Division of Cancer and Genetics
School of Medicine
Cardiff University

PARP inhibition in novel oropharyngeal cancer cell lines

By Evelyne Pirotte

2017

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Acknowledgements

I would like to thank my supervisors, Dr Ned Powell and Dr Mererid Evans for giving me the opportunity to undertake this PhD and for their guidance and support throughout this project. I would like to extend my gratitude to Dr Steve Man for the helpful advice he offered during the writing process.

Thank you to the staff of Medical Microbiology and Haematology for their technical assistance, in particular to Dr Sian Meyrick and Dr Stephen Austin. I would also like to thank the members of the HPV research team for their help and advice.

I am especially grateful for the continuous support of my friends and family over the past 4 years. I would especially like to thank my parents for their endless encouragements.

Finally, I would like to thank Cancer Research Wales for funding this project, as well as the Division of Cancer and Genetics and the Cardiff University School of Medicine.

Summary

In recent decades many developed countries have seen unprecedented increases in incidence of Human Papillomavirus (HPV) associated oropharyngeal squamous cell carcinoma (OPSCC). HPV-positive OPSCC represents a new disease entity and preclinical assessment of novel therapies is hampered by a lack of relevant *in vitro* models. It is well-established that HPV-positive OPSCC patients generally survive longer than HPV-negative patients, and this may be partly attributable to defective repair of DNA double-strand-breaks in HPV-positive tumours.

This study aimed to develop novel cell line models of HPV-OPSCC and use them, and previously validated lines, to test the hypothesis that defective DNA repair in HPV-positive cells could be exploited by synthetic lethal therapy using the Poly (ADP-ribose) polymerase (PARP) inhibitor Olaparib.

Two novel HPV-positive OPSCC cell lines were derived and characterised. mRNA sequencing confirmed expression of the HPV oncogenes and HPV integration state, but did not show consistent differences in transcript levels of genes involved in DNA repair between HPV-positive and negative lines. The effects of Olaparib were assessed in a panel of eight cell lines, including effects on colony formation, cell cycle distribution, DNA double-strand-break persistence and p53 activity. All lines were sensitive to high doses of Olaparib (10 μM), however at doses between 0.5-1 μM, the surviving fractions differed significantly between lines. Two HPV-positive lines were sensitive to Olaparib (Surviving Fraction (SF)<40%), and two were resistant (SF>80%). Neither HPV-status, nor basal levels of PARP correlated with Olaparib sensitivity. The data were not consistent with the original hypothesis, but did suggest that monotherapy with PARP inhibitors might be useful in some OPSCC patients.

The study also included an investigation into the natural history of HPV in the oropharynx. This demonstrated that HPV infection is a rare event in non-malignant tonsil tissue (prevalence of 0%: 95% confidence interval (CI) 0-0.58%).

Abbreviations

γ-H2AX	H2AX histone phosphorylated at serine 139
AP	Apurinic apyrimidinic
APE	Apurinic apyrimidinic Endonuclease
APOBEC	Apolipoprotein B mRNA editing catalytic polypeptide-like
APTX	Aprataxin
ATM	Ataxia telangiectasia mutated
ATP	Adenosine Triphosphate
ATR	Ataxia telangiectasia and RAD3 related
BER	Base Excision Repair
BP	Biological Processes
CDK	Cyclin Dependent Kinase
CC	Cellular Components
CI	Confidence Interval
DAPI	4',6-Diamidine-2'-phenylindole
DEIA	DNA EIA
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA-PK	DNA dependent protein kinase
dNTP	Deoxynucleotide
DSB	Double Strand Breaks
DSBR	Double Strand Break Repair
E	Early
EDTA	Ethylenediaminetetraacetic acid
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
EIA	Enzyme Immunoassay
ELISA	Enzyme Linked Immunosorbent Assay
FA	Fanconi Anemia
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FFPE	Formalin Fixed Paraffin Embedded
FITC	Fluoresceinisocyanate
GMEM	Glasgow Minimal Essential Medium
GO	Gene Ontology
HE	Hematoxylin and Eosin
HEKn	Human neonatal foreskin keratinocytes
HGKS	Human Keratinocyte Growth Supplement
HNSCC	Head and Neck Squamous Cell Carcinoma
HPV	Human Papillomavirus
HR	Homologous Recombination
hTERT	Human telomerase reverse transcriptase
IGV	Integrative Genomics Viewer
IHC	Immuno Histochemistry

IR	Ionising Radiation
ISH	In situ hybridisation
L	Late
LCR	Long control regions
MMR	Mismatch Repair
MF	Molecular Functions
mRNA-seq	Whole transcriptome messenger RNA sequencing
NCS	Newborn Calf Serum
NEAA	Non Essential Amino Acids
NER	Nucleotide Excision Repair
NHEJ	Non-Homologous End Joining
OPSCC	Oropharyngeal Squamous Cells Carcinoma
ORF	Open Reading Frame
PAR	Poly (ADP-ribose)
PARP	poly (ADP-ribose) polymerase
PARPi	poly (ADP-ribose) polymerase inhibitor
PBS	Phosphate Buffer Saline
PCOC	Primary Culture of Oropharyngeal Cells
PCR	Polymerase Chain Reaction
PD	Population Doubling
Pen Strep	Penicillin/Streptomycin
PNKP	Polynucleotide kinase 3'-posphatase
p-p53	Phospho-p53
pRb	Retinoblastoma protein
RIN	RNA integrity Number
RPKM	Reads Per Kilobase of transcripts per Million mapped reads
RT	Room Temperature
SCC	Squamous Cell Carcinoma
SF	Surviving Fraction
SNP	Single Nucleotide Polymorphism
SSB	Single Strand Breaks
STR	Short Tandem Repeat
TBE	Tris/Borate/EDTA
TE	Trypsin/EDTA
TIL	Tumour Infiltrating Lymphocyte
TNM	Tumour Node Metastasis
XRCC	X-ray cross-complementation

Contents

Figures

Tables

Chapter 1 - Introduction

1.1 Human Papillomavirus

Papillomaviruses are found in many species, but in higher vertebrates in particular. Over 200 types have been described in human so far (Doorbar et al. 2016). Papillomaviruses are strictly species specific and display a tropism towards epithelial cells, both in cutaneous and mucosal epithelium, often in specific locations (Egawa et al. 2015).

1.1.1 Taxonomy

Papillomaviruses were originally classified in the phylogenic group papovaviridae together with the polyomaviruses with which they share the same basic structure, being non-enveloped viruses with double stranded DNA. However their genomes and protein expression differ and two separate families, the polyamaviridae and the papillomaviridae, were classified in the 1970s (Fenner 1976). Papillomaviruses are classified genetically, based on the DNA sequence of the highly conserved L1 Open Reading Frame (ORF). The main subgroups, called genera, share less than 60% sequence similarity with one another.

Human Papillomaviruses (HPV) are divided in to Alpha, Beta, Gamma, Mu and Nu genera with the majority in Alpha and Beta. Each genera is further divided into smaller groups or species, which comprise several types of papillomaviruses. Virus types differ in at least 10% of the L1 sequence whilst subtypes are defined by sequence variations smaller than 10% (2-10%). The numbering dates back to early research on HPV typing and corresponds to the order of discovery (de Villiers et al. 2004; Egawa et al. 2015). Phylogenetic tree representations do not usually correspond to any pathological features of the viruses, although the oncogenic HPV types, including HPV16 and 18, are all members of the Alpha genus (Doorbar et al. 2012). In humans an alternative classification, developed with regards to the 40 types which can infect the genital mucosa, and referring to pathological implications, is often used. HPV are separated in two categories, high risk and low risk, which describes their oncogenic potential. The high risk group comprises at least 14 types, the main ones being HPV16 and 18 predominantly associated with anogenital cancers. The most common low risk types of HPV are HPV6 and 11, which are the main causal agents for genital warts (Muñoz et al. 2003).

1.1.2 Physical and genetic structure

HPV are non-enveloped viruses, with a 55 nm diameter icosahedral capsid. The capsid is formed of two proteins: L1 (55 kDa), the major capsid protein, comprising about 80% of the capsid, and L2 (50 kDa) comprising the remaining 20%. The capsid is composed of 72 pentameres of the L1 protein (capsomeres), at the centre of which a unit of the L2 protein can be present, there are between 12 and 72 units of L2 in a capsid. A notable feature of the HPV capsid, is the capacity for self-assembly of L1 into pentameres and then into a full capsid. This property is exploited for the production of virus-like-particle vaccines directed against the main oncogenic types of HPV and HPV6 and 11 which cause genital warts (Buck et al. 2008).

The 7900 bp genome of HPV consists of double stranded DNA. It contains 8 overlapping ORF and is divided in to three regions (Figure 1.1). The early (E) region of around 4 kb, encodes 6 proteins expressed at the beginning of the viral life cycle and involved in the replication and pathogenicity of the virus. The 3 kb late (L) region, encodes the capsid proteins L1 and L2 that are expressed towards the end of the viral cycle (Doorbar et al. 2016). The Long Control Region (LCR) is situated between the E and L sequences, and is a non-coding regulatory region involved in the control of genes expression as well as encapsidation, replication and transcription of the genome. This region also contains the promoter sequences of the E genes. High risk HPV have two promoters: P97, involved in the expression of E genes is situated upstream of the E6 open reading frame, and the P742 (P670) promoter, located in the E7 ORF which functions as the main promoter for late gene expression (Longworth and Laimins 2004).

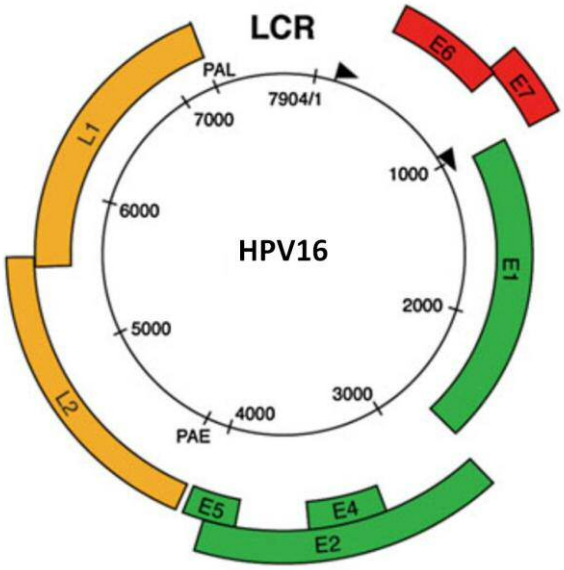


Figure 1.1: HPV16 genomic structure
HPV16 genome is composed of 6 early genes, whose products are involved in virus replication, and two late genes, encoding the capsid proteins. The LCR is involved in gene regulation.

1.1.3 HPV life cycle and infection

HPV is one of the most common sexually transmitted viruses. Most sexually active people will contract a genital HPV infection at some point in their life, however the majority of infections are transient, and few will persist and develop into a neoplasia (Moscicki et al. 2010). HPV infection and life cycle are closely related to the differentiation and maturation of epithelial cells. The viral life cycle and subsequent infection can be divided in the steps described below.

1.1.3.1 Cell entry

The virus infects cells in the basal epithelial layer which are reached through micro abrasions. Alternatively, in some cases the nature of the epithelium, such as a reticulated epithelium in the tonsils, can facilitate the virus access to the basal layer. Once the virus reaches the basal layer of the epithelium, part of the L1 protein binds to heparan sulfate proteoglycans on the keratinocytes surface or the extracellular matrix (Culp et al. 2006). Several heparan sulfate proteoglycans, which could be involved in this process, have been described. The main one is syndecan-1 which is mainly expressed in epithelial cells and up-regulated during wound healing (Elenius et al. 1991). The binding causes a conformational

change in L1 which exposes a highly conserved furin convertase site on L2 (Day et al. 2008). The exposed site leads to the furin mediated cleavage of the N-terminal region of L2. Another site in turn becomes exposed on L1, this neutralizing epitope is then able to bind to secondary cellular receptors (such as integrins α6 and β4) (Shafti-keramat et al. 2003)). Following binding to the cellular surface, the virions are internalized into the host cell. Various mechanism of entry have been described, but recently the endocytosis has been shown to be independent of clathrin, caveolin, lipid raft, flotillin, cholesterol and dynamin (Schelhaas et al. 2012). The virus uncoats in the endosome and escapes through an L2 mediated mechanism. An L2 and HPV genome complex is formed and trafficked through the cytoplasm to the nucleus where RNA transcription begins (Raff et al. 2013).

1.1.3.2 Maintenance phase

After cellular entry, the viral genome is maintained in an episomal form at a low level, around 50-100 copies per cell, with low level expression of E genes (Flores et al. 1999). Transcription of the early genes is regulated by the P97 early promoter located on the 3' end of the LCR region. Various cellular transcription factors and the E2 gene products are involved in its regulation (Kammer et al. 2000). Activation of P97 in the early stage of infection leads to the constant expression of low level E1 and E2 proteins (Wilson et al. 2005). E1 binds to the origin of replication site, and exhibits an Adenosine Triphosphate (ATP) dependent helicase activity, causing the unwinding and separation of the DNA double helix necessary for the initiation of replication (Lin et al. 2002; Longworth and Laimins 2004). E2 functions in conjunction with E1, binding to sequences in the LCR region close to the E1 binding site. The C terminal region of E2 facilitates E1 binding to the DNA (Sedman and Stenlund 1998). E2 also plays an important role in the establishment of a productive infection. During mitosis, E2 ensures the efficient segregation of viral genomes into daughter cells (Van Tine et al. 2004). Another role of E2 is the regulation, both positive and negative, of gene expression. Its binding sites overlap with the sites of various transcription factors, with which E2 interacts. This includes P97 which is positively regulated by low E2 concentration and negatively by high level of E2, enabling the self-regulation of P97. This process ensures a balanced level of viral gene products such as E6 and E7 and promotes stable genome replication in undifferentiated cells (Steger and Corbach 1997; Longworth and Laimins 2004). E4 and E5 proteins are involved in the regulation of the later events of the cycle (Fehrmann et al. 2003).

1.1.3.3 Proliferation and amplification

In epithelia, the basal layer cells divide, producing daughter cells which separate from the basal membrane, and differentiate as they progress through the layers. Fully differentiated cells are shed once they reach the epithelial surface. E6 and E7 activity is increased once this process begins. Together they enable cells to continue cycling which maintains a supply of the S-phase cellular factors necessary for genome replication, and prevents cellular differentiation.

E7 causes the degradation of retinoblastoma protein (pRb), promoting entry into S-phase (Dyson 1998). E7 is also able to interact with other types of proteins with roles related to cell proliferation such as histone deacetylases and cyclin dependent kinase inhibitors p21 and p27 (Doorbar 2005). The activities of the E6 protein are complementary to E7, and are primarily involved with preventing apoptosis (zur Hausen 2002). In order to produce viral particles in the host cell, the viral genome is amplified. This amplification process results in accumulation of a higher number of genome copies in the upper layers of the epithelium (Middleton et al. 2003). During this phase of the viral cycle, the late promoter P670 (E2 independent) is activated in terminally differentiated cells and favors the expression of the early proteins involved in DNA replication (E1, 2, 4 and 5). This results in an increase of viral copy number to about 1000 per cell (Flores et al. 1999).

1.1.3.4 Assembly

The last stages of the cycle consist of the encapsidation of the genomes to form full virions that are subsequently released. The late genes (L1 and L2) are expressed in the upper epithelial layers, followed by self-assembly of the capsid. Viral assembly involves the coordinated association between human proteins, HPV DNA, L1 L2 and E2 proteins (Doorbar et al. 2016). The viral particles are then released when the host cells naturally shed at the epithelial surface, this mechanism of release contributes to the low immune response against the virus, as there is no pathological cell death (Woodman et al. 2007).

1.1.4 Molecular biology of HPV mediated carcinogenesis

The malignant transformation of HPV infected cells is a rare and accidental event (Ho 1995). When it occurs, the pattern of HPV gene expression is disrupted and the virus loses the capacity to produce virions. This type of infection is therefore referred as non-productive (Münger and Howley 2002; Middleton et al. 2003). E6 and E7 are central to malignant transformation, their effects are described in more details below together with other factors promoting carcinogenesis.

1.1.4.1 Cell cycle progression

The c-terminal end of E6 protein comprises a PDZ binding domain enabling interaction with a range of human proteins containing a PDZ domain (Thomas et al. 2008; Pim et al. 2012). The PDZ binding activity is also involved in cell cycle regulation via the expression of cyclin B1, necessary for the G2/M transition (Russo et al. 2006). Other identified protein interactions are involved in the regulation of cell polarity, adhesion and differentiation (Elsum et al. 2012).

E7 causes the degradation of pRb and pRb-related proteins, resulting in release of transcription factor E2F, which regulates the expression of genes involved in G1/S-phase progression (Figure 1.2), leading to inappropriate cell cycle progression (Dyson 1998). This includes increased expression of cyclin A and E, as well as modifications of cyclin dependent kinase 2 (CDK2) activity (Duensing and Münger 2004). One of the genes up-regulated by E2F activity encodes p16, which is involved in regulation of cell cycle progression via a negative feedback loop suppressing cyclinD/CDK4-6 mediated phosphorylation of pRb. During HPV infection, p16 levels increase but the feedback loop is ineffective, as release of E2F is no longer cyclinD/CDK4-6 dependent. Hence p16 is often used as a biomarker of HPV infection. E7 is also able to interact with other types of proteins with roles related to cell proliferation such as histone deacetylases and CDK inhibitors p21 and p27 (Doorbar 2005). CDK and their inhibitors are essential to cell cycle regulation. p21 and p27 are involved in preventing G1/S progression via the inhibition of CDK. Hence neutralization of p21 and p27 due to E7 binding results in constantly high levels of CDK activity and unregulated G1/S progression.

