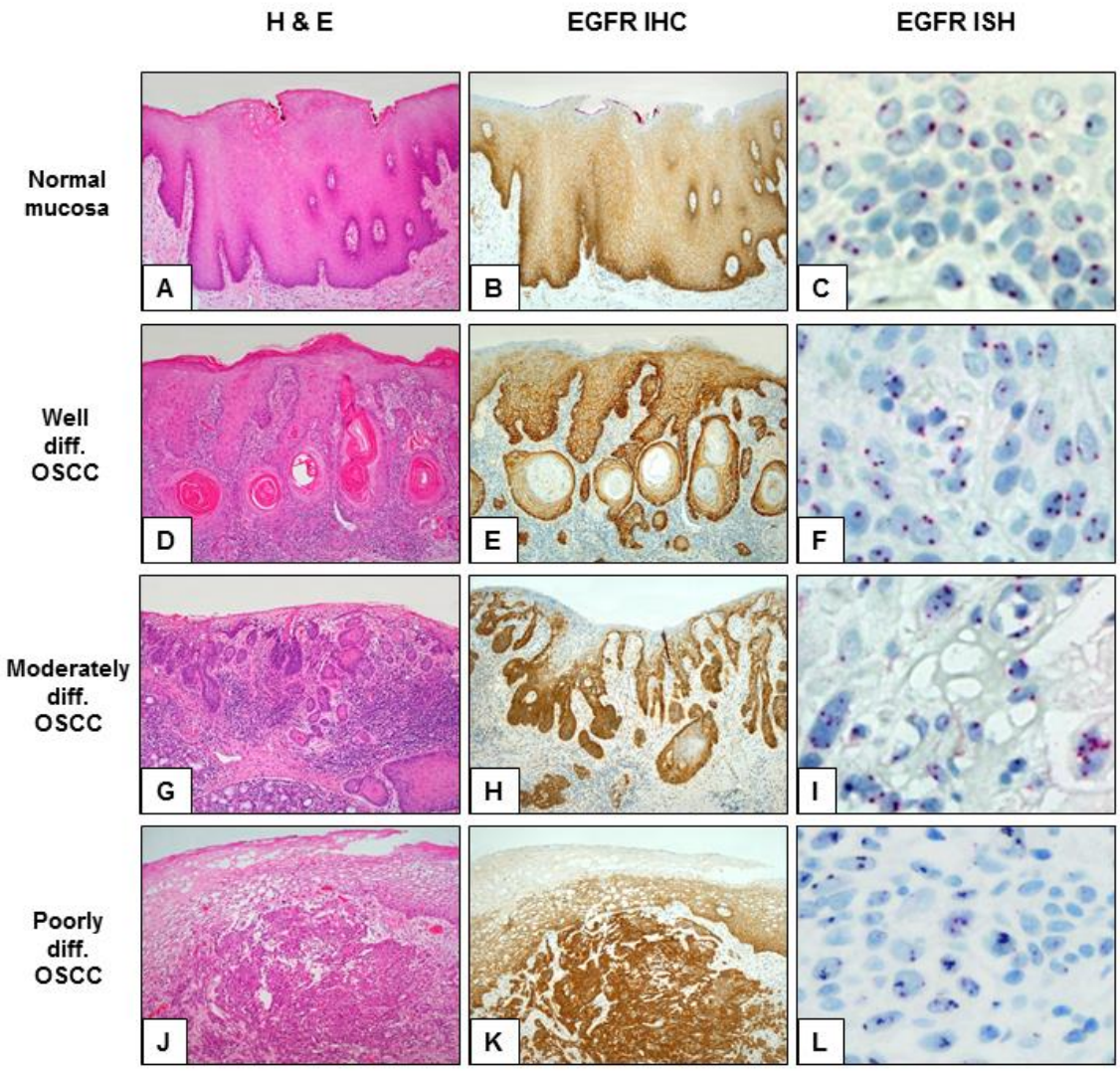


Figure 3-9 EGFR protein expression and gene copy number in normal mucosa and early-stage oral squamous cell carcinoma stratified according to histopathological differentiation

A) Normal mucosa. B) EGFR protein expression is strongest in the basal/parabasal layers. C) Nuclei of normal keratinocytes show disomy. D) Well differentiated OSCC. E) There is increased EGFR protein expression relative to the normal epithelium. F) Nuclei show low polysomy. G) Moderately differentiated OSCC. J) Poorly differentiated OSCC. H) and K) EGFR protein expression is increased relative to the normal epithelium in the corresponding EGFR stains, but intensity is similar to that of the well differentiated OSCC shown in B). I) EGFR *in situ* hybridisation shows high polysomy (3-4 nuclear signals in $\geq$40% of cells). L) Clustered gene copy number signals are seen in the example of poorly differentiated OSCC.

H&E & EGFR IHC taken at x100 magnification; EGFR ISH at x400 original magnification.

3.4 Discussion

Oral squamous cell carcinoma (OSCC) is a global healthcare problem (Jemal et al., 2011). In the UK, the incidence of OSCC rose by a third during the decade to 2010. There are now ~6500 new cases annually (Cancer Research UK, 2010). In much of the developed world, the increase in incidence has been greatest among young adults (Schantz and Yu, 2002; Garavello et al., 2010).

3.4.1 Patient characteristics and risk factors

In the present study, most patients with OPMD/early-stage OSCC were in their 6th to 7th decades. This demographic profile is consistent with the prevalence of traditional risk factors that were identified in the OPMD group. Alcohol and tobacco consumption are well recognised for their carcinogenic potential and synergistic interactions (Danaei et al., 2005; Hashibe et al., 2009).

Over the past decade, it has emerged that human papillomavirus (HPV) infection is significant in the aetiology of a subset of oropharyngeal squamous cell carcinomas (Schache et al., 2011; Robinson et al., 2012). There is emerging evidence that HPV positivity may have prognostic significance in squamous cell carcinoma presenting at non-oropharyngeal sites, including the oral cavity (Chung et al., 2014). However, the proportion of HPV-positive OSCC is low (~3%). This suggests that HR-HPV infection is not a major aetiological agent in oral carcinogenesis (Salazar et al., 2014). Interestingly, however, HPV infection has been implicated in the formation of a subset of OPMD (Angiero et al., 2010) and specific histopathological features have been described for this subset of cases (Lerman and Woo, 2014). However, to date convincing evidence for the contribution of HPV infection to OPMD and OSCC formation has yet to emerge (Lopes *et al.*, 2011; Lingen *et al.*, 2013; McCord *et al.*, 2014; Nankivell *et al.*, 2014). Given the contrasting role of HPV in the formation of OSCC and oropharyngeal squamous cell carcinoma, the exclusion of oropharyngeal lesions is an important strength of the present study. It seems plausible that other risk factors associated with the increased incidence of OSCC in young adults are yet to be identified. It is conceivable that future studies may show a reduction in the prevalence of alcohol and tobacco consumption relative to the present study.

3.4.2 Classification and grading of epithelial dysplasia

The current edition of the World Health Organisation (WHO) Classification of Head and Neck Tumours details three systems for classifying epithelial dysplasia (Barnes et al., 2005). Of these, the 2005 WHO Classification is most widely used by both diagnostic pathologists and clinicians in the UK. The WHO Classification is used in the Department of Cellular Pathology and was applied to the original diagnosis of OPMD in the present study. However, there were too few cases of carcinoma in-situ (CIS) to support valid statistical comparisons. The present study therefore applied the 'Squamous Intraepithelial Neoplasia' (SIN) classification, which combines CIS and severe epithelial dysplasia in a single category, SIN 3.

It is well recognised that the grading of oral epithelial dysplasia is prone to inter-observer variation (Kujan et al., 2007). Furthermore, it is documented that the grade of epithelial dysplasia does not accurately predict clinical outcomes, particularly the risk of malignant transformation (Dost et al., 2014). A binary system for grading epithelial dysplasia has been developed in an attempt both to reduce inter-observer variation and enhance the predictive value of the assigned dysplasia grade (Kujan et al., 2006). The present study confirms the superior prognostic value of the binary system compared to the SIN classification. However, the correlation identified between high-grade epithelial dysplasia and malignant transformation in the present study was fairly weak. Used in isolation, high-grade epithelial dysplasia had low sensitivity and specificity for detecting cases of OPMD destined to undergo malignant transformation. This supports the view that there may be fundamental limitations to the predictive value of a morphological diagnosis, irrespective of the accuracy and reliability of the classification system upon which it is based (Kujan et al., 2007; Dost et al., 2014). This in turn highlights the need for biomarkers that reliably detect the subset of OPMD that are at greatest risk of malignant transformation (Mishra, 2012).

3.4.3 Rate of malignant transformation in OPMD

There is wide variation in the documented rates of malignant transformation in OPMD (Liu et al., 2011; Warnakulasuriya et al., 2011; Dost et al., 2014). Overall, nearly one-third of OPMD in the present study underwent malignant transformation. This rate is considerably higher than the 2.6% reported by Warnakulasuriya et al (2011) in a UK-based study of more than 1300 OPMD patients. It is also higher than the 7% rate of malignant transformation recently reported by Diajil et al (2013), from whose study many of our OPMD cases were drawn. This high rate is an artefact of using two different strategies to identify OPMD cases. Strategy 2), the retrospective search of the Department of Cellular Pathology database, identified OSCC with previous OPMD biopsies. Including cases identified by this strategy therefore increased the proportion of OPMD that underwent malignant transformation. It is likely that the rate of 7% reported by Diajil et al (2013) in a cohort of consecutive cases is a more accurate indication of the rate of malignant transformation in OPMD.

The heterogeneous approach to identifying cases leads to a further limitation of the present study: variable management of OPMD patients, with some undergoing laser excision, and others being managed by surveillance. The majority of the OPMD that underwent malignant transformation were managed by surveillance. By contrast, the majority of OPMD managed by laser excision had no adverse outcome. The present study was not designed to evaluate the clinical efficacy of laser excision in reducing the risk of malignant transformation. However, it is conceivable that some cases managed by surveillance may not have proceeded to OSCC had they been managed by laser excision.

3.4.4 EGFR protein expression and gene copy number as cancer biomarkers

EGFR protein expression and gene copy number are used in both the prognostication of non-small cell lung carcinoma (Nicholson et al., 2001; Hirsch et al., 2003) and prediction of its response to EGFR-targeted chemotherapeutic agents (Takano et al., 2005). Since EGFR protein over-expression in OSCC was first documented (Grandis and Tweardy, 1993a) it has been hoped that

EGFR would prove a similarly effective biomarker and chemotherapeutic target for patients with OPMD/OSCC. However, the literature to date has consistently drawn attention to the complexity of the EGFR pathway and consequent limitations of its clinical utility as a biomarker in OSCC/OPMD (Forastiere, 2007; Gusterson and Hunter, 2009; Rosin and Califano, 2010). Recently, EGFR gene copy number has emerged as a potential biomarker that may complement EGFR protein expression (Erjala et al., 2006; Forastiere, 2007); however, the prevalence and clinical significance of EGFR gene copy number increases in OSCC are not well defined (Chung et al., 2006; Agulnik et al., 2007; Temam et al., 2007; Pectasides et al., 2011). Moreover, it remains unclear whether it is appropriate to apply criteria validated for the interpretation of EGFR gene copy number signals in non-small cell lung carcinomas to other cancers.

3.4.5 EGFR protein is up-regulated in OPMD and early-stage OSCC

Our data confirm that EGFR protein expression is increased in the majority of OPMD and OSCC (Grandis and Tweardy, 1993a; Ries et al., 2013). The ubiquity of EGFR over-expression highlights the critical role played by the EGFR pathway in oral carcinogenesis, but limits its clinical utility as a biomarker for stratifying patient management. A smaller proportion of OPMD showed EGFR over-expression compared to OSCC. The biological significance of this is uncertain, but it supports the contention that the effects of EGFR are likely to be more significant in the later stages of oral carcinogenesis (Ryott et al., 2009). Alternatively, increased EGFR expression may represent a bystander change, reflecting, but not driving, tumour progression, which would account for the lack of correlation with disease-specific clinical outcomes (Forastiere, 2007; Gusterson and Hunter, 2009; Rosin and Califano, 2010).

The high prevalence of EGFR over-expression in the current study is in contrast to data recently reported by Rössle et al (2013). The latter study reported EGFR over-expression in just 19.3% of early stage OSCC involving the tongue. However, although the study was similar in its focus on pStage I and II OSCC, it was limited by subjective semi-quantitative assessment of tissue cores. It is also possible that different thresholds were used to define EGFR over-expression to those used in the current study. Furthermore, it is our experience that EGFR

expression shows tumour heterogeneity and tissue microarray sampling may not correlate with measurements taken from whole sections.

3.4.6 EGFR protein expression correlates with grade of dysplasia but not clinical outcome in OPMD

Our data confirm the positive correlation between EGFR expression and grade of epithelial dysplasia reported by Ries et al (2013). In contrast to this earlier study, however, we were unable to identify a significant correlation between EGFR over-expression and the risk of malignant transformation. This correlation has also been documented in a recent comprehensive study of 148 OPMD (Nankivell *et al.*, 2013). Both of these studies analysed whole sections using subjective semi-quantitative assessment (Nankivell *et al.*, 2013; Ries *et al.*, 2013). The absence of a correlation in the current study may reflect the application of digital image analysis; as in the study by Rössle et al (2013), it is also possible that different thresholds were used to define EGFR over-expression. On balance, the high prevalence of increased EGFR expression in our study makes positive correlations with clinical parameters less likely.

3.4.7 Detecting significant EGFR gene copy number changes in OPMD

A fifth of OPMD in the present study showed abnormal EGFR gene copy number. Only one case showed unequivocal EGFR genomic gain as defined using the current criteria for non-small cell lung cancer. This supports data from OSCC EGFR gene copy number studies, which suggest that EGFR genomic gain is a late event in oral carcinogenesis (Freier et al., 2003; Rössle et al., 2013; Ryott et al., 2009). It is striking, however, that the majority of OPMD with abnormal EGFR gene copy had adverse clinical outcomes, and over half progressed to OSCC. This is in concordance with evidence from two recent studies, which suggest abnormal (or, ‘increased’) EGFR gene copy number is a relatively early feature of OPMD that are destined to undergo malignant transformation and may precede gross EGFR genomic gain (Benchekroun et al., 2010; Poh et al., 2012). Both of these studies used fluorescence ISH (FISH) rather than the chromogenic ISH (CISH) technique used in the current study, but the results are comparable. Benchekroun et al (2010) studied EGFR FISH in a subset of 49 OPMD, applying a definition of FISH positivity that

encompassed all EGFR gene copy number abnormalities, including trisomy and low polysomy. While one case showed EGFR gene copy number amplification using conventional criteria, a further 41% of cases showed FISH positivity according to their modified criteria. FISH-positive OPMD had significantly higher rates of malignant transformation compared to those with normal EGFR gene copy number ($p<0.001$). A recent study by Poh et al (2012) also supports the application of a lower threshold for detecting EGFR gene copy number abnormalities. In a study of 20 OPMD, one case showed clusters (i.e. EGFR genomic gain using the current criteria). However, the authors also reported that any gain in EGFR gene copy number was strongly associated with an increased risk of malignant transformation, irrespective of whether the genomic gain was low or high; EGFR genomic gain was also associated with a reduced time to malignant transformation (Poh et al., 2012).

It is a limitation of the current study that neither EGFR mutation status nor downstream targets of increased EGFR gene copy number/protein were evaluated. It is conceivable that EGFR gene copy number may represent a 'surrogate' for other genetic and molecular abnormalities and simply reflect chromosomal instability. Nevertheless, the positive correlation identified between EGFR gene copy number and protein over-expression suggests that EGFR genomic gain is likely to be functionally significant.

3.4.8 Abnormal EGFR gene copy number in OPMD may indicate an increased risk of malignant transformation

Our data suggest that the detection of EGFR gene copy number abnormalities may be a useful tool for identifying OPMD at high risk of adverse outcomes, particularly malignant transformation. The data indicate that while EGFR gene copy number abnormalities are relatively specific for identifying high-risk OPMD, they have low sensitivity. Low sensitivity may reflect the temporal relationship of the index biopsy to the transformation event, as it is conceivable that EGFR gene copy number abnormalities may have accumulated within an altered field after the index biopsy.

Data from the group of OSCC that transformed from OPMD provide some evidence to support this hypothesis. The majority of cases in this subset had a

higher EGFR gene copy number category in the OSCC relative to the index OPMD. The EGFR gene copy number signal of the OPMD was already either abnormal or increased in a further one-quarter of cases, and the signal was maintained in the subsequent OSCC. The transition from OPMD to OSCC was associated with a move to a lower EGFR gene copy number category in just two cases. Although it is conceivable that this represents either the selection of a different clone, given the marked heterogeneity of EGFR gene copy number signal, it is also possible that this represents a sampling error.

Although low sensitivity may limit the utility of abnormal EGFR gene copy number as an isolated biomarker in OPMD, its specificity means that it could have added predictive value in the context of a broader panel of more sensitive biomarkers. In isolation, grade of epithelial dysplasia does not reliably predict clinical behaviour. However, our data show that the reliability of these two markers is increased when they are combined. This finding signals the potential value of EGFR gene copy number as part of a panel of molecular biomarkers.

3.4.9 Prevalence and prognostic significance of EGFR genomic gain in OSCC

A quarter of early-stage OSCC in the present study showed EGFR genomic gain. This is higher than the prevalence of 9% reported in a recent tissue-microarray study by Rössle et al (2013). As with EGFR protein expression, however, it is our experience that OSCC exhibit heterogeneous ISH patterns. It seems plausible that this difference may also reflect a methodological artefact. Nevertheless, the proportion of cases with EGFR genomic gain in the current study is still towards the lower end of the range of values reported in earlier studies (Freier et al., 2003; Ryott et al., 2009). Interestingly, the prevalence of EGFR genomic gain in the subset of OSCC that had transformed from OPMD was similar to that of the entire early-stage OSCC group.

Our study did not identify a significant correlation between EGFR genomic gain and clinical outcome in OSCC. This is in contrast to a previous study by Temam et al (2007), which reported a 9% five-year survival rate for patients with EGFR genomic gain compared with 71% five-year survival rate for patients with no genomic gain. Although the study used quantitative real-time PCR, its findings

have been supported by FISH studies (Freier et al., 2003; Chung et al., 2006). However, all of these studies included a wide range of clinical stages (including stages III and IV). The absence of a correlation in the present study may reflect its narrower inclusion criteria and focus on early-stage OSCC. This interpretation is supported by the association between EGFR and late-stage OSCC (Grandis and Tweardy, 1993b; Grandis and Tweardy, 1993a). In their comparable study of early-stage OSCC of the tongue, Rössle et al (2013) were also unable to identify a significant correlation between EGFR gene copy number and clinical outcome.

3.4.10 EGFR-targeted chemotherapeutic agents

There is evidence to suggest that EGFR gene copy number may help to predict the response of head and neck cancers to EGFR-targeted agents. For example, high EGFR gene copy number has been shown to predict which patients have an increased likelihood of response to erlotinib therapy (Agulnik et al., 2007). The present study was not designed to investigate patients' response to EGFR-targeted agents or other clinical interventions: none of the patients received EGFR-targeted therapy and the OPMD group was heterogeneous, including cases managed by surveillance and laser excision (Diajil et al., 2013). Despite these limitations, however, our data support the view that a sub-group of OPMD and OSCC harbour gene copy number abnormalities. Whether this sub-group has differential responses to EGFR-targeted agents or other therapies remains to be tested.

3.5 Conclusion

This study highlights the potential clinical utility of EGFR gene copy number assessment for predicting malignant transformation of OPMD. Abnormal EGFR gene copy number indicates an increased risk of malignant transformation in OPMD. In the majority of cases, the transition from OPMD to OSCC is associated either with maintenance or accumulation of EGFR gene copy number abnormalities or progression to EGFR genomic gain. This supports the view that EGFR gene copy number abnormalities continue to accumulate within an altered field following an index OPMD biopsy. EGFR genomic gain is present in a quarter of early-stage OSCC, but does not correlate with their clinical outcomes. Abnormal EGFR gene copy number and EGFR genomic gain correlate with increased EGFR protein expression in both OPMD and OSCC. However, although EGFR protein expression is over-expressed in the majority of OPMD and early-stage OSCC, it does not correlate with clinical outcomes in either group.

Chapter 4. Expression of Two Novel Biomarkers in Potentially Malignant Disorders and Early-Stage Squamous Cell Carcinoma of the Oral Cavity

4.1 Introduction

The previous chapter detailed the protein expression and gene copy number profiles of epidermal growth factor receptor (EGFR) in patients with oral potentially malignant disorders (OPMD) and early-stage oral squamous cell carcinoma (OSCC). EGFR is an established prognostic and predictive biomarker in non-small cell lung carcinoma (Nicholson et al., 2001; Hirsch et al., 2003; Takano et al., 2005). EGFR has been studied as a candidate biomarker in oral carcinogenesis since EGFR protein over-expression in OSCC was first reported more than 20 years ago (Grandis and Tweardy, 1993a; Grandis and Tweardy, 1993b).

This chapter will examine the protein expression profiles of two novel biomarkers, SOX2 and PAX9, in the same two groups of patients, i.e. those with OPMD and early-stage OSCC. SOX2 and PAX9 are transcription factors with vital roles in the regulation of development and differentiation. *SOX2* is a potential oncogene in squamous cell carcinoma (SCC) of the lung. There is emerging evidence that *SOX2* has a similar role in OSCC. By contrast, *PAX9* is a potential tumour-suppressor gene in oesophageal SCC. To date, the expression profile and potential role of *PAX9* in oral carcinogenesis has yet to be described.

4.1.1 SOX2 is a potential oncogene in oral carcinogenesis

The *SOX* family of genes encode transcription factors that play critical roles in the regulation of development. These roles include the induction/suppression of stem cell proliferation and maintenance of stem cell pluripotency (Chew and Gallo, 2009).

SOX2 protein is expressed in the normal oral squamous epithelium. It promotes epithelial proliferation and the stabilisation of basal progenitor cells (Okubo et al., 2009). SOX2 protein is over-expressed in OPMD (Qiao et al., 2013) and OSCC (Freier et al., 2010). Genomic amplification of *SOX2* has also been

detected in OSCC (Freier et al., 2010; Kokalj Vokač et al., 2014). These findings suggest that *SOX2* has an oncogenic function in oral carcinogenesis.

This contention is supported by earlier studies of squamous cell carcinoma (SCC) of the lung and oesophagus (Bass et al., 2009; Hussenet et al., 2010; Hussenet and du Manoir, 2010; Lu et al., 2010). *In vitro* studies suggest that increased SOX2 expression contributes to de-differentiation in SCC (Hussenet et al., 2010; Hussenet and du Manoir, 2010). However, *SOX2* is also implicated in the differentiation pathway of SCC arising in the lung. Up-regulation of *SOX2* induces expression of markers of squamous differentiation, such as p63 and keratin 6 (Bass et al., 2009). This suggests that SOX2 may also promote squamous differentiation during carcinogenesis.

4.1.2 PAX9 is potential tumour-suppressor gene in oral carcinogenesis

The *PAX* gene family encodes transcription factors that regulate stem cell renewal, cellular proliferation, differentiation, migration, and survival. The contribution of *PAX* genes to these processes suggests they may also play significant roles in neoplasia (Chi and Epstein, 2002; Robson et al., 2006).

PAX9 protein is expressed in the squamous epithelium of the oesophagus and oral cavity. It is believed to regulate key aspects of epithelial differentiation (Jonker et al., 2004). PAX9 expression is lost/down-regulated in oesophageal squamous cell carcinoma and its precursor lesions (Gerber et al., 2002). In these oesophageal lesions there is an inverse relationship between PAX9 expression and clinical outcome. A decrease in the proportion of PAX9-positive cells correlates with increasing grade of epithelial dysplasia. PAX9 expression is lowest in invasive SCC (Gerber et al., 2002). These data suggest *PAX9* functions as a tumour-suppressor gene.

This hypothesis is supported by *in vitro* studies of lung cancer. *PAX9* is located on chromosomal region 14q13.3. A subset of lung cancer cell lines shows allelic loss of this region and loss of PAX9 expression (Harris et al., 2011). By contrast, other *in vitro* studies suggest *PAX9* has an oncogenic function. *PAX9* interacts with *c-myb*, a proto-oncogene that enhances the survival of SCC cell lines and is implicated in the pathogenesis of a range of cancers (Lee et al., 2008). *PAX9* contributes to the survival of a subset of lung cancer cell lines that

show amplification of chromosomal region 14q13.3 (Kendall et al., 2007). These findings suggest that *PAX9* may perform diverse roles in carcinogenesis, analogous to *TGF-β* (Paterson et al., 2001; Prime et al., 2004).

4.2 Aims

The aims of this chapter are:

- To describe the profiles of SOX2 and PAX9 protein expression in groups of patients with OPMD and OSCC. The patient characteristics, clinical outcomes, and clinico-pathological features of these groups have been described previously in Chapter 3
- To correlate the profiles of SOX2 and PAX9 protein expression with the patient characteristics, clinical outcomes, and clinico-pathological features of the OMPD and early-stage OSCC groups.

4.3 Results

4.3.1 Expression of SOX2 and PAX9 proteins in normal squamous epithelium

Normal squamous epithelium was selected for image analysis in 30 OPMD and 30 OSCC biopsies.

For SOX2 protein expression, the mean number of nuclei analysed per case was 2155 (range: 716 - 5753). The mean percentage of positive nuclei (PPN) was 32.5% (range: 0.4% - 72.2%). The mean percentage of strongly positive nuclei (3+PN) was 1.2% (range: 0% - 10.7%).

For PAX9 protein expression, the mean number of nuclei analysed per case was 2412 (range: 590 - 7554). The mean PPN was 39.0% (range: 0.1% - 83.2%). The mean 3+PN was 6.8% (range: 0% - 42.5%).

4.3.2 SOX2 protein expression in oral potentially malignant disorders

Dysplastic epithelium was analysed in each of the 78 OPMD. The mean number of nuclei analysed per case was 1328 (range: 714 - 2105). SOX2 had a heterogeneous expression profile in dysplastic epithelium, varying both between and within individual OPMD cases.

Overall, SOX2 expression was down-regulated in dysplastic epithelium relative to normal epithelium (Figure 4-1). The mean PPN was 23.6% (range: 0% - 88.3%). Dysplastic epithelium had PPN values that were lower than the mean PPN value of the normal epithelium in the majority (75.6%) of cases. A pairwise comparison showed that, across the group, dysplastic epithelium had a significantly lower mean PPN value relative to the normal epithelium ($p<0.05$, Independent T-test).

However, several OPMD cases showed strong focal expression of SOX2 in the dysplastic epithelium (Figure 4-1F). There was also a subset of cases that showed strong, generalised over-expression of SOX2 (Figure 4-1H). Many of the cells with strong SOX2 expression exhibited basaloid morphology. The mean 3+PN for SOX2 was 3.2% (range: 0% - 64.4%). Although the mean 3+PN value for dysplastic epithelium was higher than the mean 3+PN value of the

normal epithelium (1.2%), this was due to a small subset of cases with very high 3+PN values. This subset comprised four cases with 3+PN values greater than 10%. In fact, SOX2 3+PN values of dysplastic epithelium were actually lower than the mean 3+PN of normal epithelium in the majority of cases (80.8%). A pairwise comparison showed that the difference between the mean 3+PN values of normal and dysplastic epithelium was not statistically significant ($p>0.05$, Mann-Whitney U-test; Figure 4-2).

4.3.3 SOX2 protein expression correlates weakly with histological grade of epithelial dysplasia but not with clinical outcome

Dysplastic epithelium graded as SIN 3 had a significantly higher mean 3+PN value than dysplastic epithelium graded as SIN 1 ($p<0.05$, Mann-Whitney U-test, Bonferroni correction; Figure 4-2). However, differences between the mean 3+PC values of other SIN categories were not significant ($p>0.05$, Mann-Whitney U-test, Bonferroni correction). Differences between mean PPN values were not significant ($p>0.05$, One-way ANOVA, Bonferroni correction). Neither mean PPN nor mean 3+PN values differed significantly when SOX2 expression was compared in dysplastic epithelium graded using the binary classification (3+PN values shown in Figure 4-2).

SOX2 expression did not differ significantly according to clinical outcome. Pairwise comparisons of PPN and 3+PN values of cases that underwent malignant transformation and those which did not (i.e. cases with no adverse outcome, local recurrence, or new lesion) were not statistically significant ($p>0.05$, Independent T-test and Mann Whitney U-test). Similarly, comparisons of the PPN and 3+PN values of cases grouped according to each of the four individual outcome categories were not significant ($p>0.05$, One-way ANOVA and Kruskal Wallis tests, Bonferroni correction; Figure 4-2). Kaplan-Meier time-to-event analysis confirmed the absence of a significant correlation between either the PPN or 3+PN value and the risk of malignant transformation (Figure 4-3).

SOX2 protein expression did not differ significantly according to patients' age; sex; alcohol/tobacco habits; or mucosal subsite of the OPMD [data not shown].

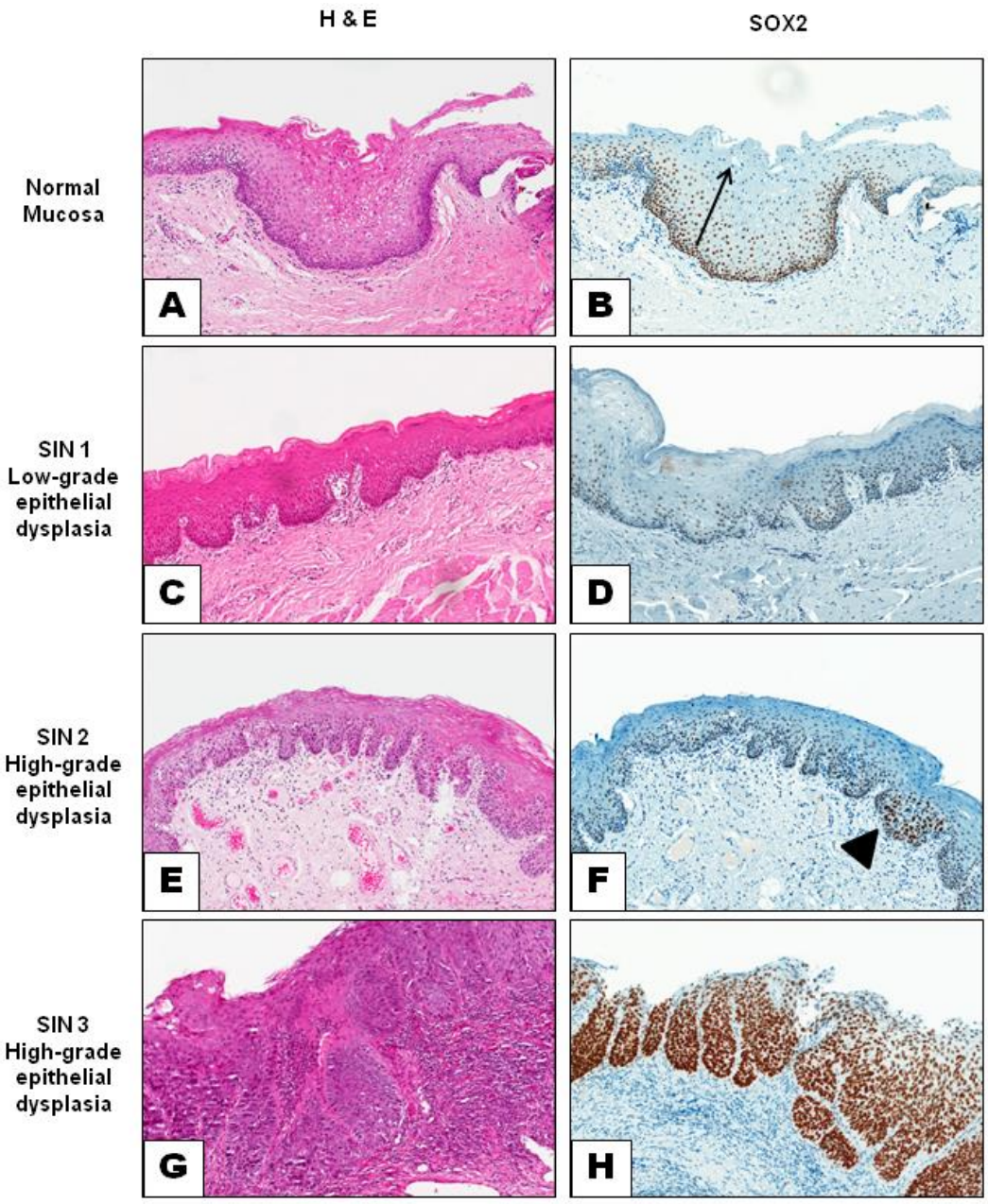


Figure 4-1 SOX2 protein expression in normal epithelium and oral potentially malignant disorders stratified according to histological grade of epithelial dysplasia

A) Normal oral mucosa. B) SOX2 expression is strongest in the basal layer and becomes weaker as keratinocytes differentiate and progress towards the surface (arrow). C) SIN 1. D) SOX2 expression is variable but generally down-regulated, most noticeably in the basal layer. E) SIN 2. F) Variable SOX2 expression is also seen in SIN 2, but there is a focus of strongly-positive cells (arrowhead). G) SIN 3. H) There is over-expression of SOX2 throughout the full thickness of the epithelium. The strongly SOX2-positive cells have a basaloid cytomorphology.

Images taken at x100 magnification.

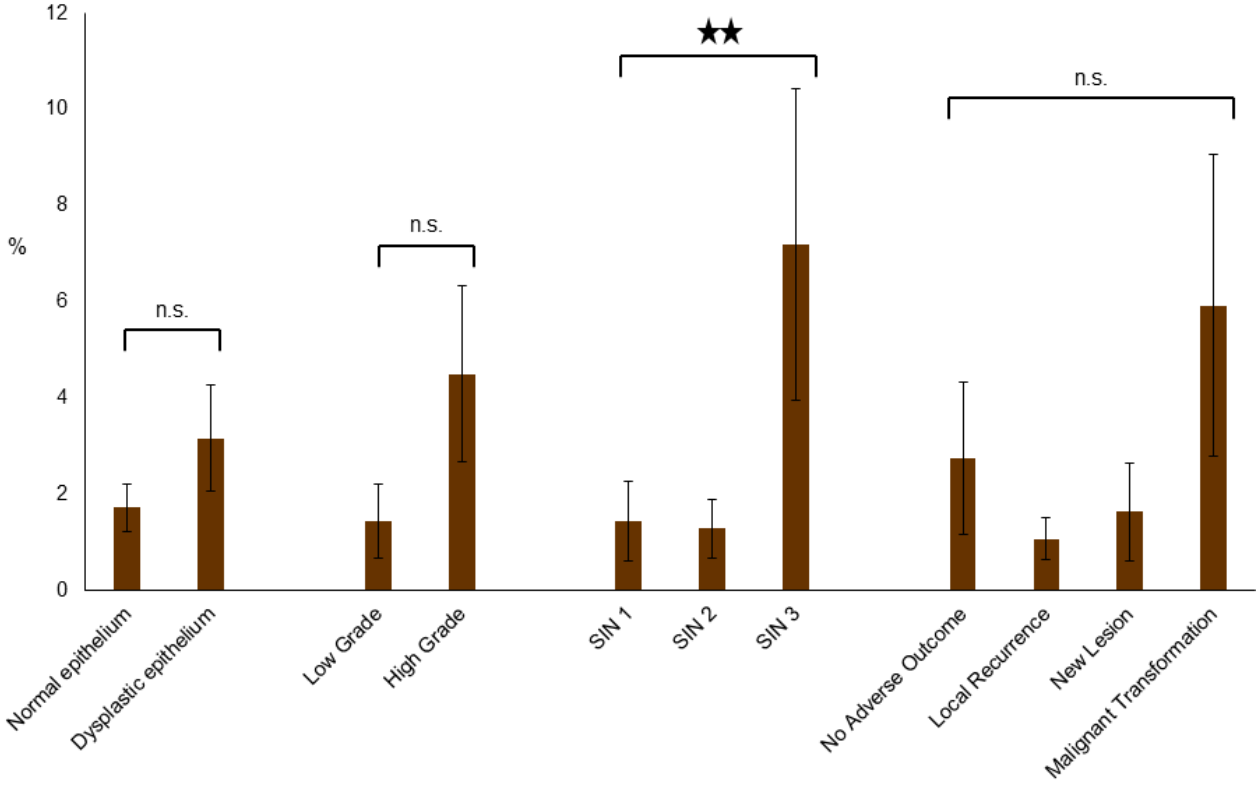


Figure 4-2 Bare chart comparing the mean SOX2 3+PN values of normal epithelium and oral potentially malignant disorders stratified according to grade of epithelial dysplasia and clinical outcome

The mean SOX2 3+PN values of normal and dysplastic epithelium did not differ significantly at $p<0.05$. Cases of OPMD with SIN 3 had a significantly higher mean 3+PN value than cases with SIN 1 ($p<0.01$). The mean SOX2 3+PN values did not differ significantly when OPMD were stratified according to clinical outcome.

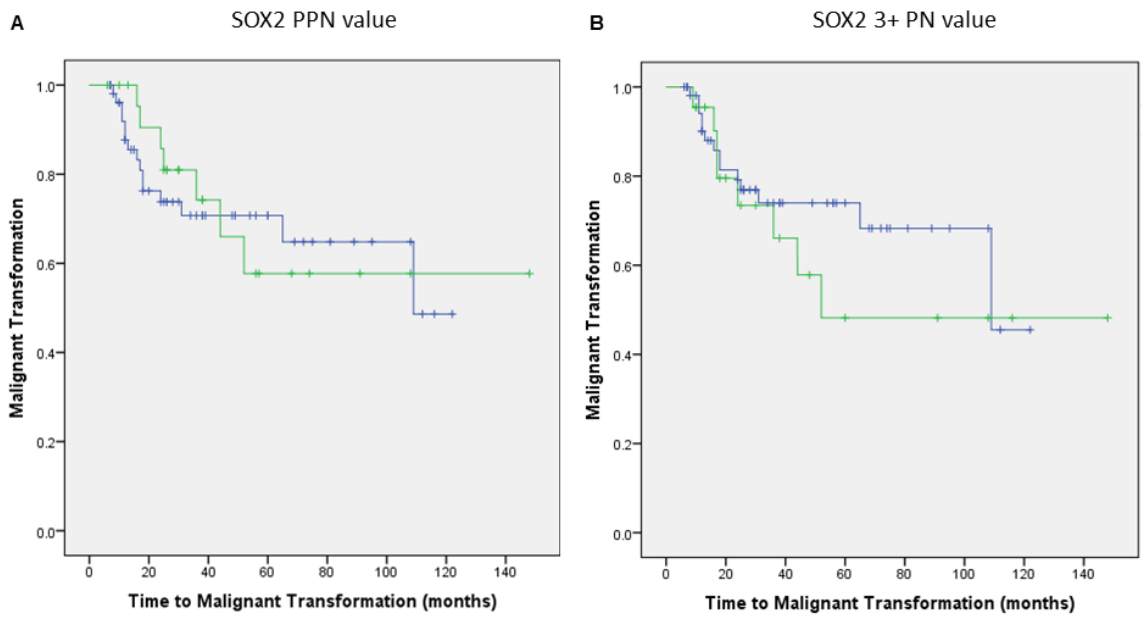


Figure 4-3 Kaplan-Meier time-to-event analysis of malignant transformation in OPMD grouped according to SOX2 PPN and 3+PN values

Colour index: Blue line - low SOX2 protein expression (i.e. PPN/3+PN value lower than the mean for the normal mucosa); green line - high SOX2 protein expression (i.e. PPN/3+PN value greater than mean for the normal mucosa.)

There was no correlation between SOX2 expression and risk of malignant transformation. A) Malignant transformation in OPMD grouped according to SOX2 PPN value ($p>0.05$, x^2 value – 0.16, 1 d.f.). B) Malignant transformation in OPMD grouped according to SOX2 3+PN value ($p>0.05$, x^2 value – 0.075, 1 d.f.).

4.3.4 PAX9 protein expression in oral potentially malignant disorders

Dysplastic epithelium was analysed in each of the 78 OMPD. The mean number of nuclei analysed per case was 1197 (range: 941 - 1736). The mean PPN was 24.6% (range: 0% - 72.8%). The mean 3+PN was 3.5% (range: 0% - 30.6%).

PAX9 protein expression was down-regulated in dysplastic epithelium relative to the normal epithelium (Figure 4-4). Dysplastic epithelium had PPN and 3+PN values lower than the corresponding mean values of normal epithelium in 75.6% and 84.6% of cases respectively. Pairwise comparisons showed that both the mean PPN and 3+PN values of dysplastic epithelium were significantly lower than the corresponding mean values of normal epithelium ($p<0.05$ Independent T-test and $p<0.01$ Mann-Whitney U-test respectively; 3+PN values shown in Figure 4-5).

4.3.5 PAX9 protein expression correlates with clinical outcome and detects cases of OPMD destined for malignant transformation

PAX9 protein expression showed a positive correlation with clinical outcome. PAX9 expression was significantly higher in the dysplastic epithelium of OPMD with adverse outcomes and, more specifically, the subset of cases that underwent malignant transformation. OPMD with adverse outcomes (i.e. local recurrence, new lesion formation, malignant transformation) had a significantly higher mean PPN value than OPMD with no adverse outcome ($p<0.01$, Independent T-test; Figure 4-5). Both the mean PPN and 3+PN values were significantly higher in OPMD that underwent malignant transformation relative to cases which did not (i.e. with no adverse outcome, local recurrence, or new lesion formation) at $p<0.0001$ (Independent T-test and Mann-Whitney U-test respectively).

In a comparison of OPMD stratified according to each of the four individual outcome groups, cases that underwent malignant transformation had significantly higher PPN values than cases with no adverse outcome ($p<0.0001$), local recurrence ($p<0.0001$), and new lesion formation ($p<0.01$, One-way ANOVA with Bonferroni correction; Figure 4-5). Corresponding 3+PN values showed the same trends. Small differences in the PPN and 3+PN values

of cases in the three non-transforming groups (i.e. no adverse outcome, local recurrence, new lesion formation) were not statistically significant.

OPMD were assigned to a binary ‘high PAX9’ or ‘low PAX9’ category according to whether PAX9 expression was either above or below the mean PAX9 PPN for normal epithelium. Kaplan-Meier time-to-event analysis confirmed that there was a significant correlation between high PAX9 expression and malignant transformation (Figure 4-6B). High PAX9 expression had a stronger correlation with malignant transformation than either high-grade epithelial dysplasia used in isolation (Figure 4-6A) or in a combined category (i.e. high-grade epithelial dysplasia combined with high PAX9 expression, Figure 4-6C). Receiver-operator curve analysis confirmed that high PAX9 expression was a more accurate predictor of malignant transformation than either high-grade epithelial dysplasia or the combined category (Figure 4-7). High PAX9 expression also had the highest positive predictive value for detecting cases that were destined to undergo malignant transformation (Table 4-1).