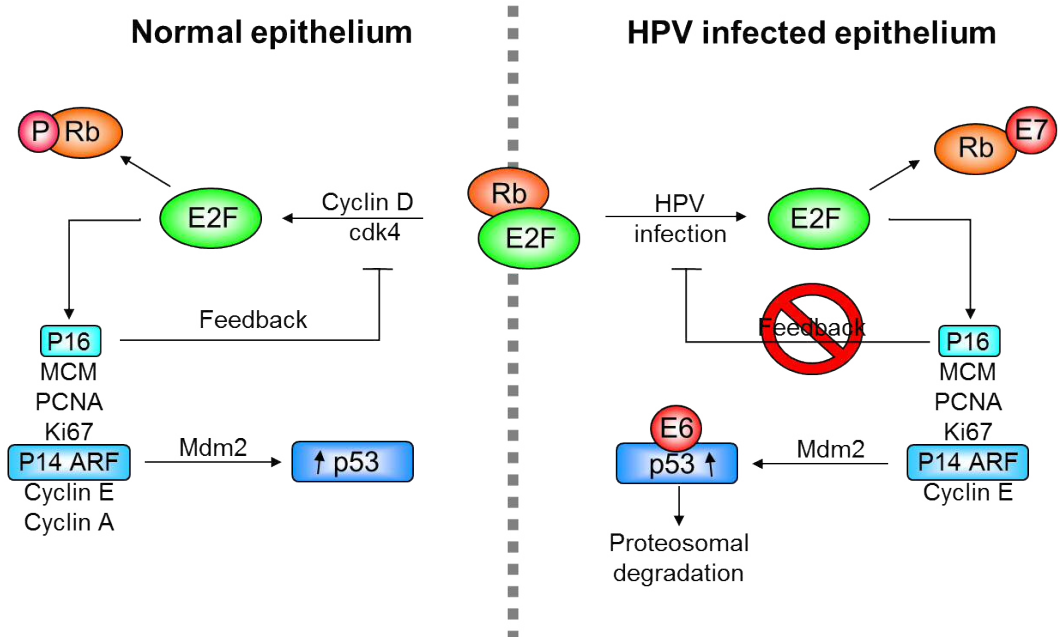


Figure 1.2: Effect of E6 and E7 on HPV infected epithelial cells

The Rb tumour suppressor protein regulates progression through the G1/S cell cycle checkpoint. In normal epithelia E2F transcription factors are only released when Rb is phosphorylated by the cyclinD/CDK4 complex. Binding of E2F transcription factors to multiple promotors results in increased expression of genes involved in S-phase progression. E2F transcription factors also activate a negative feedback loop in which increased expression of p16 inhibits the activity of the cyclinD/CDK4 complex leading to reduced expression of E2F dependent genes. Additional feedback is achieved via p14/ARF, which inhibits Mdm2-mediated degradation of p53 resulting in accumulation of p53 which promotes cell cycle arrest. In HPV infected epithelia, HPV E7 protein binds Rb and causes the release of E2F transcription factors, but negative feedback does not occur as release of E2F is no longer driven by cyclinD/CDK4.

1.1.4.2 p53 degradation

E6 associates with the endogenous ubiquitin ligase E6-Apurine apyrimidic (AP), and this complex targets p53 for ubiquitinylation eventually leading to proteosomal degradation. This inhibits the apoptotic response which should be associated with forced entry into S-phase, and also facilitates the accumulation of mutations and consequent genomic instability (zur Hausen 2002).

1.1.4.3 Telomerase activity

E6 has several p53 independent functions, and these include increasing the activity of the human telomerase reverse transcriptase (hTERT), the catalytic subunit of telomerase, by targeting its promoter repressor for proteosomal degradation. hTERT is also indirectly activated by E6 through the degradation of p53. E6 also directly interacts with the telomerase complex to increase telomerase activity. This enables the maintenance of telomere integrity following multiple divisions, and contributes to replicative immortality (Klingelhutz et al. 1996). The effects of loss of cell cycle control and p53 mediated, combined with telomerase activation are summarized in Figure 1.3.

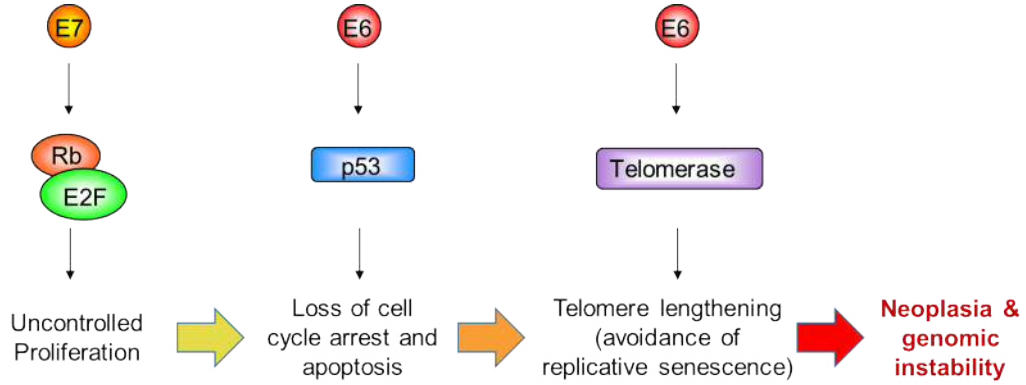


Figure 1.3: The combined effects of HPV gene expression
The E7 protein deregulates the activity of Rb resulting in uncontrolled proliferation. This would ordinarily be associated with p53 mediated cell cycle arrest and apoptosis, but this is circumvented by reduced levels of p53 due to the activities of the HPV E6 protein. Uncontrolled proliferation normally results in replicative senescence associated with telomere shortening, but telomere length is maintained due to increased telomerase activity associated with presence of HPV E6 protein.

1.1.4.4 Genomic instability

Expression of HPV E6 and E7, and consequent unregulated cell cycle progression, contribute to the accumulation of mutations and genomic instability in the host cell. E7, via its inhibiting effect on some protein kinases (e.g. cyclinE cyclinA/CDKD2 complexes), can lead to abnormal centrosome and centriole number. The joint effect of E6 and E7 can result in incorrect chromosome numbers and aberrant mitotic figures (Duensing et al. 2000; Duensing et al. 2001). Genomic instability also facilitates integration of the viral DNA into the host genome. Integration occurs mainly, but not exclusively, at common fragile sites in the human genome. Integration does not seem to be sequence specific and lots of variation can be seen in the integration sites (Thorland et al. 2003). During integration, viral DNA is often disrupted in or near the genes encoding E2 and E1 or other regions downstream of E6 and E7. This often disrupts the expression of E2 which, as mentioned earlier, is involved in the negative regulation of E6 and E7 expression. Loss of E2 control facilitates the high level and long term expression of E6 and E7, hence viral integration is thought to be an important factor in carcinogenesis (Choo et al. 1987). However, integration events are not present in all patients, implying that there are alternative reasons for dysregulation of viral gene expression (Moody and Laimins 2010).

1.1.5 HPV and cancer

Since the 1960s studies involving first animal, and then human, data indicated a probable association between papillomavirus and cancer (Olson et al. 1965). The link between HPV and cervical cancer was more firmly established in 1983, by work conducted in the laboratory of Prof. Harald zur Hausen. It showed the presence of HPV16 DNA in cervical cancers from women diagnosed in diverse geographical locations, including Germany, Kenya and Brazil (Dürst et al. 1983). HPV was subsequently identified in 99.7% of cervical cancers, with two types alone causing 70% of the lesions. HPV 16 was reported in c. 50% of cases and HPV 18 in c.20%; the remaining cases were caused by other high risks types of HPV (Walboomers et al. 1999). This lead to HPV being regarded as a “necessary, but not sufficient” factor in the aetiology of cervical cancer. Over the past two decades, other types of cancers, in both men and women have also been associated with high risk HPV infections. These include anogenital cancers (vulva, penis, anus) and head and neck cancers (IARC 2007).

1.2 Oropharyngeal squamous cells carcinoma

1.2.1 Anatomy of the oropharynx

Head and neck cancers refer to a variety of upper aerodigestive tract, craniofacial bone, skin and soft tissue neoplasms, i.e. they encompass several anatomical sites, in which the disease behaviour can differ markedly. The oropharynx is situated posterior to the oral cavity and connects it with the nasopharynx (above) and larynx (below). In the oral cavity, the oropharynx is delimited by the palatoglossal arch. The palatine tonsils, lingual tonsils and the posterior third of the tongue, soft palate and posterior pharyngeal wall are all part of the oropharynx (Figure 1.4).

The aerodigestive tract represents a significant potential entry route for pathogens and therefore contains several lymphoid tissues located around the tract, described as Waldeyer's ring. This includes the palatine and lingual tonsils, situated in the oropharynx. The tonsils are formed by a stratified squamous epithelium with invaginations that forming the crypts (Figure 1.5). The tonsil crypts are covered by a reticulated (net-like) epithelium which facilitates the interaction between antigens and immune cells. Lymphoid nodules are present under the epithelial surface and the base of the tonsil is separated from surrounding tissue by a capsule. The oropharynx itself is predominantly covered by squamous epithelia.

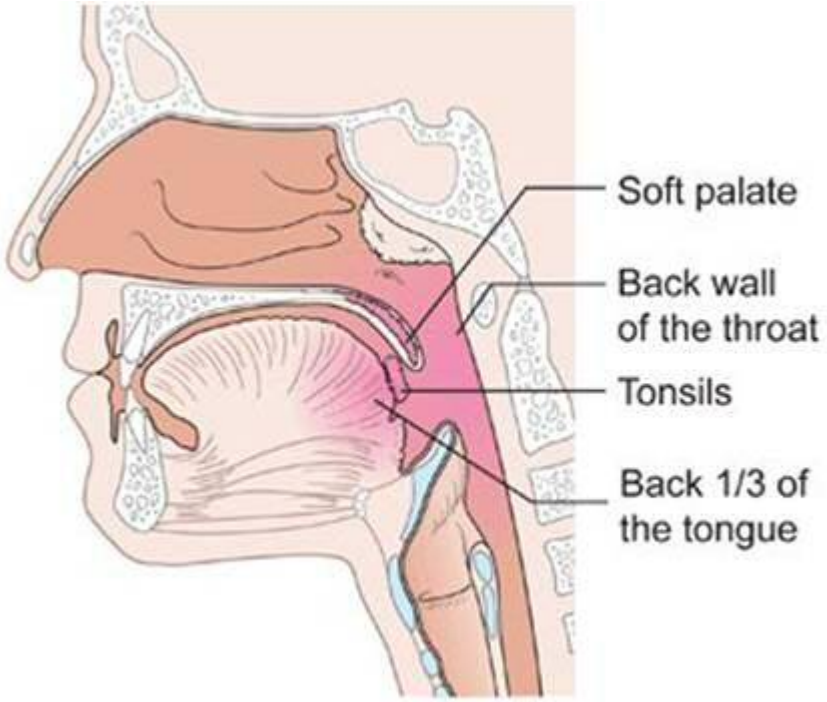


Figure 1.4: Anatomy of the oropharynx

The oropharynx is located posterior to the oral cavity and includes the soft palate (the back part of the roof of the mouth), the side and back wall of the throat, the tonsils and the back third of the tongue. Picture obtained from Cancer Research UK.

Figure 1.5: Anatomy of the palatine tonsil

The surface epithelia deeply invaginates forming crypts (highlighted in the black box). Drawing by Max Brödel, John Hopkins University.

1.2.2 Epidemiology

Head and neck cancers, which include oropharyngeal tumours, are estimated to be the 10th most common cancers worldwide, but are the 7th highest cause of cancer related deaths. In 90% of cases these tumours are squamous cell carcinomas (SCC). The incidence of head and neck SCC (HNSCC) varies with geographical location, higher incidence rates are observed in the Indian subcontinent, Australia, France, Brazil, and Southern Africa. Nasopharyngeal cancer is disproportionately common in southern China. In recent years the incidence of oral, laryngeal, and other smoking related cancers has stabilised or declined in North America and Western Europe. Incidence also correlates with gender, with higher rates observed in men. Incidence also increases with age, in Europe, 98% are over 40 years of age, and 50% are over 60 years of age (Mehanna et al. 2010; Ferlay et al. 2015).

Over recent decades, the epidemiology of HNSCC has changed. In the UK, the incidence of oral cavity, laryngeal, and hypopharyngeal cancers has remained relatively stable, and this has been attributed mostly to the tobacco prevention campaigns and the resulting decrease in its use.(Schache et al. 2016). In contrast, a significant increase in oropharyngeal cancer incidence has occurred. Similar trends have been observed in several countries including the USA, where oropharyngeal cancers rose by 2.5% a year between 2002 and 2012 (Mourad et al. 2017). In the UK, the number of oropharyngeal cancers doubled during the last decade, the age-standardized rate in 2002 was 2.1 (95% Confidence Interval (CI), 1.9–2.2) and rose to 4.1 (95% CI, 4.0–4.3) by 2011 (Schache et al. 2016).

Historically, alcohol and tobacco consumption have been the primary risk factors in the development of Oropharyngeal Squamous Cell Carcinoma (OPSCC), with a synergistic effect when combined (Franceschi et al. 1990; Smith et al. 1998).The rise in oropharyngeal cancer incidence has been also been observed in non-tobacco and alcohol consumers of a younger age group. In the past two decades, HPV has been recognised as an important risk factor for head and neck cancers, and for OPSCC especially. A causal link has been established between the presence of HPV infected cells and SCC (Smith et al. 1998; Gillison et al. 2000). Since 2007, HPV has been recognised as a risk factor for OPSCC by the World Health Organisation (IARC 2007). Distinct trends in incidence of HPV-positive compared to HPV-negative tumours have been reported depending on the geographical location, with higher rates of HPV-positive disease in western Europe compared to eastern Europe or Asia (Mehanna et al. 2016). In the

UK, the proportion of HPV-positive and negative OPSCC remained constant in the past decade (around 50% each) whilst the incidence of both doubled (Schache et al. 2016). In South Wales, the proportion of HPV-positive OPSCC was 55% during the 2001-2006 period (Evans et al. 2013). Relative to HPV-negative tumours, HPV-positive OPSCC display characteristic pathological, molecular and clinical features, and hence can be regarded as a distinct disease entity (Gillison and Shah 2001).

1.2.3 Risk factors

The main risk factors for HNSCC in general are the consumption of tobacco and alcohol. The higher the consumption of tobacco, the higher the risk of head and neck cancer (Franceschi et al. 1990). Tobacco consumption is not limited to active smoking, passive smoking as well as smokeless tobacco products also contribute to carcinogenesis (Zhang et al. 2000; Muwonge et al. 2008). Some components of tobacco smoke such as polycyclic hydrocarbons are known carcinogens, and when metabolised, some of the products formed can bind to DNA forming mutagenic DNA adducts (Gelboin 1980). Regarding ex-smokers, the associated risk of developing HNSCC decreases with cessation of smoking, but remains higher than for a non-tobacco user (Schlecht et al. 1999). Tobacco consumption is frequently associated with p53 mutations, which are a common feature of HPV-negative tumours (Brennan et al. 1995).

The second major risk factor is consumption of alcohol, which is another known carcinogen involved in many types of cancer. The oncogenic effect of alcohol is mainly due to its metabolite acetaldehyde which directly damages DNA (Baan et al. 2007). Together, the effects of tobacco and alcohol are synergistic, as alcohol acts to increase the effect of the tobacco metabolites on the cell. This could be due to the fact that it is a solvent and could increase the time of exposure of tissues to those metabolites (Pai and Westra 2009). Clearly, many people who consume tobacco and alcohol do not develop head and neck cancers, but genetic polymorphisms or mutation in the enzymes metabolising alcohol and tobacco, and in the detoxification pathways, can increase the effect of alcohol and tobacco in some individuals. For instance, polymorphisms and/or mutations in the ALDH2 gene encoding acetaldehyde dehydrogenase, which detoxifies acetaldehyde produced during alcohol metabolism, increase the risk of heavy alcohol consumers to develop HNSCC (Hakenewerth et al. 2011).

As it is the case for most cancers, genetic factors can predispose to the development of head and neck neoplasms. People with a family member who developed a head and neck neoplasm have an increased risk of developing one too. Various genetic mutations could be linked to a higher risk of HNSCC. These include mutations in DNA Base Excision Repair (BER) genes and Single Nucleotide Polymorphism (SNP) such as one in the 10q23 region identified through genome wide association studies (Kowalski et al. 2009; Yuan et al. 2013). Other factors that have been associated with increased risk of developing head and neck cancers include diets poor in fruit and vegetables, poor dental hygiene, smoke and chemicals, immune system deficiency, and body mass index. These risk factors often occur in combination (Lissowska et al. 2003; Hashim et al. 2016).

Some patients do not present with any major risk factors but nonetheless develop HNSCC. Studies investigating this population led to the identification of HPV in different stages of the disease (invasive and lymph nodes) and as a causal agent for some head and neck cancers. HPV associated tumours are mostly located in the oropharynx (57%) compared to the other regions of the head and neck. The type distribution of HPV differs between cervical and HNSCC. In HNSCC, a higher proportion (over 90%) of tumours are associated with HPV16, with other high risk types, mainly HPV 31, 33 and 18, present in the remainder (Gillison and Shah 2001). The tumours originate mostly from the lingual and palatine tonsils crypts but can also originate from other areas of the oropharynx (Begum et al. 2005). As it is the case in cervical cancer, sexual behaviour influences risk of HPV-associated OPSCC. A high number of sexual partners increases the risks of infection by high risk types of HPV, and for OPSCC, oral sex is also considered a risk factor. People with 6 or more sexual partner throughout their life and 4 or more oral sex partners have been shown to have an increased risk of developing HNSCC (Heck et al. 2010).

1.2.4 Natural history

The association between HPV and oropharyngeal cancer has only been convincingly established relatively recently, hence the natural history of HPV infection in the oropharynx and oral cavity, is not yet well defined. OPSCC appear to most commonly arise in the tonsillar crypts (Begum et al. 2005). Due to this localisation, pre-malignant and early stage lesions are very difficult to detect, hence screening protocols such as those used for cervical cancer are likely to be impractical (Franceschi et al. 2015).

OPSCC affects predominantly men, and rates are up to 4 fold higher in men compared to women. This disparity is observed for both HPV-negative and HPV-positive OPSCC. For HPV-negative disease, this has been attributed to higher consumption of tobacco and alcohol in men than women (Evans et al. 2013; Schache et al. 2016). Why this disparity is also present in HPV-positive disease is less clear, however several hypotheses have been suggested. It could partially be explained by the higher prevalence of oral high risk HPV infection in men. In the USA in the 2009-2010 period, the oral prevalence was 3.6% in women vs 10.1% in men. This included two peaks: 7.3% and 11.4% in men aged 30-34 and 60-64 respectively (Gillison et al. 2012). However, this finding is not consistent across the literature as some studies report very little variation in oral HPV prevalence between men and women (Kreimer et al. 2010). Sexual behaviour, including the number of lifetime sexual partners may also play a role. In the USA, men have been reported to have on average a higher number of partners than women (D'Souza et al. 2014). Different rates of seroconversion between men and women following HPV infection could also be relevant. Following natural genital infection with HPV, some people develop detectable antibodies against the virus while others do not. The higher the antibody titre, the higher the protection to subsequent infection compared to seronegative individuals (Safaeian et al. 2010). Several studies have shown a higher level of seroconversion in women compared to men. In one study in England, seroconversion for HPV6, 11, 16, and 18 was 16.4%, 5.7%, 14.7%, and 6.3%, respectively, among females, compared to 7.6%, 2.2%, 5.0%, and 2.0%, respectively, among males (i.e. seroconversion for all types was significantly higher in females than males: $P < 0.001$) (Desai et al. 2011).

Tonsil tissue is an immunologically active site, and may therefore be an unexpected site for HPV infection. However, the specialised epithelium lining the tonsillar crypt may actually facilitate HPV infection. In the reticulated epithelium, the epithelial cells junctions are not as strong as they involve a mix of infiltrating lymphocyte and epithelial cells. This structure facilitates the virus access to epithelial stem cells (Egawa et al. 2015). The unsuitability of the tonsil to accommodate the full viral life cycle has also been suggested as a reason why HPV infection might more often result in tumour development in the tonsils (Powell et al. 2015). The capacity to produce virions has recently been shown *in vitro* using raft cultures, but it remains to be observed in tonsils (Israr et al. 2016). Similarly to genital HPV infections, persistent oral cavity HPV infections are at risk of developing into oral or oropharyngeal SCC, and the latency period for OPSCC development has been estimated to be 10-30 years (Gillison et al. 2015).

1.2.5 Clinical and pathological features of HPV-positive vs HPV-negative tumours

HPV-positive and negative tumours differ in their histological appearance. HPV-positive tumours have a tendency to present with several features traditionally associated with poor clinical outcomes. They tend to be less differentiated, not present signs of keratinisation and have a basaloid appearance (Ang and Sturgis 2012). By contrast, HPV-positive tumours are generally smaller macroscopically at the time of diagnosis which is reflected by an average lower T grade in the Tumour Node Metastasis (TNM) grading (Gillison et al. 2012). The nodal stage also tends to differ, with a higher proportion of HPV-positive patients presenting with nodal metastasis at the time of diagnosis (Goldenberg et al. 2008). Immunological differences also occur, HPV positive tumours have a higher density of CD8+ lymphocytes infiltration compared to HPV negative lesions (Näsman et al. 2012).

The demographic characteristics of OPSCC patients also show some association with tumour HPV status. People presenting with a HPV-positive tumour tend to: be younger, have never smoked (or smoked very little), be predominantly of white race, have better overall health performance status, and show absence of anaemia, relative to patients whose tumour is negative for the presence of HPV (Ang and Sturgis 2012).

Various genetic differences between HPV-positive and negative tumours have been reported. The increasing use of next generation sequencing has enabled the characterisation of large sample cohorts and the description of distinct mutational profiles. An analysis of whole-exome sequencing data from 74 HNSCC tumour-normal pairs showed that the mutation rate of HPV-positive tumours was approximately half that found in HPV-negative HNSCC (mean number of mutations per megabase = 2.28 vs 4.83 mutations; P = 0.004) (Stransky et al. 2011). The most common mutational differences are associated with the gene coding for the protein p53. Mutations in p53 are important in the malignant transformation of many cancers but are rarely present in HPV-positive tumours. This is presumably because p53 mutation would be redundant against a background of p53 protein loss caused by the viral protein E6.

The Cancer Genome Atlas study included several head and neck sites, and identified different expression profiles for HPV-positive and negative tumours. Some genetic alterations were commonly found in both HPV-positive and negative HNSCC (e.g. region 3q26/28 which involves TP63, SOX2 and PIK3CA). The gene encoding the cell cycle transcription factor E2F1

was amplified in some HPV-positive tumours, while the CDKN2A gene, encoding p16 protein, was generally intact (but was often mutated in HPV-negative). Regarding the oropharynx in particular, several genes were up-regulated in HPV-positive OPSCC, including CDKN2A/C, RPA2, MCM2/3/5, and TAF7L. The type of genetic alterations seen in HPV-negative OPSCC was similar to the pattern observed in other smoking related tumours.

HPV-positive OPSCC also show patterns of mutations characteristic of up-regulation of APOBEC (Apolipoprotein B mRNA Editing catalytic polypeptide) activity (Stransky et al. 2011; Alexandrov et al. 2013; Lawrence et al. 2015). APOBECs are a family of 11 cytidine deaminase proteins capable of binding to RNA and single-stranded DNA. These enzymes, through the conversion of cytosine to uracil, have a variety of important biological functions (Salter et al. 2016). Activation-induced deoxycytidine deaminase is expressed in B cells and implicated in antibody diversity through its role in somatic hypermutation and class switch recombination (Muramatsu et al. 2000). APOBEC1 is mostly known for its ability to bind and regulate the stability of RNA. APOBEC2 and 4 are not well characterised, but APOBEC2 may be involved in embryonic development (Liao et al. 1999; Anant and Davidson 2000). The APOBEC3 subfamily (APOBEC3 A, B, C, D, F, G, H) have been characterised in more detail, and are involved in the innate immune response to viruses and retroviruses in particular by interfering with viral replications and clearing foreign DNA (Chiu and Greene 2008; Stenglein et al. 2010). APOBEC have been studied in relation to cancer as they could potentially lead to the acquisition of somatic mutations involved in tumour development. Mutagenic processes tend to cause a specific type of mutation which can be identified, called mutational signatures. APOBEC catalysed cytosine deamination is one mutational signatures already identified in several types of tumours (Roberts et al. 2013). Tumours in which APOBEC associated mutations have been identified include other HPV-associated tumours such as cervical cancer (Burk et al. 2017). In relation to HNSCC, HPV-positive tumours have been linked with an enriched signature of APOBEC mutations compared to HPV-negative cases (Faden et al. 2017).

1.2.5.1 HPV and prognosis

In recent decades, several clinical trials have shown that the response to therapy was better in patients with HPV-positive tumours compared to HPV-negative OPSCC. This was associated with increased survival and reduced disease progression. The better prognosis of HPV-positive patients in general has been observed with a range of treatments, including

surgery, chemotherapy and radiotherapy, or combinations thereof. In South Wales, a retrospective study demonstrated a clear association between HPV-positivity and favourable prognosis, with 3 and 5-year overall survival rates of 82.6% and 75.4% respectively in HPV-positive patients, compared to 32.2% and 25.3% in HPV-negative patients, equating to a 78% reduction in death rate associated with HPV-positivity (HR 0.220, 95% Confidence Interval (CI); 0.132-0.366, $p < 0.001$) (Evans et al. 2013). A recent meta-analysis, including clinical trial data for 1130 OPSCC patients, indicated a hazard ratio for improved overall survival of 0.49 (95% CI 0.35-0.69) associated with HPV-positive tumours (Masterson et al. 2014).

It has also been noted that survival has increased in the HPV-positive patients but has remained the same in HPV-negative patients (Chaturvedi et al., 2011). It is however notable that even within the HPV-positive tumours survival varies widely, and other factors associated with HPV-positive tumours such as young age, disease status, smoking and lower alcohol consumption could account together for up to 9% of the response of HPV-positive tumours (Ang et al. 2010). Smoking and HPV status are not exclusive, and tobacco consumption also affects the outcome of HPV-positive patients. Independently of p16 status, smoking significantly increased the risk of death and of disease progression (Gillison et al. 2012). Accordingly, Ang *et al*. proposed that patients could be divided into 3 groups: Low risk, comprising "typical" HPV-positive patients; High risk comprising HPV-negative patients; and an intermediate risk groups including HPV-positive smoker with high nodal stage and HPV-negative non-smokers with T2 and T3 stage (Ang et al. 2010).

The implications of, and the mechanisms behind, the difference in prognosis have been studied using retrospective FFPE samples and patients derived cell lines. However, few HPV-positive HNSCC cell lines have been described in the literature. As of 2012, five lines had been published and only one originated from the oropharynx (Steenbergen et al. 1995; Balló et al. 1999; Lansford et al. 1999; Ferris et al. 2005; Tang et al. 2012). In contrast to the current trend in HPV-positive patients, all of these lines were derived from heavy or ex-smokers. The available cell lines may therefore not provide the best possible representation of typical HPV-positive non-smoking patients with OPSCC. The development of new cell lines, from typical HPV-positive patients, is needed to increase our understanding of treatment response in HPV-positive smokers and non-smokers.

1.2.5.2 Tumour microenvironment

The tonsils are highly infiltrated by immune cells due to the communication with the underlying lymphoid tissue through the crypts reticulated epithelia. As a result, the oropharynx has a crucial role in the detection of, and protection against, pathogens entering the organism through the aero-digestive system. In a variety of cancers, the tumour microenvironment and the local, as well as systemic immune responses have been shown to play an important role in the control and growth of tumours (Vesely et al. 2011). In the context of HPV-positive OPSCC, this is of particular interest as the presence of viral antigens can elicit an immune response.

An HPV-specific immune response can be detected in some HPV-positive HNSCC patients mainly with oropharyngeal tumours. Detection of circulating T cells specific to HPV16 E7 epitopes occurs more frequently in HPV-positive than HPV-negative patients (Hoffmann et al. 2006). Locally, HPV16 infection can, in some patients, result in a local HPV16 specific T cell response comprising of both CD4+ and CD8+ T cells, which is absent in HPV-negative patients (Heusinkveld et al. 2012).

The number of cytotoxic T lymphocytes (CD8+) Tumour Infiltrating Lymphocytes (TIL) has been associated with clinical outcome independently of HPV status following radiotherapy. Patients with higher amounts of CD8+ TIL had better clinical outcomes than those with fewer TIL. The same tendency has been observed once patients were stratified according to HPV status (Näsman et al. 2012). Similarly, the amount of CD8+ TIL has been associated with favourable outcome in patients treated with chemotherapy. However, no significant influence of the infiltration levels of other T lymphocytes such as T helper (CD4+) and T reg (FOXP3+) has been recorded (Balermpas et al. 2014; Balermpas et al. 2016). While CD8+ TIL seems to have a positive influence on the outcome irrespective of HPV-status, the total amount of infiltrated cells is higher in HPV-positive patients. This indicates that those tumours might be more immunogenic and hence could contribute to the overall better outcome observed in these patients (Balermpas et al. 2016).

Downregulation of the PD1-PDL1 immune checkpoint, which impairs the function of TIL, has been linked with immune evasion in various tumours (Topalian et al. 2016). This is also of interest in the oropharynx as PDL1 is expressed in the tonsil crypts where many HPV-positive tumours originate. In HNSCC, PDL1 is expressed in most tumours and PD1 by TIL (Lyford-Pike et

al. 2013). Similar to CD8, PD1 and PDL1 expression is overall higher in HPV-positive patients compared to HPV-negative patients (Balermpas et al. 2017). The activation of the PD1-PDL1 checkpoint suggests a mechanism of tumour immune evasion in HPV-positive HNSCC. The use of novel therapies targeting the PD1-PDL1 checkpoint is currently being investigated for HNSCC with promising results for some patients (Seiwert 2016).

PDL1 expression has been investigated as a biomarker for overall outcome. Its influence varies greatly between studies with positive and no association with outcome reported (Kim et al. 2016; Balermpas et al. 2017). The combination of high CD8 TIL and high PDL1 expression has however been identified as a predictive marker of favourable outcome in several cohorts. While this is the case in both HPV-positive and negative status, they are more common in HPV-positive cases (Cancer Genome Atlas Network 2015; Balermpas et al. 2017).

1.2.6 Diagnosis of OPSCC and HPV testing

Patients can present with a variety of symptoms, but often have painless lumps in the neck. When the tumour originates in the tonsil crypts, it can grow undetected for a prolonged period. Diagnosis of oropharyngeal cancer is based on imaging techniques as well histological examination of biopsies obtained during panendoscopy. As it is the case for other tumours, the stage of the disease can be established according to the TNM classification where T1-T4 describes the size of the tumour, N0-N3 refers to nodal stage, and M0-M1 indicates if metastases are present (Patel and Shah 2005).

HPV-positivity in OPSCC can be an indicator of better prognosis and treatment response in some patients, hence p16 Immuno Histochemistry (IHC), as a surrogate marker of HPV infection, is now generally performed for OPSCC patients. In a research or clinical trial setting, the presence of HPV in a tumour can be shown by the presence of high risk HPV DNA, RNA or proteins or the up-regulation of human proteins (p16). Various techniques can be used, in isolation or in combination, for this purpose. The main goal of HPV detection is to identify infections in which HPV is likely to be driving the disease. The gold standard method is to establish the presence of transcriptionally active HPV by amplifying E6 or E7 transcripts as the combined actions of these two proteins is essential in carcinogenic transformation (Wiest et al. 2002; Schache et al. 2011).

When detecting HPV RNA, the best results are obtained from fresh or fresh frozen tissue and reverse transcription quantitative Polymerase Chain Reaction (PCR). However this can be logistically challenging as OPSCC biopsies are routinely processed and stored as Formalin Fixed Paraffin Embedded (FFPE) tissue blocks, and RNA obtained from FFPE samples can be unsuitable for RNA based testing (von Ahlfen et al. 2007; Robinson et al. 2012). p16 IHC on FFPE slides is therefore commonly used. p16 is widely recognised as a surrogate biomarker of HPV infection as it is strongly up-regulated following the action of the E7 protein (Doorbar 2005). However while p16 detection has a high sensitivity, it can lack specificity (Prigge et al. 2017), and some OPSCC have also been described as p16 positive but HPV-negative, highlighting that factors other than HPV can up-regulate p16 in tumours (Perrone et al. 2011). Hence p16 IHC is often used in combination with detection of HPV DNA in FFPE samples.

HPV DNA can be detected by *in situ* hybridisation (ISH) or PCR. Several DNA based tests can detect the presence of groups of high risk HPV types, while others detect HPV16 in isolation. These tend to be PCR and Enzyme Linked Immunosorbent Assay (ELISA) based techniques. Most HPV genotyping PCRs target the L1 region of the HPV genome, the amplicons are then hybridised with a cocktail of probes for the various HPV types (Robinson et al. 2012). The use of p16 and DNA based testing as standalone tests do not reach the sensitivity and specificity of mRNA detection. Combinations of those tests on the other hand provide reliable methods of HPV detection. The combination of p16 and HPV DNA by ISH has emerged as a diagnostic standard of HPV-positive OPSCC (Smeets et al. 2007; Schache et al. 2011; Prigge et al. 2017).

1.2.7 Treatment of OPSCC

Historically, OPSCC patients were often treated with open surgery such as mandibulotomy followed by extensive reconstruction. This approach led to a high number of severe and/or fatal complications. However, the locoregional control and 5-year disease specific survival was similar to that observed with radiotherapy alone or a combination of surgery and radiotherapy. This led to the reduction of extensive open surgery (Parsons et al. 2002). Contemporary treatment of OPSCC, independently of HPV status, utilises laser and robotic surgery, radiotherapy (intensity modulated) and chemotherapy (induction and/or concurrent) often combined as chemoradiotherapy (Evans and Jones 2016). Radiotherapy, even intensity modulated, can cause significant toxicity to adjacent tissues affecting the

salivary glands, swallowing function and teeth. The patient's quality of life can be substantially reduced in the long term, especially when the swallowing function is impaired (Denis et al. 2003; Eisbruch et al. 2011). There is therefore mounting interest in development of less toxic, targeted treatments for OPSCC.

1.3 DNA damage and repair

DNA damage occurs continuously in cells as a consequence of exogenous and endogenous factors. Exogenous causes include ultraviolet radiations, ionising radiations, and chemical compounds (e.g. metabolites of alcohol, tobacco and some foods). Endogenous sources include replication errors, free radicals and spontaneous loss or alteration of purine or pyrimidine bases (e.g. due to breakage of glycosidic bonds between the pentose sugar and base, or alteration such as methylation) (Lindahl 1993). While a very low mutation rate is an evolutionary necessity, the maintenance of genomic integrity is essential to allow the cells, and ultimately the organism, to function properly (Kolodner et al. 2002). Numerous mechanisms exist to ensure detection and repair of damaged DNA. These include BER, Nucleotide Excision Repair (NER), Mismatch Repair (MMR), Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ). These pathways can be broadly divided in those that deal with, or involve, DNA Single Strand Breaks (SSB) and those that address Double Strand Breaks (DSB) (Sancar et al. 2004).

BER is the predominant pathway for repair of DNA lesions caused by endogenous sources such as reactive oxygen species and depurination events (David et al. 2007). BER involves the removal of the damaged base by a glycosylase, followed by the excision of the sugar phosphate by the AP endonuclease APE1, which leaves a gap in the sequence (Krokan et al. 1997). The damaged termini are repaired by a complex including APE1, DNA polymerase β, polynucleotide kinase 3'-posphatase (PNKP) and aprataxin (APTX) (Sobol et al. 1996; Mol et al. 2000; Rass et al. 2007). The X-ray repair cross-complementation protein 1 (XRCC1) scaffolding protein, acts by stabilising and stimulating the components of the complex (Kubota et al. 1996). New nucleotides are added by DNA polymerase β and the remaining gap sealed by DNA ligase (Cotner-Gohara et al. 2008).

Direct SSB, for example arising as the result of the disintegration of deoxyribose sugar, are detected by poly (ADP-ribose) polymerase (PARP) 1. PARP1 migrates to sites of SSB and

causes rapid conversion of NAD^+ to ADP-ribose polymers (PAR), promoting the recruitment and/or activation of further proteins involved in SSB repair (Masson et al. 1998). Similarly to BER, damaged ends are repaired by a protein complex including XRCC1, polymerase β, and ligase 3 (El-Khamisy et al. 2003). During replication, unrepaired SSB can cause a collapse of the replication fork or stalled transcription which results in the formation of DSB (Branzei and Foiani 2010).

DSB repair is primarily achieved through the HR and NHEJ pathways. Both pathways include damage recognition and repair mechanisms which function together to achieve repair. Three proteins are central to the recognition stage: Ataxia Telangiectasia Mutated (ATM), Ataxia Telangiectasia and RAD3 related (ATR) and DNA-dependent Protein Kinase (DNA-PK) (Branzei and Foiani 2010). DNA-PK signalling is involved in NHEJ, ATM signalling in HR, whilst ATR is implicated in DSB arising due to stalled replication forks (Jackson 2002).

Homologous recombination is the more complex DSB repair mechanism but also the most accurate, as it utilises the sister chromatid as a template to repair the lesion. However, it only occurs in stages of the cell cycle during which a sister chromatid is present i.e. when the DNA has been replicated but before cell division, in S and G2-phases (Saleh-Gohari and Helleday 2004). The process begins with resection by an exonuclease of one of the two DNA strands at each DNA end. The resulting single stranded DNA then invade the DNA of the undamaged sister chromatid, and base pair with their complementary sequences. The intact strand is used as a template for DNA synthesis, and a DNA polymerase extends the single stranded DNA in a 5'-3' direction. In the final steps, the Holliday junctions are resolved and the DNA nicks are ligated (Weinberg 2013).

NHEJ occurs mostly in the G1-phase of the cell cycle, when sister chromatids are not present (Rothkamm et al. 2003). Hence, by comparison with HR repair, NHEJ is inevitably a relatively error-prone process. This is because the alignment between the DNA ends being joined is not informed by an undamaged template. NHEJ also begins with resection of one strand of each broken DNA end, but the two ends are then joined without reference to a template. This tends to result in deletion of short sections of DNA and also has the potential result in genomic rearrangements (Lieber 2010). Variations of NHEJ have been described, only the general process is described here. Ku proteins bind to DNA ends, enabling the recruitment of nucleases and polymerases at the break site (Ma et al. 2004). Ligation requires the presence

of two blunt ends, the ends are processed to achieve this. Proteins involved in end processing include DNA-PK and Artemis (Ma et al. 2002). The gaps are ligated by a DNA ligase 4 complex (DNA ligase 4-XRCC4) (Gu et al. 2007).