PAX9 protein expression did not correlate with the histological grade of epithelial dysplasia. Using the binary classification system, the mean PPN value of low-grade epithelial dysplasia was significantly lower than the normal epithelium ($p<0.05$, Independent T-test, Figure 4-5). However, the difference between the mean PPN values of low and high-grade-epithelial dysplasia was not significant ($p>0.05$, Figure 4-5). PPN values did not differ significantly when cases were stratified according to the SIN classification. 3+PN values did not differ significantly in either classification [data not shown].

PAX9 protein expression did not differ significantly when stratified according to patients’ age or sex; alcohol/tobacco habits, or the mucosal subsite of the OPMD [data not shown].

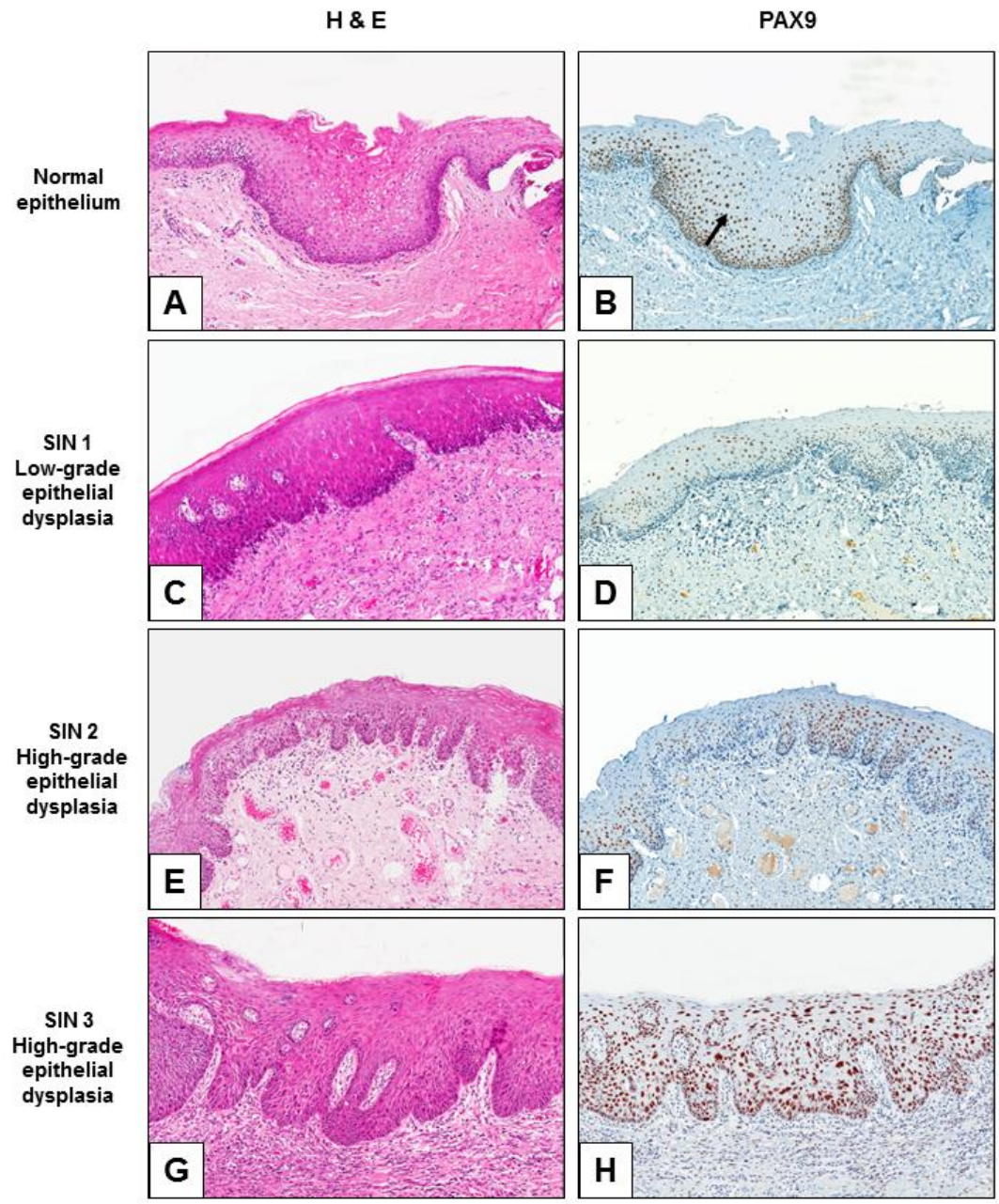


Figure 4-4 PAX9 protein expression in normal mucosa and oral potentially malignant disorders stratified according to grade of epithelial dysplasia

A) Normal oral mucosa. B) In normal epithelium, PAX9 is weakly expressed in the basal layer but becomes more intense as cells begin to exhibit squamous differentiation in the prickle layer (arrow). PAX9 expression is then progressively lost as cells migrate towards the surface.

In dysplastic epithelium, PAX9 expression was generally down-regulated relative to the normal epithelium. PAX9 expression was lowest in low-grade epithelial dysplasia (C, D). Although it was relatively increased in high-grade epithelial dysplasia, this trend was weak and not statistically significant (see Figure 4-5). This was due to variability in the PAX9 expression among cases with high-grade epithelial dysplasia. In the first example of high-grade epithelial dysplasia (E, F), PAX9 expression is heterogeneous but generally down-regulated relative to the normal epithelium. By contrast, the second example of high-grade epithelial dysplasia (G, H) shows increased expression of PAX9 relative to normal epithelium. OPMD with increased PAX9 expression relative to the normal epithelium were more likely to transform to OSCC than cases with low PAX9 expression (see Figure 4-5).

Images taken at x100 magnification.

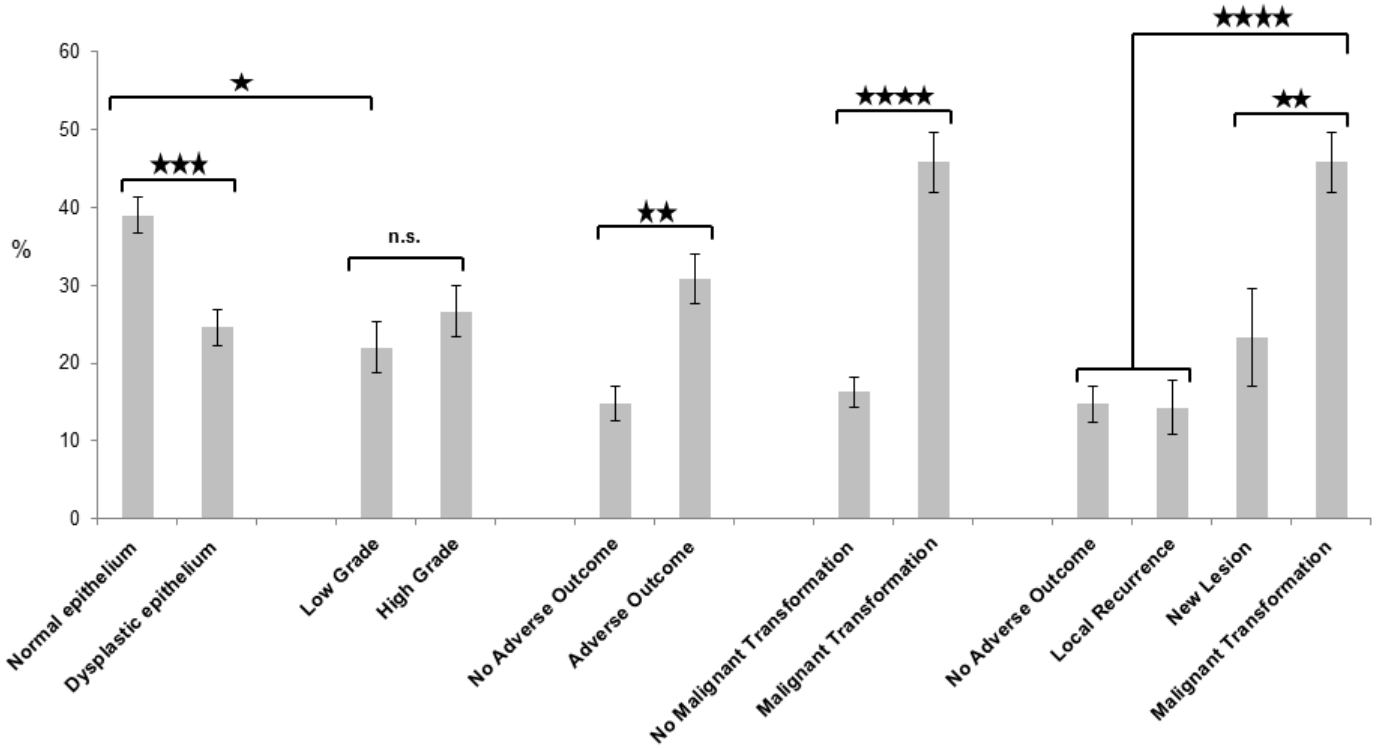


Figure 4-5 Bar chart comparing the mean PAX9 PPN values of normal epithelium and oral potentially malignant disorders stratified according to grade of epithelial dysplasia (binary classification) and clinical outcome

PAX9 expression was significantly down-regulated in dysplastic epithelium relative to normal epithelium ($p<0.001$). Cases with low-grade epithelial dysplasia had a significantly lower mean PPN value than the normal epithelium ($p<0.05$). The mean PPN values of low-grade and high-grade epithelium did not differ significantly, however ($p>0.05$). Together, OPMD with adverse outcomes had a significantly higher mean PPN value than OPMD with no adverse outcome ($p<0.01$). Specifically, OPMD that underwent malignant transformation had a significantly higher mean PPN value than cases which did not ($p<0.0001$). The same trend was detected when cases that did not transform were compared as a single group (i.e. with no adverse outcome, local recurrence, or new lesion formation combined) and in each of their individual outcome groups

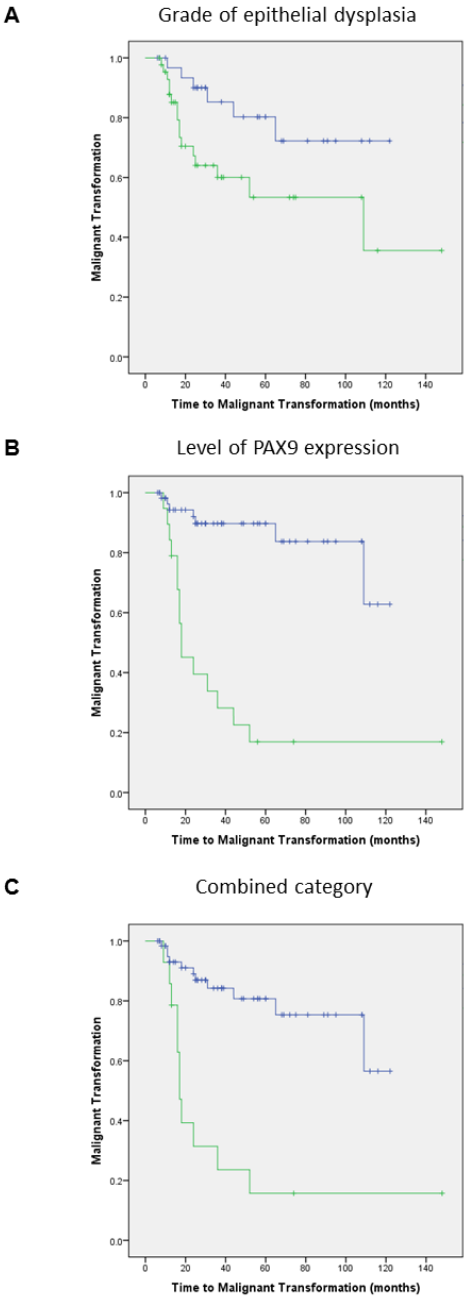
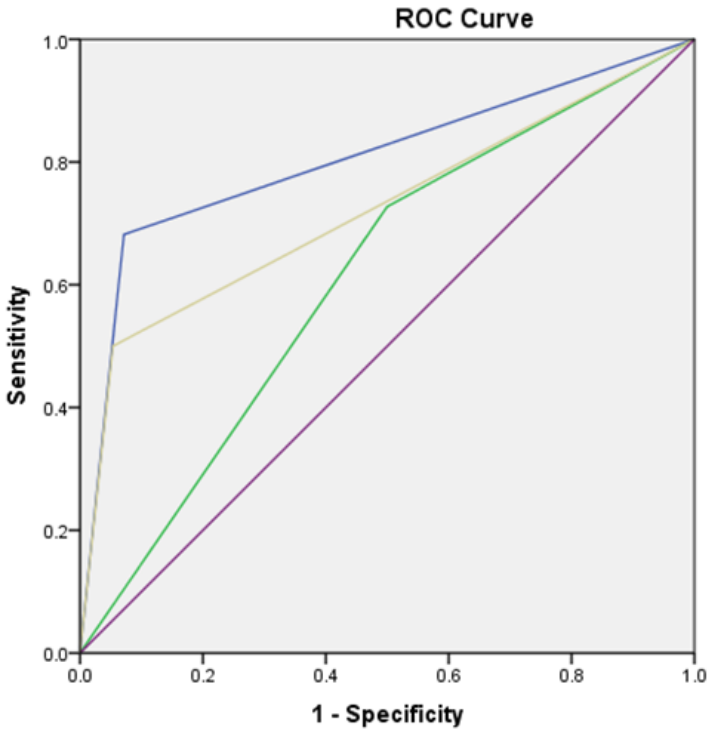


Figure 4-6 Kaplan-Meier time-to-event analysis showing malignant transformation in OPMD grouped according to grade of epithelial dysplasia, level of PAX9 protein expression, and a combined category

Colour index: Blue line - low grade epithelial dysplasia, low PAX9 protein expression, negative combined score; green line - high-grade epithelial dysplasia; high PAX9 protein expression; positive combined score.

A) There was a significant correlation between malignant transformation and high-grade epithelial dysplasia ($p<0.05$, χ^2 value - 4.974, 1 d.f.). B) There was also a significant correlation between malignant transformation and high PAX9 protein expression ($p<0.0001$, χ^2 value - 30.2, 1 d.f.) C) There was also a significant correlation between malignant transformation and a single category which combined cases with high-grade epithelial dysplasia and high PAX9 protein expression ($p<0.0001$, χ^2 value - 21.9, 1 d.f.).



Variable	Area beneath curve	Standard Error	Asymptotic Significance	Asymptotic 95% Confidence Interval	
				Lower bound	Upper bound
High-grade epithelial dysplasia	0.61	0.069	p>0.1	0.477	0.750
PAX9 over-expression	0.81	0.064	p<0.0001*	0.681	0.930
Combined PAX9 over-expression and high grade epithelial dysplasia	0.72	0.072	p<0.01*	0.582	0.864

Figure 4-7 Receiver-operator curve analysis comparing detection of malignant transformation by high PAX9 expression, high-grade epithelial dysplasia, and a combined category

Colour index: Blue line - PAX9 protein over-expression; green line - high-grade epithelial dysplasia; yellow line – combined PAX9 protein over-expression and high grade epithelial dysplasia; purple line – reference line.

The curve for high PAX9 expression is consistently to the left of the curve for high grade epithelial dysplasia. The area beneath the curve for high PAX9 expression is thus larger than the area for high grade epithelial dysplasia. Moreover, the asymptotic significance for high PAX9 expression is p<0.0001. The asymptotic value for high grade epithelial dysplasia is not significant (p>0.1). The combined category is intermediate between the two, being more accurately predictive of malignant transformation than high grade epithelial dysplasia but less so than high PAX9 expression in isolation.

Table 4-1 Detection of OPMD that underwent malignant transformation using high-grade epithelial dysplasia, high PAX9 protein expression, and a combined category

	High-grade epithelial dysplasia	High PAX9 expression	Combined category
Sensitivity	72.7	68.2	50
Specificity	50.0	92.9	94.6
Positive predictive value	36.4	79.0	78.6
Negative predictive value	82.4	88.1	82.8

4.3.6 SOX2 protein expression in OSCC transformed from OPMD cases

Malignant transformation occurred in 22 OPMD. Biopsy material was available for analysis in 21 of these cases. Three cases were among the early-stage OSCC group which is described in a following section, 4.3.8.

The mean number of nuclei analysed per case was 4610 (range: 699 – 11257). The mean PPN was 19.1 (range: 0 – 88.9%). The mean 3+PN was 2.9 (range: 0 – 26.3). The mean PPN of the transformed OSCC group was lower than the mean PPN of normal epithelium (32.5%). By contrast, the mean 3+PN of the transformed OSCC group was higher than the normal epithelium (1.2%). However, neither of these trends was statistically significant ($p>0.05$, Independent T-test (PPN) and Mann Whitney U-test (3+PN)).

The mean PPN and 3+PN values of the transformed OSCC group were lower than the corresponding mean values of dysplastic epithelium. This trend was consistent when the transformed group was compared with the dysplastic epithelium of all OPMD and the subset of cases that underwent malignant transformation. However, neither of these trends was statistically significant ($p>0.05$, Independent T-test (PPN) and Mann Whitney U-test (3+PN)).

The progression from OPMD to OSCC was not associated with a uniform change in SOX2 expression. There were cases in which SOX2 protein expression was consistent in both the OPMD and subsequent OSCC biopsy, either consistently low (Figure 4-8A, B) or high (Figure 4-8C, D). There were also cases in which the SOX2 protein expression was increased in the OSCC biopsy relative to the OPMD (Figure 4-8E, F) or decreased (Figure 4-8G, H).

The mean PPN and 3+PN values of the transformed OSCC group were lower than the corresponding values for early-stage OSCC outlined in section 4.3.8. However, neither of these trends was statistically significant ($p>0.05$, Independent T-test (PPN) and Mann Whitney U-test (3+PN)).

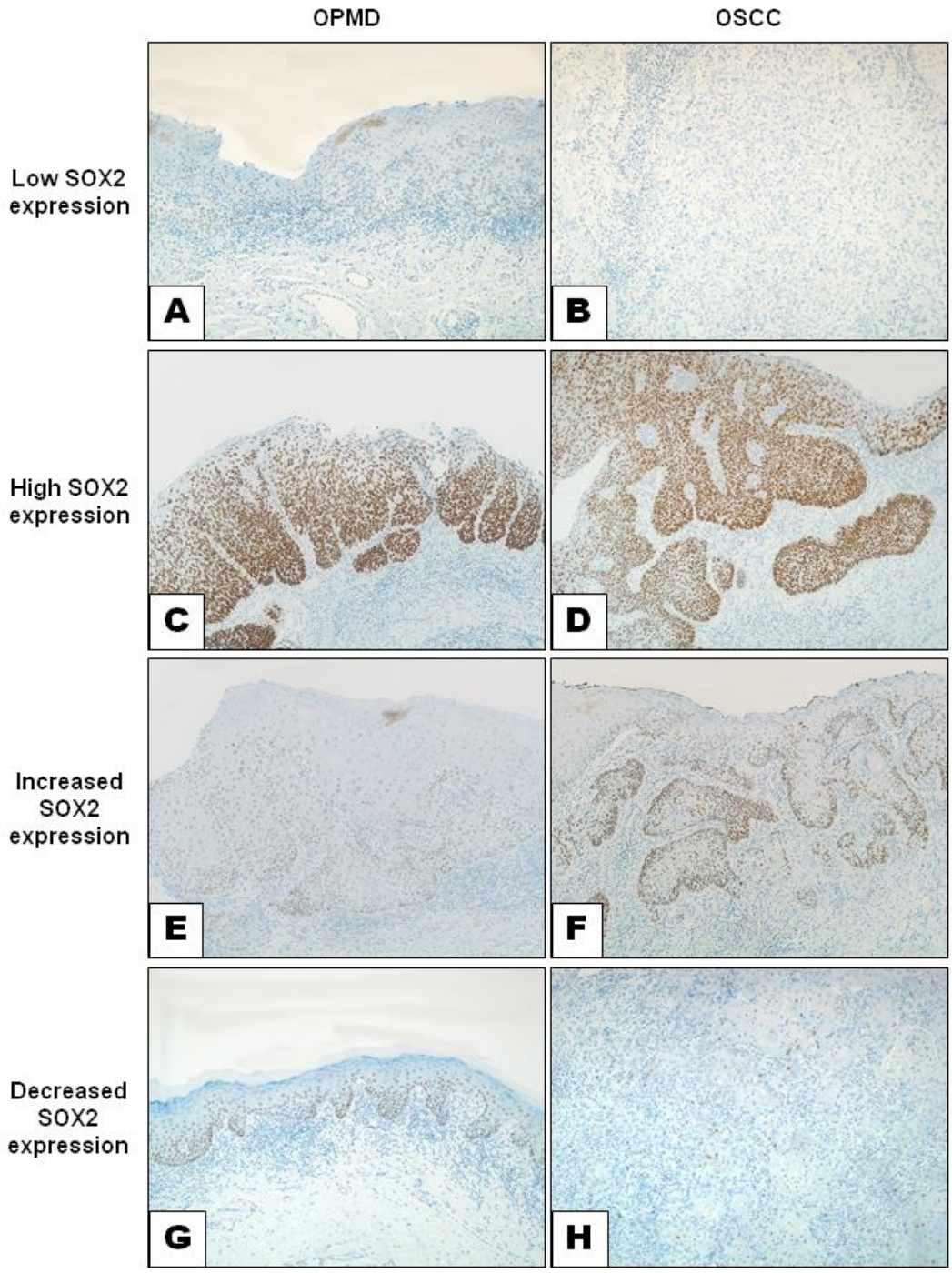


Figure 4-8 SOX2 expression in transforming oral potentially malignant disorders and subsequent oral squamous cell carcinomas

Four contrasting trends in SOX2 expression were detected. In some cases, SOX2 expression was consistently low in both the transforming OPMD (A) and the subsequent OSCC (B). In other cases, SOX2 expression was consistently high in both OPMD (C) and OSCC (D). Interestingly, the strongly SOX2-positive cells shown in C and D were characterised by a basaloid cytomorphology in both the OPMD (C) and subsequent OSCC (D). By contrast, there were also cases in which malignant transformation was associated with either an increase in SOX2 expression (E compared to F) or a decrease in SOX2 expression (G compared to H).

Images taken at x100 magnification.

4.3.7 PAX9 protein expression in OSCC transformed from OPMD

The mean number of nuclei analysed per case was 7558 (range: 1603 – 22170). The mean PPN was 46.6% (range: 0.02 – 90.1%). The mean 3+PN was 8.3% (range: 0 – 40.3%).

The mean PPN and 3+PN values of the transformed OSCC group were higher than the corresponding values for the normal epithelium (39.0% and 6.8% respectively). However, neither of these trends was statistically significant ($p>0.05$, Independent T-test and Mann-Whitney U-test; Figure 4-10).

The mean PPN and 3+PN values of the transformed OSCC group were significantly higher than the corresponding values for dysplastic epithelium. This trend was consistent when the transformed group was compared with the dysplastic epithelium of all OPMD and the subset of OPMD cases that underwent malignant transformation. When the transformed OSCC group was compared with the dysplastic epithelium of all OPMD, the difference in both PPN and 3+PN values were significant at $p<0.0001$ (Independent T-test and Mann-Whitney U-test; Figure 4-10). When the transformed OSCC group was compared with the dysplastic epithelium of the subset of OPMD that underwent malignant transformation the difference in both PPN and 3+PN was significant at $p<0.01$ (Mann Whitney U-tests).

The mean PPN and 3+PN values of the transformed OSCC group were significantly higher than the corresponding values for early-stage OSCC ($p<0.0001$, Independent T-test and Mann Whitney U-test; Figure 4-10). In the transformed OSCC group, over-expression of PAX9 protein was more evenly distributed in the invasive component compared to the early-stage OSCC group, in which PAX9 expression was heterogeneous (Figure 4-11).

4.3.8 SOX2 protein expression in early-stage OSCC

Areas of OSCC were analysed in each of the 92 cases. The mean number of nuclei analysed per case was 6962 (range: 578 - 25806).

SOX2 had a heterogeneous expression profile in OSCC. The mean PPN was 30.6% (range: 0% - 93.9%). The mean PPN value of OSCC was therefore similar to that of the normal epithelium (32.5%). Although a slight majority of early-stage OSCC cases (62.0%) had PPN values lower than the mean for normal epithelium, a binary comparison confirmed that the difference between the mean PPN values was not statistically significant ($p>0.5$, Independent T-test).

Several cases showed foci of high SOX2 expression in a background of tumour cells in which SOX2 expression was variably down-regulated relative to the normal epithelium. There were also subsets of cases that showed either generalised over-expression or down-regulation of SOX2 (Figure 4-9). The mean 3+PN was 8.9% (range: 0% - 64.7%). The mean 3+PN value of early-stage OSCC was therefore higher than the mean 3+PC of the normal epithelium (0.64%). This difference was significant at $p<0.01$ (Mann-Whitney U-test). Just over half of early-stage OSCC cases (51.1%) had 3+PN values higher than the mean 3+PN of the normal epithelium.

SOX2 expression in early-stage OSCC was increased relative to the dysplastic epithelium. A pairwise comparison showed that the mean 3+PN value for OSCC was significantly higher than the value for dysplastic epithelium ($p<0.001$, Mann-Whitney U-test). However, although the PPN value for OSCC was also higher than the PPN value for dysplastic epithelium, the difference was not statistically significant ($p>0.05$, Independent T-test).

4.3.9 SOX2 protein expression correlates with clinical outcome and differentiation of early-stage OSCC

Early-stage OSCC with adverse outcomes (i.e. local recurrence, cervical lymph node metastasis, formation of second primary tumours/pulmonary metastasis, death from disease) had lower levels of SOX2 protein expression than cases with no adverse outcome (i.e. cases which were either alive and disease-free, or who had had died free from disease). Although both the PPN and 3+PN values were lower in OSCC with adverse outcomes the difference was only statistically significant when comparing PPN values ($p<0.05$, Mann-Whitney U-test).

There was an inverse relationship between SOX2 expression and histological differentiation (Figure 4-9). Well differentiated OSCC had the highest levels of SOX2 expression, with the highest mean PPN and 3+PN values. SOX2 expression was relatively decreased in moderately differentiated OSCC and lowest in the subset of poorly differentiated OSCC. Poorly differentiated OSCC had significantly lower PPN and 3+PN values than moderately differentiated OSCC ($p<0.01$, Mann-Whitney U-tests, Bonferroni correction).

The subset of OSCC at pStage II had significantly lower PPN and 3+PN values relative to those at pStage I. The differences between both values were significant at $p<0.05$ (Mann-Whitney U-test).

SOX2 protein expression did not differ significantly according to patients' age, sex, or the mucosal subsite of the OSCC [data not shown].

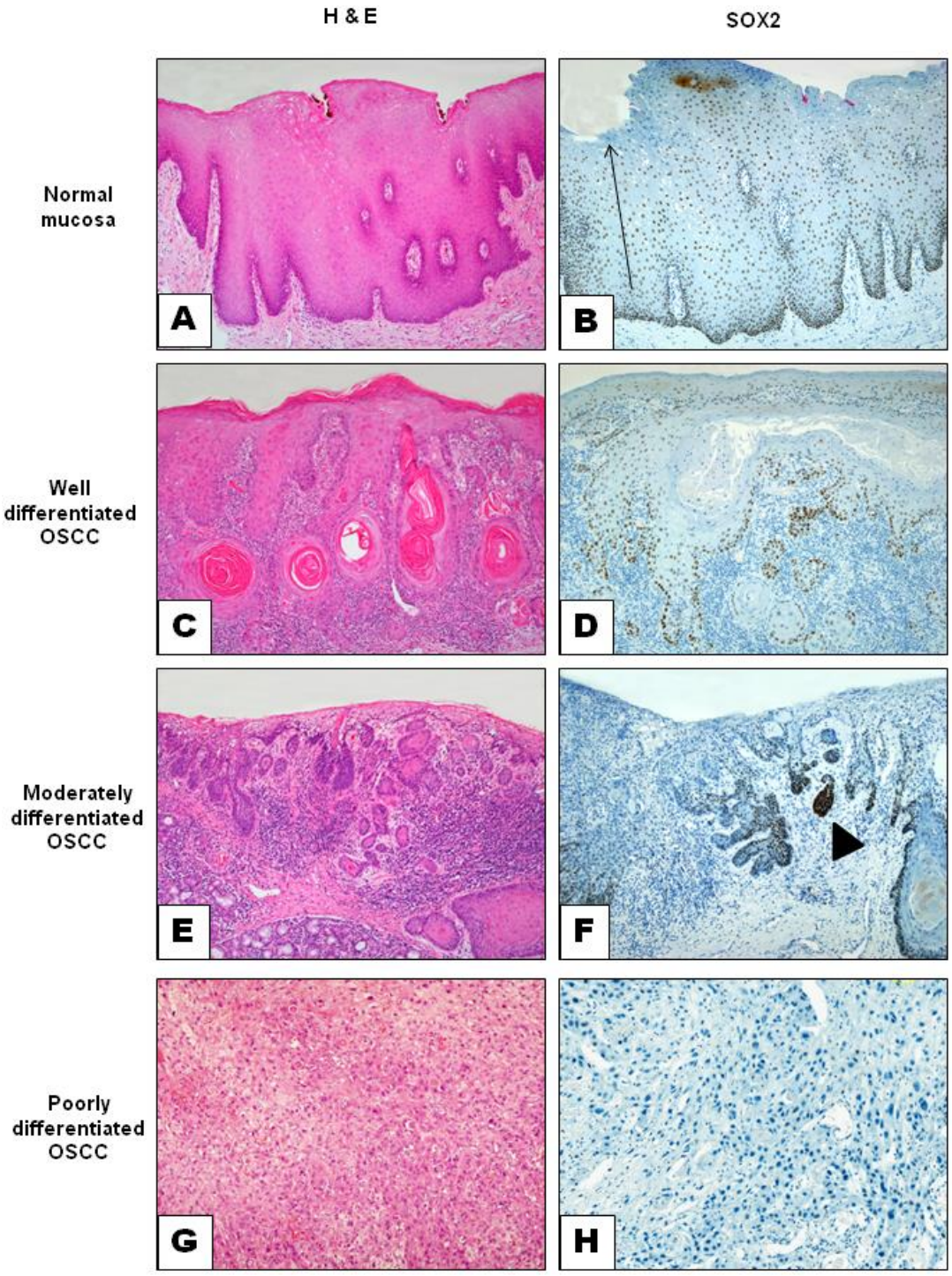


Figure 4-9 SOX2 expression in normal epithelium and early-stage oral squamous cell carcinoma stratified according to histopathological differentiation

A) Normal oral mucosa. B) In normal squamous epithelium, SOX2 expression is strongest in the basal layer, but becomes weaker as keratinocytes differentiate and progress towards the surface (arrow). SOX2 expression was highest in well differentiated OSCC (C, D), where it recapitulated the pattern of SOX2 expression in normal epithelium. Variable expression is seen in moderately differentiated OSCC (E, F) with foci of strongly SOX2-postive cells (F, arrowhead). SOX2 expression was lowest in poorly differentiated OSCC, with complete loss of SOX2 expression seen in this example (G, H).

Images taken at x100 magnification.

4.3.10 PAX9 protein expression in early-stage OSCC

Areas of OSCC were analysed in each of the 92 cases. The mean number of nuclei analysed per case was 6116 (range: 434 - 15243). The mean PPN was 11.9% (range: 0% - 78.9%). The mean 3+PN was 0.7% (range: 0% - 9.4%).

PAX9 protein expression was heterogeneous but generally down-regulated in the majority of OSCC compared to the normal epithelium (Figure 4-11). Early-stage OSCC had lower PPN and 3+PN than the corresponding mean values of normal epithelium in 94.6% and 97.8% of cases respectively. Pairwise comparisons of the PPN and 3+PN values of OSCC and normal epithelium were both significant ($p<0.0001$, Independent T-test and Mann-Whitney U-test; Figure 4-10).

PAX9 protein expression was also down-regulated in early-stage OSCC relative to the dysplastic epithelium. A pairwise comparison of the mean PPN value of OSCC and dysplastic epithelium was significant ($p<0.0001$, Independent T-test; Figure 4-10). However, mean 3+PN values did not differ significantly ($p>0.05$, Mann-Whitney U-test).

4.3.11 Correlation of PAX9 protein expression profile with clinico-pathological features and clinical outcomes of early-stage OSCC

Early-stage OSCC with adverse clinical outcomes (i.e. local recurrence, cervical lymph node metastasis, formation of second primary tumour, formation of pulmonary metastases, death from disease) had significantly lower PAX9 PPN values than early-stage OSCC with no adverse outcome ($p<0.01$, Independent T-test). A similar trend was identified when comparing 3+PN values, but the difference was not statistically significant ($p>0.05$, Mann-Whitney U-test).

OSCC had significantly lower PAX9 expression than normal epithelium irrespective of histological grade of differentiation. This trend was the same for both the PPN and 3+PN values at $p<0.0001$. Small differences in the PAX9 expression levels of each grade were not significant ($p>0.05$, One-way ANOVA and Kruskal Wallis tests with Bonferroni correction).

PAX9 expression was lower in OSCC at pStage II than OSCC at pStage I. Both the mean PPN and 3+PN values for pStage II OSCC were significantly lower

than the corresponding values for pStage I OSCC ($p<0.05$, Mann-Whitney U-test).

Neither the mean PPN nor the mean 3+PN value differed significantly according to patients' age, sex, or the mucosal subsite of the OSCC [data not shown].

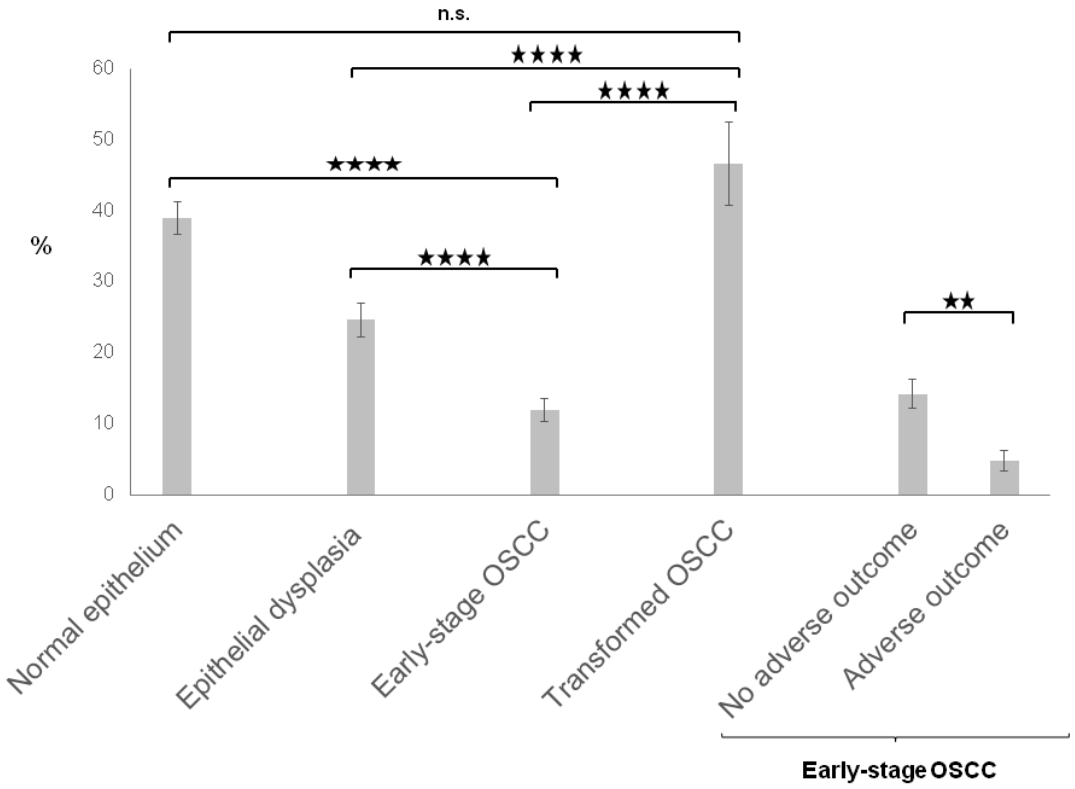


Figure 4-10 Bar chart comparing PAX9 PPN values of normal epithelium, dysplastic epithelium, OSCC arising from OPMD (mixed stages), and early-stage OSCC

Early-stage OSCC had significantly lower PPN values than either normal or dysplastic epithelium ($p<0.0001$). By contrast, the group of OSCC that arose from OPMD showed significantly higher PPN values than dysplastic epithelium and early-stage OSCC ($p<0.0001$). Early-stage OSCC with adverse outcomes had significantly lower PPN values than OSCC with no adverse outcome ($p<0.01$).

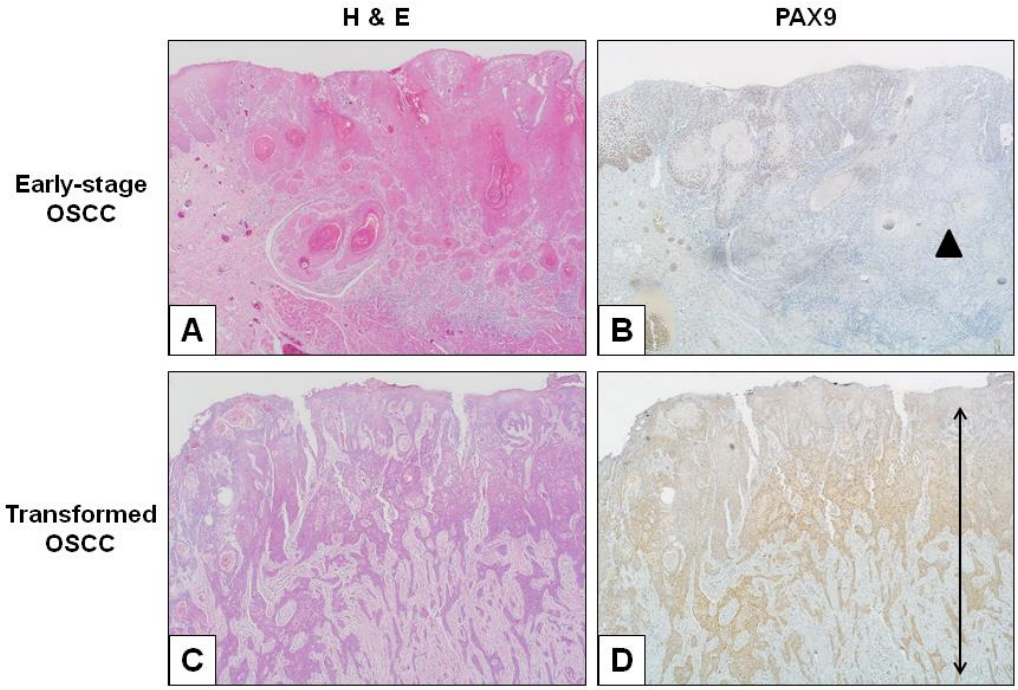


Figure 4-11 PAX9 expression in early-stage OSCC and OSCC arising from transformed OPMD

(A, B) Moderately differentiated early-stage OSCC. PAX9 expression is generally down-regulated relative to the normal epithelium at the surface, particularly at the invasive front (arrowhead). (C, D) By contrast, in a moderately differentiated OSCC arising from a transformed OPMD, there is strong expression of PAX9 throughout the full thickness of the tumour.

Images taken at x50 magnification.

4.4 Discussion

SOX2 and PAX9 are novel biomarkers in oral potentially malignant disorders (OPMD) and early-stage oral squamous cell carcinoma (OSCC). Both are transcription factors with vital roles in the regulation of development and differentiation. SOX2 is critical to the maintenance and proliferation of stem cells, functions which, if dysregulated, may be significant in tumorigenesis (Ellis et al., 2004; Masui et al., 2007; Hanahan and Weinberg, 2011). *SOX2* is a potential oncogene in squamous cell carcinoma (SCC) of the lung. There is emerging evidence that SOX2 has a similar role in oral carcinogenesis (Freier et al., 2010; Qiao et al., 2013). By contrast, PAX9 is believed to regulate key aspects of squamous epithelial differentiation (Jonker et al., 2004). Evidence from studies of oesophageal SCC suggests that *PAX9* functions as a tumour-suppressor gene (Gerber et al., 2002). This contention is supported by *in vitro* studies of PAX9 expression in lung cancer cell lines (Harris et al., 2011). However, to date there no studies of expression or potential role of *PAX9* in oral carcinogenesis.

4.4.1 SOX2 protein expression is heterogeneous but generally down-regulated in OPMD

Our data show that SOX2 has a heterogeneous expression profile in OPMD and is down-regulated in the majority of cases. To date, the profile of SOX2 expression in OPMD has been described in an isolated study by Qiao et al (2013). SOX2-positivity was detected in 90% of OPMD. Furthermore, co-expression of SOX2 and OCT4 - a feature not detected in the normal epithelium - was identified in 60% of OPMD. Similar trends were also demonstrated in an animal model (rat). However, the study was limited by a relatively small sample size (n = 20 human OPMD) and semi-quantitative scoring of stained sections. Cases were categorised either as SOX2-positive or SOX2-negative, but the intensity of staining was not described. Moreover, the authors did not document SOX2 expression in the normal human epithelium and were thus unable to comment on the relative SOX2 expression levels of OPMD (Qiao et al., 2013). It is therefore difficult to compare our data accurately with the results of this study.

4.4.2 SOX2 protein expression is up-regulated in early-stage OSCC

Our data show that SOX2 has a heterogeneous expression profile in OSCC. The proportions of SOX2-positive nuclei in OCCC and normal epithelium were similar. However, the proportion of strongly SOX2-positive nuclei was significantly higher in OSCC relative to normal epithelium. Up-regulation of SOX2 in OSCC has been reported in several studies by immunohistochemistry (IHC) (Freier et al., 2010; Huang et al., 2014), proteomic analysis (Misuno et al., 2013) and *in situ* hybridisation (ISH) (Freier et al., 2010; Kokalj Vokač et al., 2014). Up-regulation of SOX2 has also been described in squamous cell carcinomas of the lung, oesophagus, cervix, and penis (Hussenet et al., 2010; Lu et al., 2010; Maier et al., 2011).

Freier et al (2010) was the first to describe SOX2 over-expression in OSCC. However, the authors did not report SOX2 over-expression to be a common feature of OSCC: high SOX2 expression was only identified in a small subset of cases (18.1%). By contrast, our results show that the proportion of strongly SOX2-positive cells was increased in OSCC relative to the normal epithelium in over half of cases. The methodology of the previous study differed in its use of a polyclonal SOX2 antibody and semi-quantitative system for scoring stained sections. Moreover, tissue microarrays (TMA) were used rather than whole sections. The heterogeneous pattern of SOX2 expression detected in the present study is likely to compromise a TMA-based approach. Interestingly, ISH data were more sensitive, detecting SOX2 gene copy number amplification in more than half of cases (Freier et al., 2010).