1.3.1 HPV and DNA damage

Several lines of evidence show significant interactions between HPV and DNA repair proteins. The production of new virus particles is dependent upon activation of both the ATM and ATR DNA repair pathways (Moody and Laimins 2009; Reinson et al. 2013; Hong et al. 2015; Kanginakudru et al. 2015). Activation of the ATM pathway ordinarily occurs in response to the presence of DNA DSB, while activation of ATR results from replication stress and the presence of single-stranded DNA at stalled replication forks (Eyfjord and Bodvarsdottir 2005; Sulli et al. 2012). High risk HPV have been shown to regulate components of these DNA repair pathways to promote viral replication (Wallace and Galloway 2014). This is mediated mainly by the E1 and E7 proteins which recruit repair proteins for viral replication (Moody and Laimins 2009; Fradet-Turcotte et al. 2011).

Interaction between the ATM and ATR DNA repair pathways is mediated by the ATM/ATR interacting Fanconi Anemia (FA) protein FANCD2, which is recruited to viral genomes and is required for maintenance of viral genomes in undifferentiated cells (Spriggs and Laimins 2017) The FA pathway is responsible for repair of DNA inter-strand crosslinks and includes multiple DNA repair proteins including BRCA1 and γ-H2AX. HPV also recruits additional DSB repair proteins to the HPV DNA foci, including 53BP1, RAD51 and BRCA1. γ-H2AX is also found at these foci (Gillespie et al. 2012). The E7 protein also targets clapsin for proteolytic degradation which in turn inhibits the DNA damage checkpoint response (Spardy et al. 2009).

Most previous studies have addressed the relationships between HPV and DNA repair in the context of productive infections, in which the HPV genome is episomal. However, in cancers, HPV is rarely found in its episomal form, but instead is more commonly (c. 70% cases) integrated into the host genome (Parfenov et al. 2014). In this situation the predominant effects of HPV on DNA repair appears to be to cause a specific defect in repair of DNA DSB (Duensing and Münger 2004) and induce activation of APOBEC activity (Henderson et al. 2014; Vieira et al. 2014). This is consistent with the genomic instability observed during the

carcinogenesis process, in tumour cells, and with the presence of APOBEC signature mutations in HPV-positive OPSCC (Henderson et al. 2014).

Studies comparing the radio-sensitivity of a limited number of HPV-positive and negative cell lines have demonstrated that HPV-positive lines are usually less able to repair DSB. Two studies have investigated reduced repair of DSB in HPV-positive cell lines and both suggest that persistence of DSB, assessed by immunofluorescence for γ-H2AX, correlated with the radio-sensitivity of the lines (Kimple et al. 2013; Rieckmann et al. 2013). Both also agree that in HPV-positive lines, accumulation of DSB results in failure to traverse the G2/M checkpoint, causing accumulation of cells in G2. The events following G2 arrest are less well defined. Kimple et al. suggested that following irradiation, HPV-positive cells showed increased p53 and phospho-p53, and this was associated with significant increases in apoptosis. In contrast, Rieckmann et al. assessed apoptosis by caspase 3 activation and PARP cleavage in irradiated cells and found no increase in apoptosis in either HPV-positive or negative lines. Taken together, these studies suggest exploiting deficiencies in DNA damage repair in HPV-positive OPSCC could have potential as a therapeutic strategy.

1.3.2 Synthetic lethal therapy and PARP inhibition

The concept of synthetic lethality was originally used to describe a genetic interaction in which the combination of two gene knockouts is lethal. More recently, this concept has been extended to include interactions between proteins and pathways that are normally non essential but that are critical for survival in cancer cells. This principle is highly relevant in cancer therapy because many tumours harbour genetic mutations that are not present in the other cells of the body (Kaelin 2005). Such treatments could be highly beneficial due to the potential to selectively kill cancer cells while sparing normal cells (Helleday 2011). The archetypical therapeutic example of synthetic lethality is the use of PARP inhibitors (PARPi) as monotherapy for treatment of tumours with defects in repair of DSB by HR (Bryant et al. 2005; Committee for Medicinal Products for Human Use 2014).

In OPSCC some studies have indicated a possible common defect in HR. Therefore targeting SSB repair could lead to synthetic lethality in OPSCC cells. The mode of action of a PARPi in OPSCC is illustrated in Figure 1.6.

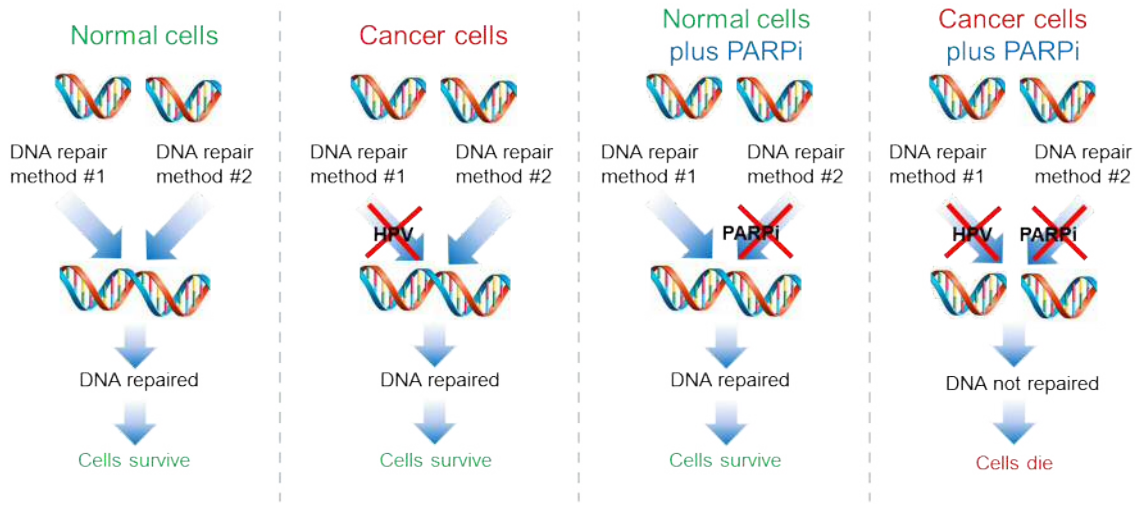


Figure 1.6: PARP inhibition leading to synthetic lethality in OPSCC
Normal cells are able to repair DNA damage via SSB and DSB repair. In HPV-positive cancer cells DSB repair is defective but SSB repair is intact. Normal cells treated with a PARPi would be able to repair the DSB resulting from collapsed replication fork due to the unrepaired SSB. However in HPV-positive cells treatment with PARPi will cause an unrepairable level of DNA damage leading to cell death.

1.3.3 PARP proteins involvement in DNA repair

PARP comprise a family of 17 proteins with a wide range of functions in biological pathways and processes including DNA repair, DNA transcription, genomic stability and cell cycle control (Krishnakumar and Kraus 2010). PARP catalyse the parylation reaction, a reversible post-transcriptional modification which entails covalent attachment of ADP ribose units to the carboxyl groups of glutamate, aspartate and lysine residues of the target proteins or itself using NAD^+ as a substrate (D'Amours et al. 1999; Altmeyer et al. 2009). Most members of the PARP family catalyse the addition of a single ADP unit whilst others such as PARP1, PARP2 and PARP5 synthesise linear or multi-branched ADP polymers. The majority of PARP are located in the cytoplasm. Their expression is ubiquitous in human cells but expression profiles are not cell type specific. However, within a cell, localisation varies throughout the cell cycle (Vyas et al. 2014).

The best known and most abundantly expressed of the PARP family is PARP1. This 113 kDa protein is composed of a DNA binding domain which contains N-terminal zinc finger motives in tandem repeat, a caspase 3 cleavage site and nuclear localisation signal, an auto-

modification domain involved in protein-protein interactions and a catalytic domain binding for NAD^+. The catalytic domain is characteristic of PARP proteins (Kameshita et al. 1984).

PARP1 is important in DNA repair, most notably in SSB repair, but also in DSB repair. PARP1 acts as a sensor for SSB and quickly binds to the break site. The zinc finger motives are essential in the recognition and binding of PARP1 to the site of DNA breaks. Zinc finger motives from two separate PARP1 molecules are able to interact with each other forming a complex of two PARP1 molecules and DNA. This allows for the recognition of a variety of DNA lesions including DSB, SSB and gaps (Ali et al. 2012). Multiple inter-domain interactions occur following PARP1 binding to DNA. As a result, the low basal level activity of the catalytic domain is increased through changes in its thermal stability. The conformational changes also enhances the catalytic domain to the protein substrates (Langelier et al. 2012).Through its parylation activity, the protein itself is modified as well as interacting with several other proteins such as DNA repair and chromatin remodelling factors to recruit them to the break site. Recruited proteins include XRCC1, DNA polymerase β and PCNA (Masson et al. 1998). PARP1's role in DSB repair is less well characterised, but includes interactions with various proteins involved specifically in NHEJ and HR such as MRE11, KU80 and NBS1, and also with proteins involved in DNA damage sensing such as ATM (Galande and Kohwi-Shigematsu 1999; Haince et al. 2007; Haince et al. 2008). PARP1 does not remain at the site of DNA damage, its removal is necessary for adequate repair (Amé et al. 2009). The auto-parylation on PARP1 as well as the parylation of downstream targets is transient due to the activity of glycohydrolases. The poly-ADP-ribose-glycohydrolase cleaves the ribose-ribose bonds leaving one ADP ribose unit on the protein. This residue is later removed by a mono-ADP-ribose glycohydrolase (Slade et al. 2011; Rosenthal et al. 2013).

Within the PARP family, PARP2 is the most similar to PARP1. This 66 kDa protein has similar localisation within the cells and also contains a nuclear localisation signal and DNA binding domain (Amé et al. 1999). Similarly to PARP1, PARP2 is involved in BER via its interaction with XRCC1, DNA polymerase β and DNA ligase 3 (Schreiber et al. 2002).

1.3.4 Olaparib

Olaparib (also known as KU-0059436 and AZD2281) was developed by KuDOS Pharmaceuticals, which has been bought by AstraZeneca. It is marketed for clinical use under

the brand name Lynparza. Olaparib displays inhibitory activity concentrated towards PARP1 and 2. However, its cellular effect is mediated mainly via PARP1 inhibition as it constitutes the majority of the total PARP proteins. Knock down of PARP2 was showed to have minimal effect on Olaparib cytotoxicity (Murai et al. 2012)

Olaparib molecular formula and chemical name are shown in Figure 1.7. It is based around a phthalazizone structure and was developed to obtain an orally bioavailable compound. It includes an amine group which replicates the interaction between NAD and PARP thus preventing the binding of the nicotinamide structure of NAD with the active site on PARP1. This involves the formation of bonds between the inhibitor and Ser904 and Gly863 on the PARP protein which are similar to the ones formed between NAD and PARP during the normal enzymatic activity (Jagtap and Szabó 2005; Menear et al. 2008). The effect of Olaparib is not mediated solely by its catalytic inhibition, another mode of action referred to as PARP trapping has been described. The use of a PARP inhibitor also inhibits the auto-phosphorylation of PARP itself, a process required for PARP dissociation form the DNA (Mortusewicz et al. 2007). As a result, Olaparib treatment causes the binding of both PARP1 and PARP2 to the chromatin forming complexes of PARP and single stand DNA intermediate. These complexes do not dissociate, during replication they are converted into more toxic lesions inducing further cytotoxicity. This process does not correlate with the catalytic inhibition capacity of PARP inhibitors (Murai et al. 2012)

In early *in vitro* studies, Olaparib was found to have the same potency against PARP1 and PARP2. The PARP inhibitory activity of Olaparib was tested in SW60, a human adenocarcinoma cell line, via the measurement of PAR (the downstream product of PARP activation). Following consideration alongside other candidates with a similar chemical structure, Olaparib was selected for clinical development (Menear et al. 2008). Another *in vitro* study using a panel of 95 cell lines, including some HNSCC lines, showed wide variation in IC50, ranging from 0.018 μM to over 10 μM. Higher sensitivity (IC50 <1 μM) correlated with BRCA1 and 2 mutation and low HR pathway gene expression. Low expression of PARP1 was also associated with reduced sensitivity (study number KTS00101, October 2007 unpublished, described in the FDA pharmacological report). Another study with a panel including 11 HNSCC lines found a correlation between sensitivity to platinum based agent carboplatin and cisplatin and Olaparib sensitivity (study number KTS0049, September 2008 unpublished).

In humans, the drug is metabolised via CYP3A4 and 3A5. The main metabolic processes are hydrolysis and oxidation with some products undergoing multiple metabolic transformations. Excreted products are found in faeces and urine. Olaparib was quickly absorbed with the peak plasma concentration reached after 1-3 h, and was mostly excreted within 48 h. No significant off-target activity was observed at clinically relevant doses (observed in rat, dog and human studies KMX032 and KKR007).

In 2014, a marketing authorization for Lynparza was granted by the European medicine agency. The 50 mg hard capsules are indicated as monotherapy for the maintenance treatment of adult patients with platinum-sensitive relapsed BRCA mutated (germline and/or somatic) high grade serous epithelial ovarian, fallopian tube, or primary peritoneal cancer who are in response (complete response or partial response) to platinum-based chemotherapy.

Olaparib is now being evaluated as monotherapy, and in combination with chemo and radiotherapy, in numerous trials around the globe. Tumour sites being investigated include breast, ovarian, lung, oesophagus, pancreas and prostate (clinicaltrials.gov). In the context of HNSCC, the US Clinical Trial database (clinicaltrials.gov) lists numerous trials evaluating Olaparib treatment in diverse HNSCC subsites (summarized in Table 1.1). However, in many cases, these clinical studies are supported by limited *in vitro* data.

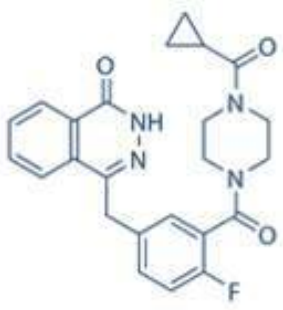


Figure 1.7: Chemical structure of Olaparib
4-(3-(1-(cyclopropanecarbonyl)piperazine-4-carbonyl)-4-fluorobenzyl)phthalazin-1(2H)-one, with the molecular formula $C_{24}H_{23}FN_4O_3$ was first described as KU-0059436, then AZD2281, then Olaparib, before being marketed under the name Lynparza.

Table 1.1: Clinical trials of Olaparib in HNSCC

Study title	Country	Treatment
Phase I Study of Olaparib Combined with Cisplatin-based chemoradiotherapy to treat locally advanced head and neck cancer (ORCA-2)	UK	Olaparib Cisplatin Radiation therapy
Preoperative administration of Olaparib with or without Cisplatin in patients who are candidates for surgery of carcinoma of head and neck (OPHELIA)	Greece	Olaparib Cisplatin
Study of Olaparib with radiation therapy and Cetuximab in advanced head and neck cancer with heavy smoking history	USA	Olaparib Cetuximab Radiation therapy
Pharmacodynamic study to assess the anti-proliferative activity of the PARPi Olaparib in patients With HPV positive and HPV negative HNSCC	USA	Olaparib
Phase I study of Olaparib with Cisplatin based chemoradiotherapy in squamous cell carcinoma of the head and neck (ORCA)	UK	Olaparib; cisplatin; Radiation therapy
A study to investigate biomarker effects of pre-surgical treatment with DNA Damage Repair (DDR) agents in patients With Head and Neck Squamous Cell Carcinoma (HNSCC)	USA, France, UK	AZD6738 Olaparib
Olaparib and radiotherapy in head and neck Cancer	Netherlands	Olaparib Radiation therapy
Selumetinib and Olaparib in solid tumours	USA	Selumetinib Olaparib
Radiotherapy and Olaparib in Combination for carcinoma of the oesophagus (ROCOCO)	UK	Olaparib Radiation therapy
Ascending doses of AZD6738 in combination With chemotherapy and/or novel anti-cancer agents	AstraZeneca - multinational	AZD6738 in combination with carboplatin; AZD673; AZD6738 in combination with Olaparib; AZD6738 in combination with MEDI4736

1.4 Study hypotheses and aims

1. Low-level high-risk HPV infection occurs and can be detected in non-malignant tonsils.
    a. Obtain a large number of tonsil samples from a population likely to harbour high-risk HPV infection.
    b. Detect HPV presence in the samples and type the virus using several tests.
2. HPV-positive and negative OPSCC lines have distinct patterns of gene expression (including pathways relating to DNA repair).
    a. Develop new HPV-positive oropharyngeal cell lines from low or non-smoker patients and validate them as models for OPSCC.
    b. Establish and characterise a panel of HPV-positive and negative cell lines representing the wide variations seen in OPSCC tumours (p53 status, treatment, smoking status).
    c. Investigate individual transcripts expression as well as processes and pathways differentially expressed within the panel.
3. A synthetic lethal approach using Olaparib would be suitable to target HPV-positive OPSCC lines.
    a. Determine the colony formation capacity of each cell line of the panel following treatment with a range of doses of Olaparib.
    b. Investigate the difference in response between lines regarding accumulation of unrepaired DSB and cell cycle distribution following treatment.
    c. Investigate the level of PARP1 and p53 expression following treatment.

Chapter 2 - Material and methods

2.1 Cell culture

This section contains a description of the cell lines, cell culture reagents and culture conditions used.

2.1.1 Established cell lines

2.1.1.1 3T3 mouse fibroblasts

The mouse embryo fibroblast cell line J2 3T3 was established in 1963 (Todaro and Green 1963). During this project, they were used to supplement the culture of the Primary Culture of Oropharyngeal Cells (PCOC) cultures and the low-density culture of other cell lines during Olaparib dosing assays. Stocks of J2 3T3 stocks (passage 10-15) cells were obtained from Dr Sally Roberts (University of Birmingham).

2.1.1.2 CaSki

The CaSki cell line was derived from a cervical carcinoma metastasised to the small bowel in a 40 year-old Caucasian female (Pattillo et al. 1977). This cell line is infected with HPV16 and contains around 600 copies of the HPV genome integrated in tandem. CaSki cells were obtained from the European Collection of Authenticated Cell Cultures (catalogue number 87020501).

2.1.1.3 SiHa

The SiHa cell line was derived from a cervical SCC in a 55 year-old woman. SiHa cells contain 1 to 2 integrated copies of HPV16. Viral integration caused the disruption of E2 and E4 open reading frames (Baker et al. 1987). This cell line was sourced from the American Type Culture Collection (catalogue number HTB-35).

2.1.1.4 HeLa

HeLa cells were derived from a cervical adenocarcinoma in a 31 year-old African American woman. Cells contain integrated HPV18 (Scherer et al. 1953). This line was obtained from the American Type Culture Collection (catalogue number CCL-2).

2.1.1.5 UPCISCC90

UPCISCC90 is an oropharyngeal squamous cell carcinoma cell line derived from a base of tongue SCC in a 44 year-old male smoker (Ferris et al. 2005). This cell line contains between 100-150 integrated copies of HPV16 at various sites in tandem. It was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen, (catalogue number ACC 670).