Three studies document a correlation between high SOX2 expression and poor clinical outcome in OSCC, particularly metastasis to cervical lymph nodes (Du et al., 2011; Michifuri et al., 2012; Huang et al., 2014). The current study identified the opposite trend: OSCC with adverse outcomes had significantly lower expression of SOX2 relative to cases with no adverse outcome. The reasons for this apparent discrepancy are unclear, but there are several possible explanations:

- Whilst the present study was confined to early-stage OSCC, the previous studies also encompassed late-stage OSCC. It is conceivable

that SOX2 expression varies between early-stage and late-stage OSCC. Our study was not designed to address this possibility. As only a small subset of cases proceeded to neck dissection, there were insufficient numbers to detect a correlation between SOX2 expression and lymph node metastases

- The previous studies used a range of manual staining methods, by contrast to the automated Ventana platform used in the present study. These manual methods may have been subject to a range of technical variations between runs
- Each of the previous studies used a different SOX2 antibody. Polyclonal antibodies were used in two studies (Michifuri et al., 2012; Huang et al., 2014). These antibodies are more likely to cross react with other SOX family proteins and result in non-specific staining than the monoclonal antibody used in the present study
- The previous studies used semi-quantitative scoring as opposed to digital image analysis. In one case, only digitised images of selected fields - rather than whole sections - were available to the scoring pathologist (Huang et al., 2014)
- In the present study, 26 cases had adverse clinical outcomes. Our data also showed that SOX2 expression was lower in poorly differentiated OSCC compared with well- and moderately differentiated OSCC, and in pStage II OSCC compared with pStage I OSCC. However, several of these groups contained relatively low numbers (poorly differentiated OSCC – 9 cases; pStage II – 17 cases). Non-parametric tests were used for these comparisons, but it is conceivable that comparisons of groups containing higher/similar numbers might generate contrasting trends.

On balance, our finding that low SOX2 expression correlates with adverse outcomes in OSCC needs to be interpreted cautiously, as it is in contrast to several published studies. Each of these has potential methodological limitations relative to the present study. However, they include a greater range of clinical stages and are thus more suitably designed to detect a relationship between SOX2 expression and clinical outcome in OSCC.

Interestingly, heterogeneous SOX2 expression was described by Michifuri et al (2012) in 80 mixed-stage OSCC. The authors grouped cases into one of two categories according to their pattern of SOX2 expression, either 1) peripheral or 2) diffuse. The present study includes cases that are consistent with these patterns, peripheral (e.g. in Figure 4-9D, well differentiated OSCC) and diffuse (e.g. in Figure 4-8D, uniformly high SOX2 expression). However, this classification system does not account for cases with generalised down-regulation/loss of SOX2 expression or the foci of strongly SOX2-postive nuclei in cases with variably-intense staining seen in the present study. Interestingly, the authors demonstrated a correlation between the diffuse pattern of SOX2 staining and metastasis to cervical lymph nodes. It is not possible to comment on the correlation between staining pattern/clinical outcome in the present study due to the small number of cases with uniform staining patterns.

SOX2 expression was significantly higher in OSCC relative to OPMD in a comparison of the two groups combined. In isolation, this finding might suggest that up-regulation of SOX2 is associated with malignant transformation in OPMD. However, analysis of SOX2 expression in the subset of transforming OPMD and their subsequent OSCC did not support this hypothesis, but only confirmed the highly variable levels of SOX2 expression at each stage of oral carcinogenesis. The striking heterogeneity of SOX2 expression identified across all groups in the present study suggests that SOX2 has limited utility as a diagnostic or prognostic biomarker in OPMD and OSCC.

Such consistent variation does however raise several hypotheses as to the biological significance of SOX2 in oral carcinogenesis. Firstly, *SOX2* may have multiple stage-dependent roles in oral carcinogenesis, with a spectrum of potential oncogenic and tumour-suppressor functions. Secondly, while some OPMD/OSCC may be SOX2-dependent, there may be subsets of cases that have acquired alternative mutations and adaptive strategies. Strong SOX2 expression is associated with a basaloid cytomorphology and several cases show foci of malignant cells with intense SOX2 expression. Together, these findings suggest a third hypothesis: SOX2 may be significant in the maintenance and functioning of cancer stem cells. This hypothesis is supported by *in vitro* analysis that has demonstrated up-regulation of SOX2 and OCT4 in

oral cancer stem cells (CSC) (Lim et al., 2011; Bourguignon et al., 2012; Misuno et al., 2013). It is also consistent with the contribution of SOX2 to stem cell functioning in normal squamous epithelium (Okubo et al., 2009). CSC are accountable for resistance to therapy, local recurrence, and metastatic spread in head and neck squamous cell carcinoma, and are therefore important therapeutic targets (Routray and Mohanty, 2014). Further work is required to delineate the contribution of SOX2 to the maintenance, survival, and proliferation of CSC. However, the evidence to date suggests that SOX2 may be an important therapeutic target in a subset of OSCC and OPMD. It is hoped that future *in vitro* studies utilising the cell lines generated in this project will be useful in testing these hypotheses.

4.4.3 PAX9 protein expression in oral carcinogenesis: two contrasting trends

Our data show that PAX9 has a heterogeneous expression in oral potentially malignant disorders (OPMD) and oral squamous cell carcinoma (OSCC). Overall, there are two contrasting trends:

1) PAX9 expression is progressively down-regulated in OPMD and OSCC
2) PAX9 expression is up-regulated in transforming OPMD and their subsequent OSCC.

4.4.4 PAX9 expression is progressively down-regulated in OPMD and early-stage OSCC

This study demonstrated reduced PAX9 expression in OPMD relative to normal epithelium, and reduced PAX9 expression in OSCC relative to OPMD. This profile of progressively down-regulated PAX9 expression is consistent both with the role of Pax9 in the differentiation of mouse oral squamous epithelium (Jonker et al., 2004) and evidence from a study of oesophageal cancer that suggests *PAX9* acts as a tumour-suppressor gene (Gerber et al., 2002).

Pax9 is critical to regulating keratin transcription in mouse oral squamous epithelium. Specifically, Pax9 determines the relative proportions of ‘hard keratins’ (e.g. Krt1-5) and ‘soft keratins’ (e.g. Krt2-1, Krt2-17) in filiform papillae. The lingual epithelium of Pax9-deficient mice shows an abnormally increased

proportion of soft keratins (Jonker et al., 2004). Soft keratins Krt2-1 and Krt2-17 are over-expressed in dysplastic lesions of the oral squamous epithelium in humans (Bloor et al., 2003). It is postulated that alterations in the differentiation pathway of squamous epithelium in Pax9-deficient mice is analogous to those occurring in human epithelial dysplasia (Jonker et al., 2004). Our data confirm that PAX9 in is also down-regulated in the majority of human OPMD and OSCC. They provide some support for the contention that down-regulation of PAX9 could induce alterations in the keratin composition of dysplastic squamous epithelium. Further analysis of the cytokeratin profile of our samples and correlation with PAX9 expression would be required to test this hypothesis. However, this overall trend does suggest that *PAX9* has a tumour-suppressor function.

Down-regulation of PAX9 expression has been documented in a study of squamous cell carcinoma (SCC) and epithelial dysplasias of the human oesophagus by Gerber et al (2002). The study demonstrated either loss or significantly reduced PAX9 expression in the majority of oesophageal SCC and dysplastic epithelial lesions. An inverse relationship between PAX9 expression and clinical course was also detected. A decrease in the proportion of PAX9-positive cells correlated with increasingly malignant behaviour of lesions. This inverse trend was also identified in the current study: OSCC with adverse outcomes had significantly lower PAX9 expression compared to cases with no adverse outcome. However, the opposite trend was identified in OPMD.

Gerber et al (2002) analysed fewer cases compared to the present study (SCC – 36 cases; epithelial dysplasia – 35 cases). The study was further constrained by its use of semi-quantitative scoring. PAX9 positivity was recorded but not the intensity of staining. Importantly, however, PAX9 immunohistochemical staining was performed using the same monoclonal antibody as in the present study, which shows nuclear expression in normal squamous epithelium of the oesophagus.

On balance, the present study confirms the overall trend of progressively-reduced PAX9 expression that has been documented in oesophageal cancer. Our data also confirm that PAX9 down-regulation is associated with adverse outcomes in OSCC. Together, these trends suggest that *PAX9* has a tumour-

suppressor function (Gerber et al., 2002). There is some evidence from *in vitro* studies to support this hypothesis. A study of lung cancer cell lines has demonstrated allelic loss at 14q13.3 and subsequent down-regulation of PAX9 and other transcription factors such as NKX2-8 in a subset of tumours (Harris et al., 2011).

4.4.5 PAX9 expression is up-regulated in transforming OPMD and their subsequent OSCC

Our results demonstrate PAX9 over-expression in OPMD that underwent malignant transformation; PAX9 over-expression was also conserved in the subsequent OSCC. This pattern is in contrast to the overall trend identified for OPMD and early-stage OSCC, and suggests two different hypotheses:

1) *PAX9* is a critical oncogene that is transiently up-regulated in OPMD immediately before, during, and shortly after transformation to OSCC
2) PAX9 is over-expressed in a subset of OPMD and OSCC, but is not a uniform hallmark of oral carcinogenesis.

Our data show PAX9 over-expression is a more reliable predictor of malignant transformation than histological grade of epithelial dysplasia. This finding suggests that PAX9 over-expression is an important genetic/molecular change during transition from *in situ* epithelial dysplasia to invasive OSCC. Given the wider context of progressively reduced PAX9 expression in OPMD and OSCC, it is conceivable that PAX9 is dynamically expressed during oral carcinogenesis. PAX9 may be down-regulated in early/low-grade lesions, with consequent loss of its tumour-suppressor function. As lesions progress to OSCC, PAX9 is over-expressed due to its critical oncogenic function. In established lesions, this oncogenic function may be redundant due to the acquisition of further oncogenic alterations, particularly those which confer an advantageous metastatic capacity. PAX9 expression is thus gradually lost as OSCC progresses beyond the initial transformation stage.

Evidence from *in vitro* studies supports the view that PAX9 performs oncogenic functions. *PAX9* interacts with *c-myb*, a proto-oncogene that enhances the survival of SCC cell lines and is implicated in the pathogenesis of a range of cancers (Lee et al., 2008). Although Harris et al (2011) identified allelic loss of

chromosomal region 14q13.3 in a subset of lung cancer cell lines, the majority actually of lung cancers show amplification of this region. *PAX9* and other transcription factor genes *TTF-1* and *NKX2-8* have been shown to contribute to the survival of lung cancer cell lines. *PAX9* may therefore act as an oncogene in lung cancer (Kendall et al., 2007).

The hypothesis that PAX9 is dynamically expressed in oral carcinogenesis must be considered cautiously, however. The present study has a cross-sectional rather than longitudinal design. This precludes accurate temporal description of PAX9 expression in individual cases. Comparisons are further hampered by variation in the numbers of OPMD biopsies for each case and time intervals between the index OPMD biopsy and subsequent OSCC biopsy. Previous OPMD biopsies were not available for the majority of early-stage OSCC cases in our study. It is open to conjecture whether these OSCC, in which PAX9 expression was significantly lower than the OSCC arising from transformed OPMD, were preceded by OPMD with low or high levels of PAX9 expression. Animal studies – which are more amenable to uniform longitudinal studies of individual cases – may help to determine whether PAX9 expression is truly stage-dependent.

It is conceivable that the majority of early stage OSCC were preceded by OPMD with low PAX9 expression. This leads to a second hypothesis, that the high PAX9 expression characterises a subset of OPMD and OSCC but is not uniformly critical to malignant transformation. Interestingly, ‘normal’ (i.e. relatively high) levels of PAX9 expression were reported in oesophageal SCC and dysplasias by Gerber et al (2002). It is possible that these oesophageal lesions with high PAX9 expression were biopsied at a time when they were undergoing malignant transformation, which would support the previous hypothesis. However, it may also provide further evidence that OPMD and OSCC can be broadly separated into two groups according to their PAX9 expression profile.

Together, our data show that the overall trends in PAX9 expression in OPMD/OSCC are very similar to those documented in oesophageal SCC and squamous epithelial dysplasia. This suggests that *PAX9* has an important tumour-suppressor function. However, there is also strong evidence to suggest

that *PAX9* also has an oncogenic function. These functions may be stage-dependent. Alternatively, either one may predominate in individual cases. It is hoped that *in vitro* studies using the cell lines generated in this project will help to determine the precise contribution of PAX9 to oral carcinogenesis. Importantly however, our data show that PAX9 over-expression may help to identify OPMD at risk of progressing to OSCC.

4.5 Conclusion

SOX2 protein has a heterogeneous expression profile in both OPMD and early-stage OSCC, which may limit its potential as a biomarker. SOX2 expression is down-regulated in the majority of OPMD and does not correlate with clinical outcome. Conversely, SOX2 expression is generally up-regulated in early-stage OSCC. This is consistent with studies of SOX2 expression in squamous cell carcinoma of the lung and oesophagus. However, whereas previous studies have documented a correlation between high SOX2 expression and poor clinical outcome, we found that OSCC with adverse outcomes had significantly lower expression of SOX2 than cases with no adverse outcome. In both OPMD and OSCC, SOX2 over-expression was associated with foci of basaloid cells. These foci may harbour oral cancer stem cells. This supports the hypothesis that SOX2 is functionally significant in oral carcinogenesis and is therefore a potential chemotherapeutic target in OPMD/OSCC.

PAX9 expression was down-regulated in OPMD relative to normal epithelium. PAX9 expression in OSCC was reduced relative to both normal epithelium and OPMD. Furthermore, OSCC with adverse outcomes had significantly lower PAX9 expression compared to cases with no adverse outcome. These findings are concordant with a study of oesophageal squamous cell carcinoma, and support the hypothesis that *PAX9* acts as a tumour-suppressor gene. Interestingly, PAX9 was over-expressed in the subset of OPMD that underwent malignant transformation and in their subsequent OSCC. This apparently paradoxical finding suggests that *PAX9* may also have an oncogenic function in oral carcinogenesis.

Chapter 5. Morphological and Initial Molecular Characterisation of a Mouse Model of Oral Carcinogenesis

5.1 Introduction

5.1.1 Longitudinal studies of oral carcinogenesis

Longitudinal studies of biomarkers in human subjects facilitate the study of disease processes at the cellular and molecular level (Munoz and Gange, 1998). Patients with oral squamous cell carcinoma (OSCC) and oral potentially malignant disorders (OPMD) may have multiple diagnostic biopsies and surgical interventions during their management. Biomarkers may be analysed in biopsy tissue taken at various stages of OSCC formation, facilitating the longitudinal study of oral carcinogenesis. However, longitudinal studies are frequently constrained by patient drop out. Increasingly, surgical intervention is regarded as the preferred treatment modality for the management of OPMD. This raises ethical and potential medico-legal issues around the management of an OPMD by surveillance and repeated biopsy (Goodson and Thomson, 2010). Animal models are therefore comparatively advantageous for the longitudinal study of oral carcinogenesis, as they facilitate the generation and analysis of tissue from each stage of OSCC formation (Herzig and Christofori, 2002).

5.1.2 Mouse models of oral carcinogenesis

Mouse models of disease are widely used because the mouse genome has a history of being successfully manipulated (Brudno et al., 2004). Spontaneous OSCC formation in mice is rare (Thurman et al., 1997). OSCC formation may be induced chemically by the use of agents such as 4-nitroquinoline 1-oxide (4-NQO) (Vered et al., 2005). 4-NQO induces a spectrum of precancerous changes alongside OSCC formation. The genetic alterations induced by 4-NQO are similar to those implicated in human OSCC (Kanojia and Vaidya, 2006).

Studies using transgenic and knockout mice have made a significant contribution to our understanding of human disease. However, genetic modifications that result in OSCC formation in mice are rare (Mognetti et al., 2006). Treating a transgenic/knockout mouse strain with 4-NQO potentially

simulates the complex interactions between environmental/genetic factors that occur during human oral carcinogenesis (Wong, 2009).

5.1.3 Pax9 is a candidate tumour-suppressor gene in oral carcinogenesis

Data from our analysis of human tissue samples, combined with evidence from the available literature, suggest that *PAX9* has a tumour-suppressor function in oral carcinogenesis. *Pax9* is conditionally inactivated when $Pax9^{flox}$ transgenic mice are crossed with K14-Cre mice (Kist et al., 2007). Aged Pax9-deficient mice develop fissures and hyperplastic epithelial lesions [unpublished data]. Whilst these lesions do not show morphological features of epithelial dysplasia and do not progress to OSCC, it is conceivable that they harbour molecular changes that are shared with those occurring in oral carcinogenesis.

We hypothesise that *Pax9* is a tumour-suppressor gene and that exposure to 4-NQO treatment will induce more extensive pre-cancerous changes and more rapid progression to OSCC in Pax9-deficient mice than in controls. This chapter outlines the findings of a feasibility study, analogous to a phase I clinical trial, (Cancer Research UK, 2013), which is the first to assess the response Pax9-deficient mice to 4-NQO treatment.

5.2 Aims

The aims of this chapter are:

1. To determine the tolerable 4-NQO dose range for Pax9-deficient mice
2. To identify potential side effects/systemic toxicity that may be induced in Pax9-deficient mice by 4-NQO treatment
3. To characterise the morphological and molecular features of the oral mucosa of Pax9-deficient and control mice following 4-NQO treatment
4. To characterise the morphological and molecular features of the oral mucosa of aged Pax9-deficient mice and compare them with those of mice that have been exposed to 4-NQO.

5.3 Results

5.3.1 Characteristics of the group of mice treated with 4-NQO: sex, genotype, genetic background, and phenotype

Sixty-eight mice were treated with 4-NQO (*n* = 68). At the start of 4-NQO treatment, the mice had a mean age of 33.8 weeks (range: 9 - 57 weeks). The other characteristics of these mice are summarised in Table 5-1.

5.3.2 Weight of mice prior to and following 4-NQO-treatment

At the start of the experiment, the mice had a mean weight of 25.7 g (range: 16.5 – 42.3). The mean weight of control mice was 27.0 g (range: 18.5 – 42.3). Pax9-deficient mice generally weighed less than controls prior to 4-NQO treatment (mean 24.5; range: 16.5 - 33.7). However, the difference between the mean weights of Pax9-deficient mice and controls was not significant ($p>0.05$, Independent T-test).

All mice were weighed following 4-NQO treatment, immediately before autopsy. The mice had a mean weight of 31.4 g (range: 11.6 – 52.8). The mean weight of the control mice was 35.4 g (range: 20.7 – 52.8). Pax9-deficient mice weighed less than controls following 4-NQO treatment (mean 25.5, range: 11.6 – 46.6). This difference was significant at $p<0.0001$ (Independent T-test). Weight loss in Pax9-deficient mice was accompanied by a range of other features, which are described further in section 5.3.4.

Table 5-1 Characteristics of the group of mice treated with 4-NQO (*n* = 68)

Characteristic	Number (%)
Sex:	
Male	31 (45.6)
Female	37 (54.4)
Genotype	
Wild-type	32 (47.1)
Pax9 mutant:	36 (52.9)
Phenotype of Pax9 mutants*	
Complete knockout	19 (52.8)
Partial knockout	11 (30.6)
Normal phenotype	6 (16.7)
Strain:	
Black 6	35 (51.5)
FVB	17 (25)
Hybrid (Black6 x FVB)	16 (23.5)
Generation:	
F1	16 (23.5)
N2F16	3 (4.4)
N3	4 (5.9)
N3F1	29 (42.6)
N5F9	16 (23.5)

(*N.B. Complete knockout – complete absence of filiform papillae; partial knockout – some areas contain normal filiform papillae; normal – no loss of filiform papillae.)

5.3.3 Duration of 4-NQO treatment

The mean duration of 4-NQO treatment was 21.1 weeks (range: 2 – 28 weeks). The majority of both Pax9-deficient mice and controls tolerated 4-NQO treatment and completed the planned duration of 4-NQO treatment (intervals between 22 to 28 weeks). However:

- Four Pax9-deficient mice became sick soon after beginning 4-NQO treatment. They were sacrificed after just two weeks. Of these:
    - Each had a Black 6 genetic background
    - Three cases were treated with 50µg/ml 4-NQO
    - One case was treated with 20µg/ml 4-NQO
- Ten cases were sacrificed following <20 weeks of 4-NQO treatment:
    - Six of these were Pax9-deficient mice sacrificed on humane grounds due to a toxic systemic response to 4-NQO
    - Four were healthy control mice. These controls were sacrificed to provide a comparison of the oral mucosa of controls and Pax9-deficient mice prior to completion of the planned 4-NQO treatment
- Twelve Pax9-deficient mice tolerated 4-NQO treatment for more than 20 weeks. However, these mice deteriorated shortly before the end of treatment. They were sacrificed for humane reasons at intervals between 22 and 28 weeks.

The length of 4-NQO treatment and other characteristics of the mice that showed a toxic systemic response to 4-NQO treatment are summarised in Table 5-2. The characteristics of these mice are detailed in the next section.

5.3.4 Toxic systemic response to 4-NQO treatment

The majority of mice (66.2%) were healthy on completion of 4-NQO treatment, and showed no evidence of a toxic systemic response to 4-NQO. However, the tolerance of mice to 4-NQO treatment varied according to sex, genetic background, dose of 4-NQO, and genotype.

A total of 23 mice showed a toxic systemic response to 4-NQO treatment (Table 5-2). These mice were characterised by excessive saliva production, which caused drooling and matting of the fur around the chin and abdomen. The

majority of these mice were otherwise healthy. However, in a subset of nine mice excessive salivation was accompanied by:

- A generalised failure to thrive
- A reduced growth rate
- Rapid, extreme weight loss
- A ‘hunched’ position at rest.

All of these mice were too sick to continue the experiment and were sacrificed for humane reasons prior to completion of their planned 4-NQO treatment. At autopsy, abdominal examination revealed they had only one kidney.

Female mice were more likely to develop a toxic systemic response to 4-NQO treatment than male mice. 25.8% of females developed excessive salivation, compared with 16.2% of males. Similarly, 16.1% of females became sick, compared with 10.8% of males. However, cross tabulation showed that the association between sex and toxic systemic response was not statistically significant ($p>0.05$, Pearson x^2 value - 1.7, 2 d.f.).

The majority of the 23 mice that developed a toxic systemic response to 4-NQO treatment had a Black 6 genetic background (18 cases, 78.3%). Four FVB mice and only one hybrid developed a toxic response. A cross tabulation showed that there was a significant association between the Black 6 genetic background and the development of a toxic response ($p<0.05$, Pearson x^2 value - 11.6, 4 d.f.).

Thirty-seven mice were treated with a 10µg/ml concentration of 4-NQO. At the end of treatment, the majority of these mice (31 cases, 83.8%) were healthy. By contrast, less than half of the mice treated with 20µg/ml and 50µg/ml concentrations of 4-NQO were healthy. Of the twelve mice treated with 50µg/ml 4-NQO, four became sick and a further three cases developed excessive salivation. A cross tabulation showed that the association between 4-NQO dosage and development of a toxic systemic response to 4-NQO treatment was significant ($p<0.01$, Pearson x^2 value - 13.6, 4 d.f.).

The majority (90.6%) of control mice were healthy on completion of 4-NQO treatment. There were three control mice among those who showed excessive salivation. However, none of the controls became sick or had to be sacrificed

prematurely due to a toxic systemic response to 4-NQO treatment. By contrast, the majority (55.6%) of Pax9-deficient mice developed a toxic systemic response to 4-NQO treatment; less than half of Pax9-deficient were healthy prior to sacrifice. Each of the nine mice that became sick as a result of the 4-NQO treatment was Pax9-deficient. A cross tabulation confirmed that the association between the Pax9-deficient genotype and development of a toxic systemic response to 4-NQO treatment was significant ($p<0.0001$, Pearson x^2 value - 17.2, 2 d.f.).

Table 5-2 Summary of the characteristics of the mice that developed a toxic systemic response to 4-NQO treatment

Mouse number.	Sex	Genotype	Strain	4-NQO		Condition at sacrifice
				Dosage (µg/ml)	Duration (weeks)	
846	F	Mutant	Black6	50	14	Drooling
925	M	Mutant	Black6	50	14	Sick
926	M	Mutant	Black6	50	23	Drooling
948	M	Mutant	FVB	10	25	Drooling
960	M	Mutant	FVB	10	27	Drooling
965	F	Mutant	FVB	10	11	Sick
971	F	Control	FVB	10	27	Drooling
977	M	Mutant	Black6	50	2	Sick
978	M	Mutant	Black6	50	8	Drooling
980	F	Mutant	Black6	50	2	Sick
981	F	Mutant	Black6	50	2	Sick
983	M	Mutant	Black6	20	22	Sick
985	M	Mutant	Black6	20	23	Drooling
988	F	Mutant	Black6	20	23	Drooling
989	F	Mutant	Black6	20	22	Drooling
990	F	Mutant	Black6	20	22	Drooling
997	F	Mutant	Hybrid	20	28	Sick
1012	M	Mutant	Black6	20	2	Sick
1015	F	Control	Black6	20	23	Drooling
1016	F	Mutant	Black6	20	4	Drooling
1018	F	Mutant	Black6	20	4	Sick
1019	F	Control	Black6	20	22	Drooling
1030	M	Mutant	Black6	10	22	Drooling

5.3.5 Autopsy features of the lingual mucosa following 4-NQO treatment

At autopsy, only six mice had macroscopically normal tongues. Each of these had a control genotype (Figure 5-1A).

Keratosis of the lateral surface of the tongue was identified in half of Pax9-deficient mice, compared with just 15.6% of controls (Figure 5-1D). The association between Pax9 genotype and keratosis of the lateral tongue was significant ($p<0.01$, Pearson x^2 value - 8.94, 1 d.f.). Well-defined lesions (plaques, papules, and nodules) were identified in the majority of both Pax9-deficient mice and controls (Figure 5-1B, D). A cross tabulation of discrete lesions and genotype was not statistically significant ($p>0.05$, Pearson x^2 value - 0.005, 1 d.f.). A total of seventeen cases showed tumours at autopsy (Figure 5-1D). The majority of these lesions were in Pax9-deficient mice (12 cases, 71%). However, a cross tabulation showed that the association between genotype and the presence of macroscopic tumours was not significant ($p>0.05$, Pearson x^2 value - 2.83, 1 d.f.).

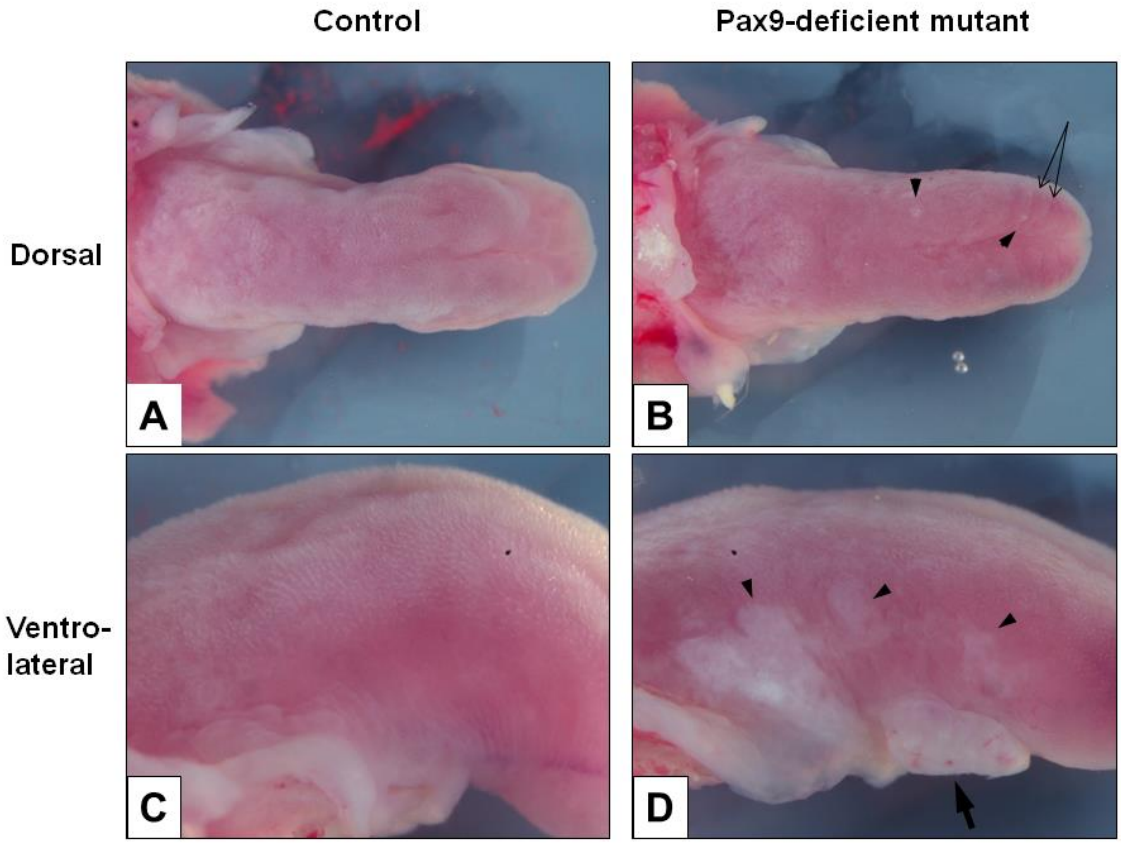


Figure 5-1 Macroscopic autopsy features of the tongues of Pax9-deficient mice and controls following 4-NQO treatment

A) Normal dorsal mucosa of a control mouse. B) The dorsal surface of the Pax9-deficient mouse shows fissuring towards the tip (arrows). There are several small, well-defined nodules (arrowheads). C) Normal ventro-lateral epithelium of a control mouse. D) The Pax9-deficient mouse shows diffuse keratosis of the lateral surface. There are several plaque-like lesions reminiscent of human leukoplakia (arrowheads). An exophytic tumour projects from the anterior ventral surface (arrow, to the right).

A) and B) taken at 60x magnification; C) and D) taken at 160x magnification.

5.3.6 Histopathological features of mouse tongues following 4-NQO treatment

Histopathological examination showed a broad spectrum of features. Individual mice were classified according to the most advanced stage of disease (Table 5-3). A cross tabulation showed that the distribution of mice according to genotype and histological diagnosis was statistically significant ($p<0.0001$, Pearson x^2 value – 75.9, 10 d.f.).

Normal epithelium was seen in 20 cases, with even numbers of Pax9-deficient mice and controls (Table 5-3). Nine of the mice with normal epithelium had been sacrificed prior to completion of the scheduled 4-NQO treatment. However, mice in this group had a range of treatment durations (2 – 28 weeks) and 4-NQO concentrations (10, 20 and 50 µg/ml).

Thirty-eight cases showed epithelial dysplasia. Epithelial dysplasia was detected in even numbers of Pax9-deficient mice and controls (Table 5-3). The scoring pathologists did not stratify these cases according to grade of epithelial dysplasia. However, a wide range of architectural and cytological features were observed in both Pax9-deficient mice and controls (Figure 5-2). The shortest time to induction of epithelial dysplasia was two weeks. This was in a Pax9-deficient mouse that had been treated with 20µg/ml 4-NQO and was sacrificed prematurely having developed a toxic systemic response to 4-NQO. Epithelial dysplasia was also detected in another Pax9-deficient mouse after four weeks of treatment with 20 µg/ml 4-NQO. By contrast, control mice sacrificed at four weeks (20 µg/ml), six weeks (10 µg/ml), and eight weeks (50 µg/ml) showed normal epithelium with no evidence of epithelial dysplasia.

Oral squamous cell carcinoma (OSCC) was detected in ten mice. The histological features and invasive patterns of these tumours were diverse (Figure 5-3). The majority arose in Pax9-deficient mice (Table 5-3). The shortest time to induction was 22 weeks in a Pax9-deficient mouse treated with 20 µg/ml 4-NQO. Interestingly, the majority of mice that developed OSCC were treated with the lowest dose (10 µg/ml 4-NQO). The shortest time to induction with 10 µg/ml 4-NQO was 24 weeks.

Table 5-3 Summary of the number of Pax9-deficient and control mice assigned to each diagnostic category following histopathological examination

Diagnostic category	Pax9-deficient mice	Control	Total
Normal microanatomy	10	10	20
Epithelial dysplasia	19	19	38
Squamous cell carcinoma	7	3	10
Total	**36**	**32**	**68**

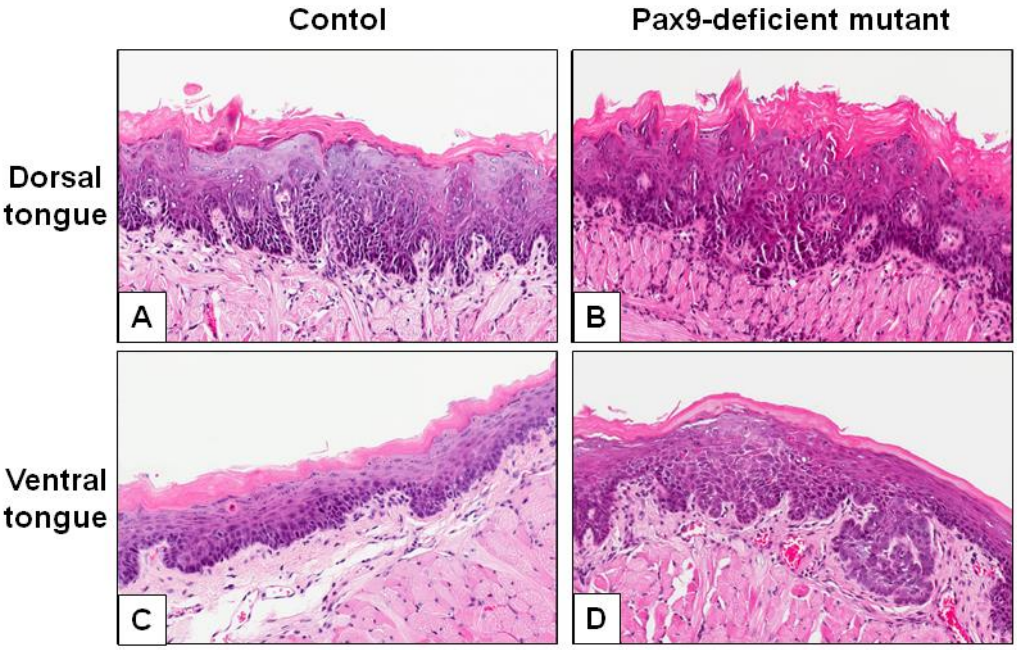


Figure 5-2 Epithelial dysplasia involving dorsal and ventral tongue mucosa of Pax9-deficient mice and controls

A, B) Both Pax9-deficient mice and controls developed epithelial dysplasia on the dorsal surface of the tongue. C, D) Mice with both genotypes also developed epithelial dysplasia on the ventral surface of the tongue. A diverse spectrum of architectural and cytological abnormalities were observed at both subsites. Epithelial dysplasia was not assigned to a histological grade during this project.

Images taken at x100 magnification.

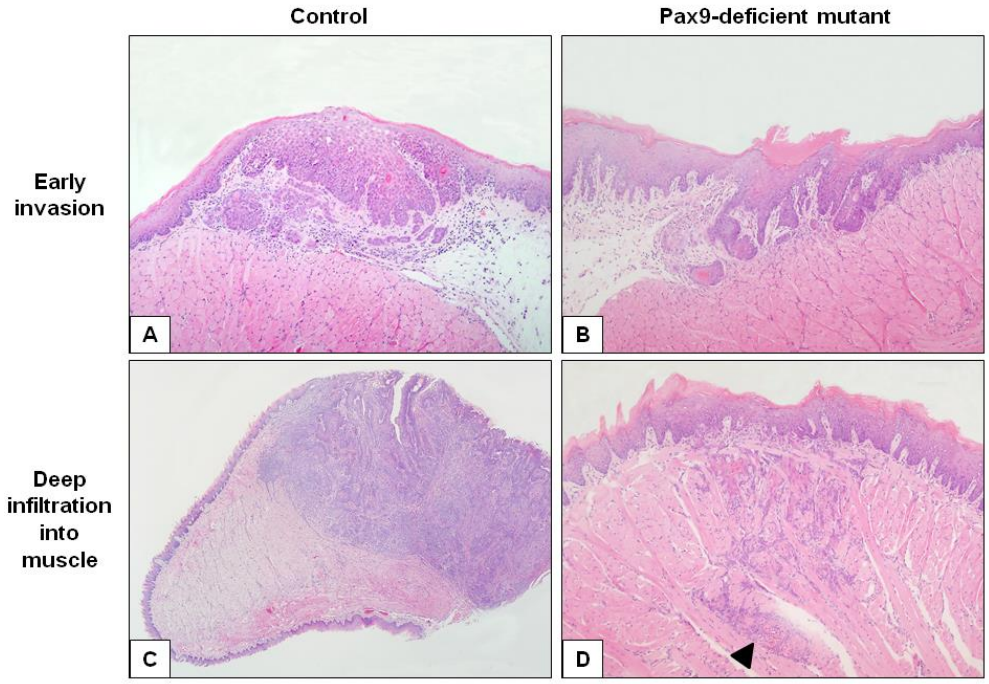


Figure 5-3 Histopathological features of OSCC involving the tongues of Pax9-deficient mice and controls

A) and B) OSCC shows early invasion of the lamina propria in a control and Pax9-deficient mouse respectively. C) In a control mouse, OSCC has infiltrated the musculature of the tongue, extending through nearly the full thickness towards the posterior aspect. D) In a Pax9-deficient mouse, although there is an intact surface OSCC infiltrates the musculature of the tongue with aggressive, non-cohesive, and sarcolemmal pattern of invasion (arrowhead).

A), B), and D) taken at x100 magnification; C) taken at x50 magnification.

5.3.7 Chronic inflammatory changes and ulceration

Chronic inflammatory ulceration was detected in four cases. Chronic inflammation without ulceration was detected in one further case. Ulceration/chronic inflammation was recorded as an incidental finding and the mice were categorised according to other 'worst' diagnosis according to changes elsewhere in the mucosa (Table 5-3).

5.3.8 Macroscopic tumours arising at other mucosal subsites

This study focused primarily on squamous cell carcinoma and epithelial dysplasia involving the tongue. However, there was a subset of 14 mice in which tumours developed at non-tongue oral subsites. The subsites and numbers of Pax9-deficient mice and controls affected are summarised in Table 5-4. An example of one tumour that has been confirmed histologically as an OSCC is shown in Figure 5-4.

5.3.9 Oesophageal squamous cell carcinoma

The oesophagus of each mouse was examined at autopsy. The majority of cases showed no oesophageal abnormality. However, in two Pax9-deficient mice the oesophagus was distended by tumour. The tumour had a pale-grey cut surface. In these cases, the oesophagus was submitted for histological examination. Macroscopically normal oesophagus from one control and one Pax9-deficient mouse was also submitted for comparison.

Histological examination confirmed that both tumours were squamous cell carcinoma (Figure 5-5). Interestingly, both mice had appeared systemically well prior to autopsy, showing no sign of weight loss. Both were treated with 10µg/ml 4-NQO and were able to complete the planned duration of treatment (27 weeks and 28 weeks). One had a FVB background and one was a hybrid. The FVB mouse had also developed OSCC of the tongue. The hybrid mouse showed epithelial dysplasia on the tongue but no evidence of OSCC.

Table 5-4 Summary of tumours identified macroscopically at non-tongue mucosal subsites

Subsite	Pax9-deficient mice	Controls	Total
Floor of mouth	3	2	5
Buccal mucosa	1	1	2
Lip	2	1	3
Gingiva	2	1	3
Oropharynx	1	0	1
Total	**9**	**5**	**14**

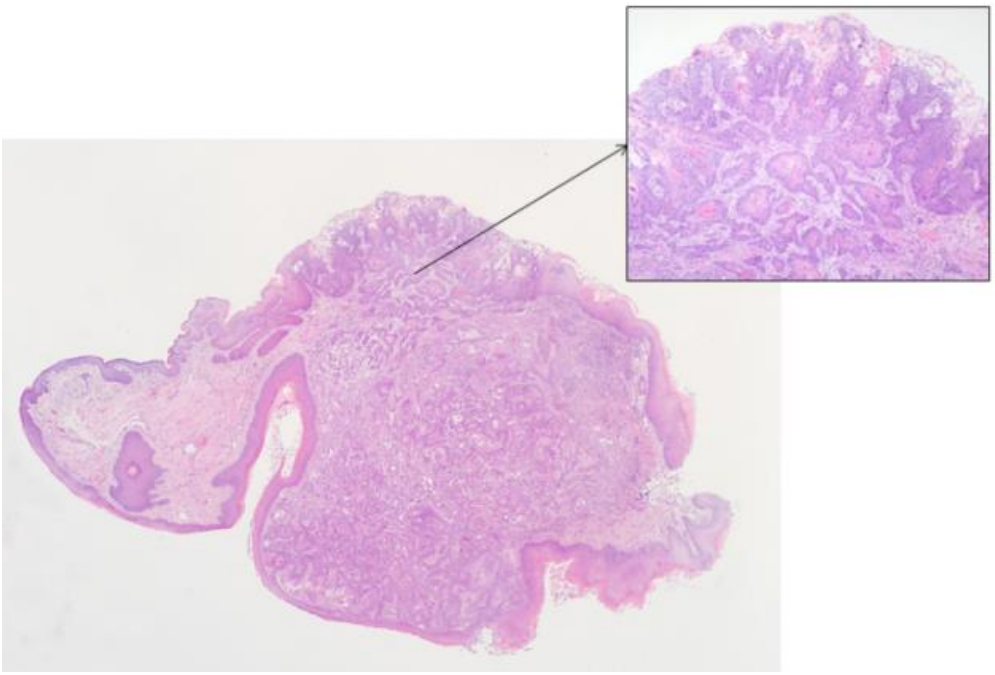


Figure 5-4 Oral squamous cell carcinoma arising on the buccal mucosa of a control mouse

This is a section through the anterior buccal mucosa. Squamous cell carcinoma extends from the mucosal surface through the full thickness of the lip to undermine the skin. The insert shows islands of atypical squamous cells that show keratin formation centrally.

Main picture taken at x50 magnification, insert at x100 magnification.

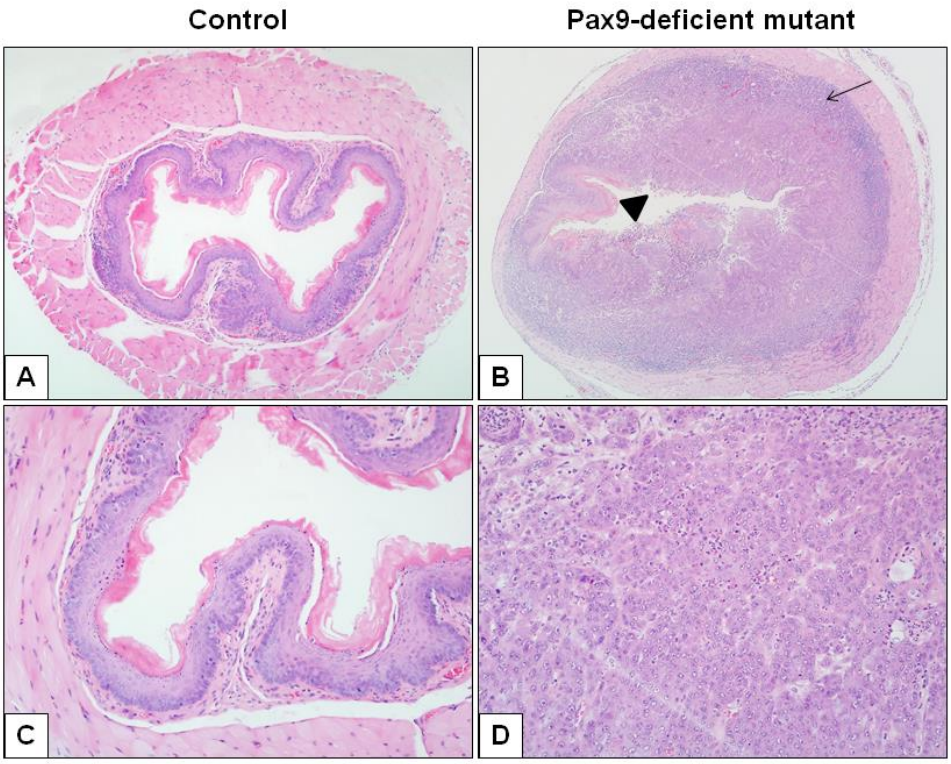


Figure 5-5 Oesophageal squamous cell carcinoma in a Pax9-deficient mouse

A) and C) The control mouse shows the normal micro-anatomy of the oesophagus. There is a thick wall of muscle and the mucosa has an undulating outline. The squamous epithelium shows orthokeratosis. B) In the Pax9 mutant, the normal micro-anatomy is distorted by squamous cell carcinoma. The oesophagus is distended with compression and invasion of the muscular wall. Carcinoma almost entirely occludes the lumen (arrowhead). D) In the mutant, the epithelium is replaced by sheets of malignant squamous cells.

A) and B) taken at x50 magnification;. C) and D) taken at x200 magnification.

5.3.10 Aged Pax9-deficient and control mice not treated with 4-NQO

Two aged mice that had not been treated with 4-NQO were sacrificed for morphological and immunohistochemical analysis. Both were females and had a Black 6 genetic background. Both were aged two years at the time of sacrifice. One was a Pax9-deficient mutant with complete knockout phenotype; the second was a control.

The control mouse showed a normal phenotype with the normal complement of filiform papillae on the dorsal surface of the tongue (Figure 5-6A). The Pax9-deficient mouse showed a smooth dorsal surface due to loss of filiform papillae. Microscopically, there were a few sparse, abortive filiform papillae (Figure 5-6B). Keratinocytes in the control showed a nuclear pattern of Pax9 expression that was most intense in the prickle cell layer (Figure 5-6C). By contrast, keratinocytes in the Pax9-deficient mouse were negative (Figure 5-6D). This confirmed site specific Pax9 knock-out in the K14 $Pax9^{flox/flox}$ mouse. The ventral surfaces of the control and mutant showed common histological features, with an undulating rete profile and light keratosis (Figure 5-6 F, G). Keratinocytes in the control mouse showed nuclear expression of Pax9, but were negative in the Pax9-deficient mouse (Figure 5-6H, I).

On the dorsal surface of the tongue, the Pax9-deficient mouse showed fissuring. In these areas, the squamous epithelium invaginated and pushed down into the lamina propria and musculature of the tongue (Figure 5-7A, B). There was some up-regulation of Sox2 in the Pax9-deficient mouse relative to the control (Figure 5-7C, D). The control mouse showed nuclear Ki67 expression in the basal keratinocytes at the tips of rete processes (Figure 5-7E). There was slight up-regulation of Ki67 in the Pax9-deficient mouse, notably at the apex of the fissure (Figure 5-7F). p53 was negative in the control, but a few p53-positive cells were identified in the Pax9-deficient mouse [data not shown].

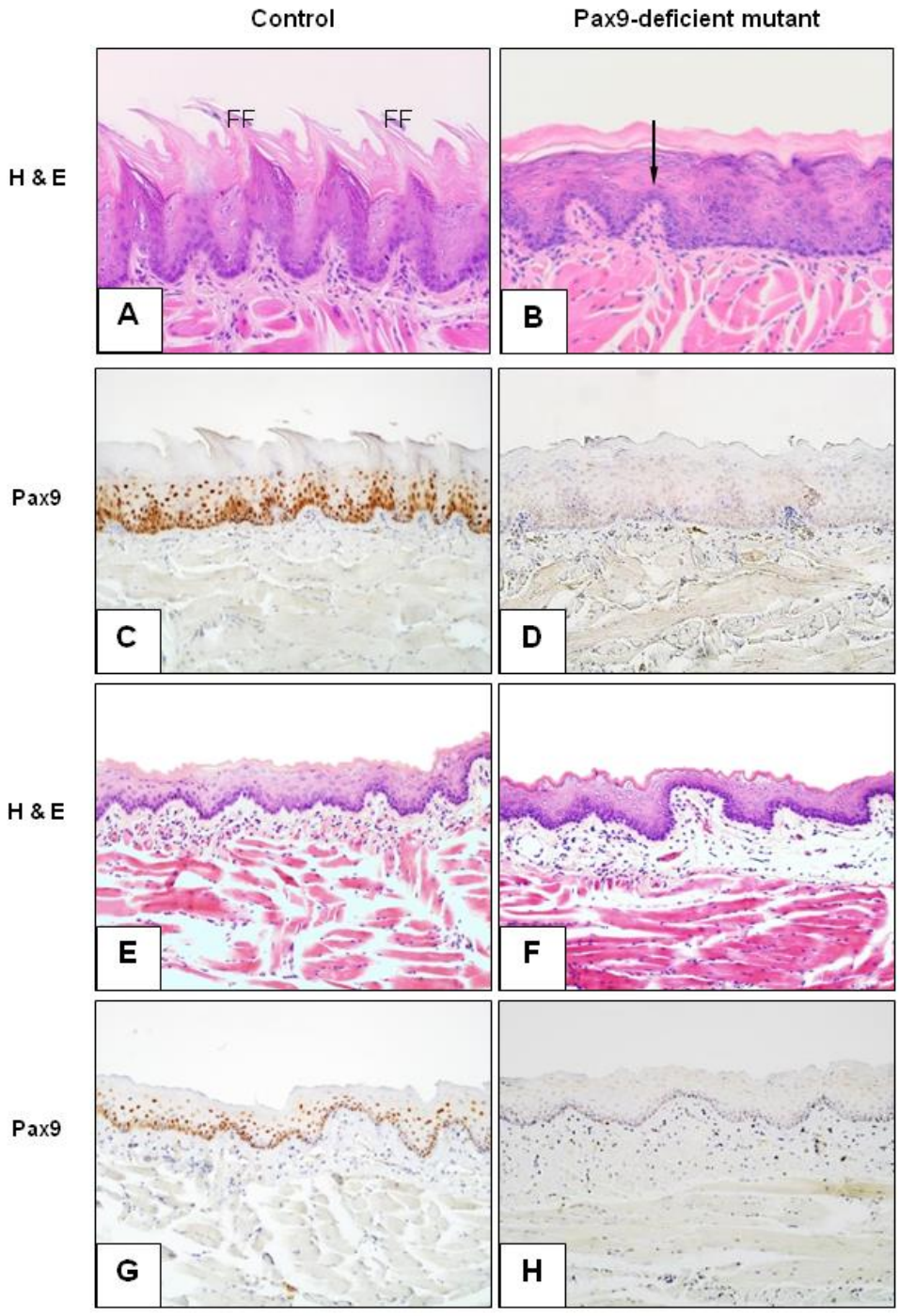


Figure 5-6 Immunohistochemical characterisation of the dorsal tongue epithelium of control and Pax9-deficient adult mice not treated with 4-NQO

A) H&E shows a normal complement of filiform papillae (FF) on the dorsal surface of the control. B) FF are absent (arrow) in Pax9-deficient dorsal tongue epithelium and only remnants of the mesenchymal papilla can be observed. C) Pax9 is expressed in basal and parabasal layers and is strongest in differentiating suprabasal keratinocytes in controls. D) Pax9 is absent in the Pax9-deficient mouse. E) and F) the ventral epithelium of the Pax9 deficient mouse and the control show similar features. However, whereas there is nuclear Pax9 expression in the control G), Pax9 is negative in the mutant.

Images taken at x100 magnification.

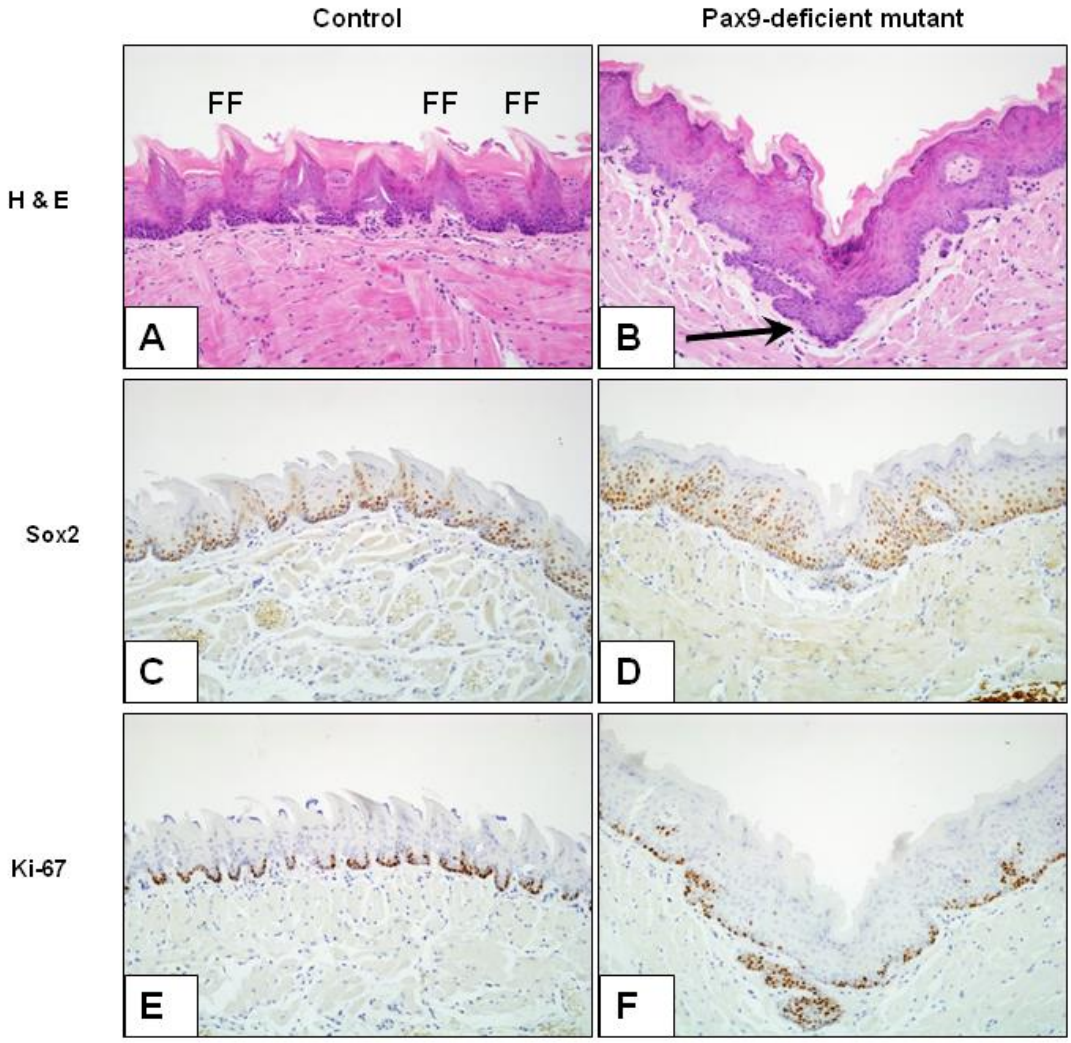


Figure 5-7 Comparison of SOX2 and Ki67 immunohistochemistry in the fissured tongue of an aged Pax9-deficient mouse and a control

A) The control mouse shows filiform papillae (FF) on the dorsal tongue. B) In the Pax9-deficient mouse, filiform papillae are absent and there is a prominent fissure (arrowhead). C) Expression of the oral keratinocyte marker Sox2 is restricted to the basal and parabasal layers in the control. D) In the Pax9-deficient mouse, Sox2 expression varies, but is up-regulated in areas. E) Expression of the proliferation marker Ki67 is restricted to the basal layer in the control. F) There is up-regulation of Ki67 at the apex of the fissure in the Pax9-deficient mouse.

Images taken at x100 magnification.

5.3.11 Selection of 4-NQO treated cases for immunohistochemical analysis

Representative Pax9-deficient and control mice were selected for initial molecular analysis by immunohistochemistry. The majority of the mice that were selected showed epithelial dysplasia. Three control mice with morphologically normal epithelium were analysed as a comparator (Table 5-5).

5.3.12 Initial molecular characterisation of mice treated with 4-NQO

Expression of Ki-67, p53, and Sox2 were analysed by immunohistochemistry in tongue epithelium from Pax9-deficient mice and controls. Ki67 and p53 were analysed as both are up-regulated in dysplastic epithelium relative to normal epithelium in human OPMD. Ki67 is a marker of proliferation, and p53 a marker of cell damage (Kovesi and Szende, 2003; Varun et al., 2014). Our analysis of human tissue samples and the literature to date suggests that SOX2 is a marker of cancer stem cells and may have an oncogenic role in OSCC formation (Boumahdi et al., 2014).

The expression of each of the three biomarkers in dysplastic epithelium was compared relative to normal epithelium in both Pax9-deficient mutants and control mice. On the ventral tongue, both p53 and Sox2 were significantly up-regulated in dysplastic epithelium relative to the normal epithelium in both Pax9-deficient mice and controls (Figure 5-8; Sox2 up-regulation shown in Figure 5-9). Ki67 was also up-regulated in dysplastic epithelium relative normal epithelium. However, the difference was not significant in either Pax9-deficient mice or controls (Figure 5-8). The same trends were identified in the dorsal tongue [data not shown].

Table 5-5 Summary of the 4-NQO-treated cases selected for immunohistochemical analysis

Mouse ID	Sex	Genotype	Age (weeks)	Strain	4-NQO (µg/ml)	Duration (weeks)	Tongue Phenotype
Normal epithelium							
1001	F	Control	36	Hybrid	10	28	Normal
1033	M	Control	30	B6	10	22	Normal
1034	M	Control	30	B6	10	22	Normal
Epithelial Dysplasia							
846	F	Mutant	45	B6	50	14	Normal
922	M	Control	38	B6	50	15	Normal
925	M	Mutant	37	B6	50	14	Complete KO
982	M	Control	34	B6	20	23	Normal
983	M	Mutant	33	B6	20	22	Complete KO
987	F	Control	33	B6	20	22	Normal
988	F	Mutant	33	B6	20	23	Complete KO
989	F	Mutant	33	B6	20	23	Partial KO
1030	M	Mutant	29	B6	10	22	Complete KO
1031	M	Mutant	30	B6	10	22	Complete KO
959	M	Control	43	FVB	10	27	Normal
960	M	Mutant	43	FVB	10	27	Complete KO
973	F	Control	43	FVB	10	27	Normal
974	F	Mutant	43	FVB	10	27	Complete KO
994	M	Mutant	36	Hybrid	10	28	Complete KO
995	F	Control	36	Hybrid	10	28	Normal
997	F	Mutant	36	Hybrid	10	28	Complete KO
998	M	Mutant	36	Hybrid	10	28	Complete KO
999	F	Control	36	Hybrid	10	28	Normal
1004	M	Mutant	36	Hybrid	10	26	Partial KO

(Mouse ID - reference identification for internal mouse and Home Office databases; Sex - F = female, M = male; Strain - genetic background of mice analysed (hybrid = BL6 x FVB crossed mice); Tongue phenotype - either complete or partial knockout or normal Pax9 expression as determined by presence or absence of filiform papillae on the dorsal tongue.

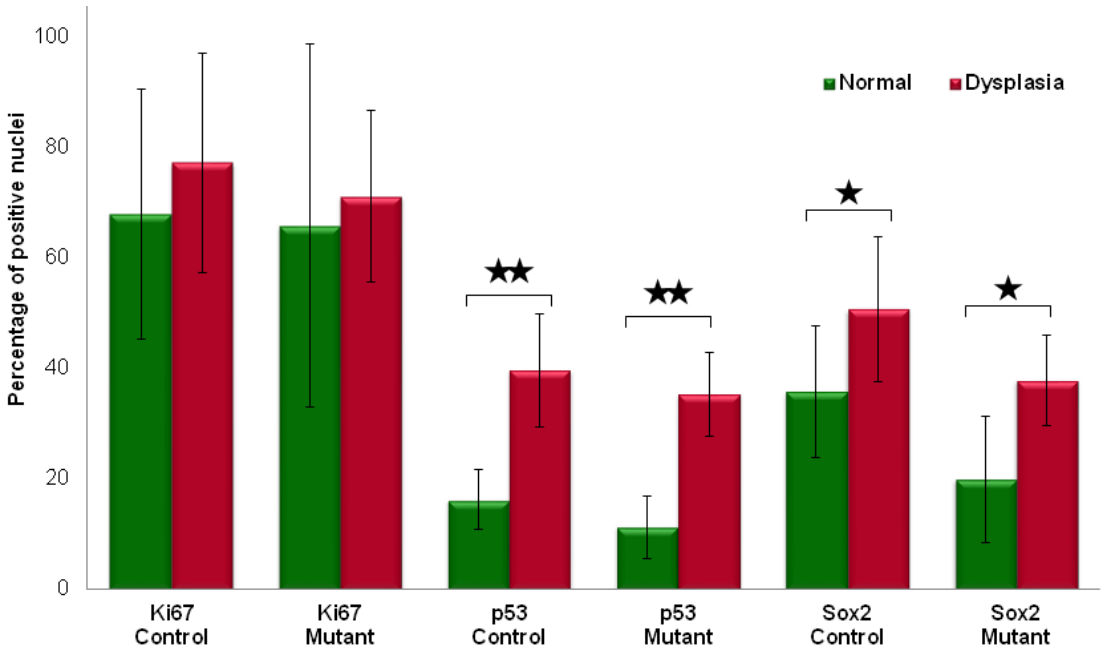


Figure 5-8 Bar chart comparing mean PPN Ki-67, p53, and Sox2 values of normal and dysplastic ventral epithelium of Pax9-deficient and control mice

The proportion of p53-positive and Sox2-positive nuclei is significantly increased in ventral epithelial dysplasia relative to histologically normal epithelium. This trend was identified in both controls and mutants ($p<0.01$ for p53; $p<0.05$ for Sox2). Ki-67 expression was also up-regulated in epithelial dysplasia relative to the normal epithelium, but the differences were not statistically significant.

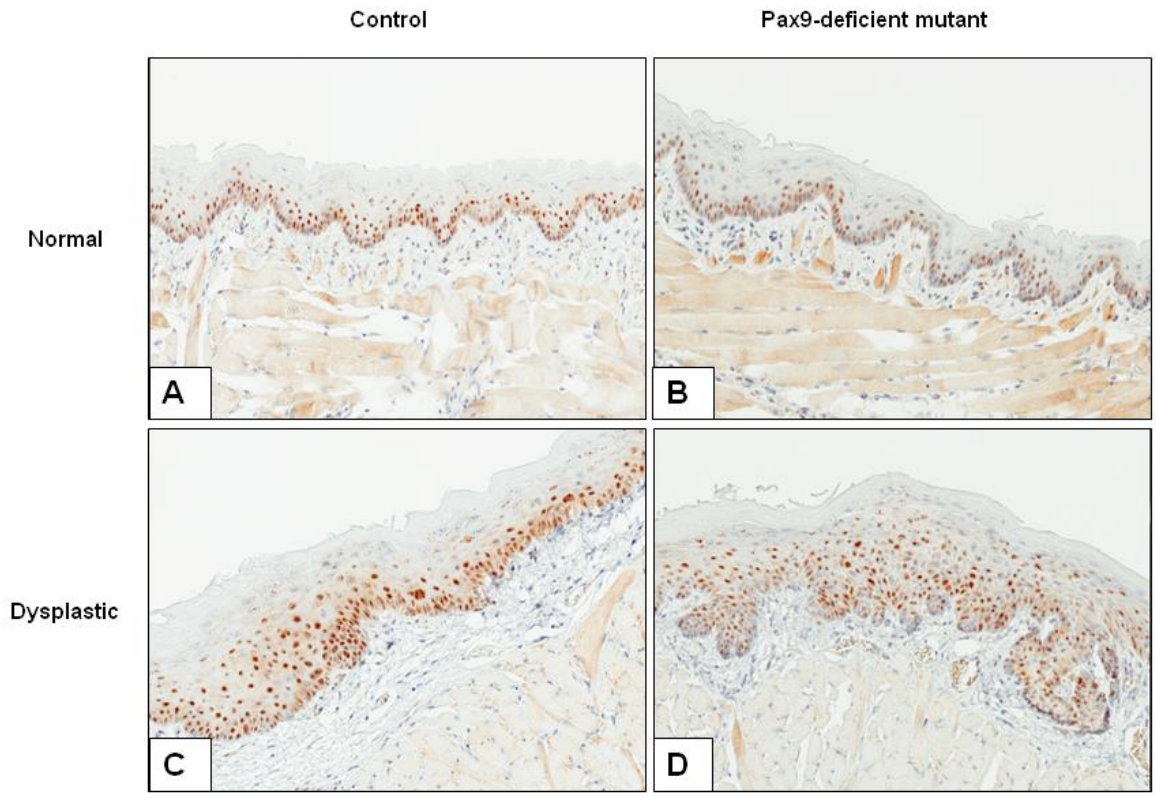


Figure 5-9 Sox2 protein expression in normal and dysplastic ventral epithelium of control and Pax9-deficient mice treated with 4-NQO

The proportion of Sox2-positive cells is increased in epithelial dysplasia relative to the normal epithelium in both control mice (A and C) and in Pax9-deficient mutants (B and D).

Images taken at x 200 magnification.

5.4 Discussion

Mouse models of oral carcinogenesis are critical to the evaluation of biomarkers and chemotherapeutic targets (Herzig and Christofori, 2002). Chemical induction of oral carcinogenesis using 4-nitroquinolone-1-oxide (4-NQO) is a well-established technique. 4-NQO has been used in the evaluation of chemotherapeutic agents in wild-type mice (Czerninski et al., 2009). It has also been used in combined models using transgenic mice to determine the contribution of specific genetic alterations to oral squamous cell carcinoma (OSCC) formation (Zhang et al., 2006). This feasibility study is the first of its kind to induce Pax9-deficient mice using 4-NQO. It is therefore analogous to a phase I clinical trial in humans. Phase I trials are designed to identify the response to a particular treatment, the safe dose range, and the possible side effects of treatment (Cancer Research UK, 2013). This study has identified important trends in the susceptibility of Pax9-deficient mice to formation of epithelial dysplasia and OSCC in response to 4-NQO treatment. Moreover, this study has detected enhanced sensitivity to the systemic effects of 4-NQO in these Pax9-deficient mice. It therefore has important recommendations for the experimental design of future studies involving larger groups of mice (Cancer Research UK, 2013; Understanding Animal Research, 2014).

5.4.1 Formation of epithelial dysplasia

Epithelial dysplasia was detected in the majority of both control and Pax9-deficient mice following 4-NQO induction. The prevalence of epithelial dysplasia across the group confirms the experimental reliability of the 4-NQO model (Vitale-Cross et al., 2009). It also confirms the inductive efficacy of the three dosages used in these experiments (i.e. 10, 20, and 50µg/ml).

Pax9-deficient mice developed epithelial dysplasia after just two weeks following the start of 4-NQO treatment. By contrast, in control mice the shortest time from 4-NQO induction to development of epithelial dysplasia was 15 weeks. Induction of epithelial dysplasia after just two weeks has been reported in rats (Nauta et al., 1996). However, the majority of studies using 4-NQO induction in mice report an eight week time-lag from the start of treatment to development of epithelial dysplasia (Tang et al., 2004; Vitale-Cross et al., 2004b).

This finding provides some support for our hypothesis that Pax9-deficient mice are more susceptible to 4-NQO due to the loss of Pax9 tumour-suppressor function. However, this finding must be interpreted with caution. A relatively small number of mice were sacrificed after a brief duration of 4-NQO treatment. In order to confirm that epithelial dysplasia is induced more rapidly in Pax9-deficient mice, a future experiment would need to schedule the sacrifice and autopsy of matched control and Pax9-deficient mice at short, predetermined intervals of treatment (e.g. weekly intervals for the first six weeks). Such experiments are likely to require considerably larger mouse populations. The incremental loss of mice from early in the experiment would otherwise reduce the number of cases exposed for a sufficient length of time to develop OSCC. In the current experiment, detection of epithelial dysplasia at two weeks was essentially an incidental finding, i.e. it was due to the unexpected loss of mice early in the experiment due to a toxic systemic response to 4-NQO rather than in accordance with the experimental design.

5.4.2 Grading of epithelial dysplasia

The scoring pathologists did not categorise cases according to a stratified grade of epithelial dysplasia. There is a risk that this approach has failed to identify important differences between and trends with Pax9-deficient mice and controls regarding the severity of epithelial dysplasia. It is conceivable that Pax9-deficient mice develop high-grade epithelial dysplasia more frequently (and/or more rapidly) than controls. Such trends would not be detected using the current binary (i.e. ‘present’ or ‘absent’) classification. However, there is no agreed classification system for grading oral epithelial dysplasia in mice. Mouse epithelium is thinner than human epithelium. This confounds the direct transference of classification systems for human epithelial dysplasia. It is well documented that the grading of human oral epithelial dysplasia is subject to inter-observer and intra-observer variability (Kujan et al., 2007). It is our experience that grading epithelial dysplasia in control mice is challenging. This is due to difficulty in distinguishing between normal epithelial and subtle basal cell hyperplasia in the thinner mouse epithelium, and the keratinisation of the ventral epithelium as part of the normal microanatomy of the mouse tongue. Grading is further complicated in Pax9-deficient mice, in which the normal

squamous epithelium shows architectural abnormalities. In the context of an already abnormal epithelial architecture, it is conceivable that architectural changes specifically reflecting epithelial dysplasia may be either overlooked or identified erroneously. These issues reflect the difficulties inherent to the study of a genetically modified mouse strain with a unique phenotype. They also highlight the need for clear guidelines to be developed for the diagnosis and grading of epithelial dysplasia in mice. Such guidelines would support researchers using the 4-NQO model and potentially bring greater uniformity and reliability to the literature.

There are also potential ethical implications if cases were stratified according to the grade of epithelial dysplasia. By using a binary 'absent' or 'present' approach, the methodology used in the present study is consistent with the principle of reducing the number of animals required in an experiment to a minimum (Understanding Animal Research, 2014). In order to support valid statistical comparisons, considerably larger numbers of mice would be required in an experiment that stratified epithelial dysplasia according to the World Health Organisation (WHO) or Squamous Intraepithelial Neoplasia (SIN) classifications (Barnes et al., 2005).

5.4.3 Formation of oral squamous cell carcinoma

4-NQO treatment induced OSCC formation in both control and Pax9-deficient mice. However, the majority of OSCC arose in Pax9-deficient mice. This provides some support for our hypothesis that Pax9 mutant mice are more susceptible to OSCC formation than controls due to the loss of the tumour-suppressor function of Pax9.

Across all diagnostic categories, the association between genotype and diagnosis was statistically significant. However, statistical comparison according to genotype within the group of mice that developed OSCC is compromised by the small number of cases (n = 10). Further serial sections will be examined as the analysis of the tongue tissue continues beyond this project. It is conceivable that further OSCC will be identified during this process. Larger numbers of OSCC cases would support further statistical analysis. However, our macroscopic analysis and initial histological assessment suggest that the number of OSCC is not likely to increase substantially as a result of further

sectioning. The present data therefore need to be verified by future experiments that use standardised conditions on greater numbers of mice.

The shortest interval from the start of 4-NQO induction to OSCC formation was 22 weeks. This is concordant with several previous 4-NQO studies in which 4-NQO treatment has continued for up to 50 weeks (Ma et al., 1999; Ide et al., 2001). However, it is longer than the eight-week induction interval reported by Hasina et al (2009). In that study, a much higher concentration of 4-NQO was used (100 µg/ml). Given the strong association with high dosage and toxic systemic effects of 4-NQO, this dose is not feasible to use in a population of Pax9-deficient mice. Our data show that the maximum dosage in the present study (50 µg/ml 4-NQO) resulted in a rapid toxic systemic effect. It is therefore likely that treatment of Pax9-deficient mice with 100 µg/ml 4-NQO would result in unnecessary suffering or even premature death. This would be in clear breach of the principle of 'refining' experiments to ensure that any suffering of the animals is kept to an absolute minimum (Understanding Animal Research, 2014). In fact, it is interesting to note that the majority of mice that developed OSCC were actually treated with the lowest concentration (10 µg/ml 4-NQO). This confirms findings from previous studies that 10 µg/ml 4-NQO is a sufficient concentration to induce OSCC formation, albeit over a longer period (Ma et al., 1999; Ide et al., 2001).

It is interesting that the majority of mice that developed dysplastic nodules were also Pax9-deficient mice. These dysplastic nodules resemble papillary squamous cell carcinoma in humans. The precise criteria required in order to establish a diagnosis of papillary squamous cell carcinoma in humans is a subject of debate (Barnes et al., 2005). Similar lesions have been described previously in mice treated with 4-NQO (Czerninski et al., 2009). As with grading epithelial dysplasia, however, established criteria for making this diagnosis in mice are not available at present. It is conceivable that some of these dysplastic nodules may show invasion of the lamina propria on examination of deeper levels. It is also possible that had there been a greater interval between the cessation of 4-NQO treatment and sacrifice, some of these lesions may have developed into conventional squamous cell carcinoma. Future experiments may therefore benefit from an increased time interval from the end of 4-NQO treatment to sacrifice/autopsy. The experimental model could be further

enhanced by examining the oral cavity of mice under anaesthesia at intervals during 4-NQO treatment. Mice with nodular lesions could then be selected for a longer interval between cessation of 4-NQO treatment and sacrifice.

5.4.4 Oesophageal squamous cell carcinoma

The majority of epithelial dysplasias and invasive squamous cell carcinomas involved the oral cavity, specifically the tongue. However, two Pax9-deficient mice developed squamous cell carcinoma of the oesophagus. This is consistent with the site affinity of 4-NQO documented in the literature to date. This affinity reflects the high concentration of diaphorase, a reductase that activates 4-NQO, in the oesophagus as well as in the oral cavity (Imaida et al., 1989; Kanojia and Vaidya, 2006). It is interesting that these mice were otherwise well and did not show extreme weight loss, which might have been expected given the extent of oesophageal obstruction. Autopsy did not identify tumours elsewhere in the gastrointestinal tract. This is consistent with the negative autopsy findings beyond the oral cavity/oesophagus documented by Tang et al (2004).

5.4.5 Non-specific inflammatory changes

Only a small subset of mice showed non-specific inflammatory changes and/or non-specific ulceration as a result of 4-NQO treatment. This is in concordance with the literature regarding the method of action for 4-NQO, which is generally less irritant than chemical carcinogens such as DMBA (Eveson and MacDonald, 1978).

5.4.6 Systemic effects of 4-NQO induction

Our data show that Pax9-deficient mutant mice are at greater risk of developing a systemic response (i.e. excessive salivation, extreme weight loss) to 4-NQO treatment relative to wild-type controls.

The precise mechanism for this increased sensitivity is unclear. Excessive salivation was an early sign in all mice that subsequently deteriorated (i.e. developed extreme weight loss and a hunched position). However, it did not portend a poor prognosis in all cases. Excessive salivation may have had a range of effects, to which mice were variably susceptible:

- Excessive salivation may result in fluid loss and dehydration
- Dehydrated mice may subsequently attempt to maintain homeostasis by drinking greater quantities of water:
    - In the context of this experiment, where 4-NQO is present in drinking water, this would in turn have increased their exposure to 4-NQO
    - This generates a cycle of deterioration: excessive salivation leads to increasing water consumption, thus increasing 4-NQO exposure and causing a further increase in salivation, and drinking
    - This cycle culminates in a relentless increase in the concentration of 4-NQO and its metabolites in the bloodstream
- Finally, it is possible that excessive salivation may have been accompanied by loss of appetite/impaired masticatory efficiency. This would contribute to weight loss independent of the other systemic effects of 4-NQO. The dental abnormalities detected in a minority of cases were mild and unlikely to have significantly compromised masticatory efficiency in isolation.

The mice that deteriorated most rapidly had only one kidney. Impaired excretion of 4-NQO may have further contributed to an increase in systemic 4-NQO concentration. It was not possible to identify which Pax9-deficient mice had only one kidney prior to beginning 4-NQO treatment. The cause of hunching in these mice was unclear. It may have reflected either visceral or muscular pain from the abdominal region. Mice with one kidney may have developed kidney failure due to 4-NQO toxicity. The hunching may also have been related to kidney failure. All organs from these mice have been formalin fixed and stored. Histological analysis of vital organs, especially the kidney, may provide greater insight into the nature of the systemic impact of 4-NQO. Future experiments using Pax9-deficient mice must be designed in order to avoid suffering and loss from the experiment of this vulnerable subset of mice with only one kidney. One possibility would be to assess the presence of both kidneys prior to commencing treatment by imaging, such as CT scanning.

5.4.7 Optimisation of a combined 4-NQO/Pax9 knockdown model

Over the past 15 years, several workers have attempted to optimise an experimental protocol for the 4-NQO model, varying parameters such as the duration, site/mechanism of carcinogen application, overall duration of

treatment, and point of sacrifice (Vitale-Cross et al., 2009). Our data suggests that experiments involving Pax9-deficient mice should include the following design considerations:

1) Genetic background: Our data showed that Pax9-deficient mice with a Black 6 genetic background are more susceptible to 4-NQO than mutants with a FVB or hybrid background. Although wild-type Black 6 mice did not demonstrate a heightened sensitivity to 4-NQO, Pax9-deficient mice with a Black 6 background are at greater risk of suffering. The Black 6 background strain should therefore be avoided in future combined 4NQO/Pax9 experiments. Other strains such as FVB and hybrids should be used where possible as they are more resilient to the systemic effects of 4-NQO
2) 4-NQO concentration: Our data also show that a concentration of 10 µg/ml 4-NQO was sufficient to induce formation of epithelial dysplasia and OSCC. Concentrations greater than 10 µg/ml are associated with increased risk of systemic effects. This suggests that there is little experimental benefit in using high dosages of 4-NQO. Loss of Pax9 mutant mice due to the systemic effects of high 4-NQO concentration may compromise the overall efficiency of the experiment through attrition of the sample population
3) Sex: Although the difference in response between males and females was not statistically significant, our data suggests that males are more resilient to 4-NQO than females. In a study of Pax9-deficient mice, males may therefore be better candidates for 4-NQO treatment
4) Interval between cessation of 4-NQO treatment and sacrifice: Future experiments may therefore benefit from an increased time interval from the end of 4-NQO treatment to sacrifice/autopsy. This would allow dysplastic lesions to progress to OSCC while mice are free from the potentially toxic systemic effects of the 4-NQO treatment
5) Oral inspection: In the present experiment, mice were sacrificed either at the end of the planned experimental period or due to toxic systemic effects of 4-NQO. The decision to sacrifice and analyse the tongue was not based upon the clinical appearances following exposure to 4-NQO. This could be modified by performing an oral examination of the mice at

intervals following the start of 4-NQO treatment. Mice could then be selected for analysis according to whether they had developed a potentially dysplastic lesion. Alternatively, mice with clinically identifiable lesions could continue 4-NQO treatment for a range of different time intervals in order to assess the development of specific types of lesion. Oral inspection would need to be performed under a brief general anaesthesia (GA). GA is performed at the FGU by a number of research groups. Although it would add to the complexity of the experiment, there are facilities and suitably trained staff to make this modification feasible

6) Tissue biopsy: Oral inspection could be complemented and enhanced by performing a tissue biopsy while the mouse is under anaesthesia. This biopsy could take the form of either a brush biopsy or a small punch biopsy. The selection of cases would reduce the numbers of mice required for the overall experiment.

5.4.8 Immunohistochemical profile of normal and dysplastic epithelium

The increase in Ki-67 and p53 observed in the ventral tongue dysplasias compared to normal epithelium is consistent with published literature (Nylander *et al.*, 2000; Scholzen and Gerdes, 2000; Kovesi and Szende, 2003; Varun *et al.*, 2014). That the increase is present in both Pax9-deficient mice and controls suggests that Pax9 inactivation does not interfere with the molecular pathways of these key genes, as there is no change between the controls and Pax9 mutants. Conservation of these pathways suggests that the combined Pax9 mutant mouse chemical carcinogenesis induction model is a suitable model of OSCC formation in humans. In both Pax9-deficient and control mice, Sox2 expression is increased in epithelial dysplasia compared to normal epithelium. This supports evidence from the literature that suggests Sox2 may have an oncogenic function in oral carcinogenesis (Freier et al., 2010; Qiao et al., 2013; Kokalj Vokač et al., 2014).

5.5 Conclusion

This feasibility study has demonstrated that Pax9-deficient mice are more susceptible to OSCC formation than controls. This may be due to loss of the tumour-suppressor function of Pax9. However, Pax9-deficient mice are also more sensitive to the systemic effects of 4-NQO treatment. Based on these data, future experiments using a combined 4-NQO/Pax9 knockdown model should be performed using mice with a FVB or hybrid genetic background, as these are more resilient to the systemic effects of 4-NQO than mice with a Black 6 genetic background. Pax9-deficient mice should be imaged prior to treatment in order to confirm the presence of two kidneys; only mice with two kidneys present should be treated with 4-NQO. A concentration of 10 µg/ml is adequate to induce OSCC and a range of precursor lesions within a 24-week period. Higher concentrations of 4-NQO have an increased risk of causing toxic systemic effects, without significantly reducing the length of the experiment or inducing either more or a wider spectrum of epithelial lesions. 28 weeks is a suitable length of 4-NQO treatment. A longer interval between cessation of 4-NQO treatment and sacrifice would allow more time for established precursor lesions and OSCC to develop, and thus add value to the experiment. Oral inspection in conjunction with a biopsy of the lingual mucosa could be used to assess the response of the mucosa to the 4-NQO treatment and guide the selection of cases for either sacrifice or a further period of 4-NQO treatment. Oral inspection could be carried out at intervals e.g. once per month.

Chapter 6. Generation of Stably-inducible Oral Squamous Cell Carcinoma Cell Lines

6.1 Introduction

Normal keratinocytes harvested from the oral cavity for primary cell culture generally senesce following three to four passages (Prime et al., 1990). Growth *in vitro* may require mesenchymal support, for example by co-culturing with mitomycin C-treated 3T3 fibroblasts (Rheinwald and Green, 1975). Malignant keratinocytes derived from oral squamous cell carcinoma (OSCC) exhibit diverse phenotypes and may show characteristics of either normal or malignant cells (Parkinson, 1989). In contrast to their behaviour *in vivo*, cells cultured from OSCC may show little or no growth *in vitro*. Lack of growth may be due to inappropriate cell culture conditions or insufficient numbers of cancer stem cells (Rheinwald and Beckett, 1981). OSCC-derived cells vary in their dependence on mesenchymal support. Mesenchymal support may actually compromise the growth of some OSCC-derived cells (Rupniak et al., 1985; Prime et al., 1990).

Despite the challenges that the primary culture of malignant oral keratinocytes presents, several well characterised OSCC-derived cell lines have been established (Prime et al., 1990; Edington et al., 1995). These cell lines are valuable tools in the study of oral carcinogenesis. They have been characterised with regard to differentiation, behaviour *in vitro* (Sugiyama et al., 1993; Prime et al., 1994a) and tumorigenicity following both subcutaneous and orthotopic transplantation in athymic mice (Prime et al., 1994b; Paterson et al., 2002). Furthermore, the cells have been genetically modified to study the functional significance of key molecules in oral carcinogenesis (Paterson et al., 2002).

The findings reported in Chapters 4 and 5 suggest that *PAX9* has a tumour-suppressor effect and *SOX2* has a tumour-promoting effect. This chapter outlines the screening of a panel of OSCC cell lines and selection of two lines for stable transfection with plasmid DNA targeting PAX9 and SOX2 expression.

6.2 Aims

The aims of this chapter are:

1. To summarise the identification and selection of a cell line with low endogenous PAX9 expression (for transfection with an inducible PAX9 expression plasmid) and a cell line with high endogenous SOX2 expression (for transfection with an inducible SOX2 ‘knock-down’ plasmid)
2. To describe the generation of stably-inducible transfectants using lentivirus vectors
3. To demonstrate proof of principle that the expression constructs are functional using an exemplar cell line.

6.3 Selection of suitable oral squamous cell lines for transfection

Nine commercially-available OSCC-derived cell lines were screened by assessing PAX9 and SOX2 RNA and protein expression. An immortalised normal oral keratinocyte line, OKF6/hTERT, was used as a comparator. A long-term aim is to test the transfected cell lines in an orthotopic mouse model of tumorigenesis (Prime et al., 2004). Consequently, the selection strategy combined assessment of relative PAX9 and SOX2 expression with consideration of the known tumourigenesis of the parent cell line.

6.3.1 Semi-quantitative RT-PCR analysis comparing relative PAX9 and SOX2 RNA expression in each cell line

RNA was extracted from each of the cell lines. Semi-quantitative RT-PCR analysis was then performed to compare the relative endogenous expression of PAX9 and SOX2 RNA in each cell line. Expression of PAX9 and SOX2 RNA was normalised relative to expression of the housekeeping gene *β-actin* (ACTB RNA).

PAX9 RNA expression was low in three of the OSCC cell lines: H103, H314, and H357. Interestingly, PAX9 expression was also low in the normal keratinocyte OKF6/hTERT control relative to several OSCC-derived cell lines (Figure 6-1). Only one OSCC cell line, H400, had higher levels of SOX2 expression than the OKF6/hTERT control (Figure 6-1). Interestingly, none of the OSCC-derived cell lines showed a combination of low PAX9 and high SOX2 expression, the expression profile anticipated in light of our data from human and mouse tissue samples.