2.1.1.6 UMSCC4, 6, 19, 47, 74a

UMSCC4 is an HPV-negative cell line derived from a tonsil SCC in a 47 year-old woman. UMSCC6 and UMSCC19 are HPV-negative lines derived from oropharyngeal SCC in a 37 year-old man and a 67 year-old man respectively. UMSCC74a is an HPV-negative cell line derived from a base of tongue SCC in a 50-year-old-man. UMSCC47 was derived from the oral cavity of a 53 year-old man and contains integrated copies of HPV16. All of these cell lines were obtained from Professor Thomas Carey (University of Michigan).

2.1.2 Human epidermal keratinocytes

Human Epidermal Keratinocytes isolated from neonatal foreskin (HEKn) were purchased from Thermo Fischer Scientific (catalogue number C0015C) (Thermo Fisher Scientific, Waltham, USA). These cells are derived from normal tissue and are not infected by HPV. HEKn were used as controls for the Olaparib dosing experiments. With the culture conditions used, they can sustain approximately 30 population doublings.

2.1.2.1 PCOC2 and PCOC3

PCOC2 and PCOC3 HPV16-positive lines were derived during this project as described in Section 2.1.5 . PCOC2 was derived from a tonsil SCC in a 52 year-old male non-smoker and PCOC3 from a tonsil SCC in a 44 year- old male ex-smoker.

2.1.3 Cell culture material

2.1.3.1 Cell culture vessels

All tissue culture flasks (T25, T80 and T175), multiwell plates (6 wells) and single plates (6 and 10 cm diameter) were obtained from Thermo Fisher or VWR (Thermo Fisher Scientific; VWR, Lutterworth, UK).

2.1.3.2 Media

Cell culture media were supplemented, these reagents and their suppliers are listed in Table 2.1. The formulations of media used are outlined in Table 2.2 and Table 2.3. All components were thawed or warmed in a 37 °C water bath before use. Glasgow Minimal Essential Medium (GMEM) was used for the culture of the PCOC lines; this media was formulated with and without Epidermal Growth Factor (EGF). Dulbecco's Modified Eagle's Medium (DMEM) (Sigma-Aldrich, Gillingham, UK) was used for the J2 3T3 cells, CaSki, SiHa and other established cell lines. EpiLife (Thermo Fisher Scientific) media supplemented with 1% Human Keratinocyte Growth Supplement (HGKS) was used for the culture of HEKn.

Table 2.1: Reagents used in media formulation and their sources

Reagents	Sources
Fetal Bovine Serum (FBS)	Sigma-Aldrich
Newborn Calf Serum (NCS)	
Penicillin/Streptomycin (Pen Strep)	
Hydrocortisone	
Cholera toxin	
L-glutamine	
Non Essential Amino Acids (NEAA)	
EGF	
HKGS	Thermo Fisher Scientific

Table 2.2: GMEM media composition

Components	Concentration
FBS	10%
Penicillin	100 U/ml
Streptomycin	100 ng/ml
Hydrocortisone	1.5 nM
Cholera toxin	0.1 nM
L-glutamine	2 mM
EGF	10 ng/ml

Table 2.3: DMEM media composition

Components	Concentration
NCS	10%
Penicillin	100 U/ml
Streptomycin	100 ng/ml
L-glutamine	2 mM
NEAA	0.2%

2.1.3.3 Cell culture solutions

The isotonic solution Phosphate Buffer Saline (PBS) was used to wash cell layers, to resuspend cells before antibody staining and to dilute solutions. A 1× solution containing 10 mM phosphate buffer, 138 mM sodium chloride and 2.7 mM potassium chloride was obtained by diluting the content of one pouch (9.07-10 g) in 1 l of deionised water (Sigma-Aldrich). The pH was adjusted to 7.4 and the solution autoclaved.

Trypsin 0.05%/ethylenediaminetetraacetic acid (EDTA) 0.02% (TE) was obtained by diluting 100 ml of 10× stock solution in 900 ml PBS, aliquoted and stored at -20 °C (Sigma-Aldrich).

2.1.4 General cell culture

2.1.4.1 General culture conditions

All cell lines were cultured in a humidified incubator at 37 °C, with 5% carbon dioxide. Copper sulphate was used to prevent bacterial growth in the water tray. Cell culture was performed in a class II safety cabinet using aseptic technique. Surfaces were cleaned using a 70% ethanol solution. Before use, cell culture reagents were warmed to 37 °C. Liquid waste

was inactivated in a sodium dichloroisocyanurate (Presept) solution, diluted according to the manufacturer's instructions, prior to being discarded.

2.1.4.2 Cell counting and viable cell counts

Cell numbers (total and viable cells) were assessed using a haemocytometer (Camlab, Cambridge, UK) and a 0.4% w/v trypan blue solution for viability (Thermo Fisher Scientific).

A single cell suspension of the cell line to be counted was prepared and 10 µl loaded on each side of a haemocytometer. Cells present within the grid were counted using a light microscope with 100X magnification. Viability counts were obtained by diluting the cell suspension 1 in 2 with trypan blue solution prior to loading the haemocytometer. Non-viable cells, stained in blue, were excluded from the count.

The number of cells contained in 1 ml was obtained by multiplying the average number of cells in the 25 squares of 1 grid (obtained from the average of 2 grids minimum) by 10000 and by the dilution factor 2 in case of viability counts.

Formula: Average number of cells per 25 squares×10000×dilution factor = number of cells/ml

2.1.4.3 Freezing and storage

Cells were counted, pelleted by centrifugation (5 min at 160×g) and resuspended in freezing media at a density of 2×10^6 cells/ml. The composition of freezing media is outlined in Table 2.4. A volume of 1 ml was dispensed in each cryovial and placed in a Nalgene Mr Frosty freezing device at -80 °C for at least 12 h (Thermo Fisher Scientific). Cryovials were then transferred to a liquid nitrogen unit for long term storage.

Table 2.4: Freezing media composition

Cell line	Freezing media composition
PCOC	FBS, 10% Dimethyl Sulfoxide (DMSO)
Keratinocytes	DMEM, 10% FBS, 10% DMSO
HEKn	EpiLife, 10% DMSO

2.1.4.4 Thawing of cryopreserved cells

Cryovials were briefly placed into a 37 °C water bath until the culture was mostly melted. The content of the vial was then transferred into a 15 ml falcon tube containing 9 ml of the appropriate media and mixed gently by hand. Cells were pelleted by centrifugation (5 min at 160×g) and resuspended in the appropriate media. A viable cell count was performed, before transferring cells to a culture vessel. PCOC cells were seeded into 10 cm plates with 2×10^6 irradiated J2 3T3 cells in 10 ml GMEM EGF- media. Already established keratinocytes were seeded into T25 flasks in 5 ml DMEM media. HEKn cells were seeded at a density of 2.5×10^3 cells/cm^2 into T175 flasks.

2.1.4.5 Subculture

Subculture was performed once 80% confluence was reached. The media was removed and the adherent cell layer was rinsed with PBS. Cells were incubated with TE solution until 80% detached from the culture vessel (the times necessary to achieve this for each line are described in Table 2.5). The culture vessels were lightly tapped to mechanically dislodge the remaining cells. TE was inactivated with an equal volume of the appropriate media and the cell suspension collected in a sterile universal container. Cells were pelleted by centrifugation (5 min at 160×g) and resuspended in the appropriate media. Cells were frozen for liquid nitrogen storage when required.

Already established keratinocytes lines were routinely split according to the ratios described in Table 2.5. The cell suspension in DMEM media was distributed in T80 or T175 flasks with a total volume of 15 or 25 ml respectively. Due to their high cell-cell adhesion, UPCISCC90 were first incubated 20 min with 2 mM PBS/EDTA then 5 min with TE.

The main modification in the PCOC subculture was related to the presence of irradiated J2 3T3 cells. Following the PBS rinse, a 2 min incubation with TE was used to remove the remaining J2 3T3 cells prior to adding the TE. PCOC cells were counted and seeded at a density of 0.8 to 1×10^6 cells per 10 cm plate with 2×10^6 irradiated J2 3T3 cells in a total of 10 ml GMEM EGF- media. Thereafter media was replaced every two days with GMEM EGF+ media and irradiated J2 3T3 cells when cover was sparse (usually twice a week).

HEKn cells were counted and seeded in T175 flasks at a density of 2.5×10^3 cells/cm^2 in a total of 25 ml of EpiLife media. Media was renewed every two days until 50% confluence was reached and every day thereafter.

Table 2.5: Trypsin/EDTA incubation time and split ratio of cell lines

Cell line	TE incubation time (min)	Split ratio
PCOC2	15	-
PCOC3	15	-
HEKn	5	-
UMSCC4	10	1/3
UMSCC6	5	1/2
UMSCC19	10	1/2
UMSCC47	10	1/3
UMSCC74a	7	1/2
UPCISCC90	5	1/2

2.1.5 Primary Culture of Oropharyngeal Cells

2.1.5.1 Sample collection and processing

Ethics and NHS R&D approval (reference 13/WA/0002) were obtained for the PCOC study and it was adopted into the NISCHR portfolio. Patients were recruited into the study during head and neck clinics and were consented by the clinical team (Dr Evans, Mr Owens and Mr Al-Husseini). Some patients were recruited at the time of diagnosis and others at the beginning of their treatments. Two routes of biopsy collection were used depending on the surgical procedure involved. In most cases an extra biopsy was taken during panendoscopy and placed directly in transport media. In some cases, when the whole tumour was resected, it was first assessed by a pathologist and any spare material used for research. The samples were placed in transport media, at room temperature until processed (for more details see the Study Protocol included in Appendix 1). The collection and processing steps for cell culture are illustrated in Figure 2.1. The samples were cut into two, half was used for nucleic acid extraction and the other half for cell culture. The tissue for cell culture was washed 3 times with PBS/Pen Strep 1×/amphotericin B 5.6 mg/l and incubated in the last wash for 15 min (Sigma-Aldrich). Tissue culture plates were coated with a thin layer of FBS. The biopsy was placed on the plate and curved scissors were used to it cut into pieces of 2 mm or less. During this process extra FBS was added if needed to avoid the biopsy drying out. The fragments of sample were then spread evenly between several 10 cm culture plates and incubated at 37 °C,

with 5% carbon dioxide, for 2 h in an inverted position. GMEM EGF- culture media (10 ml) containing 2×10^6 irradiated J2 3T3 cells was then added and the plates returned to the incubator.

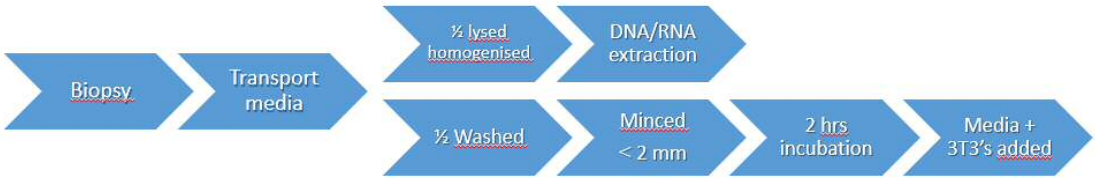


Figure 2.1: PCOC biopsy processing
PCOC biopsies were placed in transport media until processed. Half of each biopsy was lysed and homogenised than stored for nucleic acid extration. The other part of the biopsy was washed in PBS/Pen Strep/Amphotericin B, than cut into small fragments and placed into culture plates coated with FBS. Plates were incubated 2 h before EGF- GMEM and irradiated 3T3 feeder cells were added.

2.1.5.2 PCOC2 and PCOC3 samples

PCOC2 was derived from a fit 52 year-old man with no smoking history or any previous tumours. A primary tumour staged T4N2bM0 was found in the left tonsil involving the tongue base and posterior pharyngeal wall. The patient underwent a course of chemoradiotherapy which was completed in June 2013. As of April 2016, follow-up showed a positive outcome without any recurrence.

PCOC3 was derived from a 44 year-old man. He was a fit ex-smoker with a history of less than 10 pack/year and without any known previous tumours. The cut off between a light and heavy smoker is often regarded as 10 pack/year (Ang et al. 2010). A primary tumour in the right tonsil was staged T1N2aM0. The patient was treated by trans-oral laser resection and oropharynx and neck dissection. This was followed by post-operative radiotherapy, completed in June 2013. The treatment response was complete and no recurrence was observed as of 2016.

2.1.5.3 Early cell culture

The media was changed to GMEM EGF+ after two days and thereafter renewed every two days. Irradiated J2 3T3 cells were replaced when the coverage of the flask area became

sparse, usually twice a week. Photographs were taken using a Zeiss Axiovert 35M inverted bright field microscope and camera every 2-3 days in the early stages of culture to record the development of cell outgrowth (Zeiss, Cambridge, UK). Once some colonies were visible macroscopically, with a size of over 3 mm in diameter, cells were passaged for the first time. Cell passaging was performed largely as previously described. The main difference was the use of a short TE incubation of 5 to 10 min as the length of TE incubation needed to detach all cells was unknown. TE containing detached cells was removed and inactivated every 5 to 10 min and fresh TE added until cells growing out of the explants detached. The cell suspension was then centrifuged and the pellet resuspended in GMEM EGF- media. The cells were then placed in a new culture plate labelled passage 1 with 2×10^6 irradiated J2 3T3 cells. New GMEM EGF- media was added to the plates containing biopsy pieces still attached, this process was repeated over time until all the biopsy pieces detached or stopped producing cell outgrowth.

Early stage primary cultures were passaged once colonies were 50-70% confluent, from passage 2 onwards, cells were frozen down if enough cells were available, otherwise cells were cultured until consistent growth was obtained.

2.1.6 Culture of J2 3T3 cells

2.1.6.1 Thawing and seeding

J2 3T3 cells were thawed as described previously and plated into a T25 flask containing 5 ml of DMEM media. Once the cells reached 80% confluence they were transferred into a T80 flask containing 15 ml media then into several T175 flasks at a 0.5 to 1×10^6 cells density, using the method previously described.

2.1.6.2 Subculture

When 80% confluence was reached in T75 flasks, cells were collected for irradiation and re-seeded for further use. Cells were used until passage 25 than discarded and replaced by a new batch. Cells were harvested (using a 5 min TE incubation) and counted. Several T175 flasks were seeded at a density of 0.8 to 1×10^6 cells in a total of 25 ml DMEM media. The remaining cell suspension was placed in a new sterile universal container and stored at 4 °C until irradiation.

2.1.6.3 Irradiation

The J2 3T3 cells were irradiated using a caesium-137 source until a dose of 60 Gy was achieved (Best Theratronics, Ottawa, Canada). Irradiation rendered 3T3 cells post-mitotic i.e. unable to divide but capable of secretion of extra cellular matrix proteins. In further sections, irradiated J2 3T3 cells are referred to as 3T3 feeder cells or 3T3's when used to support other cells culture.

2.1.7 Growth curves

Cells were grown until 80% confluent, harvested and counted. The cells were plated in five 6-well plates with and without 3T3 feeder cells. The densities of PCOC and 3T3 feeder cells were adapted from 0.8×10^6 PCOC cells and 2×10^6 3T3 feeder cells cultured in a 10 cm plate. Every 24 h, cells contained in 3 wells were trypsinised. The cell suspension for each well was counted 3 times. One count corresponded to one grid of a haemocytometer. The average number of cells per well and per ml was calculated.

2.1.8 Chamber slide growth

Cells were harvested and counted before plating them at double the density used in normal condition in 3 ml media to ensure enough cells attached to the glass slides. Cells were fixed once they reached 50% confluent or over. An equivalent volume of 4% formaldehyde was added to the media to pre-fix cells for 2 min. The solution was then discarded and 4% formaldehyde added for 20 min. The solution was discarded and the slides washed with PBS. The chambers were removed and slides stored in PBS.

2.1.9 Mycoplasma detection

2.1.9.1 Reagents preparation and storage

A PCR based assay was used to screen cell lines for mycoplasma contamination. The primer set used in Venor GeM mycoplasma detection kit is specific to the highly conserved 16s rRNA coding region which allows for the detection of a range of mycoplasma species (Minerva Biolabs, Berlin, Germany). The reagents were rehydrated with PCR grade water according to the manufacturer's instructions and stored at -20 °C.

2.1.9.2 Sample collection and storage

Cultures were grown to at least 90% confluence and 500 µl of supernatant transferred to a microfuge tube. The samples were incubated at 95 °C for 10 min then cellular debris were pelleted by centrifugation (5 s at 8000×g), the supernatant was used in the PCR reaction mix.

The reaction mix was assembled in a designated PCR hood in sterile conditions according to Table 2.6. A volume of 23 µl was dispensed in each tube, 2 µl of negative control, positive control or sample were added to the designated tubes. DNA polymerase (5 U/µl) was not included in the kit and purchased separately (Minerva Biolabs). The amplification was performed as described in Table 2.7 using a Nexus GSX1 thermocycler (Eppendorf, Stevenage, UK).

Table 2.6: Composition of mycoplasma PCR reaction

Reagent	Volume n=1 (µl)
PCR grade water	15.3
10× reaction buffer	2.5
Primer/nucleotide mix	2.5
Internal control	2.5
Polymerase	0.2
Total volume	23

Table 2.7: Mycoplasma thermocycler conditions

Phase	Temperature (°C)	Time	Cycles
Initial denaturation	94	2 min	1
Denaturation	94	30 s	39
Primer annealing	55	30 s	
Extension	72	30 s	
Final extension	72	30 s	1

2.1.9.3 Gel electrophoresis and result interpretation

A 1.5% agarose gel was made and the samples loaded in the wells as described in section 2.2.5.2 . The electrophoresis was run at 100 V for 20 min. The gel was observed under UV light. The presence of a 191 bp internal control in all sample lanes indicated the success of

the reaction. Mycoplasma DNA was indicated by the presence of an extra band of around 265-278 bp.

2.2 Nucleic acid extraction

This section describes the various samples, their processing and the different methods of DNA and RNA extractions used.

2.2.1 DNA extraction from tonsil homogenate

2.2.1.1 Samples

Tonsil homogenates in a sucrose solution (n=524) were obtained from the National Anonymous Tonsil Archive. Samples were stored at -80 °C and thawed at room temperature before DNA extraction.

2.2.1.2 Phenol Chloroform DNA extraction

Microfuge tubes (2 ml volume) were filled with 150 µl of pre-warmed Tris/EDTA lysis buffer (composition outlined in Table 2.8), 500 µl tonsil homogenate and 100 µl of 20 mg/ml Proteinase K (Roche, Burgess Hill, UK). A blank extraction control was included every 23 samples. Samples were incubated overnight at 56 °C with gentle rotation (180 RPM).