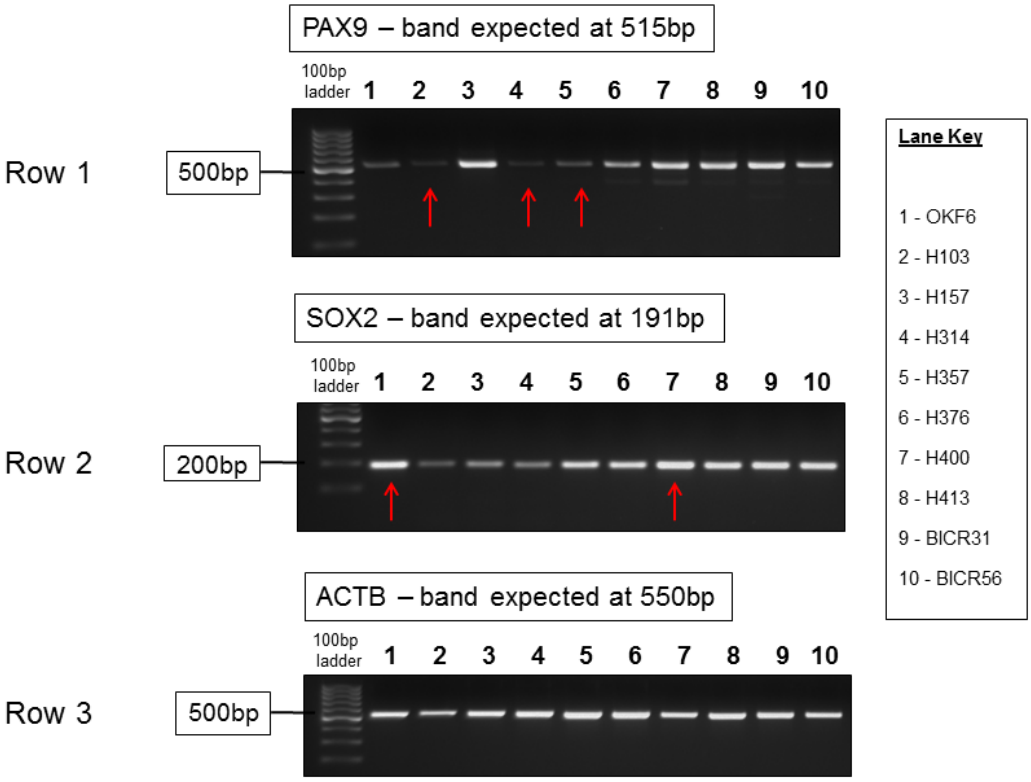


Figure 6-1 Semi quantitative RT-PCR analysis comparing endogenous PAX9 and SOX2 RNA expression in a range of cell lines

Row 1) PAX9 RNA expression was lowest in the H103, H314, and H376 cell lines (red arrows). Interestingly, PAX9 expression in the OKF6/hTERT control (lane 1) was low relative to several of the OSCC-derived cell lines. Row 2) SOX2 expression was highest in the OKF6/Htert control and H400 (red arrows). H400 was the only cell line with SOX2 RNA expression that was higher than the OKF6/Htert control after normalising against the ACTB housekeeper (shown in Row 3). There were similar band intensities for the ACTB housekeeper, indicating consistent quantities of total RNA in the samples.

H357 was not considered to be an ideal candidate for transfection with the PAX9 over-expression plasmid. Although it had low endogenous PAX9 expression, it does not form tumours following orthotopic transplantation to athymic mice (Prime et al., 2004). By contrast, both H103 and H314 are tumorigenic in athymic mice (Prime et al., 2004). Although H103 does not form metastases following transplantation, both H103 and H314 were considered to be potentially suitable candidates for transfection with the PAX9 over-expression plasmid. In order to confirm the results of the semi-quantitative RT-PCR analysis, immunohistochemistry was performed on H103 and H314 to determine their relative endogenous expression of PAX9 protein.

For transfection with the SOX2 knockdown plasmid, the results of the semi-quantitative RT-PCR analysis suggested that H400 was most suitable due to its high endogenous expression of SOX2 RNA. H400 was also suitable in terms of its documented experimental behaviour. *In vitro*, it is the most invasive of the OSCC cell lines (Robinson et al., 2003). Moreover, it is also tumorigenic following orthotopic transplantation in athymic mice (Prime et al., 2004).

6.3.2 Immunohistochemical analysis comparing PAX9 and SOX2 protein expression in H103, H314, H400, and OKF6/hTERT cell lines

Expression of PAX9 and SOX2 protein was analysed by immunohistochemistry in each of the three candidate OSCC cell lines and in the OKF6/hTERT control. Immunohistochemistry was initially performed on cell pellets. However, repeated attempts were unsuccessful, resulting in either weakly stained or completely negative sections [data not shown]. Key variables in the manual immunohistochemical staining protocol were manipulated, including the conditions used for antigen retrieval and the length/temperature of incubation with primary/secondary antibodies. Unfortunately, staining was consistently weak, particularly for SOX2 protein. The reason for this was unclear. It is possible that PAX9 and SOX2 antigens were damaged during either trypsinisation or preparation of the cell pellets.

As an alternative strategy, immunocytochemistry was performed on cells prepared using the Surepath™ technique. Cells were trypsinised and transferred in suspension to the Department of Cellular Pathology, where

Surepath™ preparations were performed by Biomedical Scientists. Stained slides were scanned using the Aperio CS2 Scanscope™ platform and underwent digital image analysis.

Using the Surepath™ preparations, immunostaining for both PAX9 and SOX2 was stronger than that of the corresponding cell pellets. It confirmed the relative expression of PAX9 and SOX2 RNA identified by semi-quantitative RT-PCR analysis. PAX9 protein expression was low in H103 and H314 relative to both OKF6/hTERT and H400. PAX9 expression was lowest in the H314 cell line (Figure 6-2A). SOX2 staining was weaker than PAX9 staining across all the cell lines. However, SOX2 expression was appreciably stronger in OKF6/hTERT and H400 relative to H103/H314. SOX2 expression was almost entirely absent from H314 cells (Figure 6-2B).

Digital image analysis of the Surepath™ stained slides confirmed that PAX9 expression was lower in the H314 cell line than in H103 (Table 6-1). By contrast to the results of semi-quantitative RT-PCR analysis, SOX2 expression was actually higher in OKF6/hTERT relative to H400. However, SOX2 expression in the H400 cell line was consistently higher than in either H103 or H314 (Table 6-1).

Table 6-1 Summary of digital image analysis of PAX9 and SOX2 protein expression by immunohistochemistry in OKF6/hTERT, H400, H103, and H314 cell lines

Cell line	PAX9			SOX2		
	PPN	3+PN	Rank*	PPN	3+PN	Rank**
OKF6	99.1	37.2	4	50.0	1.06	1
H400	89.0	35.5	3	18.1	0.371	2
H103	83.2	1.12	2	4.54	0.0486	3
H314	77.8	0.289	1	0.550	0.322	4

*Ranked low to high (1 = lowest; 4 = highest)

**Ranked high to low (1 = highest; 4 = lowest)

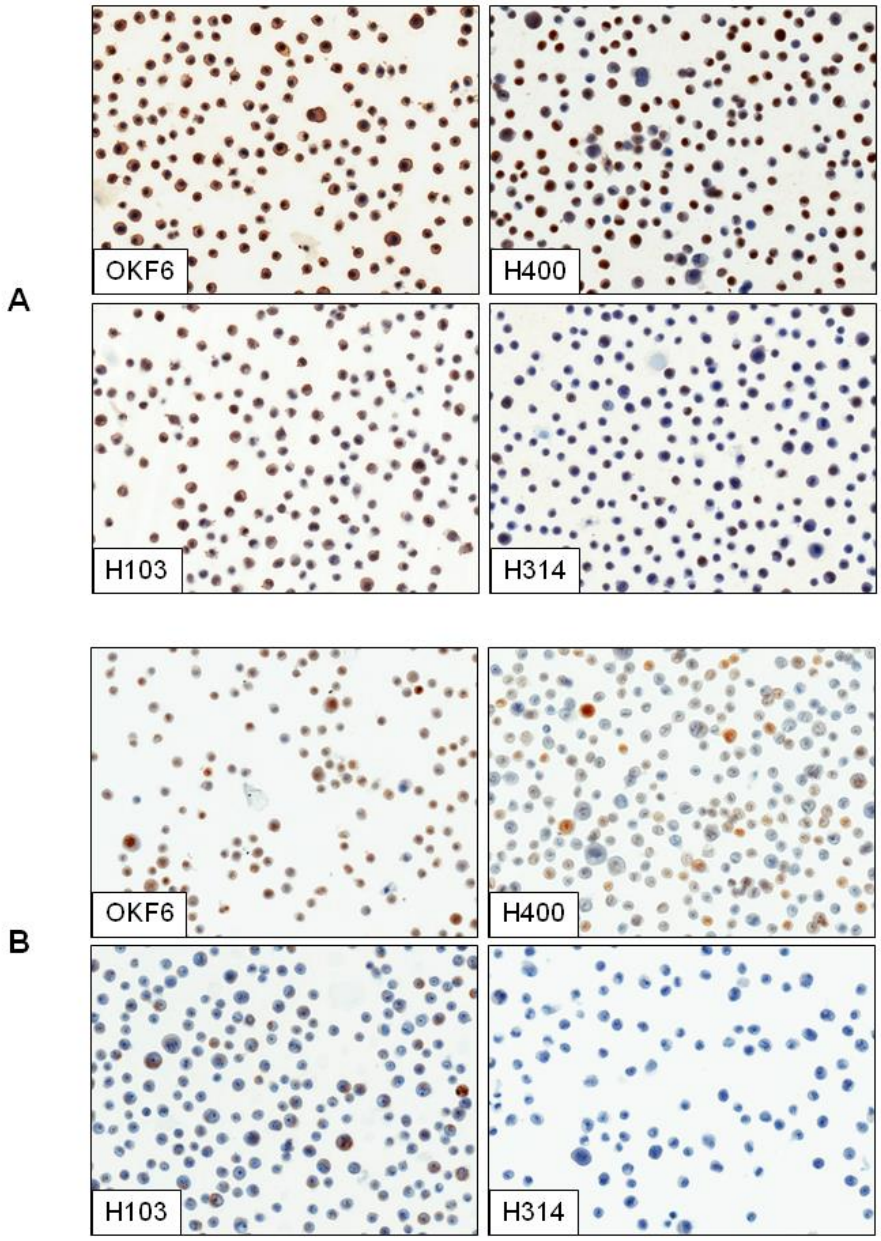


Figure 6-2 PAX9 and SOX2 protein expression by immunohistochemistry in OKF6/hTERT, H103, H314, and H400 cell lines

A) PAX9 protein expression was low in both H103/H314 relative to the OKF6/hTERT control line and H400. PAX9 expression was lowest in the H314 cell line. B) Across all the cell lines, SOX2 immunostaining was weak compared to PAX9 staining. However, SOX2 expression was appreciably higher in OKF6/hTERT and H400 than in either H103 or H314. SOX2 expression was almost entirely absent from the H314 cell line.

Images taken at x100 magnification.

6.3.3 Immunohistochemical confirmation of squamous differentiation of the cultured cells

The cell lines were cultured for several months while the RNA and protein assays were performed. The cell lines exhibited differences in cytomorphology. In particular, H314 showed a spindle-shaped morphology that was dissimilar to that of the OKF6/hTERT control. H400 and H103 were polygonal and more closely resembled normal keratinocytes. There was concern that H314 may have undergone epithelial-mesenchymal transition during *in vitro* culture. Loss of squamous differentiation would make it an unsuitable candidate for transfection with the PAX9 over-expression plasmid. Squamous differentiation was therefore verified by immunohistochemistry. Staining was performed on cell pellets using the automated Ventana platform in the Department of Cellular Pathology. Assays were performed for two markers of squamous differentiation, cytokeratin 5/6 (CK5/6) and p63.

Interestingly, automated immunostaining of the cell pellets was more successful than the manual immunostaining for either PAX9 or SOX2. Each of the cell lines exhibited strong cytoplasmic expression of the squamous marker CK5/6. Each cell line also showed strong nuclear expression of p63 (Figure 6-3). This confirmed that despite its spindle-shaped morphology, H314 continued to exhibit squamous differentiation.

6.3.4 Final selection of parent OSCC lines for transfection with PAX9 and SOX2 constructs

It was decided to use H314 for transfection with the PAX9 over-expression plasmid, as it had lower endogenous PAX9 expression than H103. H314 was also preferable in that it forms metastases following orthotopic transplantation to athymic mice, whereas H103 does not (Prime et al., 2004). H400 is tumorigenic in athymic mice, and both assays confirmed that it has high endogenous expression of SOX2 relative to other OSCC cell lines. H400 was therefore selected for transfection with the SOX2 knockdown plasmid. The characteristics of the selected cell lines are summarised in Table 6-2.

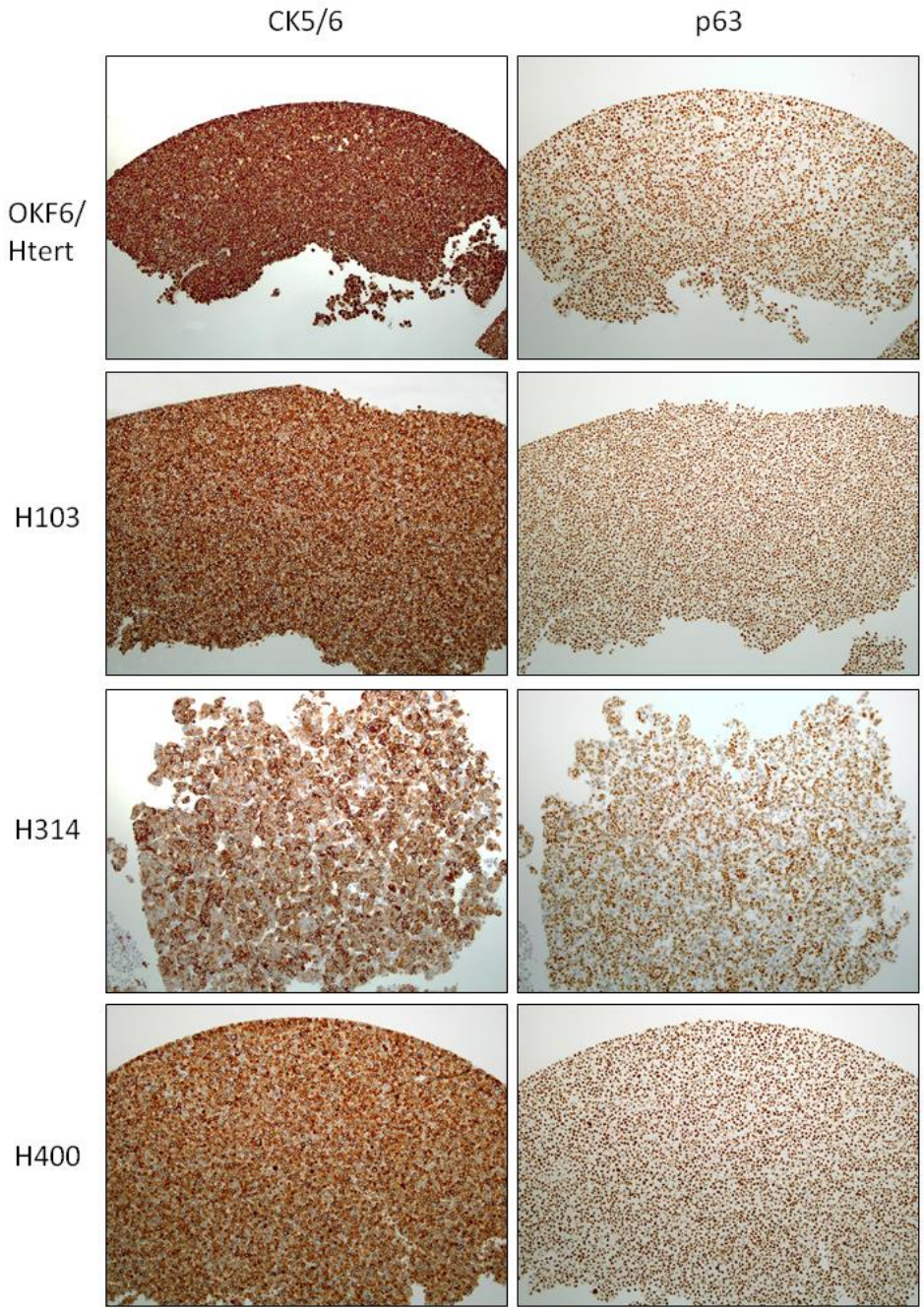


Figure 6-3 Immunohistochemical confirmation of squamous differentiation in OKF6/hTERT, H103, H314, and H400 cell lines

Each of the four cell lines showed strong cytoplasmic expression of CK5/6 and strong nuclear expression of p63. The intensity of staining in the H314 line is similar to that in the other OSCC-derived lines (H103 and H400) and the OKF6/hTERT control. This indicates that H314 retains squamous differentiation despite its spindle-shaped growth pattern in vitro.

Images taken at x100 magnification.

Table 6-2 Summary of the selected cell lines and their PAX9/SOX2 expression profiles, donor characteristics, and behaviour following orthotopic transplantation into athymic mice

Cell line	Experimental aim	Inducible construct design	mRNA		Protein		Donor characteristics			Orthotopic transplantation		
			PAX9	SOX2	PAX9	SOX2	Site	Lymph node metastasis	Distant metastasis	Tumour formation	Response to TGF-β	Metastasis
H314	PAX9 tumour suppressor function	PAX9 over-expression	↓	↓	↓	↓	FOM	+	-	+++	Refractory	+++
H400	SOX2 oncogenic function	SOX2 knockdown	↑	↑	↑	↑	AP	-	-	++	Marked inhibition	-

(FOM – floor of mouth; AP – anterior palate)

N.B. Both lines were derived from pStage II moderately differentiated squamous cell carcinoma.

6.4 Generation of tetracycline-inducible plasmid constructs

Generation of the PAX9 over-expression cell line was initially planned using the T-REx™ system (Invitrogen, Life Technologies, Paisley, UK). The T-REx™ system is a tetracycline-regulated mammalian expression system in which the chosen cell line is serially transfected with two separate plasmids. The first is the regulator plasmid, pcDNA6/TR©, which generates high levels of tetracycline-repressor (TetR) protein under the control of a cytomegalovirus (CMV) promoter sequence. The second plasmid, pcDNA™4/TO/lacz, contains the gene of interest located downstream to a 'Tetracycline-on' (TetO) sequence.

In the absence of tetracycline, the pcDNA6/TR© regulator produces TetR molecules which subsequently form a homodimer. The homodimer binds to the TetO sequence in the main pcDNA™4/TO/lacz plasmid, blocking transcription of the gene of interest. When added, tetracycline binds to the TetR molecules and induces a conformational change, forcing cleavage of the homodimer and dissociation from the pcDNA™4/TO/lacz plasmid. The gene of interest is therefore de-repressed and subsequently transcribed.

The experimental pcDNA™4/TO/lacz plasmid was constructed in three stages (Figure 6-4). In the first stage, human PAX9 cDNA was ligated into the Zero Blunt® pCR vector (Invitrogen, Life Technologies, Paisley, UK) alongside a Kozak sequence. The plasmid product was expanded through transformation into TOP10 chemically competent E. coli (Invitrogen Life Technologies, Paisley, UK). In the next stage, the final cassette was constructed by ligating the PAX9-kozak sequence with internal ribosomal entry site (IRES) and enhanced green fluorescent protein (EGFP) sequences in the second vector. The cassette was verified by analytical restriction digests and then inserted into the T-REx™ pcDNA™4/TO/lacz plasmid in the third stage (Figure 6-4).

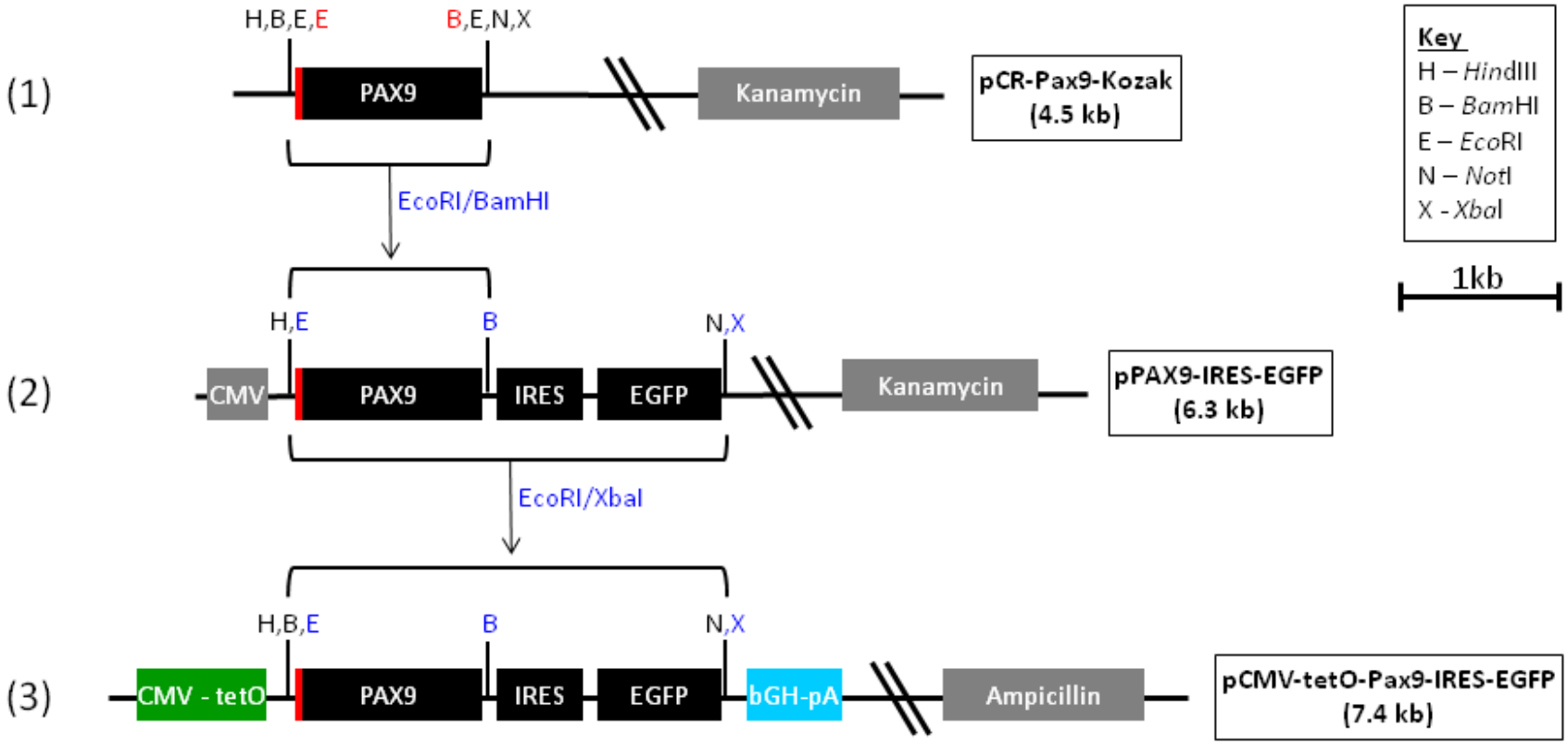


Figure 6-4 Generation of the T-REx™ tetracycline-inducible PAX9 pcDNA™4/TO/lacz plasmid

The PAX9/Kozak sequence was expanded and purified, then inserted alongside an IRES-EGFP sequence in the second vector to construct the final cassette. After being verified by analytical restriction digests, the cassette was then inserted into the pcDNA™4/TO/lacz plasmid.

(IRES – internal ribosome entry site; EGFP – enhanced green fluorescence protein; CMV – strong human cytomegalovirus; bGH-pA - reverse primer sequence. The red line at the start of the PAX9 sequence indicates the position of the Kozak sequence. Kanamycin and ampicillin antibiotics were used to select successfully transformed clones.)

6.5 Plasmid transfection of selected cell lines

The cell line transfections were out sourced to Dundee Cell Products Ltd (DCP). In the T-Rex™ system, the first of the serial transfections introduces the regulatory tetracycline-repressor plasmid, DNA vector pcDNA6/TR©, into the chosen cell line. In order to increase the transfection efficiency, DCP scientists first linearized the pcDNA6/TR© vector using the SapI restriction enzyme. Further analytical digests were performed to verify that the correct vector was being used for the transfections. Approximately 150-200 µg of plasmid vector pcDNA6/TR© DNA was used for each transfection. DCP scientists used two transfection methods: Lipofectamine® 2000 Transfection Reagent (Life Technologies, Paisley, UK) (Felgner and Ringold, 1989) and calcium phosphate (Chen and Okayama, 1987).

6.5.1 Low transfection efficiency of H314 and H400 cell lines

A control plasmid was used to determine the transfection efficiency of both the H314 and H400 cell lines. DCP routinely use a commercially-available control plasmid to determine the transfection efficiency of a range of different mammalian cell lines. The control plasmid includes a green fluorescent protein sequence (Vivid Colors™, Life Technologies, Paisley, UK) that enables transfected cells to be detected by immunofluorescence.

Immunofluorescence following transfection with the control plasmid showed low transfection efficiency for both H314 (~1-2%) and H400 (~10-20%). DCP reported that a large number of cells were used for the transfections in order to increase the probability of transfecting some of the cells.

6.5.2 Blasticidin selection of transfected clones

Following transfection with the regulatory pcDNA6/TR© plasmid, the cell lines were grown in the antibiotic blasticidin to select the transfected clones. Prior to selection, DCP performed triplicate kill curve analysis of the H314 and H400 cell lines to determine their blasticidin tolerance. The manufacturer's recommended range of concentrations for use in mammalian cell lines is 1 – 10 µg/ml (InvivoGen, Toulouse, France). Scientists at DCP initially established the kill curve at a higher range of between 2 and 60 µg/ml.a

Subsequent to their kill curve analysis, DCP used blasticidin at a concentration of 10 µg/ml for selection. However, selection with 10 µg/ml blasticidin following the first attempt at the transfections was cytotoxic. Following later attempts at the transfections, a lower concentration of 5 µg/ml blasticidin was used. However, selection at this lower concentration was also cytotoxic. The absence of blasticidin-resistant clones, combined with low transfection efficiency of the control plasmid, suggested that the transfections had failed to introduce the regulator pcDNA6/TR© plasmid into either of the chosen cell lines, consequently, it was decided to terminate the transfection experiments.

6.6 Lentivirus transfection of selected cell lines

Lentivirus transfections are known to have higher transfection efficiency than liposome-based delivery systems (Ambrosini et al., 1999; Wiznerowicz and Trono, 2003). Lentivirus transfections were carried out in a level 3 containment tissue culture facility in the School of Medial Sciences, Newcastle University, in collaboration with the Dermatology Research group. The original experimental design was modified in order to maximise the long-term value of this collaboration. It was decided to transfect both parent cell lines (H314, H400) with both the PAX9/SOX2 knockdown particles and PAX9/SOX2 over-expression particles.

Generation of the PAX9 and SOX2 knockdown cell lines involved a single stage transfection with pTRIPZ™ lentivirus particles (Thermo Scientific, Life Technologies, UK). The pTRIPZ™ particles have a ‘tetracycline-on’ (TetO) design. Custom made pTRIPZ™ lentivirus particles were generated by Thermo Scientific. The custom made particles contained PAX9 and SOX2 short hair-pin inhibitory RNA (shRNAi) sequences downstream to a TetO sequence. In the pTRIPZ™ system, addition of tetracycline to transfected cells induces transcription of shRNAi to the gene of interest. The shRNAi subsequently inhibits transcription of the gene of interest, resulting in knockdown of endogenous protein expression. The pTRIPZ™ particles include a puromycin resistance cassette to facilitate selection and a red fluorescence protein (RFP) cassette to help visualise the transfected clones.

Generation of the PAX9 and SOX2 over-expression cell lines involved serial transfections with two lentivirus particles. In the first stage, the cell lines were transfected with ready-made Tetracycline-repressor lentiviral particles (Amsbio, AMS Biotechnology, UK). The lentiviral particles included a blasticidin-RFP fusion gene vector that facilitates selection and visualisation of transfected clones. The transfected clones subsequently produce tetracycline-repressor (TetR) protein. In the second stage, these TetR protein producing clones are transfected with custom-made lentiviral particles which contain a tetracycline/cytomegalovirus (Tet/CMV) sequence upstream of a PAX9/SOX2 sequence (PAX9 lentiviral particles - Amsbio, AMS Biotechnology, UK; SOX2 and non-targeting control particles – addgene, Cambridge, MA, USA). In the absence of tetracycline, transcription of PAX9/SOX2 is inhibited by binding of the TetR protein to the Tet/CMV sequence. Addition of tetracycline induces a conformational change in the TetR protein, which then dissociates from the Tet/CMV sequence. The Tet/CMV sequence subsequently promotes transcription of the target gene. The custom-made particles also contain a puromycin resistance cassette to facilitate selection following the second transfection.

The experimental outline for the lentivirus transfections is summarised in the appendix. The cell lines were transfected with non-targeting controls in both the knockdown and over-expression experiments.

6.6.1 Tolerance of parent H314 and H400 to puromycin and blasticidin

The parent H314 and H400 cell lines were thawed down in the Dermatology Research tissue culture facility. When sufficiently expanded, the cells were seeded into 24-well plates at a density of 18,000 cells/well. In order to determine the tolerance of the parent lines to puromycin and blasticidin, cells were treated with variable concentrations of each antibiotic (puromycin 0 – 4 µg/ml; blasticidin 0 – 10 µg/ml). The cells were treated in two groups: one was treated with antibiotic at seeding, the other was treated 24 hours after seeding. Antibiotic culture medium was replenished at intervals of 48 and 72 hours.

Puromycin was required for selection of transfected cells in both knockdown and over-expression experiments. H314 was more sensitive to puromycin than

H400, with a more rapid cytotoxic response at high concentrations. However, at 4 day following puromycin treatment, both H314 and H400 cell lines showed a cytotoxic response at concentrations at/above 0.5 µg/ml puromycin (H400 shown in Figure 6-5). This concentration was therefore used for selection following transfections.

Both the H314 and H400 cell lines showed similar tolerance to blasticidin, with a cytotoxic response at concentrations above 5 µg/ml [data not shown]. It was therefore decided to use 5 µg/ml blasticidin to select transfected clones following stage one of the over-expression experiment. This was the concentration that had been used by DCP following transfection with pcDNA6/TR© plasmid.

Both antibiotics showed a similar cytotoxic effect at day four, irrespective of whether treatment started at seeding or at 24 hours after seeding.

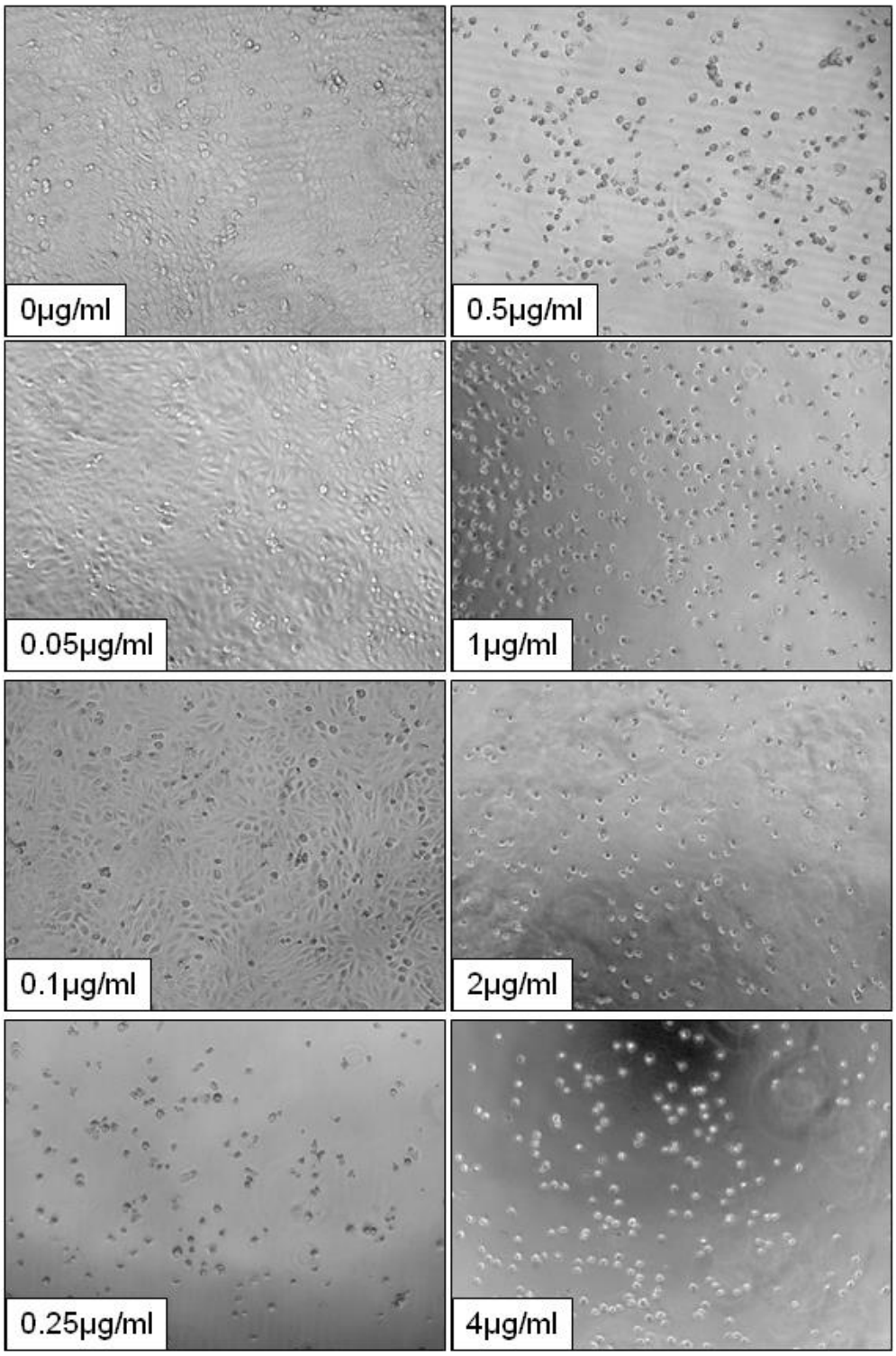


Figure 6-5 H400 cells at day four following puromycin treatment at seeding

At day four, untreated cells and those treated with the two lowest concentrations of puromycin (0.05 µg/ml and 0.1 µg/ml) were confluent. Cell growth was reduced at a concentration of 0.25 µg/ml puromycin; however, there were a few viable cells at this concentration. Puromycin was cytotoxic at concentrations of 0.5 µg/ml and above.

Images taken at x100 magnification.

6.6.2 PAX9 and SOX2 knockdown transfections

H314 and H400 parent cell lines were transfected with PAX9, SOX2, and non-targeting control pTRIPZ™ lentivirus particles according to the standard protocol used in the Dermatology Research tissue culture facility (Chapter 2, section 2.9). The transfected cells were seeded into T175 flasks and treated with 0.5 µg/ml puromycin. Unfortunately, the cells did not survive following the first attempt. There were a number of possible reasons for this: 1) the cells may have been trypsinised too soon following transfection; 2) the seeding density may have been too low; 3) puromycin selection may have been started too early.

The transfections were repeated under the same conditions. However, following transfection the conditions were adjusted in order to give the cells more time to recover and increase their probability of surviving puromycin selection. Specifically, the cells were 1) incubated for 24 hours in the six-well transfection plates prior to being trypsinised; 2) seeded into smaller T25 flasks after passaging; and 3) puromycin selection was started 24 hours following seeding.

The second attempt at transfecting both parent cell lines was successful. At day two following the start of puromycin selection, examination under an inverse light microscope showed a mixed population of cells that included adherent, viable cells and detached, non-viable cells. After one week of puromycin selection, the cells were expanded in normal culture medium. At the next passage, the transfected cells were seeded into six-well plates at a density of $1x10^5$ cells per well. At 24 hours, the culture media was replaced and the cells treated with doxycycline at six different concentrations (ranging from 0 to 1 µg/ml). When the majority of wells were confluent (H400 – day four; H314 – day six), they were examined using an inverse fluorescence microscope. Transfected cells treated with diluent alone were negative. By contrast, each of the transfected lines treated with doxycycline at a concentration of 25 ng/ml or more showed a red fluorescent signal (H400 SOX2 TRIPZ ™ shown in Figure 6-6). Cells treated with the highest doxycycline concentration (1 µg/ml) showed the most intense signal, but were generally less confluent than cells treated at concentrations at/below 500 ng/ml. This suggests that high doxycycline concentrations may have a mild cytotoxic effect.

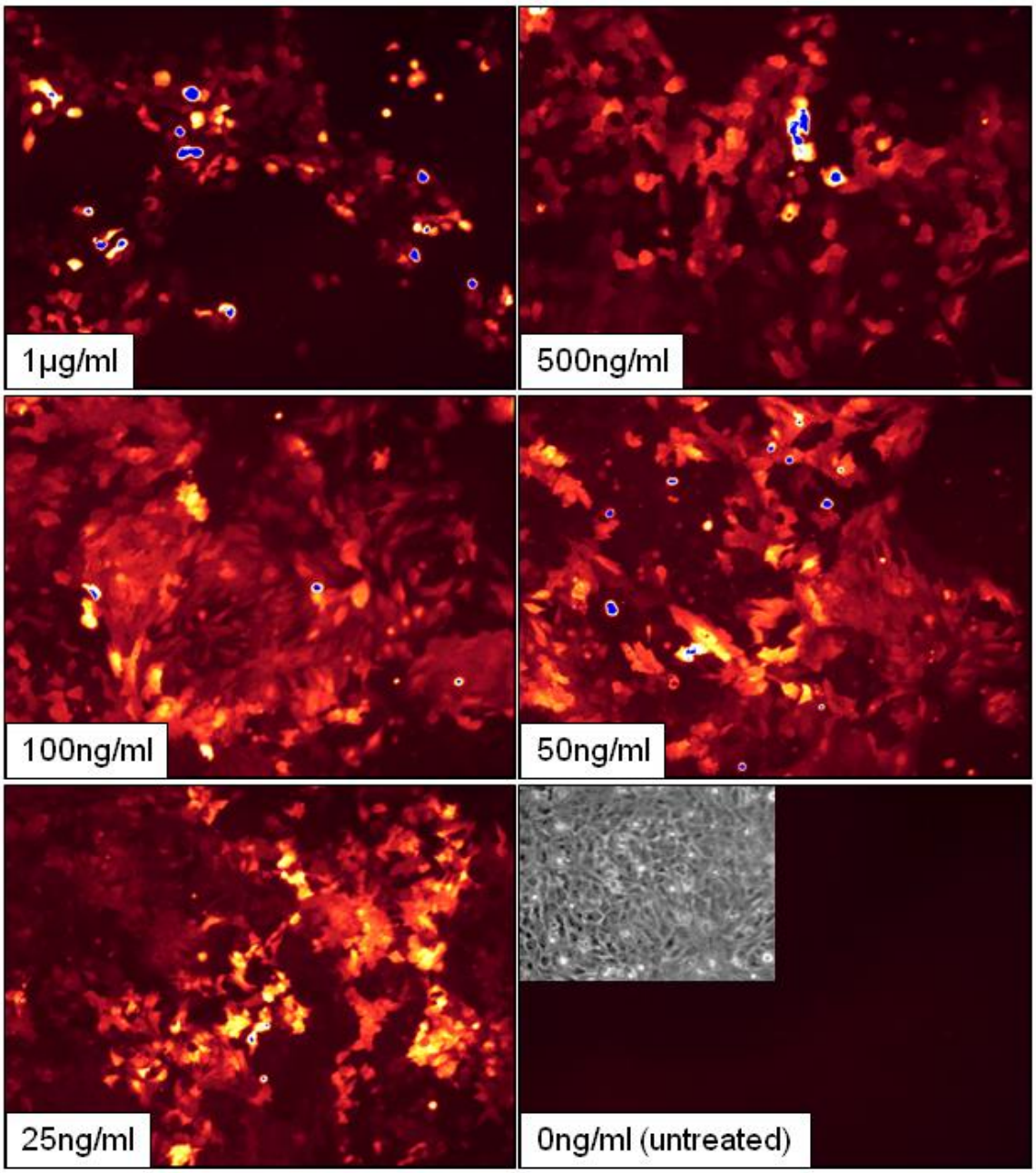


Figure 6-6 Immunofluorescence of H400 SOX2 pTRIPZ™ cells after four days of induction with doxycycline

H400 SOX2 pTRIPZ™ cells showed a positive RFP signal at day four following administration of doxycycline. Cells treated with the highest concentration (1 µg/ml, top left panel) showed the most intense signal, but were less confluent than cells treated at/below a concentration of 500 ng/ml doxycycline. Cells that were not treated with doxycycline failed to show a red fluorescent signal (lower right). The light microscopy image (inset) confirms that this untreated group were confluent at the time of imaging.

Images taken at x100 magnification.

6.6.3 PAX9 and SOX2 over-expression transfections

Amsbio were unable to offer guidelines as to the quantity of the TetR lentiviral solution (LVP017-Bsd-RFP) required for the first stage transfection of our chosen cell lines. The solution contained a high concentration of lentiviral particles (1x 10^7 IFU/mL). As there had been a high transfection efficiency using the pTRIPZ™ system, it was decided to treat each well with a small amount (5 µl) of the lentiviral particles during the first attempt. H314 and H400 cells were seeded into four wells of a six-well plate (100,000 cells per well) for transfection with TetR lentiviral particles. The transfection was performed under standard conditions as described in Chapter 2, section 2.9.

Following the transfection, the cells were grown in culture containing 5 µg/ml blasticidin to select the transfected cells and induce the RFP cassette. However, at day four examination on the immunofluorescence microscope was negative [data not shown]. The cells were viable and seemed to be resistant to blasticidin. Selection was continued and the cells incubated under standard conditions. However, further examination at days six and ten also failed to show a positive red fluorescent signal [data not shown]. The negative results were interpreted to mean that the transfections had failed due to treatment with an insufficient quantity of the lentivirus solution.

The transfections were repeated using a larger quantity of the TetR lentiviral solution. H314 and H400 cells were seeded into two wells of a 24-well plate (1.8 x 10^4 cells/well). When suitably confluent, the cells were treated with 40 µl of the lentivirus solution. The transfection was otherwise performed according to the standard protocol. However, following several days of blasticidin selection the cells were again viable in the absence of a positive fluorescence signal.

Amsbio were contacted for further advice. They reported that the RFP sequence was detectable on their microscopes using a modified DS Red filter at an excitation of 545 nm and emission of 620 nm. A TRITC filter had been used to detect the positive fluorescence signal from the pTRIPZ™ lines. It was therefore possible that a positive signal was being missed due to a different filter being used.

Unfortunately, a DS red filter was not available for use on our inverse microscope. The transfected cells were grown as monolayers in chamber slides. After three days, the cell monolayers were fixed and mounted with cover slips. The slides were then viewed using a DS Red filter in a Nikon confocal light microscope system. This detected a weak red signal in each of the transfected lines, including the clones originally treated with just 5 µl of the TetR lentivirus solution [data not shown].

The transfected H314TetR and H400TetR clones subsequently underwent the second stage in which they were transfected with lentivirus particles encoding PAX9, SOX2, and non-targeting sequences. Following the second transfection, cells were selected in puromycin at a concentration of 0.5 µg/ml. After day two of puromycin selection, there were mixed populations of cells including adherent, viable cells and detached, non-viable cells. Following one week of puromycin selection, the cells were expanded, frozen down and archived under liquid nitrogen.

6.7 Characterisation of H400 SOX2 pTRIPZ cell line

Full characterisation of the transfected cell lines was outside the remit of the thesis; however, proof of principle that the expression constructs were functional was explored using one exemplar cell line. This section outlines the initial characterisation of H400 SOX2 pTRIPZ™ cell line.

6.7.1 Semi-quantitative reverse transcription PCR

The H400 SOX2 pTRIPZ™ line was seeded into a six-well plates and treated with doxycycline at the concentrations summarised in section 6.6.1. When confluent, RNA was extracted from the cells using Trizol. The RNA was then reverse transcribed (Chapter 2, section 2.10.3) and cDNA samples used for semi-quantitative PCR analysis. For each sample, two PCRs were performed: one using primers that amplified SOX2 and a second using primers that amplified the housekeeper β-actin (ACTB RNA).

The samples were run on a 2% agarose gel. Each showed bands at the expected lengths (SOX2 - 191 base pairs; ACTB – 550 base pairs). For SOX2, there was an inverse relationship between band intensity and doxycycline concentration. The band intensity increased as the doxycycline concentration decreased. The untreated H400 SOX2 pTRIPZ™ sample showed the highest band intensity, which was similar to that of the non-transfected parent line (Figure 6-7A). There was some variation in the intensity of the ACTB band, but this did not show a specific correlation with doxycycline concentration (Figure 6-7B). However, variable ACTB band intensity suggests the samples may have contained different quantities of total RNA.

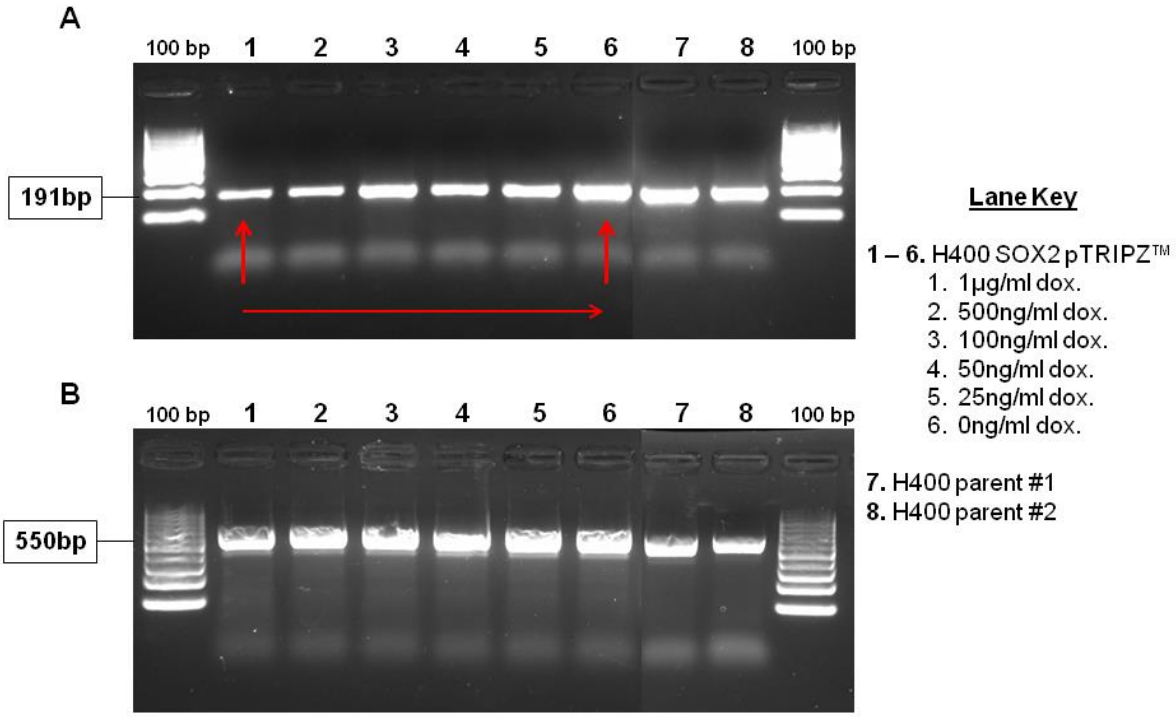


Figure 6-7 Gel Doc™ image showing RT-PCR product of H400 SOX2 pTRIPZ™ lines treated with doxycycline and two non-transfected parent lines

A) SOX2 assay. Lanes 1 – 6 show an inverse relationship between SOX2 band intensity and doxycycline concentration. The band intensity is lowest in the sample treated with 1 µg/ml doxycycline (lane 1, arrowed) and increases as the concentration of doxycycline decreases. Band intensity is highest in the untreated sample (lane 6, arrowed), which is similar to that of the native H400 cell line samples (lanes 7, 8).B) ACTB assay. The samples show some variation in the ACTB band intensity, suggesting they contain variable quantities of total RNA.

6.7.2 Quantitative PCR

Following semi-quantitative RT-PCR analysis, the cDNA samples were amplified by real-time quantitative PCR (qPCR) to provide a more accurate assessment of SOX2 knock down in the H400 SOX2 pTRIPZ™ cell line. A 96-well plate was set up with triplicate repeats for each sample. The plate was run on an Applied Biosystems® 7500 Real-Time PCR machine (Life Technologies, Paisley, UK). Cycle threshold values were used to calculate mean fold values as described in Chapter 2 (section 2.10.5). Samples were then ranked according to their mean fold value (Table 6-3).

The non-transfected H400 parent cell lines had the highest ranking, followed by the non-treated H400 SOX2 pTRIPZ™ sample. Overall, the mean fold values showed the same trend that had been identified by semi-quantitative RNA analysis, i.e. there was an inverse relationship between the mean fold value and doxycycline concentration. However, several of the mean fold values did not fit this relationship in a continuous pattern. Samples 2 and 5 (treated with 500 and 25 ng/ml doxycycline) had the same mean fold value, as did samples 3 and 4 (treated with 100 and 50 ng/ml). These outliers may have been caused by variation in the quantity of ACTB RNA.

Table 6-3 Ranked qPCR mean fold values for H400 SOX2 TRIPZ™ samples and two non-transfected H400 parent cell lines

Sample number*	Mean fold value	Rank**
1	0.07	6
2	0.17	4
3	0.11	5
4	0.11	5
5	0.17	4
6	0.22	3
7	1.00	2
8	1.02	1

*As per lane key in Figure 7; ** Ranked high =1; low = 8

6.7.3 SOX2 protein quantification by Western blot

The transfected lines were seeded into six-well plates and treated with doxycycline at the same concentrations as those used for RNA preparation. Protein was extracted from the transfected lines using RIPA buffer (Sigma Aldrich, UK; Chapter 2, section 2.10.6). Western blotting was performed on each of the samples using antibodies to both SOX2 and GAPDH (glyceraldehyde 3-phosphate dehydrogenase).

Analysis of band intensity confirms the inverse relationship between SOX2 band intensity (expected band width – 35 kDa) and doxycycline concentration. Band intensity is weakest for the sample treated with the highest concentration of doxycycline (1 µg/ml) and increases as the concentration of doxycycline decreases. Untreated H400 SOX2 pTRIPZ™ had the highest band intensity, similar to that of the non-transfected parent lines (Figure 6-8A). The Western blot for GAPDH (expected band length of 37 kDa) shows a similar band intensity for each sample (Figure 6-8B). This indicates that there were similar quantities of total protein in each of the samples. Overall, comparison of relative SOX2 protein expression by Western blot confirmed the trends identified by comparison of SOX2 RNA. Together, the results indicates that doxycycline administration induces knockdown of SOX2 expression in H400 SOX2 pTRIPZ™.

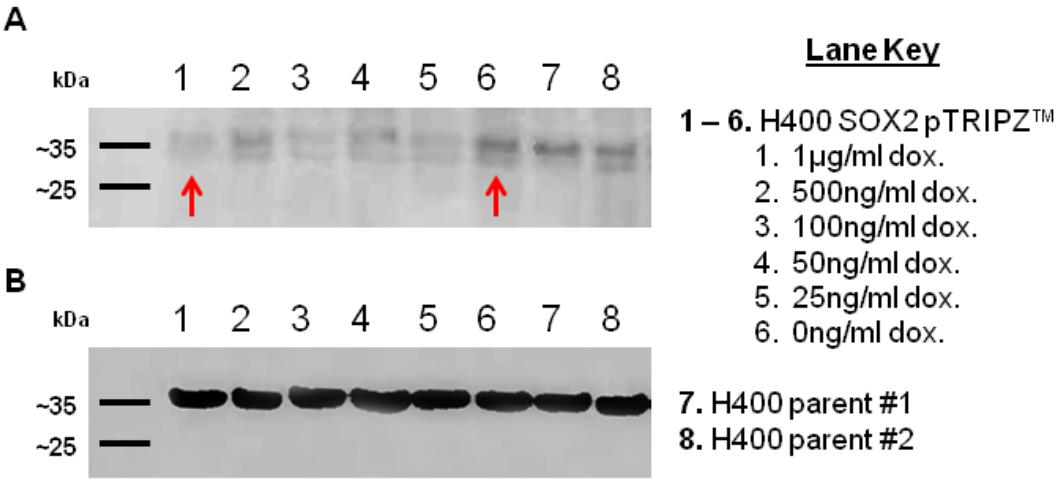


Figure 6-8 Western blot showing SOX2 and GAPDH protein levels in H400 SOX2 pTRIPZ™ cells treated with doxycycline and non-transfected parent lines

A) SOX2 assay. At the expected band length of 35 kDa, there was an inverse relationship between band intensity and doxycycline concentration. Band intensity is lowest in the sample treated with 1 µg/ml doxycycline (lane 1, arrowed) and increases as the concentration of doxycycline reduces. Intensity is highest in the untreated sample (lane 6, arrowed), which shows a similar band intensity to that of the parent H400 cells (lanes 7, 8).There are non-specific bands at ~45 and 55 kDa, which show the opposite relationship to doxycycline concentration.

B) GAPDH assay. The Western blot shows similar GAPDH band intensity (expected band length 37 kDa) for each sample, indicating that they contained similar quantities of total protein.

6.8 Discussion around the limitations of the work carried out to date and suggestions for future work

6.8.1 Endogenous expression of PAX9 and SOX2 in parent OSCC lines

Identification of suitable OSCC cell lines for transfection relied on two assays, 1) semi-quantitative analysis of reverse transcription PCR (RT-PCR) data to compare relative RNA expression, and 2) manual immunohistochemistry to compare relative protein expression. These techniques were used as they were already established in our laboratory and had been optimised on a range of other samples (e.g. the protocol for manual PAX9 immunohistochemistry had been performed successfully on both human and mouse tissue).

Quantitative real-time PCR (qPCR) analysis of RNA and Western blot analysis of protein levels may have provided more accurate data and helped to distinguish more reliably between the endogenous expression of PAX9/SOX2 in each cell line. However, protocols for these techniques were not developed at that stage in the experiment. To do so may have delayed starting the transfections. The protocols used in the initial characterisation of the pTRIPZ™ cell lines still require modification (see next points). When these protocols are optimised, qPCR and Western blots may be performed on each of the candidate OSCC cell lines in order to verify our selection of H314/H400.

6.8.2 Failure of PAX9 and SOX2 immunohistochemistry on cell pellets

It remains unclear why PAX9 and SOX2 immunohistochemical staining was unsuccessful using cell pellets. This finding is particularly unusual given that the cell pellets stained positively for AE1/AE3, CK5/6, and p63 using the automated Ventana platform. It may be that the PAX9/SOX2 antigens are more sensitive than cytokeratin/p63 either to trypsinisation or specific steps in the preparation of cell pellets.

Since the initial screening of the cell lines, SOX2 immunohistochemistry has been optimised on the automated Ventana platform. In future, it will therefore be possible to perform SOX2 immunohistochemistry on cell pellets using the Ventana platform. Similarly, a new PAX9 antibody that may be suitable for use

on the Ventana platform has been now identified. Further work could also be carried out to optimise the preparation of the cell pellets.

6.8.3 Cytoplasmic immunostaining using Surepath™ preparations

Stronger immunostaining was achieved using Surepath™ preparations compared to the cell pellets. However, the expected nuclear signal was accompanied by non-specific cytoplasmic staining. This was particularly noticeable for PAX9. It may have contributed to the high PAX9 PPN and 3+PN values generated by digital image analysis. This non-specific staining may reflect cellular disruption caused during trypsinisation or preparation of the Surepath™ slides.

The protocol for preparing the Surepath™ slides has already been optimised for use on human cells. It is used regularly in the Department of Cellular Pathology for diagnostic immunocytochemistry. It is therefore unlikely that the existing protocol can be further optimised to reduce the cytoplasmic staining. If improved immunostaining of cell pellets can be achieved by using the Ventana platform, it may not be necessary to continue using Surepath™ in future. If Surepath™ slides are used again, the immunostaining protocol could be adjusted by reducing the antigen and/or DAB incubation times.

6.8.4 Housekeeper controls for RT-PCT, qPCR, and Western blotting

During initial characterisation of the H400 pTRIPZ™ line, the housekeeper β-actin (ACTB RNA) was used as a control for RT-PCR/qPCR whereas GAPDH was used for Western blotting. ACTB was chosen as the control for RT-PCR and qPCR as the ACTB RT-PCR protocol had already been optimised in our laboratory. It had also been used during the screening of candidate OSCC cell lines. However, there is anecdotal evidence that ACTB is sensitive to doxycycline [Martina Elias, personal communication]. Sensitivity to doxycycline may account for the contrast between the consistent ACTB band intensity detected while screening native OSCC lines, and the variable band intensity detected while characterising doxycycline-treated H400 SOX2 pTRIPZ™ cells. It may also account for the aberrant mean fold values detected by qPCR. Together, the RT-PCR and qPCR results suggests that doxycycline impacts

ACTB transcription and is therefore an unsuitable control for induction experiments.

Western blot analysis showed consistent GAPDH protein band intensity. This suggests that the samples had even quantities of total protein and that GAPDH is more resistant to doxycycline than ACTB. Further work is required to optimise a protocol for RT-PCR and qPCR using GAPDH primers. However, GAPDH may be a more suitable housekeeping control in future experiments.

6.8.5 Quantity and quality of RNA/protein samples

There has been some difficulty in extracting RNA from the cell lines prior to RT-PCR and qPCR. The RNA yield from the H314 cell lines has been particularly low. The current RNA extraction technique TRIzol® Reagent (Life Technologies, Paisley, UK). TRIzol® is used to lyse and extract RNA from tissue samples. It may be too strong to extract RNA from cultured cells. An alternative reagent designed specifically for extracting RNA from cultured cells, such as Reliaprep™ RNA Cell Miniprep System (Promega, UK), may yield greater quantities of RNA. Cells could also be grown in larger wells and left until they had achieved 100% confluence prior to extracting RNA.

6.8.6 PAX9 antibody for Western blotting

SOX2 Western blotting has successfully used the same antibody used for manual immunohistochemistry. By contrast, PAX9 Western blotting has been unsuccessful to date. The PAX9 antibody used for manual PAX9 immunostaining is a monoclonal rat antibody produced from hybridoma cell lines (Abcam, Cambridge, UK). As it is a cell supernatant, the antibody's precise concentration is unknown. This makes it difficult to calculate appropriate dilutions for use in Western blotting. In immunohistochemistry, is effective at a 1:40 concentration. However, similarly high concentrations have not worked in the Western blot. An alternative rabbit monoclonal PAX9 antibody has been now identified (D9F1N, Cell Signalling Technology®, US). Further work is required to optimise the protocol for this antibody before PAX9 knockdown can be verified in either the H314 PAX9 pTRIPZ™ or H400 PAX9 pTRIPZ™ cell lines.

6.8.7 Polyclonal population of transfected cells

The transfected cell lines currently comprise a polyclonal population of cells in which lentiviral DNA has integrated at different sites in the host genome. Depending on the experimental model, it may be preferable to work with a polyclonal population as problems arising from specific genome integration sites are avoided. However, induction/knockdown of gene expression in a polyclonal population can be inefficient. There may also be 'leaky' non-specific expression in the absence of the inducing agent [Prof. Josef Heidenreich, personal communication].

As the initial characterisation is incomplete, it is too early to say whether either of these potential problems will have a significant impact on our cell lines. However, it may prove necessary to work with defined clones in the future. Individual clones could be selected by seeding the cells in petri dishes at low density. After being selected in a blasticidin/doxycycline culture, cloning rings can then be applied around colonies with strongly-positive fluorescent signals. When colonies are confluent within the cloning ring, they can then be trypsinised and transferred to a 24-well plate for further expansion. It has been recommended that a minimum of three clones should be characterised and used for functional assays. Using an isolated clone may yield aberrant results that reflect the impact of the integration site as opposed to the impact of modifying expression of the gene of interest [Prof. Josef Heidenreich, personal communication].

6.8.8 Functional assays

Once the polyclonal cell lines have been characterised it will be possible to begin functional assays. We plan to analyse basic cellular parameters of the transfected cells and compare these with native, non-transfected control lines. These parameters include determination of cellular viability, using the MTT assay, and cellular proliferation and apoptosis, using fluorescence-activated cell sorting (FACS) analysis. In the longer term, the genetic pathways regulating the molecular changes identified in the inducible PAX9 and SOX2 OSCC cell lines could be delineated by DNA microarray analysis. This would generate genome

wide expression profiles, which can be analysed by bioinformatics. Significant changes in gene expression can be further validated by qPCR.

6.9 Conclusion

This study identified cell lines with low endogenous PAX9 expression (for transfection with an inducible PAX9 over-expression plasmid) and high endogenous SOX2 expression (for transfection with an inducible SOX2 'knock-down' plasmid). H314 was selected for inducible PAX9 expression and H400 was selected for SOX2 knockdown. Poor transfection efficiency using plasmid DNA lead to the adoption of an alternative transfection strategy using lentivirus vectors. The experimental plan was further modified so that both cell lines were transfected with PAX9 over-expression and SOX2 knockdown constructs.

Using an exemplar cell line, H400 SOX2 pTRIPZ, we have demonstrated proof of principle that the expression constructs are functional. Knockdown of SOX2 expression using doxycycline induction has been confirmed by RNA and protein assays. However, the initial characterisation was limited by the use of sub-optimal housekeeper genes to act as controls, and technical difficulties which compromised the quantity and quality of RNA that was extracted from some cell lines. The transfected cell lines are currently polyclonal, which potentially limits their utility for functional assays and future *in vivo* experiments. Further work is therefore required to optimise the analysis of the transfected cells lines and to establish a range of characterised clones for each cell line.

Chapter 7. Discussion

7.1 Introduction

Oral squamous cell carcinoma (OSCC) is a major global healthcare problem (Jemal et al., 2011). The incidence of OSCC in the UK is increasing, a trend which is likely to continue as the population ages (Cancer Research UK, 2010). OSCC has devastating consequences for many patients diagnosed with the disease. Cure is often precluded either by late presentation (McGurk et al., 2005) or formation of second primary tumours (Day and Blot, 1992; Tsou et al., 2007). The current treatment modalities for OSCC - surgical resection and radiotherapy – are associated with complications and functional impairments that may result in psychological problems and even non-cancer-related mortality (Zwahlen et al., 2008; Shah and Gil, 2009; Campos et al., 2014; Szczesniak et al., 2014).

Outcomes for patients with OSCC may be improved if the disease is identified in its precursor stages, termed oral potentially malignant disorders (OPMD) (van der Waal, 2009; Goodson and Thomson, 2010). However, histological assessment of OPMD is subjective and does not reliably predict which cases will progress to OSCC (Lodi et al., 2006; Kujan et al., 2007). Several candidate biomarkers have emerged in recent decades but, as yet, none are validated as reliable biomarkers for use in routine diagnostic practice (Nylander et al., 2000; Kovesi and Szende, 2003; Varun et al., 2014).

This study sought to address the continuing need for biomarkers that stratify OPMD according to their risk of malignant transformation and thus facilitate early detection and management of OSCC (Mishra, 2012). This discussion highlights the key findings of this study regarding the classification of epithelial dysplasia, and the potential diagnostic utility/functional significance of each biomarker evaluated: EGFR, SOX2, and PAX9.

7.2 Key findings

7.2.1 The binary classification of epithelial dysplasia is more predictive of clinical course than the SIN classification

Our data show that the binary classification of epithelial dysplasia is more predictive of clinical outcome than the Squamous Intra-epithelial Neoplasia (SIN) classification. High-grade epithelial dysplasia correlated with malignant transformation in OPMD, whereas SIN 3 did not. The absence of a correlation with SIN 3 is likely to reflect the subset of cases that underwent malignant transformation but were classified with SIN 2 on morphological assessment.

There is evidence that removal of the intermediate SIN 2 category, equivalent to moderate epithelial dysplasia in the World Health Organisation (WHO) classification, reduces inter-observer variation and enhance the predictive value of the assigned dysplasia grade (Barnes et al., 2005; Kujan et al., 2006). Our data support this contention, but must be interpreted with some caution. The correlation between high-grade epithelial dysplasia and malignant transformation was relatively weak compared with two of the biomarkers studied (see sections 7.2.2 and 7.2.9). Used in isolation, high-grade epithelial dysplasia had relatively low sensitivity and specificity for detecting cases destined to undergo malignant transformation. This supports the view that there are fundamental limitations to the predictive value of a morphological diagnosis, irrespective of the accuracy and reliability of the classification system upon which it is based (Kujan et al., 2007; Dost et al., 2014).

Our analysis of the mouse model of oral carcinogenesis did not classify cases according to a grade of epithelial dysplasia. Rather, cases were categorised simply according to whether epithelial dysplasia was present or absent. This methodology may have failed to detect subtle differences between Pax9-deficient mice and controls. For example, it is possible that Pax9-deficient mice developed high-grade epithelial dysplasia more frequently/rapidly than controls, a trend that our methodology would not have detected. However, there is no agreed classification system for grading oral epithelial dysplasia in mice. Morphological differences between the oral squamous epithelium of mice and humans confounds the direct transference of the human classification system to

mice, mainly due to the relatively thin epithelium in the mouse. Initial attempts by the scoring pathologists to assign grades of epithelial dysplasia in control mice were challenging and not reproducible. In Pax9-deficient mice, dysplasia grading was further complicated as the architecture of normal (i.e. non-dysplastic) squamous epithelium is already altered. For these mice, it is conceivable that architectural changes indicative of epithelial dysplasia may be either overlooked or identified erroneously. On balance, the risk that our methodology resulted in 'lost' data was outweighed by the enhanced reliability and reproducibility of our simple classification system. Moreover, the approach was in keeping with the principle of reducing the number of mice required for the experiment to a minimum (Understanding Animal Research, 2014). Many more mice would have been required to support valid statistical comparison of epithelial dysplasia stratified according to the WHO/SIN classification (Barnes et al., 2005).

7.2.2 EGFR gene copy number abnormalities detect cases of OPMD destined to undergo malignant transformation

OPMD with abnormal EGFR gene copy number were at greater risk of malignant transformation than cases with normal EGFR gene copy number. Although abnormal EGFR gene copy number used in isolation had relatively low sensitivity for identifying cases destined to undergo malignant transformation, its specificity was relatively high. The high specificity of abnormal EGFR gene copy number indicates that EGFR *in situ* hybridisation (ISH) may be a useful as part of a broader panel of biomarkers. Our algorithms show that when abnormal EGFR gene copy number and high-grade epithelial dysplasia were combined they were more reliably predictive than either used in isolation.

Our findings are in concordance with two recent studies which suggested that abnormal EGFR gene copy number is an early feature of OPMD destined to undergo malignant transformation and may precede gross EGFR genomic gain (Benchekroun et al., 2010; Poh et al., 2012). Evidence from both studies supported the inclusion of cases with trisomy and low polysomy alongside those with EGFR genomic gain as currently defined in the interpretation of EGFR ISH signals in non-small cell carcinoma of the lung (Nicholson et al., 2001; Hirsch et al., 2003). It remains unclear whether it is appropriate to apply criteria validated

for the interpretation of EGFR ISH signals in lung cancer to OSCC and OPMD. However, our data show that including cases with trisomy and low polysomy provide valuable prognostic information in the context of OPMD. Further work is required to establish agreed criteria for the interpretation of EGFR ISH signals in cancers and pre-cancerous lesions outside the lungs (section 7.3).

7.2.3 EGFR genomic gain is present in a quarter of early-stage OSCC but does not correlate with clinical outcome

EGFR genomic gain was detected in a quarter of early-stage OSCC in this study. This is higher than the rate reported in a comparable study of early-stage OSCC of the tongue (Rössle et al., 2013). The discrepancy may reflect methodological differences between the studies, specifically the use of tissue-microarrays (TMA) rather than whole sections, which risks failure to detect EGFR genomic gain in heterogeneous cases (Rössle et al., 2013). Our rate is actually towards the lower end of the range reported in previous studies of mixed-stage and late-stage OSCC (Freier et al., 2003; Chung et al., 2006; Agulnik et al., 2007; Temam et al., 2007; Pectasides et al., 2011; Ryott et al., 2009). This suggests that EGFR genomic gain is a relatively late event in oral carcinogenesis. However, the majority of OPMD with abnormal EGFR gene copy in the present study progressed to OSCC, and most of the corresponding OSCC biopsies showed a higher EGFR gene copy number category relative to the index OPMD. This suggests that the progressive accumulation of EGFR gene copy number aberrations is functionally significant in the transition from epithelial dysplasia to OSCC. Given the heterogeneity of EGFR gene copy number signals, it is conceivable that the low rate identified in early-stage OSCC in this study reflects a sampling error. Alternatively, it may also represent the selection of different clones as oncogenic genetic changes accumulate within the tumour.

Our analysis did not identify a correlation between EGFR genomic gain and clinical outcome in early-stage OSCC. By contrast, a previous study reported dramatically reduced five-year survival rates in patients with EGFR genomic gain compared to patients with a normal EGFR signal (Temam et al., 2007). Although the previous study used quantitative real-time PCR to detect EGFR genomic gain, its conclusions have been have been supported by other studies

that have assayed for EGFR genomic gain by ISH (Freier et al., 2003; Chung et al., 2006). However, each of these studies has examined OSCC at a wide range of clinical stages. The absence of a correlation between EGFR genomic gain and clinical outcome in the present study may therefore reflect its narrower inclusion criteria and focus on early-stage OSCC. Rössle et al (2013), in a study that was also confined to early-stage OSCC, failed to detect a correlation between EGFR gene copy number and clinical outcome.

7.2.4 EGFR protein over-expression is prevalent in OPMD and early-stage OSCC, limiting its clinical utility

This study confirms that EGFR protein is over-expressed in the majority of OPMD and early-stage OSCC (Grandis and Tweardy, 1993a; Ries *et al.*, 2013). Our results also show a positive correlation between EGFR protein expression and grade of epithelial dysplasia in OPMD. However, we were unable to replicate the significant correlation between EGFR over-expression and risk of malignant transformation that has been reported in two previous studies (Nankivell *et al.*, 2013; Ries *et al.*, 2013). Both earlier studies benefitted from whole section rather than TMA analysis. However, they relied upon subjective semi-quantitative assessment of EGFR protein expression rather than digital image analysis. Different thresholds may also have been used to define EGFR over-expression (Nankivell *et al.*, 2013; Ries *et al.*, 2013). On balance, the prevalence of EGFR over-expression in OPMD identified in the current study makes a positive correlation with clinical parameters unlikely.

EGFR over-expression was more prevalent in early-stage OSCC than in OPMD. The biological significance of this is uncertain. It supports the contention that EGFR is more likely to be functionally significant in the later stages of oral carcinogenesis (Ryott et al., 2009). However, the literature to date consistently draws attention to the complexity of the EGFR pathway. Increased EGFR expression may be a bystander change, reflecting, but not driving, oral carcinogenesis. This would account for the lack of correlation with disease-specific clinical outcomes in either OPMD or early stage OSCC. On balance, our findings support the view that EGFR protein has limited clinical utility as a biomarker for stratifying patient management (Forastiere, 2007; Gusterson and Hunter, 2009; Rosin and Califano, 2010).

7.2.5 SOX2 protein expression is heterogeneous in OPMD but does not correlate with clinical outcome

SOX2 protein has a heterogeneous expression profile in OPMD, but is down-regulated in the majority of cases. No correlation was identified between SOX2 protein expression and clinical outcome in OPMD.

To date, the profile of SOX2 expression in OPMD has been described in an isolated study (Qiao et al., 2013). SOX2-positivity was reported in 90% of OPMD. Co-expression of SOX2 and Oct4 - a feature not detected in the normal epithelium – was also identified in the majority of OPMD. Similar trends were detected in an animal model. However, the study was limited by several factors. These include small sample size, semi-quantitative scoring of stained sections, and categorisation of cases as either SOX2-positive or SOX2-negative, without indicating the intensity of staining. Moreover, the level of SOX2 expression in normal human epithelium was not reported as a control, precluding comparison of the relative SOX2 expression of OPMD (Qiao et al., 2013). It is therefore difficult to compare our data with the results of the previous study meaningfully. Further work is required to confirm the findings of our study. However, our analysis suggests that SOX2 has limited potential value as a prognostic biomarker in OPMD.

7.2.6 SOX2 protein expression is up-regulated in early-stage OSCC

SOX2 protein had a heterogeneous expression profile in early-stage OSCC. However, in the majority of cases there was a significantly increased proportion of strongly SOX2-positive nuclei relative to the normal epithelium.

Up-regulation of SOX2 has been documented previously in squamous cell carcinoma (SCC) of the lung, oesophagus, cervix, and penis (Hussenet et al., 2010; Lu et al., 2010; Maier et al., 2011). Increased SOX2 expression in OSCC has been reported in several studies by immunohistochemistry (IHC) (Freier et al., 2010; Huang et al., 2014), proteomic analysis (Misuno et al., 2013) and *in situ* hybridisation (ISH) (Freier et al., 2010; Kokalj Vokač et al., 2014).

Interestingly, Freier et al (2010) reported that SOX2 was over-expressed in only a small subset of cases. The study used a TMA-based approach to tissue

analysis; the heterogeneous pattern of SOX2 expression that we observed suggests that such data may be unreliable due to sampling limitations. Our data show that the proportion of strongly SOX2-positive cells was increased in OSCC relative to the normal epithelium in over half of cases, suggesting a 'gain of function' or an oncogenic effect in at least a proportion of oral cancers. Freier et al (2010) reported that SOX2 gene copy number amplification was present in more than half of cases. This suggests SOX2 genomic gain may be a more homogeneous feature in OSCC and not subject to the limitations of TMA analysis and classification. The proportion of OSCC with SOX2 gene amplification is consistent with our SOX2 protein expression assessed by IHC.

Our analysis shows that OSCC with adverse outcomes had significantly lower expression of SOX2 relative to cases with no adverse outcome. This is in contrast to three previous studies that show high SOX2 expression correlates with poor clinical outcome, particularly an increased risk of metastasis to cervical lymph nodes (Du et al., 2011; Michifuri et al., 2012; Huang et al., 2014). There are several possible explanations for this discrepancy: inclusion of late-stage OSCC in previous studies; use of manual staining methods rather than the automated platform used in the present study; use of different SOX2 antibodies, including polyclonal antibodies (Michifuri et al., 2012; Huang et al., 2014); finally, the use of semi-quantitative scoring as opposed to digital image analysis. As an isolated finding, however, the correlation between low SOX2 expression and adverse outcomes in OSCC detected in our study should be interpreted with caution.

Michifuri et al (2012) also highlighted the heterogeneous profile of SOX2 expression and proceeded to categorise cases according to whether SOX2 was expressed peripherally or diffusely throughout the tumour. Our study includes cases consistent with both patterns. However, this binary classification system does not account for several cases in our study in which SOX2 expression was generally down-regulated/lost or showed a variable staining that was not consistent with either pattern. The authors of the previous study reported a correlation between the pattern of SOX2 staining and specific clinical outcomes. It is not possible to comment on a potential correlation between staining pattern

and clinical outcome in our cases due to the small number with uniform patterns and the generalised heterogeneity of SOX2 staining identified.

7.2.7 SOX2 may be significant in the maintenance and functioning of cancer stem cells

In humans, SOX2 expression was significantly increased in OSCC relative to OPMD. In both Pax9-deficient and control mice, Sox2 expression was increased in epithelial dysplasia compared to normal epithelium. These findings are concordant with the literature suggesting SOX2/Sox2 has an oncogenic function in oral carcinogenesis (Freier et al., 2010; Qiao et al., 2013; Kokalj Vokač et al., 2014). However, this hypothesis was not borne out by analysis of SOX2 expression in the subset of human OPMD that underwent malignant transformation and their subsequent OSCC. Analysis of these groups of cases further highlighted the heterogeneity of SOX2 expression at each stage of oral carcinogenesis, and the limits of its potential diagnostic utility.

Although the heterogeneous expression profile of SOX2 limits its potential as a biomarker, it remains plausible that SOX2 is functionally significant in oral carcinogenesis. Its heterogeneous expression has several explanations. It may be hypothesised that SOX2 has multiple stage-dependent roles in oral carcinogenesis, with a spectrum of oncogenic and tumour-suppressor functions. Alternatively, subsets of OPMD/OSCC may be variably dependent on the oncogenic function of SOX2, and thus show variable expression of SOX2 protein.

Interestingly, previous in vitro studies show that SOX2, along with OCT4, is up-regulated in oral cancer stem cells (CSC) (Lim et al., 2011; Bourguignon et al., 2012; Misuno et al., 2013). SOX2 contributes to stem cell functioning in normal oral squamous epithelium (Okubo et al., 2009). We noted that SOX2 expression was often strongest in foci of tumour cells with a basaloid cytomorphology; it is conceivable that these foci contain CSC. Further work is required to delineate the contribution of SOX2 to the maintenance, survival, and proliferation of CSC. CSC are responsible for resistance to therapy, local recurrence, and metastatic spread in head and neck squamous cell carcinoma, and therefore represent important therapeutic targets (Routray and Mohanty, 2014). The hypothesis that

SOX2 is a marker of CSC suggests it may therefore be an important chemotherapeutic target in OPMD/OSCC. The stably-inducible cell lines generated in this project will be useful in testing this hypotheses.

7.2.8 PAX9 has a potential tumour-suppressor function in oral carcinogenesis

This study demonstrated reduced PAX9 expression in OPMD relative to normal epithelium. PAX9 protein expression in OSCC was reduced relative to both normal epithelium and OPMD. This profile of progressively down-regulated PAX9 expression is consistent with the role of Pax9 in the differentiation of mouse oral squamous epithelium (Jonker et al., 2004) and evidence from a study of oesophageal SCC that suggests *PAX9* acts as a tumour-suppressor gene (Gerber et al., 2002).

The study by Gerber et al (2002) demonstrated that PAX9 expression was either lost or significantly reduced in the majority of oesophageal SCC and dysplastic epithelial lesions. There was an inverse relationship between PAX9 expression and clinical course: a decrease in the proportion of PAX9-positive cells correlated with increasingly malignant behaviour. This correlation was also identified in the current study: OSCC with adverse outcomes had significantly lower PAX9 expression compared to cases with no adverse outcome.

Pax9-deficient mice developed epithelial dysplasia more rapidly than controls. One Pax9-deficient mutant mouse developed epithelial dysplasia after just two weeks of chemical induction. This is less than the eight-week minimum period reported in previous studies (Tang *et al.*, 2004; Vitale-Cross *et al.*, 2004b). The finding that Pax9-deficient mice are more susceptible to 4-NQO supports the hypothesis that Pax9 has a tumour-suppressor function in oral carcinogenesis. However, it must be interpreted with caution. Only a small number of mice were sacrificed following less than eight weeks of 4-NQO treatment. The detection of epithelial dysplasia at two weeks was essentially an incidental finding due to the toxic systemic response to 4-NQO in Pax9-deficient mice (see 7.2.10).

Pax9-deficient mice were more likely to develop OSCC following chemical induction than controls. This further supports the hypothesis that Pax9 has a tumour-suppressor function. However, statistical comparison is limited by the

small number of cases that developed OSCC. Further serial sections will be examined as the analysis continues in future projects. More cases of OSCC may be identified during this process, supporting statistical analysis. There is also scope to verify our findings in future experiments using standardised conditions on greater numbers of mice (section 7.3).

Pax9-deficient mice also had higher rates of dysplastic nodules than controls. Dysplastic nodules bear a resemblance to human papillary squamous cell carcinoma. However, they lacked the definitive evidence of invasion required to group them with the OSCC cases. The precise criteria required in order to establish a diagnosis of papillary squamous cell carcinoma in humans is a subject of debate (Barnes et al., 2005). Lesions similar to papillary squamous cell carcinoma have been described in previous 4-NQO studies (Czerninski et al., 2009). However, criteria for making the diagnosis in mice have yet to be established. Some of these dysplastic nodules may show invasion of the lamina propria on examination of further serial sections. It is also conceivable that had there been a greater interval between the cessation of 4-NQO treatment and sacrifice, some of these lesions may have progressed to conventional OSCC (section 7.3).

Two Pax9-deficient mice developed oesophageal SCC. This is consistent with the site affinity of 4-NQO documented in the literature (Kanojia and Vaidya, 2006). The number of cases is too small to support statistical analysis. However, the finding is consistent with the previous study of human oesophageal SCC and lends further weight to the hypothesis that *Pax9* has a tumour-suppressor function (Gerber et al., 2002). Interestingly, these mice were otherwise well and did not show weight loss, which might have been expected given the extent of oesophageal obstruction. Consistent with the negative autopsy findings documented by Tang et al (2004) there was no evidence of tumour formation elsewhere in the gastrointestinal tract. The site specificity identified in the present study support the suitability of the 4-NQO model of oral carcinogenesis for use in future experiments.

7.2.9 PAX9 is over-expressed in transforming OPMD and their subsequent OSCC

Paradoxically, PAX9 was increased in OPMD that underwent malignant transformation and high levels were also maintained in the resultant OSCC. This contrasts with the evidence outlined in Section 7.2.8, which supports the hypothesis that PAX9 has a tumour-suppressor function. There are two possible interpretations for this finding.

Firstly, it may be hypothesised that PAX9 has an oncogenic function and is transiently up-regulated in OPMD prior to/shortly after transformation to OSCC. Over-expression of PAX9 was more reliably predictive of malignant transformation than high-grade epithelial dysplasia. This suggests PAX9 over-expression is an important event during the transition from epithelial dysplasia to invasive OSCC. It is conceivable that PAX9 is dynamically expressed during oral carcinogenesis: down-regulation and loss of its tumour-suppressor function may be critical to the development of early dysplastic lesions; up-regulation and gain of its oncogenic function may be vital to established lesions as they progress towards invasion. The oncogenic function of PAX9 may be redundant in established OSCC due to acquisition of further genetic changes following invasion. This would account for the down-regulation of PAX9 expression that we have documented in early-stage OSCC.

The view that *PAX9* has potential oncogenic functions is supported by previous *in vitro* studies (Kendall et al., 2007; Lee et al., 2008; Harris et al., 2011). However, the hypothesis that PAX9 is dynamically expressed in oral carcinogenesis must be considered cautiously. The present study of human tissue is observational and has a cross-sectional rather than longitudinal design. This precludes accurate temporal description of PAX9 expression in individual cases. Comparison of PAX9 expression in selected cases is also confounded by variable time intervals between biopsies, and the variable numbers of biopsies performed for each lesion. Previous OPMD biopsies do not exist for the majority of early-stage OSCC cases in our study. The latter showed significantly lower PAX9 expression relative to the group of OSCC that transformed from OPMD. It is open to conjecture whether these early-stage OSCC were preceded by OPMD with low or high levels of PAX9 expression.

An alternative hypothesis is that PAX9 is over-expressed in a subset of OPMD and OSCC, but that over-expression is not uniformly critical to oral carcinogenesis. Interestingly, 'normal' (i.e. relatively high) levels of PAX9 expression were reported in a subset of oesophageal SCC and dysplasias by Gerber et al (2002). It is possible that these oesophageal lesions with high PAX9 expression were biopsied at a time when they were undergoing malignant transformation, which would support the previous hypothesis. However, it also suggests that OPMD and OSCC may be separated into two groups according to whether their PAX9 expression profile is low or high relative to normal epithelium; it is thus conceivable that the majority of early stage OSCC in this study were preceded by OPMD with low PAX9 expression.

In conclusion, our data show similar trends in PAX9 expression to those documented in oesophageal SCC and squamous epithelial dysplasia. They support the hypothesis that PAX9 has a tumour-suppressor function. However, there is also compelling evidence to suggest that PAX9 has an oncogenic function. These functions may be stage-dependent or, alternatively, either function may predominate in individual cases. The cell lines generated in this project may help to determine the precise contribution of PAX9 to oral carcinogenesis. Importantly, however, our data indicate that PAX9 may be a useful biomarker that detects OPMD at high risk of progressing to OSCC.

7.2.10 Pax9-deficient mice are more susceptible to the toxic systemic effects of 4-NQO treatment

Our data show that Pax9-deficient mice have enhanced sensitivity to the toxic systemic effects of 4-NQO than controls. These included extreme weight loss and excessive salivation.

Several mice with excessive salivation deteriorated rapidly. It is possible that fluid loss and dehydration were gradually compounded by further intake of 4-NQO as the mice drank more in an attempt to quench thirst, culminating in a cycle of progressive decline. This would have caused a relentless increase in the concentration of 4-NQO and its metabolites in the bloodstream. Excessive salivation may also have been accompanied by loss of appetite and impaired

masticatory efficiency. This would further contribute to weight loss and exacerbate the toxic systemic effects of 4-NQO as body mass declined.

The precise mechanism for the enhanced sensitivity of the Pax9-deficient mice is unclear. Toxic systemic effects of 4-NQO have not been documented previously. The Pax9-deficient mice that deteriorated most rapidly had only one kidney. Impaired excretion of 4-NQO may have further contributed to an increased 4-NQO concentration in the bloodstream. The present study was not designed to identify which Pax9-deficient mice had only one kidney prior to beginning 4-NQO treatment. The cause of hunching in these mice was unclear. It may have reflected either visceral or muscular pain from the abdominal region. Mice with one kidney may have developed kidney failure due to 4-NQO toxicity; the hunching may also have been related to kidney failure.

Our data confirm the time interval from start of 4-NQO induction to OSCC formation reported in several previous studies (Ma et al., 1999; Ide et al., 2001). Although it was longer than the interval reported in a study which used a higher concentration of 4-NQO (Hasina et al., 2009), our decision to use lower concentrations of 4-NQO was justifiable given the greater sensitivity of Pax9-deficient mice to the toxic systemic effects of 4-NQO. Treating Pax9-deficient mice with 100 µg/ml 4-NQO would have resulted in the unnecessary suffering or even death of many mice, in breach of the principle of refining experiments to ensure that suffering is kept to an absolute minimum (Understanding Animal Research, 2014). Importantly, the majority of the mice that developed OSCC were actually treated with the lowest concentration of 4-NQO (10 µg/ml). This confirms findings from previous studies that 10 µg/ml 4-NQO is sufficient to induce OSCC formation, albeit over a longer period (Ma et al., 1999; Ide et al., 2001).

7.2.11 Future combined 4-NQO and Pax9 knockdown models of oral carcinogenesis may be optimised to reduce suffering and improve experimental efficiency

The methodology used in this study drew on key studies in the literature in which workers sought to optimise the 4-NQO protocol by varying parameters such as the treatment duration and mechanism of application (Vitale-Cross et al., 2009). However, our data show that 4-NQO induction in Pax9-deficient mice requires specific considerations and modifications.

Pax9-deficient mice with a Black 6 genetic background were more susceptible to the toxic systemic effects of 4-NQO than mice with a FVB or hybrid background. The Black 6 genetic background should therefore be avoided in future experiments and other, more resilient, strains such as FVB/hybrids should be used. For the same reason, male Pax9-deficient mice should be used rather than females. Pax9-deficient mice should be imaged to determine the number of kidneys present prior to the start of chemical induction. Only those with two kidneys present should be treated with 4-NQO.

A concentration of 10 µg/ml is sufficient to induce OSCC and a range of precursor lesions within a 24-week period. Higher concentrations of 4-NQO have an increased risk of causing toxic systemic effects, without significantly reducing the length of the experiment or inducing either more or a wider spectrum of epithelial lesions. There is therefore little experimental benefit in using high dosages of 4-NQO. Loss of Pax9-deficient mice due to the systemic effects of a high 4-NQO concentration may compromise the overall efficiency of the experiment through attrition of the sample population.