Table 2.8: Composition of 5x 250 mM Tris/5 mM EDTA lysis buffer

Component	Final Concentration (M or %)
Tris	0.5 M
HCl	32%
NaCl	2 M
EDTA	0.2 M
SDS	10%
De-ionised water	300 ml, pH adjusted to 8, fill to 500 ml

The phenol chloroform extraction was performed in a safety cabinet with external venting. Following the addition of 750 µl of Phenol chloroform isoamyl alcohol pH 8 (Sigma-Aldrich), the tubes were vortexed, rotated for 5 min at room temperature (RT) then centrifuged 5 min at 8000×g. The aqueous phase was transferred to newly labelled tubes and the same process repeated following the addition of 750 µl chloroform isoamyl alcohol. The

volume of each sample was estimated and 1/10 of the volume of sodium acetate 3 M and an equal volume of isopropanol were added (Sigma-Aldrich). The samples were vortexed briefly and incubated at room temperature for 1 h then centrifuged at 8000×g for 10 min. A 70% ethanol solution was added to each tube (500 µl). The samples were vortexed, rotated gently for 10 min and centrifuged 5 min at 8000×g before discarding the supernatant. A further centrifugation of 10 s at 8000×g was used to remove any residual supernatant and the pellets allowed to dry at 37 °C for 10 min. Pellets were resuspended in 500 µl of Tris 10 mM/EDTA 1 mM pH 8 and incubated at 56 °C for 1h with gentle rotation at 160 RPM. Samples were stored at -20 °C. After 24 h a 3 µl aliquot was taken from each sample to quantify the DNA.

2.2.2 DNA and RNA extraction from cell culture

DNA and RNA were purified separately using QIAamp DNA Mini and RNeasy Fibrous Tissue kits (QIAGEN, Hilden, Germany). The extractions were performed according to the manufacturer's recommendation; the general processes are summarised below.

2.2.2.1 Harvesting cells for DNA/RNA extraction

Cells were harvested and counted once 80% confluence was reached. A maximum of $5x10^6$ cells were placed in a 2 ml tube and centrifuged for 5 min at 300×g. The supernatant was discarded and the pellet lysed and resuspended according to the manufacturer's recommendation.

2.2.2.2 DNA extraction using QIAamp DNA Mini kit

The pellet was resuspended in 200 µl PBS, mixed with 200 µl of AL buffer and 20 µl of Proteinase K. The mixture was incubated 10 min at 56 °C. Next, 200 µl of 100% ethanol were mixed with the sample, prior to loading on a QIAamp DNA mini spin column. The sample was centrifuged through the column for 1 min at 6000×g which bound the DNA to the membrane. Washes were performed with 500 µl of AW1 buffer followed by a 1 min centrifugation at 6000×g then with 500 µl AW2 buffer followed by a 3 min centrifugation at 20000×g. The flow-through was discarded after each centrifugation. The column was transferred to a new 2 ml collection tube and centrifuged at 20000×g for 1 min to eliminate any risk of buffer AW2 carry over. DNA was eluted with twice 50 µl of elution buffer followed by a 1 min centrifugation at 6000×g and store at -20 °C.

2.2.2.3 Total RNA extraction using RNeasy fibrous tissue mini kit

The columns and reagent contained in this kit were the same as the RNeasy mini kit. The protocol used was the recommended protocol for total RNA extraction from cell lines contained in the RNeasy mini kit. The working area and the equipment used were cleaned with RNaseZap (Sigma-Aldrich) to eliminate RNase which could degrade the RNA. The cell pellet was lysed with 600 µl of RLT buffer, containing 6 µl of 14.3 M β-mercaptoethanol (Sigma-Aldrich). All steps of the extraction were performed at RT, the RNA was kept on ice after the elution prior to long term storage.

The lysate was homogenised by passing at least 5 times through a sterile 20-gauge needle. The lysate was mixed with 70% ethanol and centrifuged through the provided RNeasy mini column 700 µl at a time 15 s at 8000×g, this bound the RNA to the silica membrane. The column was washed with 350 µl of RW1 buffer, centrifuged for 15 s at 8000×g and the flow-through discarded.

The samples were treated with DNase to eliminate DNA contamination in downstream applications. Prior to use, the DNase stock powder was dissolved in 550 µl of RNase-free water according to the manufacturer's instruction, aliquoted and stored at -20 °C. The working DNase solution was made by mixing 10 µl of DNase I with 70 µl of RDD buffer. The 80 µl of DNase I incubation mix were added directly on the column membrane and incubated at RT for 15 min. This was followed by another wash with 350 µl RW1 buffer and two washes with 500 µl RPE buffer, a longer centrifugation time of 2 min was used to dry the membrane for the last wash. A 2 min centrifugation step (8000×g) in a new collection tube was used to eliminate RPE carry over. The RNA was eluted using twice 30 µl RNase free water (centrifuged 1 min at 8000×g) and stored at -80°C.

2.2.3 Genomic DNA and total RNA extraction from oropharyngeal cancer biopsies

A method of simultaneous DNA and RNA extraction combining RNeasy fibrous tissue mini kit and QIAamp DNA Mini kit based on a protocol obtained from the Wales Cancer Bank was used. This section describes the variation from the protocols used for of each kit individually.

Half of each biopsy was homogenised using a TissueRuptor (QIAGEN) for 20 s in 600 µl RLT buffer containing 6 µl of 14.3 M β-mercaptoethanol (Sigma-Aldrich) and reagent DX (QIAGEN) to a final concentration of 0.5% v/v to prevent foaming.

Each sample was divided into two 300 µl volumes which were processed in parallel. The sample was mixed with 20 µl of Proteinase K and 708 µl of RNase free water then incubated for 10 min at 56 °C. The digestion solution was centrifuged 3 min at 8000×g and 650 µl of supernatant were removed to be processed for RNA extraction. The pellet was resuspended in the remaining volume and placed at 55 °C for later DNA purification.

The tube containing the clear lysate was processed for RNA purification by mixing in half a volume of 100% ethanol. The sample was then processed as previously described in the RNeasy protocol.

Following the RNA purification, the tube containing the homogenised biopsy kept at 55 °C was processed for DNA purification. The sample was mixed with 350 µl of buffer AL and incubated at 70 °C for 10 min before mixing in 350 µl 100% ethanol. This sample was loaded onto the membrane of a QIAamp DNA mini spin column 600 µl at a time and centrifuged 1 min at 8000×g before discarding the flow-through. This step was repeated until the whole sample passed through the column. The following washes and elution were performed as described previously in the QIAamp protocol.

2.2.4 Nucleic acid quantification

The concentration and purity of the nucleic acids were evaluated using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific) according to the manufacturer's instructions. A blank measure was established by loading 3 µl of the elution solution. The absorbance at 230 nm and 260 nm wavelengths of 2 µl of each samples was recorded. The absorbance measures provided the concentration in ng/µl and the ratio between 230/260 nm. High purity was indicated by ratios between 1.8 and 2.2.

2.2.5 DNA integrity

PCR amplification of a 209 bp or 1020 bp fragment of the housekeeping gene β-globin was used to assess the integrity of extracted DNA and its suitability for PCR base techniques.

Primers, shown in Table 2.9, were purchased from Sigma-Aldrich (Saiki et al. 1988; de Roda Husman et al. 1995). This section describes the overall PCR technique and gel electrophoresis used in all further PCR based assays.

Table 2.9: β-globin primers sequences

Primers	Forward primer	Reverse primer	Amplicon (bp)
PCO3/PCO5	ACACAACTGTGTTCACTAGC	GAAACCCAAGAGTCTTCTCCT	209
1kF/1kR	ATCCAAATAAGGAGAAGATATGC	AAACATCAAGCGTCCCATAG	1020

2.2.5.1 General reaction set up

PCR reactions were set up in a designated PCR hood. A master mix was made, which included PCR grade water, primers, buffers and Taq polymerase (QIAGEN and Sigma-Aldrich). This was distributed as 20 µl per labelled PCR tubes. Five microliters of diluted DNA, water for the negative control or 5 µl of HPV-positive sample for positive control was then added to the corresponding tubes. The reagents and volumes used are shown in Table 2.10. The sealed tubes were placed in a Nexus GSX1 thermocycler (Eppendorf) and the amplification reaction performed according to conditions in Table 2.11. The samples were then stored at 4 °C for short term usage or at -20°C for long term storage.

Table 2.10: Composition of β-globin PCR reaction

Reagent	Concentration	Volume n=1 (µl)
Primers (each)	5 µM	2.5
Deoxynucleotide (dNTP)	2 mM	2.5
10X reaction buffer	1×	2.5
Hotstar Taq polymerase	5U/µl	0.1
PCR grade water	-	9.9
DNA	1/10	5
Total volume	-	25

Table 2.11: β-globin thermocycler conditions

Phase	Temperature (°C)	Time	Cycles
Initial denaturation	94	15 min	1
Denaturation	94	30 s	40
Primer annealing	55	30 s	
Extension	72	60 s	
Final extension	72	4 min	1

2.2.5.2 Agarose gel electrophoresis

A 2% w/v agarose gel was prepared using 1x Tris/Borate/EDTA (TBE), agarose and 1 µl 10 mg/ml ethidium bromide per 100 ml of gel (Sigma-Aldrich). Samples were mixed with 5 µl of orange G loading dye (Sigma-Aldrich). The first well from each row of the gel was loaded with 10 µl of FullRanger DNA ladder for amplicons up to 5000 bp or PCR Ranger ladder for fragments up to 1000 bp (Norgen Biotek, Thorold, Canada). The other wells were loaded with 10 µl of the combined sample/orange G. The gel was run 1 h at 100 V in an electrophoresis tank containing 1×TBE to allow separation of the products. The bands were visualised and photographed under UV light using a GelDoc-It imaging system (UVP, Upland, USA). A product size of 209 bp or 1020 bp β-globin indicated the presence of intact DNA fragments and the suitability of the sample for further PCR analysis.

2.2.6 RNA integrity

RNA samples were analysed for integrity with an Agilent RNA 6000 Nano kit combined with Agilent 2100 Bioanalyser (Agilent Technologies, Santa Clara, USA) by Central Biotechnology Services (Cardiff University). During this chip electrophoresis assay, the RNA molecules were separated by size in the chip's micro-channels and detected by laser induced fluorescence. The molecular weight was determined by comparison with a ladder of known fragment sizes; the amount of fluorescence measured correlated with the amount of RNA. The results were represented either as gel like images or electropherograms. The RNA integrity was assessed by the ratio of the 18s and 28s ribosomal fractions. An RNA Integrity Number (RIN) was automatically calculated using an algorithm developed by the manufacturer. A RIN of 10 corresponded to un-degraded RNA whereas 0 showed the presence of only degraded RNA. RIN of 8 or over are generally accepted as good quality for sequencing analysis.

With the chip placed in the priming station, 9 µl of premade gel-dye mix equilibrated at room temperature was loaded at the bottom of a marked well. The plunger was set to 1 ml and the priming station closed. The plunger was pressed down, held for 30 s and released, to distribute the gel-dye mix. The chip was removed from the priming station and 9 µl of gel-dye mix loaded to other designated wells. The remaining wells were used for sample and the ladder (previously thawed, denatured 2 min at 70 °C and stored on ice), 5 µl of RNA 6000 nano marker were loaded in each well with 1 µl of sample or 1 µl of ladder. The chip was vortexed

for 60 s then inserted into the Agilent Bioanalyser and analysed according to manufacturer's instructions.

2.3 Short Tandem Repeat typing

This assay was used to compare the Short Tandem Repeat (STR) profiles (DNA fingerprints) of the newly developed oropharyngeal cancer cell lines and the biopsies they were derived from. It is an STR multiplex assay which amplifies 15 tetranucleotide loci. It was performed with assistance from Dr Sian Meyrick and Dr Steve Austin (Cardiff University Haematology Department) using an AmpFlSTR Identifiler PCR amplification kit according to the manufacturer's instructions (Thermo Fisher Scientific).

A single multiplex PCR was used to amplify all 15 loci. The locus designation and their chromosome location are described in Table 2.12. PCR primers bind to flanking regions of the STR sites resulting in different product sizes depending on the number of repeats present. One primer of each pair has a fluorescent label resulting in fluorescently labelled products. The PCR master mix was assembled by combining for each sample 10.5 µl of AmpFlSTR PCR reaction mix, 0.5 µl of AmpliTaq Gold DNA polymerase and 5.5 µl of AmpFlSTR identifiler primer set. Fifteen microliters of the master mix were dispensed to each tube and 10 µl of sample DNA, diluted to a concentration of 005-0125 ng/µl, was added. The amplification was done using a GeneAmp PCR system 9700 according to the cycle described in Table 2.13 (Thermo Fisher Scientific).

The PCR products were separated by electrophoresis in a capillary filled with performance optimised polymer 4 and detected on an ABI Prism 310 Genetic Analyser (Thermo Fisher Scientific). A system of 5 fluorescent dyes was used for detection. Following excitation, the dyes were distinguished using virtual filters and CCD imaging and the light intensities collected by the ABI Prism 310 data collection software. Following analysis an electopherogram was generated and the peaks used to compare similarities between samples at the different loci.

Table 2.12: Designation of Loci used in STR typing and their chromosome location

Locus Designation	Chromosome Location
D8S1179	8
D21S11	21q11.2-q21
D7S820	7q11.21-22
CSF1PO	5q33.3-34
D3S1358	3p
TH01	11p15.5
D13S317	13q22-31
D16S539	16q24-qter
D2S1338	2q35-37.1
D19S433	19q12-13.1
vWA	12p12-pter
TPOX	2p23-2per
D18S51	18q21.3
Amelogenin	X: p22.1-22.3 Y:p11.2
D5S818	5q21-31
FGA	4q28

Table 2.13: STR typing thermocycler conditions

Phase	Temperature (°C)	Time (min)	Cycles
Initial denaturation	95	11	1
Denaturation	94	1	28
Primer annealing	59	1	
Extension	72	1	
Final extension	60	1	1

2.4 Amplification of HPV16 genes E1, E2 and E6

The amplification of a 161 bp fragment of the HPV16 E6 gene was used confirm the presence of HPV16. The HPV16 E1 and E2 genes are often disrupted when the virus integrates within the human genome, their disruption was used as a surrogate marker for HPV integration. Tiling PCRs amplifying 2 overlapping fragments of HPV16 E1 and 5 overlapping fragments of HPV16 E2 were used. Details of the primers are illustrated in Table 2.14 and amplification conditions in Table 2.15. The reaction composition was the same as the β-globin PCR, described in section 2.2.5.1 .

Table 2.14: E1, E2 and E6 primers sequence and amplicon size

Primer	Forward primer	Reverse Primer	Amplicon (bp)
E1	CTAGGAATTGTGTGCCCATCTG	ACGTTGGCAAAGAGTCTCCATC	1954
E1.1	CTAGGAATTGTGTGCCCATCTG	CTTTCTATCCATTCTGGCGTGTCT	1069
E1.2	GATAGAGCCTCCAAAATTGCGT	ACGTTGGCAAAGAGTCTCCATC	995
E2	TTAAGTTTGCACGAGGACGA	CGCCAGTAATGTTGTGGATG	1167
E2.1	AGGACGTGGTCCAGATTAAG	TCAAACTGCACTTCCACTGT	414
E2.2	TAACTGCACCAACAGGATGT	GCCAAGTGCTGCCTAATAAT	340
E2.3	ATCTGTGTTTAGCAGCAACG	TAAATGCAGTGAGGATTGGA	224
E2.4	ACAGTGCTCCAATCCTCACT	TCACGTTGCCATTCACTATC	244
E2.5	GGCATTGGACAGGACATAAT	CAAAAGCACACAAAGCAAAG	207
E6	GAACAGCAATACAACAAACC	GATCTGCAACAAGACATACA	161

Table 2.15: E1, E2 and E6 PCR thermocycler conditions

Assay	Phase	Temperature (°C)	Time	Cycles
E1	Initial denaturation	95	15 min	1
	Denaturation	94	30 s	30
	Primer annealing	59	30 s	
	Extension	72	3 min	
	Final extension	72	7 min	1
E2	Initial denaturation	95	15 min	1
	Denaturation	94	30 s	40
	Primer annealing	58	30 s	
	Extension	72	30 s	
	Final extension	72	5 min	1
E6	Initial denaturation	94	15 min	1
	Denaturation	94	30 s	40
	Primer annealing	55	30 s	
	Extension	72	60 s	
	Final extension	72	4 min	1

2.5 HPV typing

2.5.1 GP5+/GP6+

GP5+/GP6+ PCR-EIA is a PCR based assay which uses solid phase enzyme immunoassay (EIA) to detect several types of HPV (Jacobs et al. 1997). This technique can detect 14 high risk types: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68 and 6 low risk types: 6, 11, 40, 42, 43, 44. A fragment of the HPV L1 open reading frame is amplified using consensus primers targeting well conserved sequences. The forward primer is termed GP5+ and the biotinylated reverse primer GP6+ (Table 2.16). The biotinylated PCR products are captured on streptavidin coated wells. After denaturation, the amplicons are hybridised with digoxigenin labelled HPV oligonucleotides probes which are specific for individual HPV types.

Table 2.16: GP5+/GP6+ primers sequence

Primer	Sequence	Amplicon (bp)
GP5+	TTTGTTACTGTGGTAGATACTAC	142
GP6+	GAAAAATAAACTGTAAATCATATTC 5'Biotinylated	

2.5.1.1 GP5+/GP6+ PCR EIA

DNA was extracted from tonsil homogenates samples (as described in section 2.2.1 and diluted 1/10. The reaction components were assembled according to Table 2.17and the PCR conditions are shown in Table 2.18.

Two EIA were performed, the first used a cocktail of probes for the 14 high risk HPV types and the second used a cocktail of probes for the 6 low risk types. The PCR products were bound to a 96 Streptawell plate (Roche), 5 µl of amplicon were added per well with 50 µl of wash buffer. The plates were incubated 1h at 37 °C then washed 3 times with 1× wash buffer. The wash buffer was made as a 5× stock, comprising of 43.8 g of NaCl, 22 g of sodium citrate and 25 ml of tween 20 in 1 l of sterile water. DNA bound to the plates was denatured using 100 µl of 0.2 M NaOH per well, followed by 3 washes.

The digoxigenin labelled probe cocktail (Roche) was added, 50 µl of 10 µM probe solution per well and left to hybridize with the amplicons for 1 h at 37 °C. The plates were washed 3 times to remove unbound probes and 50 µl anti-Digoxigenin conjugate (Roche) were

added per well. This was followed by 1 h incubation at 37 °C and 5 washes. The substrate solution was added, 100 µl per well and the plates were incubated 1 h at 37°C in the dark.

The absorbance at 415 and 630 nm was recorded after 1, 2 and 24 h using a Bio-Rad 550 microplate reader (Bio-Rad, Watford, UK). A sample was considered positive when its absorbance values were at least 3 times higher than the negative control or background value.