7.3 Future work

Since the start of this project, additional cases of OPMD that have transformed to OSCC have been diagnosed in the Department of Cellular Pathology. It has not been possible to add all of these to the group of OPMD analysed in this study. In future, however, the assays carried out in this study could be performed on these additional cases. There have also been further cases of OPMD that have not, according to our database, subsequently transformed to OSCC. There is therefore potential to liaise with clinical colleagues to continue with long-term follow-up of these cases, work that was started by Diajil et al (2013). Alternatively, our initial findings could be validated by testing a set of samples collected and characterised at another unit. This would increase the numbers in each outcome group, underpinning further statistical analysis. Longitudinal analysis of biomarker expression in specific lesions would facilitate testing of the hypothesis that certain biomarkers, including PAX9, are dynamically expressed during oral carcinogenesis.

The clinicians served by the Department of Cellular Pathology are familiar with the WHO classification of epithelial dysplasia. It is also the practice of several Pathology Consultants to include the corresponding SIN grade in the histopathology report. This causes little conceptual ambiguity as the two classification are closely related. The main difference is the SIN 3 grade, which combines severe epithelial dysplasia and carcinoma in-situ. Our finding that the binary classification has greater predictive value suggests that clinicians may benefit from inclusion of a binary grade on the histopathology report. This is potentially a major change to clinical practice. It risks causing confusing clinicians and possibly patients. With this in mind, any such change to practice would need to be piloted cautiously following careful consultation with end users and stakeholders. The change would also need to be audited in order to determine its impact on clinical decisions.

Our work highlights the need for clear guidelines for the interpretation and grading of epithelial dysplasia in mice. Such guidelines would support researchers using the 4-NQO model of chemical induction and bring greater uniformity and reliability to the literature. Agreed guidelines would also support studies of chemically-induced oral carcinogenesis in genetically modified mice.

Ideally, however, there is also a need to develop specific criteria for grading epithelial dysplasia in each group of genetically-modified mice in which the phenotype is characterised by disturbed epithelial morphology.

Further work is required to develop standardised criteria for the interpretation of EGFR gene copy number signals in OPMD and OSCC. Our data support the use of a broader ‘abnormal’ EGFR gene copy number category rather than the current criteria used to define EGFR genomic gain in OPMD. It remains unclear whether it may also be appropriate to use a broader definition in OSCC. EGFR gene copy number may help to predict the response of head and neck cancers to EGFR-targeted agents. For example, high EGFR gene copy number has been shown to predict which patients have an increased likelihood of response to erlotinib therapy (Agulnik et al., 2007). The present study was not designed to investigate response to EGFR-targeted agents. Whether the sub-groups of cases with abnormal EGFR gene copy number or EGFR genomic gain have differential responses to EGFR-targeted agents or other therapies remains to be tested.

Future experiments in which Pax9-deficient mice are treated with 4-NQO need to consider a modified protocol as outlined in section 7.2.11. The experiment may also benefit from an increased time interval from the end of 4-NQO treatment to sacrifice/autopsy. This would allow dysplastic lesions to progress to OSCC while mice are free from the potentially toxic systemic effects of 4-NQO treatment. Mice could also be subject to an intra-oral examination under general anaesthesia at specific intervals. This would inform decisions around when to sacrifice mice in order to examine analyse lesions. The surveillance could be further enhanced by performing oral brush biopsy or a small punch biopsy.

All organs from the mice sacrificed in this experiment have been formalin fixed and archived. Histological analysis of vital organs, especially the kidneys, may provide greater insight into the nature of the toxic systemic effects of 4-NQO. This work could be carried out in collaboration with either veterinary pathologists or specialist histopathologists in the Department of Cellular Pathology.

The stably-inducible cell lines generated during this project need to be fully characterised. The cell lines are currently polyclonal. Specific clones will therefore need to be selected and characterised. These defined clones will then be used for functional assays. We plan to analyse basic cellular parameters of the transfected cells and compare these with native, non-transfected control lines. These parameters include determination of cellular viability, using the MTT assay, and cellular proliferation and apoptosis, using fluorescence-activated cell sorting (FACS) analysis. In the longer term, the genetic pathways regulating the molecular changes identified in the inducible PAX9 and SOX2 OSCC cell lines could be delineated by DNA microarray analysis. This would generate genome wide expression profiles, which can be analysed by bioinformatics. Significant changes in gene expression can be further validated by qPCR.

7.4 Conclusion

EGFR protein is over-expressed in OPMD and early-stage OSCC, but does not correlate with clinical outcome in either group. By contrast, EGFR gene copy number abnormalities are associated with malignant transformation in OPMD.

SOX2 has a heterogeneous expression profile in both OPMD and OSCC. This limits its potential clinical utility as a biomarker. However, the pattern of expression supports the hypothesis that SOX2 is a marker of oral cancer stem cells. SOX2 may therefore be highly significant in oral carcinogenesis and should be considered as a candidate chemotherapeutic target.

PAX9 is down-regulated in OPMD and early-stage OSCC. Pax9-deficient mice are more likely than controls to develop OPMD and OSCC following chemical induced. These findings support the hypothesis that PAX9 has a tumour-suppressor function. Paradoxically, PAX9 is over-expressed in the subset of OPMD that underwent malignant transformation. This suggests PAX9 may be dynamically expressed and play diverse roles during oral carcinogenesis, including an oncogenic function.

Pax9-deficient mice are more susceptible to the systemic effects of 4-NQO treatment. This finding has informed important modifications to the protocol for future experiments that involve 4-NQO induction in Pax9-deficient mice.

The cell lines developed during this project are potentially valuable tools that will facilitate further assessment of the role of PAX9, SOX2, and their downstream targets in oral carcinogenesis.

Chapter 8. References

Agulnik, M., da Cunha Santos, G. and Hedley, D. (2007) 'Predictive and pharmacodynamic biomarker studies in tumour and skin tissue samples of patients with recurrent or metastatic squamous cell carcinoma of the head and neck treated with erlotinib', *J Clin Oncol*, 25(16), pp. 2184-2190.

Allsopp, R.C., Vaziri, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W. and Harley, C.B. (1992) 'Telomere length predicts replicative capacity of human fibroblasts', *Proc Natl Acad Sci U S A*, 89(21), pp. 10114-8.

Ambrosini, E., Ceccherini-Silberstein, F., Erfle, V., Aloisi, F. and Levi, G. (1999) 'Gene transfer in astrocytes: comparison between different delivering methods and expression of the HIV-1 protein Nef', *J Neurosci Res*, 55(5), pp. 569-77.

Amit, M., Yen, T.C., Liao, C.T., Chaturvedi, P., Agarwal, J.P., Kowalski, L.P., Ebrahimi, A., Clark, J.R., Kreppel, M., Zoller, J., Fridman, E., Bolzoni, V.A., Shah, J.P., Binenbaum, Y., Patel, S.G., Gil, Z. and Cancer, I.C.f.O.R.I.i.H.a.N. (2013) 'Improvement in survival of patients with oral cavity squamous cell carcinoma: An international Collaborative Study', *Cancer*, 119(24), pp. 4242-8.

Amornphimoltham, P., Patel, V., Sodhi, A., Nikitakis, N., Sauk, J., Sausville, E., Molinolo, A. and Gutkind, J. (2005) 'Mammalian target of rapamycin, a molecular target in squamous cell carcinomas of the head and neck', *Cancer Res*, 65(21), pp. 9953-61.

Amornphimoltham, P., Sriuranpong, V., Patel, V., Benavides, F., Conti, C.J., Sauk, J., Sausville, E.A., Molinolo, A.A. and Gutkind, J.S. (2004) 'Persistent activation of the Akt pathway in head and neck squamous cell carcinoma: a potential target for UCN-01', *Clin Cancer Res*, 10(12 Pt 1), pp. 4029-37.

Angiero, F., Gatta, L.B., Seramondi, R., Berenzi, A., Benetti, A., Magistro, S., Ordesi, P., Grigolato, P. and Dessy, E. (2010) 'Frequency and role of HPV in the progression of epithelial dysplasia to oral cancer', *Anticancer Res*, 30(9), pp. 3435-40.

Atienza, J.A.S. and Dasanu, C.A. (2012) 'Incidence of second primary malignancies in patients with treated head and neck cancer: a comprehensive review of the literature', *Current Medical Research and Opinion*, 28(12), pp. 1899-1909.

Bani-Yaghoub, M., Tremblay, R.G., Lei, J.X., Zhang, D., Zurakowski, B., Sandhu, J.K., Smith, B., Ribecco-Lutkiewicz, M., Kennedy, J., Walker, P.R. and Sikorska, M. (2006) 'Role of Sox2 in the development of the mouse neocortex', *Dev Biol*, 295(1), pp. 52-66.

Barnes, L., Eveson, J.W., Reichart, P. and Sidransky, D. (2005) *Pathology & Genetics of Head and Neck Tumours*. Geneva: WHO Press.

Bass, A.J., Watanabe, H., Mermel, C.H., Yu, S., Perner, S., Verhaak, R.G., Kim, S.Y., Wardwell, L., Tamayo, P., Gat-Viks, I., Ramos, A.H., Woo, M.S., Weir, B.A., Getz, G., Beroukhim, R., O'Kelly, M., Dutt, A., Rozenblatt-Rosen, O., Dziunycz, P., Komisarof, J., Chirieac, L.R., Lafargue, C.J., Scheble, V., Wilbertz, T., Ma, C., Rao, S., Nakagawa, H., Stairs, D.B., Lin, L., Giordano, T.J., Wagner, P., Minna, J.D., Gazdar, A.F., Zhu, C.Q., Brose, M.S., Cecconello, I., Jr, U.R., Marie, S.K., Dahl, O., Shivdasani, R.A., Tsao, M.S., Rubin, M.A., Wong, K.K., Regev, A.,

Hahn, W.C., Beer, D.G., Rustgi, A.K. and Meyerson, M. (2009) 'SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas', *Nat Genet*, 41(11), pp. 1238-42.

Bedi, G.C., Westra, W.H., Gabrielson, E., Koch, W. and Sidransky, D. (1996) 'Multiple head and neck tumors: evidence for a common clonal origin', *Cancer Res*, 56(11), pp. 2484-7.

Benchekroun, M., Saintigny, P., Thomas, S., El-Naggar, A., Papadimitrakopoulou, V., Ren, H., Lang, W., Fan, Y.-H., Huang, J., Feng, L., Lee, J., Kim, E., Hong, W., Johnson, F., Grandis, J. and Mao, L. (2010) 'Epidermal Growth Factor Receptor Expression and Gene Copy Number in the Risk of Oral Cancer', *Cancer Prev Res*, 3(7), pp. 800 - 809

Bloor, B.K., Tidman, N., Leigh, I.M., Odell, E., Dogan, B., Wollina, U., Ghali, L. and Waseem, A. (2003) 'Expression of keratin K2e in cutaneous and oral lesions: association with keratinocyte activation, proliferation, and keratinization', *Am J Pathol*, 162(3), pp. 963-75.

Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S. and Wright, W.E. (1998) 'Extension of life-span by introduction of telomerase into normal human cells', *Science*, 279(5349), pp. 349-52.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986) 'Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila', *Cell*, 47(6), pp. 1033-40.

Bornstein, S., White, R., Malkoski, S., Oka, M., Han, G., Cleaver, T., Reh, D., Andersen, P., Gross, N., Olson, S., Deng, C., Lu, S.L. and Wang, X.J. (2009) 'Smad4 loss in mice causes spontaneous head and neck cancer with increased genomic instability and inflammation', *J Clin Invest*, 119(11), pp. 3408-19.

Bosetti, C., Scelo, G., Chuang, S.-C., Tonita, J.M., Tamaro, S., Jonasson, J.G., Kliewer, E.V., Hemminki, K., Weiderpass, E., Pukkala, E., Tracey, E., Olsen, J.H., Pompe-Kirn, V., Brewster, D.H., Martos, C., Chia, K.-S., Brennan, P., Hashibe, M., Levi, F., La Vecchia, C. and Bofetta, P. (2011) 'High constant incidence rates of second primary cancers of the head and neck: a pooled analysis of 13 cancer registries', *Int J Cancer*, 129(1), pp. 173-179.

Boukamp, P., Petrussevska, R.T., Breitkreutz, D., Hornung, J., Markham, A. and Fusenig, N.E. (1988) 'Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line', *J Cell Biol*, 106(3), pp. 761-71.

Boumahdi, S., Driessens, G., Lapouge, G., Rorive, S., Nassar, D., Le Mercier, M., Delatte, B., Caauwe, A., Lenglez, S., Nkusi, E., Brohee, S., Salmon, I., Dubois, C., del Marmol, V., Fuks, F., Beck, B. and Blanpain, C. (2014) 'SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma', *Nature*, 511(7508), pp. 246-50.

Bourguignon, L., Wong, G., Earle, C. and Chen, L. (2012) 'Hyaluronan-CD44v3 interaction with Oct4-Sox2-Nanog promotes miR-302 expression leading to self-renewal, clonal formation, and cisplatin resistance in cancer stem cells from head and neck squamous cell carcinoma', *J Biol Chem*, 287(39), pp. 32800-24.

Bowles, J., Schepers, G. and Koopman, P. (2000) 'Phylogeny of the SOX Family of Developmental Transcription Factors Based on Sequence and Structural Indicators', *Developmental Biology*, 227(2), pp. 239-55.

Braakhuis, B.J., Bloemena, E., Leemans, C.R. and Brakenhoff, R.H. (2010) 'Molecular analysis of surgical margins in head and neck cancer: more than a marginal issue', *Oral Oncol*, 46(7), pp. 485-91.

Braakhuis, B.J., Leemans, C.R. and Brakenhoff, R.H. (2004) 'A genetic progression model of oral cancer: current evidence and clinical implications', *J Oral Pathol Med*, 33(6), pp. 317-22.

Braakhuis, B.J., Tabor, M.P., Kummer, J.A., Leemans, C.R. and Brakenhoff, R.H. (2003) 'A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications', *Cancer Res*, 63(8), pp. 1727-30.

Brennan, J.A., Boyle, J.O., Koch, W.M., Goodman, S.N., Hruban, R.H., Eby, Y.J., Couch, M.J., Forastiere, A.A. and Sidransky, D. (1995) 'Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck', *New England Journal of Medicine*, 332(11), pp. 712-7.

Brudno, M., Poliakov, A., Salamov, A., Cooper, G.M., Sidow, A., Rubin, E.M., Solovyev, V., Batzoglou, S. and Dubchak, I. (2004) 'Automated whole-genome multiple alignment of rat, mouse, and human', *Genome Res*, 14(4), pp. 685-92.

Califano, J., van der Riet, P., Westra, W., Nawroz, H., Clayman, G., Piantadosi, S., Corio, R., Lee, D., Greenberg, B., Koch, W. and Sidransky, D. (1996) 'Genetic progression model for head and neck cancer: implications for field cancerization', *Cancer Res*, 56(11), pp. 2488-92.

Campos, M.I., Campos, C.N., Aarestrup, F.M. and Aarestrup, B.J. (2014) 'Oral mucositis in cancer treatment: Natural history, prevention and treatment', *Mol Clin Oncol*, 2(3), pp. 337-340.

Cancer Research UK (2010) *Cancer Stats: Key Facts*. Available at: http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/ (Accessed: October 15th).

Cancer Research UK (2013) *Phases of trials* (Accessed: September 1st).

Caulin, C., Nguyen, T., Longley, M., Zhou, Z., Wang, X. and Roop, D. (2004) 'Inducible activation of oncogenic K-ras results in tumor formation in the oral cavity.', *Cancer Res*, 64(15), pp. 5054-8.

Chen, C. and Okayama, H. (1987) 'High-efficiency transformation of mammalian cells by plasmid DNA', *Mol Cell Biol*, 7(8), pp. 2745-52.

Chew, L.J. and Gallo, V. (2009) 'The Yin and Yang of Sox proteins: Activation and repression in development and disease', *J Neurosci Res*, 87(15), pp. 3277-87.

Chi, N. and Epstein, J.A. (2002) 'Getting your Pax straight: Pax proteins in development and disease', *Trends Genet*, 18(1), pp. 41-7.

Chuang, S.C., Scelo, G., Tonita, J.M., Tamaro, S., Jonasson, J.G., Kliewer, E.V., Hemmick, K., Weiderpass, E., Pukkala, E., Tracey, E., Friis, S., Pompe-Kirn, V., Brewster, D.H., Martos, C., Chia, K.S., Bofetta, P., Brennan, P. and Hashibe, M. (2008) 'Risk of second primary cancer among patients with head and neck

cancers: a pooled analysis of 13 cancer registries', *Int J Cancer*, 123(10), pp. 2390-6.

Chung, C., Zhang, Q., Kong, C., Harris, J., Fertig, E., Harari, P., Wang, D., Redmond, K., Shenouda, G., Trotti, A., Raben, D., Gillison, M., Jordan, R. and Le, Q. (2014) 'p16 Protein Expression and Human Papillomavirus Status As Prognostic Biomarkers of Nonoropharyngeal Head and Neck Squamous Cell Carcinoma', *J Clin Oncol*, 32(35), pp. 3930-8.

Chung, C.H., Ely, K., McGavran, L. and al, e. (2006) 'Increased epidermal growth factor receptor gene copy number is associated with poor prognosis in head and neck squamous cell carcinomas', *J Clin Oncol*, 24, pp. 4170 - 4176.

Citri, A. and Yarden, Y. (2006) 'EGF-ERBB signalling: towards the systems level', *Nat Rev Mol Cell Biol*, 1(7), pp. 505–16.

Close, L.G., Brown, P.M., Vuitch, M.F., Reisch, J. and Schaefer, S.D. (1989) 'Microvascular invasion and survival in cancer of the oral cavity and oropharynx', *Arch Otolaryngol Head Neck Surg*, 115(11), pp. 1304-9.

Cregger, M., Berger, A. and Rimm, D. (2006) 'Immunohistochemisry and quantitative analysis of protein expression', *Archives of Pathology and Laboratory Medicine*, 130, pp. 1026-1030.

Czerninski, R., Amornphimoltham, P., Patel, V., Molinolo, A.A. and Gutkind, J.S. (2009) 'Targeting mammalian target of rapamycin by rapamycin prevents tumor progression in an oral-specific chemical carcinogenesis model', *Cancer Prev Res (Phila)*, 2(1), pp. 27-36.

Dahl, E., Koseki, H. and Balling, R. (1997) 'Pax genes and organogenesis', *Bioessays*, 19(9), pp. 755-65.

Danaei, G., VanderHoorn, S., Lopez, A.D., Murray, C. and Ezzati, M. (2005) 'Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors', *Lancet*, 366(9499), pp. 1784-93.

Davis, L.K., Meyer, K.J., Rudd, D.S., Librant, A.L., Epping, E.A., Sheffield, V.C. and Wassink, T.H. (2008) 'Pax6 3' deletion results in aniridia, autism and mental retardation', *Hum Genet*, 123(4), pp. 371-8.

Day, G.L. and Blot, W.J. (1992) 'Second primary tumors in patients with oral cancer', *Cancer*, 70(1), pp. 14-9.

Dayan, D., Hirshberg, A., Kaplan, I., Rotem, N. and Bodner, L. (1997) 'Experimental tongue cancer in desalivated rats', *Oral Oncol*, 33(2), pp. 105-9.

de Bont, J.M., Kros, J.M., Passier, M.M., Reddingius, R.E., Sillevis Smitt, P.A., Luider, T.M., den Boer, M.L. and Pieters, R. (2008) 'Differential expression and prognostic significance of SOX genes in pediatric medulloblastoma and ependymoma identified by microarray analysis', *Neuro Oncol*, 10(5), pp. 648-60.

Diajil, A., Robinson, C., Sloan, P. and Thomson, P. (2013) 'Clinical Outcome Following Oral Potentially Malignant Disorder Treatment: A 100 Patient Cohort Study', *Int J Dent*, 2013, pp. 1 - 8.

Dickson, M.A., Hahn, W.C., Ino, Y., Ronfard, V., Wu, J.Y., Weinberg, R.A., Louis, D.N., Li, F.P. and Rheinwald, J.G. (2000) 'Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span

become immortal yet retain normal growth and differentiation characteristics', *Mol Cell Biol*, 20(4), pp. 1436-47.

Dost, F., Lê Cao, K., Ford, P., Ades, C. and Farah, C. (2014) 'Malignant transformation of oral epithelial dysplasia: a real-world evaluation of histopathologic grading', *Oral Surg Oral Med Oral Pathol Oral Radiol*, 117(3), pp. 343-52.

Du, L., Yang, Y., Xiao, X., Wanga, C., Zhang, X., Wang, L., Zhang, X., Li, W., Zheng, G., Wang, S. and Dong, Z. (2011) 'Sox2 nuclear expression is closely associated with poor prognosis in patients with histologically node-negative oral tongue squamous cell carcinoma', *Oral Oncol*, 47(8), pp. 709-13.

Eagle, H. (1955) 'Propagation in a fluid medium of a human epidermoid carcinoma, strain KB', *Proc Soc Exp Biol Med*, 89(3), pp. 362-4.

Edington, K.G., Loughran, O.P., Berry, I.J. and Parkinson, E.K. (1995) 'Cellular immortality: a late event in the progression of human squamous cell carcinoma of the head and neck associated with p53 alteration and a high frequency of allele loss', *Mol Carcinog*, 13(4), pp. 254-65.

El-Mofty, S.K. (2007) 'Human papillomavirus (HPV) related carcinomas of the upper aerodigestive tract', *Head Neck Pathol*, 1(2), pp. 181-5.

Ellis, P., Fagan, B.M., Magness, S.T., Hutton, S., Taranova, O., Hayashi, S., McMahon, A., Rao, M. and Pevny, L. (2004) 'SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult', *Dev Neurosci*, 26(2-4), pp. 148-65.

Erjala, K., Sundvall, M. and Junttila, T. (2006) 'Signalling via ErbB2 and ErbB3 associates with resistnace and epidermal growth factor receptor (EGFR) amplification with sensitivity to EGFR inhibitor gefitinib in head and neck squamous cell carcinoma cells', *Clin Cancer Res*, 12, pp. 4103-4111.

Eversole, L. and Sapp, J. (1993) 'c-myc oncoprotein expression in oral precancerous and early cancerous lesions', *Eur J Cancer B Oral Oncol*, 29(2), pp. 131-5.

Eveson, J.W. and MacDonald, D.G. (1978) 'Quantitative histological changes during early experimental carcinogenesis in the hamster cheek pouch', *Br J Dermatol*, 98(6), pp. 639-44.

Fahey, M.S., Paterson, I.C., Stone, A., Collier, A.J., Heung, Y.L.M., Davies, M., Patel, V., Parkinson, E.K. and Prime, S.S. (1996) 'Dysregulation of autocrine TGF-beta isoform production and ligand responses in human tumour-derived and Ha-ras transfected keratinocytes and fibroblasts', *British Journal of Cancer*, 74(1074-1080).

Fearon, E.R. and Vogelstein, B. (1990) 'A genetic model for colorectal tumorigenesis', *Cell*, 61(5), pp. 759-67.

Felgner, P.L. and Ringold, G.M. (1989) 'Cationic liposome-mediated transfection', *Nature*, 337(6205), pp. 387-8.

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J. and et al. (1995) 'The RNA component of human telomerase', *Science*, 269(5228), pp. 1236-41.

Field, A. (2013) *Discovering Statistics using IBM SPSS Statistics*. London: Sage Publications.

Field, J.K. and Spandidos, D.A. (1987) 'Expression of oncogenes in human tumours with special reference to the head and neck region', *J Oral Pathol*, 16(3), pp. 97-107.

Forastiere, A. (2007) 'Epidermal Growth Factors Receptor Inhibition in Head and Neck Cancer - More Insights, but More Questions', *J Clin Oncol*, 25(16), pp. 2152 - 2155.

Forastiere, A., Koch, W., Trotti, A. and Sidransky, D. (2001) 'Head and neck cancer', *N Engl J Med*, 345(26), pp. 1890-900.

Fracalossi, A.C., Silva Mde, S., Oshima, C.T. and Ribeiro, D.A. (2010) 'Wnt/beta-catenin signalling pathway following rat tongue carcinogenesis induced by 4-nitroquinoline 1-oxide', *Exp Mol Pathol*, 88(1), pp. 176-83.

Freier, K., Joos, S. and Flechtenmacher, C. (2003) 'Tissue microarray analysis reveals sitespecific prevalence of oncogene amplifications in head and neck squamous cell carcinoma', *Cancer Res*, 63, pp. 1179 - 1182.

Freier, K., Knoepfle, K., Flechtenmacher, C., Pungs, S., Devens, F., Toedt, G., Hofele, C., Joos, S., Lichter, P. and B, R. (2010) 'Recurrent copy number gain of transcription factor SOX2 and corresponding high protein expression in oral squamous cell carcinoma', *Genes Chromosomes Cancer*, 49(1), pp. 9-16.

Frijhoff, A.F., Conti, C.J. and Senderowicz, A.M. (2004) 'Advances in molecular carcinogenesis: current and future use of mouse models to screen and validate molecularly targeted anticancer drugs', *Mol Carcinog*, 39(4), pp. 183-94.

Fronza, G., Campomenosi, P., Iannone, R. and Abbondandolo, A. (1992) 'The 4-nitroquinoline 1-oxide mutational spectrum in single stranded DNA is characterized by guanine to pyrimidine transversions', *Nucleic Acids Res*, 20(6), pp. 1283-7.

Fujino, H., Chino, T. and Imai, T. (1965) 'Experimental production of labial and lingual carcinoma by local application of 4-nitroquinoline N-oxide', *J Natl Cancer Inst*, 35(6), pp. 907-18.

Garavello, W., Bertuccio, P., Levi, F., Lucchini, F., Bosetti, C., Malvezzi, M., Negri, E. and La Vecchia, C. (2010) 'The oral cancer epidemic in central and Eastern europe', *Int J Cancer*, 127(1), pp. 160-71.

Garraway, L.A. and Sellers, W.R. (2006) 'Lineage dependency and lineage-survival oncogenes in human cancer', *Nat Rev Cancer*, 6(8), pp. 593-602.

Gasco, M. and Crook, T. (2003) 'The p53 network in head and neck cancer', *Oral Oncol*, 39(3), pp. 222-31.

Gerber, J., Richter, T., Kremmer, E., Adamski, J., Hofler, H., Balling, R. and Peters, H. (2002) 'Progressive loss of PAX9 expression correlates with increasing malignancy of dysplastic and cancerous epithelium of the human oesophagus', *J Pathol*, 197(3), pp. 293-7.

Gnepp, D. (2009) *Diagnostic Surgical Pathology of the Head and Neck*. 2nd edn. Philadelphia, PA, USA: Saunders, Elsevier.

Goessel, G., Quante, M., Hahn, W.C., Harada, H., Heeg, S., Suliman, Y., Doebele, M., von Werder, A., Fulda, C., Nakagawa, H., Rustgi, A.K., Blum, H.E. and Opitz, O.G. (2005) 'Creating oral squamous cancer cells: a cellular model of oral-esophageal carcinogenesis', *Proc Natl Acad Sci U S A*, 102(43), pp. 15599-604.

Goodson, M. and Thomson, P. (2010) 'Management of oral carcinoma: benefits of early precancerous intervention', *Br J Oral Max Surg*, 49(2011), pp. 88 – 91.

Grandis, J. (1998) 'Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival', *J Nat Cancer Inst*, 90 pp. 824 - 832.

Grandis, J. and Tweardy, D. (1993a) 'Elevated levels of transforming growth factor alpha and epidermal growth factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer', *Cancer Res*, 53(15), pp. 3579 - 3584.

Grandis, J. and Tweardy, D. (1993b) 'TGF-alpha and EGFR in head and neck cancer', *J Cell Biochem*, 17, pp. 188-191.

Gusterson, B. and Hunter, K. (2009) 'Should we be surprised at the paucity of response to EGFR inhibitors?', *Lancet Oncology*, 10, pp. 522 - 527.

Hanahan, D. and Weinberg, R. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell*, 144(5), pp. 646-74.

Harris, T., Pan, Q., Sironi, J., Lutz, D., Tian, J., Sapkar, J., Perez-Soler, R., Keller, S. and Locker, J. (2011) 'Both gene amplification and allelic loss occur at 14q13.3 in lung cancer', *Clin Cancer Res*, 17(4), pp. 690-9.

Hashibe, M., Brennan, P., Chuang, S.C., Boccia, S., Castellsague, X., Chen, C., Curado, M.P., Dal Maso, L., Daudt, A.W., Fabianova, E., Fernandez, L., Wunsch-Filho, V., Franceschi, S., Hayes, R.B., Herrero, R., Kelsey, K., Koifman, S., La Vecchia, C., Lazarus, P., Levi, F., Lence, J.J., Mates, D., Matos, E., Menezes, A., McClean, M.D., Muscat, J., Eluf-Neto, J., Olshan, A.F., Purdue, M., Rudnai, P., Schwartz, S.M., Smith, E., Sturgis, E.M., Szeszenia-Dabrowska, N., Talamini, R., Wei, Q., Winn, D.M., Shangina, O., Pilarska, A., Zhang, Z.-F., Ferro, G., Berthiller, J. and Boffetta, P. (2009) 'Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium', *Cancer Epidemiology, Biomarkers, and Prevention*, 18(2), pp. 541-550.

Hasina, R., Martin, L.E., Kasza, K., Jones, C.L., Jalil, A. and Lingen, M.W. (2009) 'ABT-510 is an effective chemopreventive agent in the mouse 4-nitroquinoline 1-oxide model of oral carcinogenesis', *Cancer Prev Res (Phila)*, 2(4), pp. 385-93.

Hecht, S.S. (2003) 'Tobacco carcinogens, their biomarkers and tobacco-induced cancer', *Nat Rev Cancer*, 3(10), pp. 733-44.

Heniford, B.W., Shum-Siu, A., Leonberger, M. and Hendler, F.J. (1993) 'Variation in cellular EGF receptor mRNA expression demonstrated by in situ reverse transcriptase polymerase chain reaction', *Nucleic Acids Res*, 21(14), pp. 3159-66.

Herbst, R. (2004) 'Review of epidermal growth factor receptor biology', *Int J Rad Oncol Biol Phys*, pp. 521 - 526.

Herzig, M. and Christofori, G. (2002) 'Recent advances in cancer research: mouse models of tumorigenesis', *Biochim Biophys Acta*, 1602(2), pp. 97-113.

Hewitt, S.M., Hamada, S., Monarres, A., Kottical, L.V., Saunders, G.F. and McDonnell, T.J. (1997) 'Transcriptional activation of the bcl-2 apoptosis suppressor gene by the paired box transcription factor PAX8', *Anticancer Res*, 17(5A), pp. 3211-5.

Hillen, W. and Berens, C. (1994) 'Mechanisms underlying expression of Tn10 encoded tetracycline resistance', *Annu Rev Microbiol*, 48, pp. 345-69.

Hillen, W., Gatz, C., Altschmied, L., Schollmeier, K. and Meier, I. (1983) 'Control of expression of the Tn10-encoded tetracycline resistance genes. Equilibrium and kinetic investigation of the regulatory reactions', *J Mol Biol*, 169(3), pp. 707-21.

Hirsch, F., Varella-Garcia, M., Bunn, P., Franklin, W. and Dziadziusko, R. (2003) 'Epidermal growth factor receptor in non-small cell lung carcinomas: correlation between gene copy numbers and protein expression and impact on prognosis', *J Clin Oncol*, 21, pp. 3798-3807.

Hoffman, D. and Wynder, E.L. (1986) *Chemical constituents and bioactivity of tobacco smoke*. IARC Scientific Publications.

Hsu, D.S., Acharya, C.R., Balakumaran, B.S., Riedel, R.F., Kim, M.K., Stevenson, M., Tuchman, S., Mukherjee, S., Barry, W., Dressman, H.K., Nevins, J.R., Powers, S., Mu, D. and Potti, A. (2009) 'Characterizing the developmental pathways TTF-1, NKX2-8, and PAX9 in lung cancer', *Proc Natl Acad Sci U S A*, 106(13), pp. 5312-7.

Huang, C., Xu, X., Wu, T., Sun, Z. and Zhang, W. (2014) 'Correlation of ALDH1, CD44, OCT4 and SOX2 in tongue squamous cell carcinoma and their association with disease progression and prognosis', *J Oral Pathol Med*.

Hunter, K.D., Parkinson, E.K. and Harrison, P.R. (2005) 'Profiling early head and neck cancer', *Nat Rev Cancer*, 5(2), pp. 127-35.

Hussenet, T., Dali, S., Exinger, J., Monga, B., Jost, B., Dembele, D., Martinet, N., Thibault, C., Huelsken, J., Brambilla, E. and du Manoir, S. (2010) 'SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas', *PloS One*, 5(1), p. e8960.

Hussenet, T. and du Manoir, S. (2010) 'SOX2 in squamous cell carcinoma: amplifying a pleiotropic oncogene along carcinogenesis', *Cell Cycle*, 9(8), pp. 1480-6.

Ide, F., Oda, H., Nakatsuru, Y., Kusama, K., Sakashita, H., Tanaka, K. and Ishikawa, T. (2001) 'Xeroderma pigmentosum group A gene action as a protection factor against 4-nitroquinoline 1-oxide-induced tongue carcinogenesis', *Carcinogenesis*, 22(4), pp. 567-72.

Imaida, K., Sato, H., Okamiya, H., Takahashi, M. and Hayashi, Y. (1989) 'Enhancing effect of high fat diet on 4-nitroquinoline 1-oxide-induced pulmonary tumorigenesis in ICR male mice', *Jpn J Cancer Res*, 80(6), pp. 499-502.

Indian Council of Medical Research (2010) *An atlas of cancer in India*. Available at: http://www.canceratlasindia.org/.

International Agency for Research on Cancer (2010) *Cancer Incidence in Five Continents* Available at: http://ci5.iarc.fr/CI5i-ix/ci5i-ix.htm.

Iwasaki, S. (2002) 'Evolution of the structure and function of the vertebrate tongue', *J Anat*, 201(1), pp. 1-13.

Jakobsson, P.A., Eneroth, C.M., Killander, D., Moberger, G. and Martensson, B. (1973) 'Histologic classification and grading of malignancy in carcinoma of the larynx', *Acta Radiol Ther Phys Biol*, 12(1), pp. 1-8.

Jayalekshmi, P.A., Gangadharan, P., Akiba, S., Nair, R.R., Tsuji, M. and Rajan, B. (2009) 'Tobacco chewing and female oral cavity risk in Karunagapally cohort, India ', *British Journal of Cancer*, 100, pp. 848-852.

Jemal, A., Bray, F., Centre, M., Ferlay, J., Ward, E. and Forman, D. (2011) 'Global Cancer Statistics', *CA Cancer J Clin*, 61(2), pp. 69-90.

Jonker, L., Kist, R., Aw, A., Wappler, I. and Peters, H. (2004) 'Pax9 is required for filiform papilla development and suppresses skin-specific differentiation of the mammalian tongue epithelium', *Mech Dev*, 121(11), pp. 1313-22.

Kanojia, D. and Vaidya, M.M. (2006) '4-nitroquinoline-1-oxide induced experimental oral carcinogenesis', *Oral Oncol*, 42(7), pp. 655-67.

Kearsley, J.H. and Thomas, S. (1993) 'Prognostic markers in cancers of the head and neck region', *Anticancer Drugs*, 4(4), pp. 419-29.

Kendall, J., Liu, Q., Bakleh, A., Krasnitz, A., Nguyen, K.C., Lakshmi, B., Gerald, W.L., Powers, S. and Mu, D. (2007) 'Oncogenic cooperation and coamplification of developmental transcription factor genes in lung cancer', *Proc Natl Acad Sci U S A*, 104(42), pp. 16663-8.

Khan, I., Kumar, N., Pant, I., Narra, S. and Kondaiah, P. (2012) 'Activation of TGF-β pathway by areca nut constituents: a possible cause of oral submucous fibrosis', *PLoS One*, 7(12).

Kist, R., Greally, E. and Peters, H. (2007) 'Derivation of a mouse model for conditional inactivation of Pax9', *Genesis*, 45(7), pp. 460-4.

Kist, R., Watson, M., Wang, X., Cairns, P., Miles, C., Reid, D.J. and Peters, H. (2005) 'Reduction of Pax9 gene dosage in an allelic series of mouse mutants causes hypodontia and oligodontia', *Hum Mol Genet*, 14(23), pp. 3605-17.

Klanrit, P., Sperandio, M., Brown, A.L., Shirlaw, P.J., Challacombe, S.J., Morgan, P.R. and Odell, E.W. (2007) 'DNA ploidy in proliferative verrucous leukoplakia', *Oral Oncol*, 43(3), pp. 310-6.

Kokalj Vokač, N., Cizmarević, B., Zagorac, A., Zagradišnik, B. and Lanišnik, B. (2014) 'An evaluation of SOX2 and hTERC gene amplifications as screening markers in oral and oropharyngeal squamous cell carcinomas', *Mol Cytogenet*, 7(1), p. 5.

Kovesi, G. and Szende, B. (2003) 'Changes in apoptosis and mitotic index, p53 and Ki67 expression in various types of oral leukoplakia', *Oncology*, 65, pp. 331-6.

Krause, C.J., Carey, T.E., Ott, R.W., Hurbis, C., McClatchey, K.D. and Regezi, J.A. (1981) 'Human squamous cell carcinoma. Establishment and

characterization of new permanent cell lines', *Arch Otolaryngol*, 107(11), pp. 703-10.

Kujan, O., Khattab, A., Oliver, R.J., Roberts, S.A., Thakker, N. and Sloan, P. (2007) 'Why oral histopathology suffers inter-observer variability on grading oral epithelial dysplasia: an attempt to understand the sources of variation', *Oral Oncol*, 43(3), pp. 224-31.

Kujan, O., Oliver, R., Khattab, A., Roberts, S., Thakker, N. and Sloan, P. (2006) 'Evaulation of a new binary system of grading oral epitheial dysplasia for prediction of malignant transformation', *Oral Oncol*, 42(10), pp. 987 – 993.

Kumar, B., Cordell, K. and Lee, J. (2008) 'EGFR, p16, HPV Titer, Bcl-xL and p53, Sex and Smoking As Indicators of Response to Therapy and Survival in Oropharyngeal Cancer', *J Clin Oncol*, 26(19), pp. 3128 - 3137.

Kumar, V., Abbas, A., Fausto, N. and Mitchell, R. (2005) *Robbins and Cotran. Pathologic Basis of Disease*. (1 vols). Philadelphia: Elsevier Saunders.

Lee, J.C., Sharma, M., Lee, Y.H., Lee, N.H., Kim, S.Y., Yun, J.S., Nam, S.Y., Hwang, P.H., Jhee, E.C. and Yi, H.K. (2008) 'Pax9 mediated cell survival in oral squamous carcinoma cell enhanced by c-myb', *Cell Biochem Funct*, 26(8), pp. 892-9.

Leeman-Neill, R.J., Seethala, R.R., Singh, S.V., Freilino, M.L., Bednash, J.S., Thomas, S.M., Panahandeh, M.C., Gooding, W.E., Joyce, S.C., Lingen, M.W., Neill, D.B. and Grandis, J.R. (2011) 'Inhibition of EGFR-STAT3 signaling with erlotinib prevents carcinogenesis in a chemically-induced mouse model of oral squamous cell carcinoma', *Cancer Prev Res (Phila)*, 4(2), pp. 230-7.

Leemans, C., Braakhuis, B. and Brakenhoff, R. (2011) 'The molecular biology of head and neck cancer', *Nat Rev Cancer*, 11(1), pp. 9 - 22.

Lerman, M. and Woo, S. (2014) 'Histopathologic features of high risk human papillomavirus-associated oral epithelial dysplasia', *Oral Surg Oral Med Oral Pathol Oral Radiol*, 117(1), p. 120.

Lim, Y., Oh, S., Cha, Y., Kim, S., Jin, X. and Kim, H. (2011) 'Cancer stem cell traits in squamospheres derived from primary head and neck squamous cell carcinomas', *Oral Oncol*, 47(2), pp. 83-91.

Lingen, M., Xiao, W., Schmitt, A., Jiang, B., Pickard, R., Kreinbrink, P., Perez-Ordonez, B., Jordan, R. and Gillison, M. (2013) 'Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas', *Oral Oncol*, 49(1), pp. 1-8.

Liu, M.C. and Gelmann, E.P. (2002) 'P53 gene mutations: case study of a clinical marker for solid tumors', *Semin Oncol*, 29(3), pp. 246-57.

Liu, W., Bao, Z.X., Shi, L.J., Tang, G.Y. and Zhou, Z.T. (2011) 'Malignant transformation of oral epithelial dysplasia: clinicopathological risk factors and outcome analysis in a retrospective cohort of 138 cases', *Histopathology*, 59(4), pp. 733-40.

Livak, K. and Schmittgen, T. (2001) 'Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method', *Methods*, 25, pp. 402-8.

Lodi, G., Sardella, A., Bez, C., Demarosi, F. and Carrassi, A. (2006) *Interventions for treating oral leukoplakia*.

Lopes, V., Murray, P., Williams, H., Woodman, C., Watkinson, J. and Robinson, M. (2011) 'Squamous cell carcinoma of the oral cavity rarely harbours oncogenic human papillomavirus', *Oral Oncol*, 47(8), pp. 698-701.

Lu, Y., Futtner, C., Rock, J.R., Xu, X., Whitworth, W., Hogan, B.L. and Onaitis, M.W. (2010) 'Evidence that SOX2 overexpression is oncogenic in the lung', *PloS One*, 5(6), p. e11022.

Ma, G., Sano, K., Ikeda, H. and Inokuchi, T. (1999) 'Promotional effects of CO(2) laser and scalpel incision on 4-NQO-induced premalignant lesions of mouse tongue', *Lasers Surg Med*, 25(3), pp. 207-12.

Maier, S., Wilbertz, T., Braun, M., Scheble, V., Reischl, M., Mikut, R., Menon, R., Nikolov, P., Petersen, K., Beschorner, C., Moch, H., Kakies, C., Protzel, C., Bauer, J., Soltermann, A., Fend, F., Staebler, A., Lengerke, C. and Perner, S. (2011) 'SOX2 amplification is a common event in squamous cell carcinomas of different organ sites.', *Hum Pathol*, 42(8), pp. 1078-88.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., Ko, M.S. and Niwa, H. (2007) 'Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells', *Nat Cell Biol*, 9(6), pp. 625-35.

McCord, C., Xu, J., Xu, W., Qiu, X., Muhanna, N., Irish, J., Leong, I., McComb, R.J., Perez-Ordonez, B. and Bradley, G. (2014) 'Association of human papilloma virus with atypical and malignant oral papillary lesions', *Oral Surg Oral Med Oral Pathol Oral Radiol*, 117(6), pp. 722-32.

McGurk, M., Chan, C., Jones, J., O'Reagan, E.O. and Sherriff, M. (2005) 'Delay in diagnosis and its effect on outcome in head and neck cancer ', *British Journal of Oral and Maxillofacial Surgery*, 43(4), pp. 281-4.

Meyer, T. and Hart, I.R. (1998) 'Mechanisms of tumour metastasis', *Eur J Cancer*, 34(2), pp. 214-21.

Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q., Bacchetti, S., Haber, D.A. and Weinberg, R.A. (1997) 'hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization', *Cell*, 90(4), pp. 785-95.

Michifuri, Y., Hirohashi, Y., Torigoe, T., Miyazaki, A., Kobayashi, J., Sasaki, T., Fujino, J., Asanuma, H., Tamura, Y., Nakamori, K., Hasegawa, T., Hiratsuka, H. and Sato, N. (2012) 'High expression of ALDH1 and SOX2 diffuse staining pattern of oral squamous cell carcinomas correlates to lymph node metastasis', *Pathol Int*, 62(10), pp. 684-9.

Mishra, R. (2012) 'Biomarkers of oral premalignant epithelial lesions for clinical application', *Oral Oncol*, 48(7), pp. 578-84.

Misuno, K., Liu, X., Feng, S. and Hu, S. (2013) 'Quantitative proteomic analysis of sphere-forming stem-like oral cancer cells', *Stem Cell Res Ther*, 4(6), p. 156.

Mithani, S.K., Mydlarz, W.K., Grumbine, F.L., Smith, I.M. and Califano, J.A. (2007) 'Molecular genetics of premalignant oral lesions', *Oral Dis*, 13(2), pp. 126-33.

Mognetti, B., Di Carlo, F. and Berta, G.N. (2006) 'Animal models in oral cancer research', *Oral Oncol*, 42(5), pp. 448-60.

Molinolo, A., Amornphimoltham, P., Squarize, C., Castilho, R., Patel, V. and Gutkind, J. (2009) 'Dysregulated molecular networks in head and neck carcinogenesis', *Oral Oncol*, 45(4-5), pp. 324-34.

Moore, A.E., Sabachewsky, L. and Toolan, H.W. (1955) 'Culture characteristics of four permanent lines of human cancer cells', *Cancer Res*, 15(9), pp. 598-602.

Mostowska, A., Kobielak, A. and Trzeciak, W.H. (2003) 'Molecular basis of non-syndromic tooth agenesis: mutations of MSX1 and PAX9 reflect their role in patterning human dentition', *Eur J Oral Sci*, 111(5), pp. 365-70.

Mouth Cancer Foundation (2013) *Mouth Cancer Action Month* Available at: http://www.mouthcancerfoundation.org/events/mouth-cancer-action-month (Accessed: October 16th).

Munoz, A. and Gange, S. (1998) 'Methodological issues for biomarkers and intermediate outcomes in cohort studies ', *Epidemiologic Reviews*, 20(1), pp. 29-42.

Muratovska, A., Zhou, C., He, S., Goodyer, P. and Eccles, M.R. (2003) 'Paired-Box genes are frequently expressed in cancer and often required for cancer cell survival', *Oncogene*, 22(39), pp. 7989-97.

Nakahara, W., Fukuoka, F. and Sugimura, T. (1957) 'Carcinogenic action of 4-nitroquinoline-N-oxide', *Gan*, 48(2), pp. 129-37.

Nankivell, P., Williams, H., McConkey, C., Webster, K., High, A., MacLennan, K., Senguven, B., Rabbitts, P. and Mehanna, H. (2013) 'Tetraspanins CD9 and CD151, epidermal growth factor receptor and cyclooxygenase-2 expression predict malignant progreession in oral epithelial dysplasia', *Br J Cancer*, 109(11), pp. 2864-74.

Nankivell, P., Williams, H., Webster, K., Pearson, D., High, A., MacLennan, K., Senguven, B., McConkey, C., Rabbitts, P. and Mehanna, H. (2014) 'Investigation of p16(INK4a) as a prognostic biomarker in oral epithelial dysplasia', *J Oral Pathol Med*, 43(4), pp. 245-9.

Nauta, J.M., Roodenburg, J.L., Nikkels, P.G., Witjes, M.J. and Vermey, A. (1996) 'Epithelial dysplasia and squamous cell carcinoma of the Wistar rat palatal mucosa: 4-NQO model', *Head Neck*, 18(5), pp. 441-9.

Neubuser, A., Koseki, H. and Balling, R. (1995) 'Characterization and Developmental Expression of Pax9, a Paired-Box-Containing Gene-Related to Pax1', *Developmental Biology*, 170(2), pp. 701-716.

Neubuser, A., Peters, H., Balling, R. and Martin, G.R. (1997) 'Antagonistic interactions between FGF and BMP signaling pathways: a mechanism for positioning the sites of tooth formation', *Cell*, 90(2), pp. 247-55.

Neville, B., Damm, D., Allen, C. and Bouquot, J. (2009) *Epithelial Pathology*. 3rd edn (1 vols). St. Louis: Saunders Elsevier.

Nicholson, R., Gee, J. and Harper, M. (2001) 'EGFR and cancer prognosis', *Eur J Cancer* 4, pp. 9-15.

Normanno, N., De Luca, A., Bianco, C., Strizzi, L., Mancino, M. and Maiello, M. (2006) 'Epidermal growth factor receptor (EGFR) signaling in cancer', *Gene*, 366(1), pp. 2–16.

Nylander, K., Dabelsteen, E. and Hall, P.A. (2000) 'The p53 molecule and its prognostic role in squamous cell carcinomas of the head and neck', *J Oral Pathol Med*, 29(9), pp. 413-25.

Oda, K., Matsuoka, Y., Funahashi, A. and Kitano, H. (2005) 'A comprehensive pathway map of epidermal growth factor receptor signaling', *Mol Syst Biol*, 1, p. 2005 0010.

Odell, E.W., Jani, P., Sherriff, M., Ahluwalia, S.M., Hibbert, J., Levison, D.A. and Morgan, P.R. (1994) 'The prognostic value of individual histologic grading parameters in small lingual squamous cell carcinomas. The importance of the pattern of invasion', *Cancer*, 74(3), pp. 789-94.

Okubo, T., Clark, C. and Hogan, B.L. (2009) 'Cell lineage mapping of taste bud cells and keratinocytes in the mouse tongue and soft palate', *Stem Cells*, 27(2), pp. 442-50.

Opitz, O.G., Quante, M., von Werder, A., Heeg, S. and Blum, H.E. (2005) 'A mouse model of oral-esophageal carcinogenesis', *Onkologie*, 28(1), pp. 44-8.

Park, K., Yang, J.H., Choi, Y., Lee, C., Kim, S.Y. and Byun, Y. (2005) 'Chemoprevention of 4-NQO-induced oral carcinogenesis by co-administration of all-trans retinoic acid loaded microspheres and celecoxib', *J Control Release*, 104(1), pp. 167-79.

Parkinson, E. (1989) *The transformation of human epidermal keratinocytes by carcinogens and viruses in vitro*. New York: Raven Press.

Partridge, M., Li, S.R., Pateromichelakis, S., Francis, R., Phillips, E., Huang, X.H., Tesfa-Selase, F. and Langdon, J.D. (2000) 'Detection of minimal residual cancer to investigate why oral tumors recur despite seemingly adequate treatment', *Clin Cancer Res*, 6(7), pp. 2718-25.

Paterson, I.C., Davies, M., Stone, A., Huntley, S., Smith, E., Pring, M., Eveson, J.W., Robinson, M., Parkinson, E.K. and Prime, S.S. (2002) 'TGF-b1 acts as a tumor suppressor of human malignant keratinocytes independently of Smad 4 expression and ligand-induced G1 arrest', *Oncogene*, 21(10), pp. 1616-1624.

Paterson, I.C., Matthews, J.B., Huntley, S., Robinson, C.M., Fahey, M., Parkinson, E.K. and Prime, S.S. (2001) 'Decreased expression of TGF-beta cell surface receptors during progression of human oral squamous cell carcinoma', *J Pathol*, 193(4), pp. 458-67.

Pectasides, E., Rampias, T., Kountourakis, P., Sasaki, C., D, K., Fountzilas, G., Zaramboukas, T., Rimm, D., Burtness, B. and Psyrri, A. (2011) 'Comparative prognostic value of epidermal growth factor quantitative protein expression compared with FISH for head and neck squamous cell carcinoma', *Clin Cancer Res*, 17(9), pp. 2947 - 54.

Peters, H., Neubuser, A., Kratochwil, K. and Balling, R. (1998) 'Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities', *Genes Dev*, 12(17), pp. 2735-47.

Pevny, L.H. and Nicolis, S.K. (2010) 'Sox2 roles in neural stem cells', *Int J Biochem Cell Biol*, 42(3), pp. 421-4.

Phi, J.H., Park, S.H., Kim, S.K., Paek, S.H., Kim, J.H., Lee, Y.J., Cho, B.K., Park, C.K., Lee, D.H. and Wang, K.C. (2008) 'Sox2 expression in brain tumors: a reflection of the neuroglial differentiation pathway', *Am J Surg Pathol*, 32(1), pp. 103-12.

Poh, C., Zhu, Y., Chen, E., Berean, K., Wu, L., Zhang, L. and Rosin, M. (2012) 'Unique FISH patterns associated with cancer progression of oral dysplasia', *J Dent Res*, 91(1), pp. 52 - 57.

Postle, K., Nguyen, T.T. and Bertrand, K.P. (1984) 'Nucleotide sequence of the repressor gene of the TN10 tetracycline resistance determinant', *Nucleic Acids Res*, 12(12), pp. 4849-63.

Prime, S.S., Eveson, J.W., Stone, A.M., Huntley, S.P., Davies, M., Paterson, I.C. and Robinson, C.M. (2004) 'Metastatic dissemination of human malignant oral keratinocyte cell lines following orthotopic transplantation reflects response to TGF-beta 1', *J Pathol*, 203(4), pp. 927-32.

Prime, S.S., Game, S.M., Matthews, J.B., Stone, A., Donnelly, M., Yeudall, W.A., Patel, V., Sposto, R., Siverstone, A. and Scully, C. (1994b) 'Epidermal growth factor and transforming growth factor alpha charactersitics of human oral carcinoma cell lines', *British Journal of Cancer* 69, pp. 8-15.

Prime, S.S., Matthews, J.B., Patel, V., Game, S.M., Donnelly, M., Stone, A., Paterson, I.C., Sandy, J.R. and Yeudall, W.A. (1994a) 'TGF-beta receptor regulation mediates the response to exogenous ligand but is independent of the degree of cellular differentiation in human oral keratinocytes', *Int J Cancer*, 56, pp. 406-412.

Prime, S.S., Nixon, S.V.R., Crane, I.J., Stone, A., Matthews, J.B., Maitland, N.J., Remnant, L., Powell, S.K., Game, S.M. and Scully, C. (1990) 'The behaviour of human oral squamous cell carcinoma in cell culture.', *Journal of Pathology*, 160, pp. 259-269.

Qiao, B., He, B., Cai, J. and Yang, W. (2013) 'The expression profile of Oct4 and Sox2 in the carcinogenesis of oral mucosa', *Int J Clin Exp Pathol*, 7(1), pp. 28-37.

Que, J., Luo, X., Schwartz, R.J. and Hogan, B.L. (2009) 'Multiple roles for Sox2 in the developing and adult mouse trachea', *Development*, 136(11), pp. 1899-907.

Que, J., Okubo, T., Goldenring, J.R., Nam, K.T., Kurotani, R., Morrisey, E.E., Taranova, O., Pevny, L.H. and Hogan, B.L. (2007) 'Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm', *Development*, 134(13), pp. 2521-31.

Ragge, N.K., Lorenz, B., Schneider, A., Bushby, K., de Sanctis, L., de Sanctis, U., Salt, A., Collin, J.R., Vivian, A.J., Free, S.L., Thompson, P., Williamson, K.A., Sisodiya, S.M., van Heyningen, V. and Fitzpatrick, D.R. (2005) 'SOX2 anophthalmia syndrome', *Am J Med Genet A*, 135(1), pp. 1-7; discussion 8.

Rahima, B., Shingaki, S., Nagata, M. and Saito, C. (2004) 'Prognostic significance of perineural invasion in oral and oropharyngeal carcinoma', *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*, 97(4), pp. 423-31.

Redman, R.S., Katuri, V., Tang, Y., Dillner, A., Mishra, B. and Mishra, L. (2005) 'Orofacial and gastrointestinal hyperplasia and neoplasia in smad4+/- and elf+/-/smad4+/- mutant mice', *J Oral Pathol Med*, 34(1), pp. 23-9.

Rheinwald, J.G. and Beckett, M.A. (1980) 'Defective terminal differentiation in culture as a consistent and selectable character of malignant human keratinocytes', *Cell*, 22(2 Pt 2), pp. 629-32.

Rheinwald, J.G. and Beckett, M.A. (1981) 'Tumorigenic keratinocyte lines requiring anchorage and fibroblast support cultured from human squamous cell carcinomas', *Cancer Res*, 41(5), pp. 1657-63.

Rheinwald, J.G. and Green, H. (1975) 'Serial cultivation of strains of human epidermal kerinocytes: the formation of keratinizing colonies from single cells ', *Cell*, 6, pp. 331-344.

Ries, J., Vairaktaris, E., Agaimy, A., Bechtold, M., Gorecki, P., Neukam, F. and Nkenke, E. (2013) 'The relevance of EGFR overexpression for the prediction of the malignant transformation of oral leukoplakia.', *Oncol Rep*.

Robinson, C.M., Stone, A.M., Shields, J.D., Huntley, S., Paterson, I.C. and Prime, S.S. (2003) 'Functional significance of MMP-2 and MMP-9 expression by human malignant oral keratinocyte cell lines ', *Archives of Oral Biology*, 48(11), pp. 779-86.

Robinson, M., Schwartz, S.M., Sloan, P. and Thavaraj, S. (2012) 'HPV specific testing: a requirement for oropharyngeal squamous cell carcinoma patients', *Head and Neck Pathology*, 6(S), pp. 83-90.

Robson, E.J., He, S.J. and Eccles, M.R. (2006) 'A PANorama of PAX genes in cancer and development', *Nat Rev Cancer*, 6(1), pp. 52-62.

Roland, N.J., Caslin, A.W., Nash, J. and Stell, P.M. (1992) 'Value of grading squamous cell carcinoma of the head and neck', *Head Neck*, 14(3), pp. 224-9.

Rose, J.S. (1966) 'A survey of congenitally missing teeth, excluding third molars, in 6000 orthodontic patients', *Dent Pract Dent Rec*, 17(3), pp. 107-14.

Rosin, M. and Califano, J. (2010) 'The Epidermal Growth Factor Receptor Axis: Support for a New Target for Oral Premalignancy', *Cancer Prev Res*, 3(7), pp. 797 - 799.

Rössle, M., Weber, C., Züllig, L., Graf, N., Jochum, W., Stöckli, S., Moch, H. and Huber, G. (2013) 'EGFR expression and copy number changes in low T-stage oral squamous cell carcinomas', *Histopathology*, 63(2), pp. 271 - 278.

Routray, S. and Mohanty, N. (2014) 'Cancer Stem Cells Accountability in Progression of Head and Neck Squamous Cell Carcinoma: The Most Recent Trends!', *Mol Biol Int*, 375325.

Rupniak, H.T., Rowlatt, C., Lane, E.B., Steele, J.G., Trejdosiewicz, L.K., Laskiewicz, B., Povey, S. and Hill, B.T. (1985) 'Characteristics of four new human cell lines derived from squamous cell carcinomas of the head and neck', *J Natl Cancer Inst*, 75(4), pp. 621-35.

Ryott, M., Wangsa, D., Heselmeyer-Haddad, K. and al, e. (2009) 'EGFR protein overexpression and gene copy number increases in oral tongue squamous cell carcinoma', *Eur J Cancer*, 45, pp. 1700 - 1708.

Salaric, I., Povrzanovic, I., Brajdic, D., Luksic, I. and Macan, D. (2015) 'Potentially malignant oral disorders and high-risk habits in liver cirrhosis and lung cancer patients', *Oral Dis*, 21(3), pp. 373-7.

Salazar, C., Smith, R., Garg, M., Haigentz, M., Schiff, B., Kawachi, N., Anayannis, N., Belbin, T., Prystowsky, M., Burk, R. and Schlecht, N. (2014) 'Human papillomavirus-associated head and neck squamous cell carcinoma survival: a comparison by tumor site and initial treatment', *Head Neck Pathol*, 8(1), pp. 77-87.

Salley, J.J. (1954) 'Experimental carcinogenesis in the cheek pouch of the Syrian hamster', *J Dent Res*, 33(2), pp. 253-62.

Sankaranarayanan, R., Ramadas, K., Thara, S., Muwonge, R., Thomas, G., Anju, G. and Matthew, B. (2013) 'Long term effect of visual screening on oral cancer incidence and mortality in a randomized trial in Kerala, India', *Oral Oncol*, 49(4), pp. 314-21.

Schache, A.G., Liloglou, T., Risk, J.M., Filia, A., Jones, T.M., Sheard, J., Woolgar, J.A., Helliwell, T.R., Triantafyllou, A., Robinson, M., Sloan, P., Harvey-Woodworth, C., Sisson, D. and Shaw, R. (2011) 'Evaluation of Human Papilloma Virus Diagnostic Testing in Oropharyngeal Squamous Cell Carcinoma: Sensitivity, Specificity, and Prognostic Discrimination', *Clin Cancer Res*, 17(9), pp. 6262-71.

Schantz, S.P. and Yu, G.P. (2002) 'Head and neck cancer incidence trends in young Americans, 1973-1997, with a special analysis for tongue cancer', *Arch Otolaryngol Head Neck Surg*, 128, pp. 268-274.

Scholes, A.G., Woolgar, J.A., Boyle, M.A., Brown, J.S., Vaughan, E.D., Hart, C.A., Jones, A.S. and Field, J.K. (1998) 'Synchronous oral carcinomas: independent or common clonal origin?', *Cancer Res*, 58(9), pp. 2003-6.

Scholzen, T. and Gerdes, J. (2000) 'The Ki-67 protein: from the known and the unknown', *J Cell Physiol*, 182(3), pp. 311-22.

Sedivy, J.M. (1998) 'Can ends justify the means?: telomeres and the mechanisms of replicative senescence and immortalization in mammalian cells', *Proc Natl Acad Sci U S A*, 95(16), pp. 9078-81.

Shah, J.P. and Gil, Z. (2009) 'Current concepts in management of oral cancer - surgery', *Oral Oncol*, 45(4-5), pp. 394-401.

Shipley, G.D., Pittelkow, M.R., Wille, J.J., Jr., Scott, R.E. and Moses, H.L. (1986) 'Reversible inhibition of normal human prokeratinocyte proliferation by type beta transforming growth factor-growth inhibitor in serum-free medium', *Cancer Res*, 46(4 Pt 2), pp. 2068-71.

Slaughter, D., Southwick, H. and Smekjal, W. (1953) 'Field cancerization in oral stratified squamous epithelium: clinical implications of multicentric origin ', *Cancer*, 6(5), pp. 963 - 968.

Slootweg, P.J., Hordijk, G.J., Schade, Y., van Es, R.J. and Koole, R. (2002) 'Treatment failure and margin status in head and neck cancer. A critical view on the potential value of molecular pathology', *Oral Oncol*, 38(5), pp. 500-3.

Smith, J.R. and Pereira-Smith, O.M. (1996) 'Replicative senescence: implications for in vivo aging and tumor suppression', *Science*, 273(5271), pp. 63-7.

Sobin, L., Gospodarowicz, M. and Wittekind, C. (eds.) (2009) *TNM Classification of Malignant Tumours*. 7th edn. Wiley-Blackwell and Union for International Cancer Control.

Stockton, D.W., Das, P., Goldenberg, M., D'Souza, R.N. and Patel, P.I. (2000) 'Mutation of PAX9 is associated with oligodontia', *Nature Genetics*, 24(1), pp. 18-9.

Strati, K., Pitot, H.C. and Lambert, P.F. (2006) 'Identification of biomarkers that distinguish human papillomavirus (HPV)-positive versus HPV-negative head and neck cancers in a mouse model', *Proc Natl Acad Sci U S A*, 103(38), pp. 14152-7.

Stuart, E.T., Haffner, R., Oren, M. and Gruss, P. (1995) 'Loss of p53 function through PAX-mediated transcriptional repression', *EMBO J*, 14(22), pp. 5638-45.

Sugiyama, M., Speight, P.M., Prime, S.S. and Watt, F.M. (1993) 'Comparison of integrin expression and terminal differentiation capacity in cell lines derived from oral squamous cell carcinomas.', *Carcinogenesis*, 14, pp. 2171-2176.

Szczesniak, M.M., Maclean, J., Zhang, T., Graham, P.H. and Cook, I.J. (2014) 'Persistent Dysphagia after Head and Neck Radiotherapy: A Common and Under-reported Complication with Significant Effect on Non-cancer-related Mortality', *Clin Oncol (R Coll Radiol)*, 26(11), pp. 697-703.

Tada, M. (1976) 'Main binding sites of the carcinogen, 4-nitroquinoline 1-oxide in nucleic acids', *Biochim Biophys Acta*, 454(3), pp. 558-66.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007) 'Induction of pluripotent stem cells from adult human fibroblasts by defined factors', *Cell*, 131(5), pp. 861-72.

Takano, T., Ohe, Y., Sakamoto, H., Tsuta, K., Matsuno, Y. and Tateishi, U. (2005) 'Epidermal growth factor receptor gene mutatin sand increased copy numbers predict gefitinib sensitivity in patients with recurrent non-small cell lung cancer ', *J Clin Oncol*, 23, pp. 6829-37.

Takeda, T., Sugihara, K., Hirayama, Y., Hirano, M., Tanuma, J.I. and Semba, I. (2006) 'Immunohistological evaluation of Ki-67, p63, CK19 and p53 expression in oral epithelial dysplasias', *J Oral Pathol Med*, 35(6), pp. 369-75.

Tanaka, T., Makita, H., Ohnishi, M., Mori, H., Satoh, K. and Hara, A. (1995) 'Chemoprevention of rat oral carcinogenesis by naturally occurring xanthophylls, astaxanthin and canthaxanthin', *Cancer Res*, 55(18), pp. 4059-64.

Tang, X., Knudsen, B., Bemis, D., Tickoo, S. and Gudas, L. (2004) 'Oral cavity and esophageal carcinogenesis modeled in carcinogen-treated mice', *Clin Cancer Res*, 10(1 Pt 1), pp. 301-13.

Temam, S., Kawaguchi, H., El-Naggar, A., Jelinek, J., Tang, H., Liu, D., Lang, W., Issa, J., Lee, J. and Mao, L. (2007) 'Epidermal growth factor receptor copy

number alterations correlate with poor clinical outcome in patients with head and neck squamous cancer', *J Clin Oncol*, 25, pp. 2164-2170.

Thariat, J., Badoual, C., Faure, C., Butori, C., Marcy, P.Y. and Righini, C.A. (2010) 'Basaloid squamous cell carcinoma of the head and neck: role of HPV and implication in treatment and prognosis', *J Clin Pathol*, 63(10), pp. 857-66.

Thavaraj, S., Stokes, A., Guerra, E., Bible, J., Halligan, E., Long, A., Okpokam, A., Sloan, P., Odell, E. and Robinson, M. (2011) 'Evaluation of human papillomavirus testing for squamous cell carcinoma of the tonsil in clinical practice', *J Clin Pathol*, 64(4), pp. 308-12.

Thurman, J.D., Greenman, D.L., Kodell, R.L. and Turturro, A. (1997) 'Oral squamous cell carcinoma in ad libitum-fed and food-restricted Brown-Norway rats', *Toxicol Pathol*, 25(2), pp. 217-24.

Torres-Rendon, A., Stewart, R., Craig, G.T., Wells, M. and Speight, P.M. (2009) 'DNA ploidy analysis by image cytometry helps to identify oral epithelial dysplasias with a high risk of malignant progression', *Oral Oncol*, 45(6), pp. 468-73.

Tsou, Y.A., Hua, C.H., Tseng, H.C., Lin, M.H. and Tsai, M.H. (2007) 'Survival study and treatment strategy for second primary malignancies in patients with head and neck squamous cell carcinoma and nasopharyngeal carcioma ', *Acta Otolaryngol*, 127(6), pp. 651-7.

Tziaferi, V., Kelberman, D. and Dattani, M.T. (2008) 'The role of SOX2 in hypogonadotropic hypogonadism', *Sex Dev*, 2(4-5), pp. 194-9.

Understanding Animal Research (2014) *The three Rs*. Available at: http://www.understandinganimalresearch.org.uk/how/the-three-Rs (Accessed: September 1st).

van der Waal, I. (2009) 'Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification, and present concepts of management', *Oral Oncol*, 45(4 - 5), pp. 317 – 323.

Varun, B.R., Ranganathan, K., Rao, U.K. and Joshua, E. (2014) 'Immunohistochemical detection of p53 and p63 in oral squamous cell carcinoma, oral leukoplakia, and oral submucous fibrosis', *J Investig Clin Dent*, 5(3), pp. 214-9.

Vasioukhin, V., Degenstein, L., Wise, B. and Fuchs, E. (1999) 'The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin', *Proc Natl Acad Sci U S A*, 96(15), pp. 8551-6.

Vered, M., Yarom, N. and Dayan, D. (2005) '4NQO oral carcinogenesis: animal models, molecular markers and future expectations', *Oral Oncol*, 41(4), pp. 337-9.

Vitale-Cross, L., Amornphimoltham, P., Fisher, G., Molinolo, A. and Gutkind, J. (2004a) 'Conditional expression of K-ras in an epithelial compartment that includes the stem cells is sufficient to promote squamous cell carcinogenesis', *Cancer Res*, 64(24), pp. 8804-7.

Vitale-Cross, L., Amornphimoltham, P., Fisher, G., Molinolo, A.A. and Gutkind, J.S. (2004b) 'Conditional expression of K-ras in an epithelial compartment that

includes the stem cells is sufficient to promote squamous cell carcinogenesis', *Cancer Res*, 64(24), pp. 8804-7.

Vitale-Cross, L., Czerninski, R., Amornphimoltham, P., Patel, V., Molinolo, A.A. and Gutkind, J.S. (2009) 'Chemical carcinogenesis models for evaluating molecular-targeted prevention and treatment of oral cancer', *Cancer Prev Res (Phila)*, 2(5), pp. 419-22.

Wallenius, K. and Lekholm, U. (1973) 'Oral cancer in rats induced by the water-soluble carcinogen 4-nitrochinoline N-oxide', *Odontol Revy*, 24(1), pp. 39-48.

Wang, Q., Fang, W.H., Krupinski, J., Kumar, S., Slevin, M. and Kumar, P. (2008) 'Pax genes in embryogenesis and oncogenesis', *J Cell Mol Med*, 12(6A), pp. 2281-94.

Wang, Z., Martin, D., Molinolo, A.A., Patel, V., Iglesias-Bartolome, R., Sol Degese, M., Vitale-Cross, L., Chen, Q. and Gutkind, J.S. (2014) 'mTOR co-targeting in cetuximab resistance in head and neck cancers harboring PIK3CA and RAS mutations', *J Natl Cancer Inst*, 106(9).

Warnakulasuriya, S. (2009) 'Global epidemiology of oral and oropharyngeal cancer', *Oral Oncol*, 45, pp. 309-16.

Warnakulasuriya, S., Harris, C.K., Scarrot, D.M., Watt, R., Gelbier, S., Peters, T.J. and Johnson, N.W. (1999) 'An alarming lack of public awareness towards oral cancer ', *British Dental Journal*, 187(6), pp. 319-22.

Warnakulasuriya, S., Kovacevic, T., Madden, P., Coupland, V.H., Sperandio, M., Odell, E. and Moller, H. (2011) 'Factors predicting malignant transformation in oral potentially malignant disorders among patients accrued over a 10-year period in South East England', *J Oral Pathol Med*, 40(9), pp. 677-83.

Wilbers, J., Kappelle, A.C., Kessels, R.P., Steens, S.C., Meijer, F.J., Kaanders, J.H., Haast, R.A., Versteeg, L.E., Tuladhar, A.M., de Korte, C.L., Hansen, H.H., Hoebers, F.J., Boogerd, W., van Werkhoven, E.D., Nowee, M.E., Hart, G., Bartelink, H., Dorresteijn, L.D. and van Dijk, E.J. (2014) 'Long term cerebral and vascular complications after irradiation of the neck in head and neck cancer patients: a prospective cohort study: study rationale and protocol', *BMC Neurol*, 14, p. 132.

Williamson, K.A., Hever, A.M., Rainger, J., Rogers, R.C., Magee, A., Fiedler, Z., Keng, W.T., Sharkey, F.H., McGill, N., Hill, C.J., Schneider, A., Messina, M., Turnpenny, P.D., Fantes, J.A., van Heyningen, V. and FitzPatrick, D.R. (2006) 'Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome', *Hum Mol Genet*, 15(9), pp. 1413-22.

Wiznerowicz, M. and Trono, D. (2003) 'Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interferenceq', *J Virol*, 77(16), pp. 8957-61.

Wollnik, B., Tukel, T., Uyguner, O., Ghanbari, A., Kayserili, H., Emiroglu, M. and Yuksel-Apak, M. (2003) 'Homozygous and heterozygous inheritance of PAX3 mutations causes different types of Waardenburg syndrome', *Am J Med Genet A*, 122A(1), pp. 42-5.

Wong, D. (1987) 'Amplification of the c-erb B1 oncogene in chemically-induced oral carcinomas', *Carcinogenesis*, 8(12), pp. 1963-5.

Wong, K.K. (2009) 'Oral-specific chemical carcinogenesis in mice: an exciting model for cancer prevention and therapy', *Cancer Prev Res (Phila)*, 2(1), pp. 10-3.

Woolgar, J.A., Rogers, S., West, C.R., Errington, R.D., Brown, J.S. and Vaughan, E.D. (1999) 'Survival and patterns of recurrence in 200 oral cancer patients treated by radical surgery and neck dissection', *Oral Oncol*, 35, pp. 257-265.

Wu, J.Y., Yi, C., Chung, H.R., Wang, D.J., Chang, W.C., Lee, S.Y., Lin, C.T., Yang, Y.C. and Yang, W.C. (2010) 'Potential biomarkers in saliva for oral squamous cell carcinoma', *Oral Oncol*, 46(4), pp. 226-31.

Xiang, R., Liao, D., Cheng, T., Zhou, H., Shi, Q., Chuang, T.S., Markowitz, D., Reisfeld, R.A. and Luo, Y. (2011) 'Downregulation of transcription factor SOX2 in cancer stem cells suppresses growth and metastasis of lung cancer', *Br J Cancer*, 104(9), pp. 1410-7.

Yamaoka, T., Yano, M., Yamada, T., Matsushita, T., Moritani, M., Ii, S., Yoshimoto, K., Hata, J. and Itakura, M. (2000) 'Diabetes and pancreatic tumours in transgenic mice expressing Pa x 6', *Diabetologia*, 43(3), pp. 332-9.

Yao, F., Svensjo, T., Winkler, T., Lu, M., Eriksson, C. and Eriksson, E. (1998) 'Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells', *Hum Gene Ther*, 9(13), pp. 1939-50.

Yarden, Y. and Schlessinger, J. (1987) 'Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor', *Biochemistry*, 26(5), pp. 1443-51.

Yeudall, W.A., Paterson, I.C., Patel, V. and Prime, S.S. (1995) 'Presence of human papillomavirus sequences in tumour-derived human oral keratinocytes expressing mutant p53', *Oral Oncology European Journal of Cancer*, 31B, pp. 136-143.

Yeudall, W.A., Torrance, L.K., Elsegood, K.A., Speight, P.M., Scully, C. and Prime, S.S. (1993) 'ras gene point mutation is a rare event in premalignant tissues and malignant cells and tissues from oral mucosal lesions', *Oral Oncology European Journal of Cancer*, 28B, pp. 63-68.

Yuen, P.W., Chow, V., Choy, J., Lam, K.Y., Ho, W.K. and Wei, W.I. (2001) 'The clinicopathologic significance of p53 and p21 expression in the surgical management of lingual squamous cell carcinoma', *Am J Clin Pathol*, 116(2), pp. 240-5.

Zhang, Z., Wang, Y., Yao, R., Li, J., Lubet, R.A. and You, M. (2006) 'p53 Transgenic mice are highly susceptible to 4-nitroquinoline-1-oxide-induced oral cancer', *Mol Cancer Res*, 4(6), pp. 401-10.

Zhou, G., Hasina, R., Wroblewski, K., Mankame, T.P., Doci, C.L. and Lingen, M.W. (2010) 'Dual inhibition of vascular endothelial growth factor receptor and epidermal growth factor receptor is an effective chemopreventive strategy in the mouse 4-NQO model of oral carcinogenesis', *Cancer Prev Res (Phila)*, 3(11), pp. 1493-502.

Zwahlen, R.A., Dannemann, C., Gratz, K.W., Studer, G., Zwahlen, D., Moergeli, H., Drabe, N., Buchi, S. and Jenewein, J. (2008) 'Quality of life and psychiatric

morbidity in patients successfully treated for oral cavity squamous cell cancer and their wives', *J Oral Maxillofac Surg*, 66(6), pp. 1125-32.

Appendix

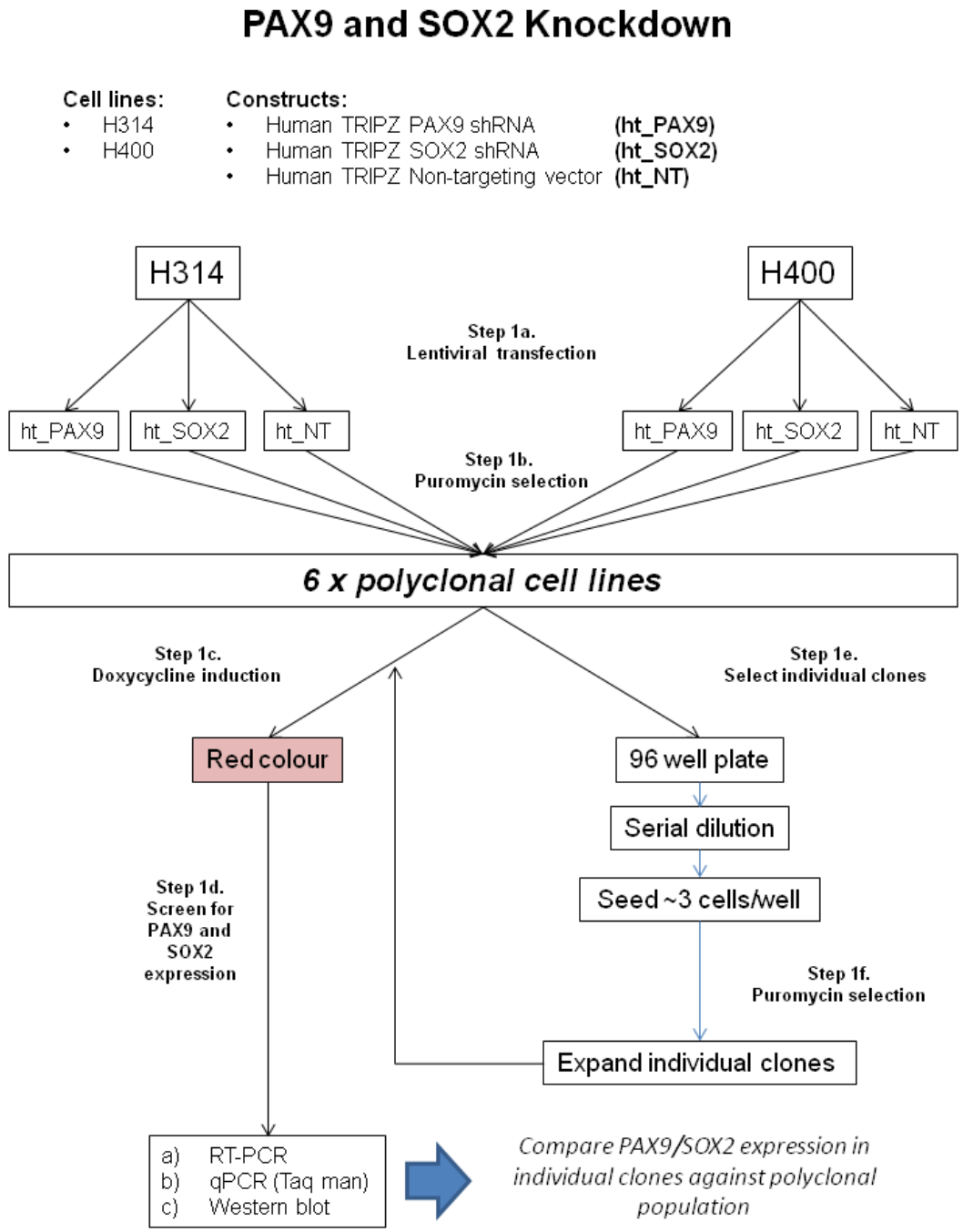
Emergency procedures for dealing with a 4-NQO spillage

4-NQO phase	Action
Powder	Remove all sources of ignition Dampen solid material with acetone Wearing gloves, soak up using paper towels Place damp towels in a clinical waste bag Wash all contaminated surfaces with acetone Wash again with soap and water Seal all contaminated towels gloves, etc, in a clinical waste bag
Stock and experimental solutions	Remove all sources of ignition Wipe the conatminated area with paper towels Wash all contaminated surfaces with acetone Wash again with soap and water Seal all contaminated towels gloves, etc, in a clinical waste bag

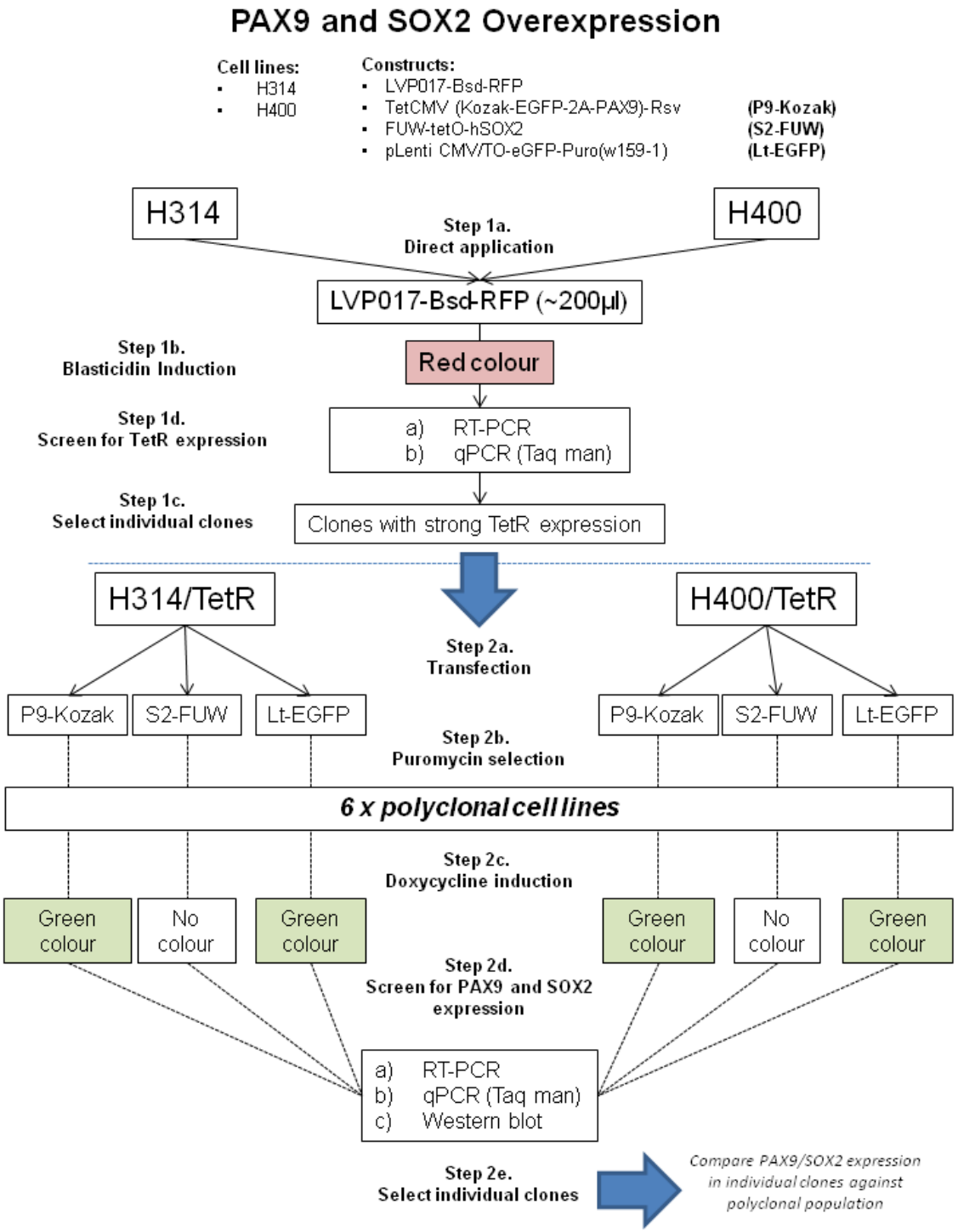
First aid measures for dealing with human exposure to 4-NQO

Site of contact	First Aid
Eyes	Check for contact lenses and remove if present Flush eyes with normal saline (or water if saline is unavailable) for between 20 and 30 minutes
Skin	Immediately flood the affected area with water for at least 15 minutes Remove contaminated clothing. Isolate in a sealed clinical waste bag. Gently wash with soap and water
Inhalation	Leave the contaminated area Take deep breaths of fresh air Monitor for symptoms such as wheezing, coughing, shortness of breath, or a burning sensation in the oral cavity or throat
Ingestion	Do not attempt to induce vomiting Drink 1-2 glasses of water to dilute the 4-NQO
Following administration of first aid seek urgent medical advice	

Experimental outline for the knockdown lentivirus transfections



Experimental outline for the over-expression lentivirus transfections



Summary of reagents not otherwise specified in the text

Reagent/ Buffer	Source	Information
100bp DNA ladder	Promega	100bp DNA marker for agarose gel electrophoresis
Go-Taq Polymerase and Buffer (15mM McCl2)	Promega	-
10x dNTP	Promega	dATP, dTTP, dGTP and dCTP mix
Antibody Diluent	Dako-Envision Kit	Used to dilute antibodies to desired dilutions
Peroxidase blocking solution	Dako-Envision Kit	Blocking reagent for IHC
HRP-labelled polymer	Dako-Envision Kit	Conjugated with secondary antibodies
DAB substrate buffer	Dako-Envision Kit	1 drop DAB chromagen stain into 1ml DAB substrate buffer
DAB chromagen stain	Dako-Envision Kit	
Trizol reagant	Invitrogen	Used for RNA extraction
Random Primer (500ng/ml)	Promega	Used for reverse transcription
Citric acid buffer	Sigma Aldrich	4.2g Sigma Aldrich citric acid powder in 2l of distilled H_20, corrected to pH 6.0 using 5M NaOH
10 x TBS buffer	Sigma Aldrich	87.60g Nacl, 24.24g Trizma base, 900ml H_20, adjusted to pH 7.4 with HCl
1M EDTA antigen retrieval buffer	Sigma Aldrich	0.372g Sigma Aldrich EDTA powder, 1l distilled H_2O, corrected to pH 8.0 using 5NM NaOH
Scotts' bluing solution	Sigma Aldrich	2.0g sodium bicarbonate, 20.0g magnesium sulphate, 1000ml dH_2O