Table 2.17: Composition of GP5+/GP6+ PCR reaction

Reagent	Concentration	Volume n=1 (µl)
Primers (each)	5 µM	2.5
dNTP	2 mM	2.5
10X reaction buffer	1×	2.5
$MgCl_2$	50 mM	1.75
Hotstar Taq polymerase	5U/µl	0.1
PCR grade water	-	11.15
DNA	1/10	2
Total volume	-	25

Table 2.18: GP5+/GP6+ PCR thermocycler conditions

Phase	Temperature (°C)	Time	Cycles
Initial denaturation	94	4 min	1
Denaturation	94	30 s	40
Primer annealing	40	90 s	
Extension	72	1 min	
Final extension	72	4 min	1

2.5.2 HPV detection using Inno DNA EIA SPF10 kit

This is an *in vitro* PCR DNA EIA (DEIA) which can detect 40 types of HPV with a high sensitivity (Innogenetics, Gent, Belgium). Due to the small size of the fragments amplified this assay is suitable for use with degraded or archival material (Tjalma et al. 2013). The HPV types detected are: HPV 3-8, 11, 13, 16, 18, 26, 27, 30-35, 37, 39, 40, 42, 43, 44, 45, 51-56, 58-62, 64-71 and 74. The protocol recommended by the manufacturer was used without modifications. A PCR step amplified a 65 bp fragments in the L1 region, one of the primers is biotinylated resulting in the labelling of one strand of the product. The amplicons were captured on the wells of a streptavidin coated plate. An incubation with a denaturating solutions and a washing step were used to remove the non-biotinylated strand. The amplicons were incubated with a cocktail of digoxygenin-labelled HPV-specific probes. The label was then detected using an

anti-digoxygenin enzyme conjugate which was visualised by a chromogenic substrate. The absorbance of each well was measured and the presence of HPV DNA determined by comparing the optical density of the PCR product to that of a cut-off value. To avoid contaminations two rooms were used: one for the PCR set up and the second for the ELISA.

2.5.3 Data analysis

The validation criteria of the assays were as follows:

1. The optical density of the HPV-positive sample preparation control was over 4 times the level of optical density of the HPV borderline DEIA control.
2. The optical density of the HPV-negative sample preparation control was over 4 lower than the level of optical density of the HPV borderline DEIA control.
3. The optical density of the HPV-positive control was over 4 times the optical density of the HPV borderline DEIA control.
4. The optical density of the HPV-negative control was over lower than the optical density of the HPV borderline DEIA control.
5. The optical density of the HPV-positive DEIA control was over or equal to 3 times the optical density of the HPV borderline DEIA control.
6. The optical density of the HPV borderline DEIA control was over or equal to 0.1 times the optical density of the HPV borderline DEIA control.
7. The optical density of the HPV-negative DEIA control was lower or equal to 0.6 times the optical density of the HPV borderline DEIA control.

If all of those criteria were met, the results were analysed. Samples were considered HPV-positive if their optical density was greater than the optical density of the HPV borderline DEIA control. Samples were considered HPV-negative if their optical density was equal or less then the optical density of the HPV borderline DEIA control. Samples with an optical density between 0.75 or 1 times of the optical density of the HPV borderline DEIA control were re-tested.

2.6 Detection of TP53 mutation by direct sequencing

PCOC2 and PCOC3 were analysed for TP53 mutation using the IARC protocol, 2010 update (http://p53.iarc.fr/Download/TP53_DirectSequencing_IARC.pdf). This technique was based on the amplification of 13 fragments of the TP53 gene within exons 2-11 followed by direct sequencing, the resulted sequences were analysed to detect any mutations.

2.6.1 Sample preparation

Cells were harvested, processed for DNA extraction and the DNA analysed with a NanoDrop 1000 Spectrophotometer to assess the concentration and purity. The samples were diluted to 50 ng/µl using PCR grade water (Sigma-Aldrich).

2.6.2 PCR reaction

The 13 primer sets were obtained from Sigma Aldrich, their sequences and amplicons sizes are described in Table 2.19. The reactions components were assembled in a PCR hood in sterile conditions as described in Table 2.20, the mix was transferred to labelled tubes. Depending on the primer sets, 4 different amplifications cycles were used as detailed in Table 2.21 and runned on a Techne TC-512 thermocycler (Bibby Scientific, Stone, UK).

Table 2.19: TP53 PCR primers and corresponding programs

Primers/direction	Sequence	Region Amplified	Amplicon (bp)	PCR Program
P-559 (F) P-E3Ri (R)	tctcatgctggatccccact agtcagaggaccaggtcctc	Exons 2-3	344	B
P-329 (F) P-330 (R)	tgctcttttcacccatctac atacggccaggcattgaagt	Exon 4	353	B
P-326 (F) P-327 (R)	tgaggacctggtcctctgac agaggaatcccaaagttcca	Exon 4	413	B
P-312 (F) P-271 (R)	ttcaactctgtctccttcct cagccctgtcgtctctccag	Exon 5	248	B
P-239 (F) P-240 (R)	gcctctgattcctcactgat ttaacccctcctcccagaga	Exon 6	181	B
P-236 (F) P-240 (R)	tgttcacttgtgccctgact ttaacccctcctcccagaga	Exons 5-6	467	B
P-333 (F) P-313 (R)	cttgccacaggtctccccaa aggggtcagaggcaagcaga	Exon 7	237	C
P-237 (F) P-238 (R)	aggcgcactggcctcatctt tgtgcagggtggcaagtggc	Exon 7	177	B
P-316 (F) P-319 (R)	ttccttactgcctcttgctt aggcataactgcacccttgg	Exon 8	231	B
P-314 (F) P-315 (R)	ttgggagtagatggagcct agtgttagactggaaacttt	Exons 8-9	445	B
9F (F) 9R (R)	gacaagaagcggtggag cggcattttgagtgttagac	Exon 9	215	E
P-E10Li (F) P-562 (R)	caattgtaacttgaaccatc ggatgagaatggaatcctat	Exon 10	260	D
P-E11L2 (F) P-E11Re (R)	agaccctctcactcatgtga tgacgcacacctattgcaag	Exon 11	245	B

Table 2.20: Composition of TP53 PCR reaction

Reagent	Concentration	Volume n=1 (µl)
Primers (each)	10 µM	0.8
dNTP	5 mM	0.8
10× reaction buffer	1×	2
Hotstar Taq polymerase	5U/µl	0.1
PCR grade water	-	14.5
DNA	50 ng/µl	1
Total volume	-	20

Table 2.21: TP53 thermocycler conditions

Name	Phase	Temperature (°C)	Time	Cycles
B	Initial denaturation	94	2 min	1
	Denaturation	94	30 s	20 -0.3 °C every 3 cycles
	Primer annealing	63	45 s	
	Extension	72	1 min	
	Denaturation	94	30 s	30
	Primer annealing	60	45	
	Extension	72	1 min	
	Final extension	72	10 min	1
C	Initial denaturation	95	15 min	1
	Denaturation	94	30 s	50
	Primer annealing	60	30 s	
	Extension	72	30 s	
	Final extension	72	10 min	1
D	Initial denaturation	94	2 min	1
	Denaturation	94	30 s	20 -0.5 °C every 3 cycles
	Primer annealing	58.5	45 s	
	Extension	72	1 min	
	Denaturation	94	30 s	30
	Primer annealing	55	45 s	
	Extension	72	1 min	
	Final extension	72	10 min	1
E	Initial denaturation	94	2 min	1
	Denaturation	94	30 s	50
	Primer annealing	57	30 s	
	Extension	72	30 s	
	Final extension	72	10 min	1

2.6.3 DNA purification

Following the PCR reactions a gel electrophoresis using 5 µl of product was used to verify the size of the products and the remainder of the samples were stored at -20 °C until further use. The Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, Amersham, UK) was used to purify and concentrate DNA according to the manufacturer's protocol. Once defrosted, a gel electrophoresis was run with each sample. Each band was viewed under low setting UV light and cut out using a clean scalpel. Gel bands were placed in labelled tube of known weight and the weight of each gel slice calculated.

For each 10 mg of gel band, 10 µl of capture buffer 3 were added in each tube. The agarose was melted with a 60 °C incubation, mixing the tubes by inversion until the agarose

was fully melted. The pH of the samples was checked using their colour and adjusted with the addition of 3M NaOH pH5 until the required colour (yellow/pale orange) was achieved.

The samples were centrifuged briefly then transferred onto a GFX microspin column placed in a collection tube800 µl at a time. Following 1 min incubation at room temperature the columns and collection tubes were centrifuged 30 s at 16000×g and the flow-through discarded. This step was repeated if the sample volume exceeded 800 µl.

The membrane was washed by adding 500 µl of wash buffer type 1 to each sample, centrifuged 30 s at 16000×g and the flow-through was discarded. This step was repeated followed by a further 30 s centrifugation at 16000xg to ensure there was no carryover of the capture buffer.

The columns were transferred to new labelled DNase free tubes. The DNA was eluted by the addition of 10 to 50 µl of elution buffer 6, incubated at room temperature for 1 min then centrifuged 1 min at 16000×g. The samples were either processed immediately or stored at –20 C until further use. Prior to freezing a 5 µl aliquot was taken from each sample to verify the purification process and estimate the DNA concentration.

Gel electrophoresis was performed as previously described and a picture taken. For each sample the presence of a single band was checked. The concentration of each PCR product to be sequenced was estimated by comparing the brightness of each band to the ladder of known concentration. The samples were then diluted using PCR grade water to obtain a concentration of 1 µg/ml in a 20 µl volume. The PCR primers were also diluted to a concentration of 3.2 pmol/µl in 20 µl. Both the diluted samples and primers were shipped for Sanger sequencing (Source BioScience, Nottingham, UK).

2.6.4 Data analysis

The chromatograms of the Sanger sequencing reactions were analysed manually using BioEdit software (Ibis Biosciences, Carlsbad, USA) to check that each base corresponded to the peak in the chromatogram before the generation of fasta files for each sequence. Using the fasta format files, parts of the sequence of poor quality were removed, usually at either end of the sequence. The remaining test sequence was aligned with the genomic sequence NC_000017.10 (Homo sapiens chromosome 17, which includes the TP53 gene). In cases where

the alignment between the test and reference sequence was less than 100%, mismatches were compared with the chromatograms to check whether this came from an error or a true mismatch. The fasta format sequence was adjusted accordingly. The portions of the adjusted fasta format sequence containing mismatches were aligned to the IARC TP53 sequence and the position of the mismatches recorded. The positions of mismatches were then entered in the NCBI SNP database (https://www.ncbi.nlm.nih.gov/snp) to check whether they corresponded to known variations.

2.7 Olaparib dosing assay

2.7.1 Drug preparation

Olaparib (AZD2281) was purchased as 10 mg powder (Selleckchem, Munich, Germany). Due to low water solubility, DMSO was used in the first dilution step. A 10 mM stock solution was made by adding 2.3017 ml of DMSO and mixing until all powder was dissolved. A second stock solution of 1000 µM was made following a 1/10 dilution of stock 1 in cell culture media (DMEM, GMEM or EpiLife). Both stock solutions 1 and 2 were aliquoted into single use doses and stored at -80 °C for a maximum of 1 year.

2.7.2 Crystal violet staining

The media was pipetted out of each plate. The remaining media and 3T3 feeder cells were removed with a PBS wash. The plates were incubated for 2 min with 5 ml of 50% v/v PBS/methanol. This solution was discarded and replaced with 5 ml of 100% methanol per plate. Following a 10 min incubation at RT, the methanol was discarded and the plates left to dry. Crystal violet, 5 ml per plate, was added and the plates were incubated for a maximum of 10 min (Thermo Fisher Scientific). The crystal violet was removed and filtered to be reused, the plates were rinsed with water and left to dry at RT.

2.7.3 Plating density assay

Cells were harvested once 80% confluence was reached and counted. The cell suspension was diluted to obtain solutions of various cell densities and plated in 6 cm plates. For each condition triplicates were used both in the presence of and without 3T3 feeder cells

at a density of 1×10^6 per plate. The plates were then placed in the incubator and the media replaced after 7 days. After 10-15 days, the plates were stained with crystal violet.

2.7.4 Olaparib dosing experiment

These experiments investigated the effect of a single dose of Olaparib on the colony forming capacity of a cell line panel. The concentrations of Olaparib used were: untreated, vehicle only (containing 0.1% DMSO), 0.1 µM, 0.5 µM, 1 µM, 5 µM and 10 µM. The DMSO concentration was adjusted to 0.1% in all conditions, this corresponded to the amount of DMSO contained in the 10 µM condition. For each condition triplicates were used.

Cells were harvested and counted and the resulting cell suspension diluted if necessary. Stock solutions were made in 3 x 50 ml tubes each containing enough cells for eight 6 cm plates, in a total of 40 ml media and containing $8x10^6$ 3T3 feeder cells if necessary. The cell densities used for each cell line were determined as described in the previous section. The cell suspension was pipetted into 21 culture plates which were incubated for 24 h to allow cells to attach before treatment with the drug Olaparib.

A 10 µM solution of Olaparib was made following a 1/100 dilution of the 1000 µM stock 2 solution (350 µl of stock 2 into 34.65 ml media). A serial dilution in 20 ml universal tubes was used to obtain the other Olaparib solutions. The required amount of DMSO was added to each tube to adjust the concentration to 0.1% before vortexing and transferring solution to the next tube. The vehicle solution was obtained by adding 20 µl DMSO to 20 ml of media.

After the 24 h incubation, the plates were checked under the microscope to ensure cells adhered to the plates and if present that the 3T3 feeder cells covered to whole plate. The media was pipette out and discarded. The plates were labelled in triplicates with each condition and 5 ml of treated solutions were added to the corresponding plates. The plates were returned to the incubator and left undisturbed for 7 days. At day 8, the plates were observed and the media was pipetted out and replaced (without Olaparib). The plates were incubated until the end of the experiment, as determined by the density tests. Pictures were taken using a bright field microscope and attached camera at several magnifications and the plates were stained with crystal violet as described previously.

2.7.5 Colony counting and data analysis

Picture of the plates were taken using a Canon EOS Rebel T3 camera coupled with a Colony Doc-It imaging station (UVP). For each plate, 3 pictures were taken. The plate was rotated 90 ° between each image acquisition. Colonies were counted using Colony Doc-It software (UVP). For each experiment a template was created by selecting a panel of pixel colours. The template was designed to recognise colonies and minimise the background. The template was then used to count colonies on 3 pictures per plate, for each plate within the triplicate of each condition. The average number of colonies was calculated for each plate and for each condition. Using this number, the plating efficiency and the surviving fraction were calculated.

2.8 DNA DSB and cell cycle analysis

This section describes the final protocol used for the serine 139 phosphorylates histone H2AX (γ-H2AX) and 4', 6-Diamidine-2'-phenylindole (DAPI) staining of nuclei and flow cytometry analysis. Several variations of the treatment conditions, timing and reagents were used in method development experiments. These will be described in the result section.

2.8.1 Cell seeding and treatment

Cells were harvested once 80% confluence was reached and counted. The required amount of cells was seeded in fourteen 6 cm or wells of 6 well plates in 5 ml of media according to Table 2.22. Six well plates were used for PCOC2 and PCOC3 to scale down the amount of cells needed and these were grown with the support of 3T3 feeder cells. Plates were treated once at least 40% confluence was reached (Table 2.22). In each experiment, 4 plates were left untreated with the media renewed. The other conditions were each applied to 2 plates: vehicle (0.1% DMSO), 0.1 1 and 10 µM Olaparib, 25 µM Etoposide (Sigma-Aldrich). As was described in previous experiments using Olaparib, the DMSO level was adjusted to 0.1% in each condition.

Table 2.22: Cells seeding densities and days until Olaparib treatment for DSB and cell cycle analysis

Cell line	Seeding density	Days until treatment
PCOC2	95 000 + 250 000 3T3 feeder cells (6 well plates)	5
PCOC3	95 000 + 250 000 3T3 feeder cells (6 well plates)	7
UMSCC47	175 000	2
UPCISCC90	100 000	3
UMSCC4	100 000	2
UMSCC19	150 000	3
UMSCC74a	800 000	2
HEKn	55 000	7

2.8.2 Cell collection

After 24 h and 48 h cells from two untreated plates and one of each treated conditions were photographed then harvested. A viability count was performed for each condition and 200 000 cells transferred to labelled 1.5 ml tubes to be used for nuclei isolation and staining (Table 2.23).

Table 2.23: Content of tubes analysed by flow cytometry

Conditions	Tube name
untreated	PBS
untreated	γ-H2AX single stain
untreated	DAPI single stain
untreated	Untreated double stain
vehicle	Vehicle double stain
0.1 µM Olaparib	0.1 µM Olaparib double stain
1 µM Olaparib	1 µM Olaparib double stain
10 µM Olaparib	10 µM Olaparib double stain
25 µM Etoposide	25 µM Etoposide double stain

2.8.3 Nuclei preparation and staining

The 1.5 ml tubes were centrifuged 5 min at 160×g at RT to pellet the cells. The supernatant was discarded and the cells resuspended in 1 ml of cold lysis buffer (Table 2.24). Following the lysis buffer addition, the tubes were directly centrifuged 5 min at 3.3 RPM and 4 °C. The supernatant containing cellular debris was discarded and the nuclei gently resuspended in PBS or staining solution before being transferred to round bottom tube (BD biosciences, Oxford, UK). The nuclei preparations were incubated for 20 min before flow

cytometry analysis. Tubes were kept on ice protected from light the whole time. DAPI was purchased as a powder solution and diluted in double distilled water to obtain a main stock solution of 1 mg/ml (Sigma-Aldrich). The Anti-γ-H2AX antibody clone JBW301 fluorescein isothiocyanate (FITC) conjugate was raised in mice (Merck-Millipore, Watford, UK; catalogue 16-202A).

Table 2.24: Composition of 1× cell lysis buffer

Components	Concentration	Source
Sucrose	320 mM	Sigma-Aldrich
Tris-HCl pH8	10 mM	
$MgCl_2$	2.5 mM	
Triton X-100	0.5%	

2.8.4 Flow cytometry analysis

Nuclei were analysed using a BD FACSCanto II flow cytometer (BD biosciences). DAPI was excited with the 405 nm violet laser and the FITC fluorochrome with the 488 nm blue laser. One drop of BD CompBeads Negative Control was added to the single stain tubes to facilitate compensation (BD biosciences). UPCISCC90 and PCOC3 cells had a tendency to clump resulting in the presence of aggregates in the nuclei preparation. To avoid cytometer blockage, the nuclei preparation was passed through 20 μm CellTrics filters (Sysmex, Milton Keynes, UK). The data analysis was performed using FlowJo (FlowJo LLC, Ashland, USA). The cell cycle pattern of single nuclei was analysed using the Watson pragmatic algorithm.

2.9 Western blotting

2.9.1 Materials

All materials used for the protein extraction, western blotting and their sources are outlined in Table 2.25 and Table 2.26. Chemicals used to make any additional solutions were purchased from Sigma-Aldrich.

Table 2.25: Material used in the protein extraction and their source

Material	Source
cOmplete mini EDTA-free protease inhibitor cocktail	Sigma-Aldrich
Protein standard 2 mg protein/ml	
Sodium orthovanadate (200 mM)	
DNase 1mg/ml	

Table 2.26: Western blot material and their sources

Material	Sources
Mini gel tank	Thermo Fisher Scientific
PowerEase 90W power supply	
Novex Bolt transfer buffer 20×	
Bolt 4-12% Bis-Tris plus gels 15 wells	
Novex Bolt Mops SDS running buffer 20×	
Novex MagicMark XP Western protein Standard (ladder)	
Novex Nitrocellulose Pre-Cut Blotting Membranes	
Novex Bolt LDS sample buffer 4×	
Novex Bolt Antioxidant	
Novex Bolt sample reducing agent 10×	
p53 antibody (rabbit) cat: 9282	Cell Signaling, Danvers, USA
Phospho-p53 (Ser15) antibody (rabbit) cat:9284	
PARP1 antibody (rabbit) cat:9542	
β-Actin (13E5) antibody (rabbit) cat: 4970	
Goat polyclonal Secondary Antibody to Rabbit IgG adsorbed cat: ab97080	Abcam,
Amersham ECL prime blocking reagent	GE Healthcare
Amersham ECL prime detection reagent	
De-ionised water	Pharmacy

2.9.2 Sample preparation

Cells were seeded and treated as described for the flow cytometry experiments (section 2.8.1). The only variation was the replacement of the Etoposide condition by a 10 Gy irradiated sample. Cells were collected 24 h and 48 h after treatment. Cells from each condition were harvested and counted. Afterwards, cells were washed by centrifugation in 1 ml of cold TBS. The TBS was discarded and the pellet was snap frozen in liquid nitrogen before storage at -80 °C for at least 24h before protein extraction. TBS solution contained 50 mM TrisHCl and 150 mM NaCl made up to 1 l in water.

2.9.3 Protein extraction

Samples, reagents and newly labelled tubes were kept on ice throughout the extraction. Pellets were thawed on ice in the presence of DNase (1 µl of 1 µg/ml solution per 1×10^6 cells). Homogenisation buffer (50 µl per 1×10^6 cells) was added to each tube and mixed by vortexing. Tubes were incubated on ice for 30 min, vortexing them every 10 min. The samples were centrifuged 5 min at 8000×g and 4 °C. The supernatant was transferred to new tubes and the volume recorded. Samples were stored at -80 °C.

2.9.4 Bradford assay

Samples were thawed on ice. In a 96 wells plate, a stock solution was made for each sample using a 1/20 dilution in water. Samples were pipetted in duplicates, 10 µl each, before the addition of 190 µl Bradford reagent (diluted 1/2). The plate was incubated 25 min at RT protected from light and the absorbance at 595 nm was recorded. Using a standard curve, the protein concentration of each sample was calculated.

2.9.5 Protein sample preparation

Protein samples for all cells lines were prepared to give a final loading concentration of 1000 µg/ml. LDS loading buffer and reducing agent were added to the sample to obtain a 1× final concentration. Water was then added up to the desired volume. Samples were mixed by vortexing before a 10 min incubation at 70 °C. A brief centrifugation was used to remove condensation from the caps and the samples kept on ice until the electrophoresis.

2.9.6 Electrophoresis

The gel tank was rinsed with tap water than assembled according to the manufacturer's instructions and filled half way with 1× running buffer (made by diluting 50 ml of Mops SDS running buffer 20× in 950 ml of water). A pre-cast gel cassette was removed from its packaging and rinsed with de-ionised water. The tape covering the slot at the back was peeled off and the comb gently removed. The wells and the cassette were rinsed twice with 1× running buffer. The cassette was inserted into the electrophoresis tank and clamped into place leaving the wells out of the buffer. Twenty microliters of sample were loaded per well whilst one well on either side of the samples was used for the ladder (5 µl). The gel was lowered into the

buffer and the tank filled up with 1× running buffer. The electrophoresis was run for 35 min at 200 V (constant) and 100 mA.

2.9.7 Transfer

During the electrophoresis, the transfer pads and nitrocellulose membrane were soaked in transfer buffer. The transfer buffer composition is detailed in Table 2.27. The filter paper was soaked briefly just before use. The anode and cathode were rinsed with transfer buffer then 5 ml of transfer buffer placed in each. At the end of the electrophoresis, the plastic cassette was separated using a gel knife. One side was removed so that the gel was left adhering on the other side and the wells were cut out. One piece of filter paper was placed on the gel before it was detached from the remaining cassette side and transferred filter paper down onto blotting pads in the cathode. The rest of the transfer sandwich was assembled according to Figure 2.2, the membrane was placed on the gel followed by another sheet of filter paper and blotting pads. The Anode was place on top to close the module before placing it in the electrophoresis tank. The centre of the module was filled with transfer buffer as well as the tank just below the electrodes. A 10 V constant voltage and 160 mA were applied for 1 h.

Table 2.27: Composition of 1× transfer buffer

Component	Volume (ml)
Bolt transfer buffer (20×)	50
Bolt antioxidant	1
Methanol	100
Water	849

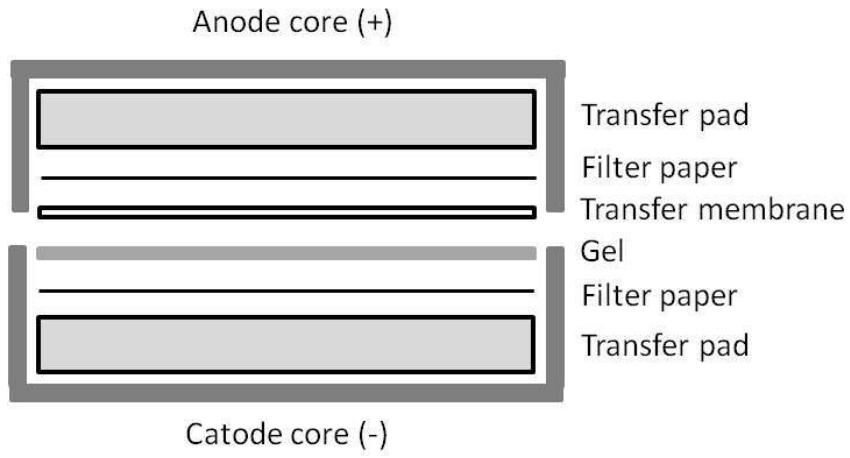


Figure 2.2: Western blotting transfer sandwich
Assembly of transfer pads, filter paper, gel and transfer membrane in the blotting module for protein transfer.

2.9.8 Blocking

A 5% blocking solution was prepared by dissolving 0.5 g ECL blocking solution in 10 ml TBST per membrane. TBST was formulated like TBS, adding 1 ml of Tween20. The membrane was washed twice for 5 min in 20 ml water in a shaker set at 1 revolution per second. The blocking solution was applied and the membrane incubated 1 h in a shaker at the same speed.

During the blocking, the primary antibody solutions were prepared in 10 ml of 5% blocking solution. Primary antibodies were all used at 1/1000 dilution. After the blocking, the membrane was rinsed twice with TBST then washed in TBST on the shaker 15 min followed by 3 times 5 min. The membrane was incubated overnight in the primary antibody solution in a cold room on a shaker. If the primary antibody used was targeting a phosphorylated protein, a 1/1000 dilution of phosphatase was added.

The next day the membrane was rinsed twice with TBST and washed on a shaker at room temperature 15 min then 3 times 5 min in TBST. The secondary antibody solution was prepared using a 1/5000 dilution in 10 ml of 2% blocking solution per membrane. The membrane was incubated 1 h in the secondary antibody on the shaker. It was then rinsed twice with TBST and washed 15 min then 3 times 5 min in TBST on the shaker.

2.9.9 Chemiluminescence

The ECL Prime chemiluminescence substrate solution was prepared by mixing 2 ml of solution A and B equilibrated at RT and keeping the solution in the dark. The membrane was drained from excess liquid and placed on a clean plastic sheet. The chemiluminescence substrate was spread evenly on the membrane and incubated in the dark for 5 min. The solution was drained and another plastic sheet placed on top of the membrane. Image capture was performed using a LAS-3000 imager (Fujifilm) using an initial 30 s exposure than 2 min increments. The analysis and quantification was performed using Image Studio Lite (LI-COR Biotechnology, Cambridge, UK).

2.9.10 Reprobing

Following image capture, the membrane was rinsed with TBST then washed twice with TBST for 5 min in the shaker. The membranes were left to dry overnight in the dark then stored at 4 °c wrapped in plastic film. The steps described above were then followed starting after the blocking.

2.10 RNA sequencing

RNA sequencing enables the accurate profiling of the entire transcriptome of a cell or tissue (Wang et al. 2009). RNA was extracted, quantified and quality tested as described in sections 2.2.2, 2.2.4 and 2.2.6. A sequencing library was then prepared using the TruSeq mRNA library preparation kit (Illumina, San Diego, USA). The library was sequenced using the Illumina 2000 sequencing technology. Both the library preparation and the sequencing were carried out as a commercial service provided by the Beijing Genomics Institute. Sequencing data was returned via an FTP server and was analysed with support from Drs Kevin Ashelford and Peter Giles of the Wales Gene Park. Data analysis included quality assessment, analysis of viral gene expression and integration, differential gene expression and Gene Ontology (GO) analysis. Comparison of expression levels of individual genes were also performed.

2.10.1 Library preparation and cluster amplification

A short insert sequencing library was prepared using the TruSeq Stranded mRNA library preparation kit (Illumina). This enabled capture of the coding transcriptome with precise

strand identification (achieved using oligo-dT bead-based capture of poly-A tailed mRNA). Library generation was performed by the Beijing Genomics Institute in accordance with the manufacturer's instructions. mRNA was first captured, then fragmented by the use of high temperature and divalent cations. The fragments were reverse transcribed into cDNA and adapters ligated to both ends. The library was amplified by PCR using a primer cocktail which annealed to the adapters.

Cluster amplification was performed on glass flow cells coated with two types of oligonucleotides, each complimentary to one of the adapters. Following the binding of a cDNA strand to an oligonucleotide on the flow cell, a complementary strand was synthesized by a polymerase. The double stranded molecules were denatured and the original template washed away. The DNA strand then folded over to bind to the other oligonucleotide complimentary to its 3' end. The complimentary strand was synthesized by the polymerase and the double strands denatured. This process resulted in two complimentary strands each bound to an oligonucleotide. This process was repeated and resulted in the clonal amplification of the cDNA strands. This occurred at different places or clusters on the chip for all fragments present in the library.

2.10.2 Sequencing reactions

Sequencing was performed using a 101 bp paired-end strategy. This approach facilitated sequencing of both ends of each fragment to generate high-quality alignable sequence. Paired-end RNA sequencing enables discovery applications such as detection of gene fusions in cancer and characterisation of novel splice isoforms. Sequencing was performed at the Beijing Genomics Institute using a HiSeq 2000 (Illumina). Sequencing primers were bound to the end of the template and nucleotides, each labelled with a different fluorescent dye, were added. Only the nucleotide corresponding to the sequence was incorporated. Following each nucleotide addition, the fluorochromes were excited by a light source and the signal recorded. Sequencing data were recorded in pairs, each fragment was sequenced both ways. Following this, FastQ files were generated and were the basis of subsequent analysis.

2.10.3 Sequence data analysis

2.10.3.1 Quality control

For each sample, standard indicators of sequence quality were assessed using a combination of FastQC and a variety of custom codes. Metrics assessed included (but were not limited to) total number of reads, median insert size, per base quality assessment and sequence length distribution.

2.10.3.2 Sequence mapping

Reads were mapped against human reference sequence hg19 (genome build 19) and HPV16 reference sequence NC_001526 using Tophat2 (Center for Computational Biology, USA). The number and proportion of mapped reads were determined.

2.10.3.3 Gene expression

Raw read counts were calculated for each feature in the UCSC hg19 and HPV16 gene model (returning NCBI RefSeq IDs). A read was counted if it had the same orientation (strandedness) and at least 20 bases of a read were mapped to the exon. The output generated is a text file containing the raw read count for each transcript and transcript feature.

2.10.3.4 Differential gene expression

The gene expression text files were read in to R and a matrix created of transcript read counts for all samples. Differentially expressed genes were identified using an edgeR analysis on normalised count data (Robinson et al. 2010). The edgeR analysis comprises three distinct stages. A normalisation, estimate of common dispersion, followed by a test for a difference in means between two groups of data. The resulting p-values were corrected for multiple testing and false discovery issues using the False Discovery Rate (FDR) method (Benjamini and Hochberg 1995). The results of this analysis were represented using heatmaps. Data was log2 transformed and median centred before hierarchical clustering (using average linkage and Pearson's correlation as the similarity metric) to group similar data together and reveal any patterns within the dataset. The resulting values were then converted into colours, where green is under expressed (compared to median), black is around the median expression and red is over expressed.

2.10.3.5 Gene ontology over-representation analysis

Over-representation analysis against GO classifications was applied to investigate biological themes within the resultant lists of differentially expressed genes. Analysis was performed in R using the Goseq library against GO categories of biological process, molecular function and cellular components (Young et al. 2010). The results were sorted by the p-value of a hypergeometric based test yielding categories where more genes have been classified with that category than would be expected by chance. The resultant data was corrected for multiple testing and false discovery using the FDR method.

2.10.3.6 Human-viral fusion transcript plots (Circos plots)

Circos plots were used to visually represent the location and abundance of human-viral mRNA fusion transcripts. Circos is a software package for visualizing genomic data in a circular layout (Krzywinski et al. 2009). A circos links track was generated from the BAM file for each sample by filtering the reads in a stepwise fashion. First duplicate reads were removed (using picard tools), the reads were filtered to remove good pairs (reads on the same region of the same chromosome) using samstools and finally filtering to only include reads where one read pair mapped to NC_001526. The generated tracks file was plotted onto a combined hg19 and HPV16 ideogram.

2.10.3.7 Integrative genomics viewer

Levels of viral gene expression, mismatches with the reference sequence and splicing patterns were visualised using the Integrative Genomics Viewer (IGV), available from the Broad Institute (Robinson et al. 2011).

2.10.3.8 Genview

Histograms showing expression levels for individual genes were produced using the "Genview" software designed by Dr Peter Giles of the Wales Gene Park.

Chapter 3 - HPV prevalence in non-malignant tonsils

3.1 Introduction

The incidence of oropharyngeal cancers has sharply increased over recent years (Schache et al. 2016; Mourad et al. 2017). This has been associated with a higher proportion of HPV-related diseases originating in the tonsils and base of tongue. However, the natural history and pathobiology of HPV infection in the oropharynx is still very poorly described. The Cardiff HPV Research Group therefore undertook a study to investigate the biology of HPV in non-malignant tonsils. The primary objective was to determine the prevalence of HPV infection in non-malignant tonsil tissue in both sexes across all age groups. The secondary objectives were to determine the contribution of different routes of transmission, confirm the site of HPV infection within the tonsil, establish whether productive infections (i.e. resulting in the generation of infectious virus particles) are present in the tonsils, and explore whether HPV infection in the tonsil is associated with neoplasia.

The results were intended to inform cost-benefit analyses of the potential of HPV vaccination against specific HPV types and support decisions regarding the appropriate age at which to vaccinate. They were also envisaged as having potential to provide baseline data against which the effects of HPV vaccination on tonsillar HPV infection could be measured.

3.1.1 Background

In a study of oral rinse samples from 5579 people in the USA, HPV DNA was identified in 10.1% of men and 3.6% of women (Gillison et al. 2012). However, it was important to investigate HPV infection in actual tonsil tissue, as the assessment of HPV prevalence in the tonsils using oral rinse or exfoliated cell samples is unreliable. Oral rinse samples are very poorly representative of tonsillar tissue, as shown by the weak correlation between HPV positivity in OPSCC tumour tissue and oral rinse samples taken from the same patients prior to surgery (Schwartz et al. 1998; D'Souza et al. 2007). This is likely due to the site of infection/neoplasia being located within the reticulated crypts, rather than on the surface epithelium of the palatine tonsils (Begum et al. 2005). This emphasises that HPV infection in the tonsils can only be adequately assessed using biopsy material including both surface and crypt epithelia.

Unlike the stratified epithelium in the cervix, the reticulated tonsil crypt epithelium might not allow a complete HPV life cycle, associated with the production of virions and the slow development of pre-malignant lesions. Few studies have investigated HPV prevalence in non-malignant tonsils and whether pre-malignant lesions occur and can be detected is currently unknown.

3.1.2 Study design and samples

Tonsil tissue was obtained from the National Anonymous Tonsil Archive. This archive was originally established to investigate the prevalence of disease-related prion proteins in the British population (Clewley et al. 2009). Paired palatine tonsils were obtained, with consent, from children and adults of all ages undergoing elective tonsillectomies. During the 2004-2008 period, over 100000 tonsils pairs were collected from 138 hospitals throughout England and Scotland. One tonsil from each pair was stored in formalin then embedded in paraffin. The other was stored as fresh tissue at -4 °C then homogenised and kept at -80 °C. Cases were selected from the archive, stratified for age, sex and geographical distribution. A total of 4095 FFPE samples were equally distributed in age groups ranging from 0 to over 50 years old.

In the initial phase of the study, completed before the start of this PhD, 3377 FFPE non-malignant tonsil samples had been analysed for the presence of HPV DNA using two PCR-based techniques: HPV16 E6 PCR and GP5+/6+ EIA. No HPV was detected in any of those samples. This result was unexpected and raised concerns regarding the adequacy of the sampling method, and the sensitivity of the assays used. The DNA for the HPV detection assays was extracted from thin sections, cut using a microtome, from the centre of the resected tonsils. It was possible that this method may have missed small foci of infection. The main assay used in the initial phase was the GP5+/6+ EIA, the sensitivity of which has been aligned with clinically significant infection in cervical cytology samples (i.e. containing a relatively high viral load). It was possible that this was insufficient to detect very low levels of clinically insignificant infection. These concerns were addressed in the second phase of the project, which was conducted as part of this PhD.

In Phase 2, a second panel of samples was selected from whole homogenised tonsils, and HPV tested for both high and low risk HPV types using 3 different techniques, including the

most sensitive commercially available assay. The results of this second part of the study are presented in this chapter and discussed in the context of the overall study.

3.1.3 Aim and hypothesis

The aim was to determine the prevalence of HPV infection in non-malignant tonsil tissue from men likely to harbour infection. The hypothesis tested was that low-level high risk HPV infection occurs and can be detected in non-malignant tonsils.

3.2 Materials and methods

3.2.1 Clinical samples

The purpose of the second part of the study was to ensure that the presence of small foci of HPV infection was not missed in FFPE samples. To account for this possibility, the analysis was centred on a subset of 524 tonsil homogenates selected from the original 4095 cases (fixed and homogenised tonsils were available for most patients within the tonsil archive). Two age groups were chosen to represent patients most likely to harbour HPV infection: men aged 25-34 years (n=335) and >44 years (n=189). Based on oral rinse samples, HPV prevalence is higher in these groups (Gillison et al. 2012).

3.2.2 HPV testing

DNA was extracted from the homogenised tonsil samples using a phenol-chloroform based method. The suitability of the extracted DNA for PCR-based analysis was assessed by amplification of a 209 bp fragment of the human β-globin gene (this was longer than the amplicons generated by the HPV tests). The presence of HPV DNA was assessed using three different HPV assays, each with distinct performance characteristics (listed below). Two of the assays (HPV16 E6 PCR and GP5+/6+ PCR-EIA) had previously been used on the original set of samples.

HPV16 E6 PCR: This assay amplified the E6 region which is maintained in both productive and oncogenic infections. Only HPV16, the most common HPV type causing squamous cell carcinoma in the tonsil was targeted. HPV16 E6 PCR was rapid, cost effective and enabled a first evaluation of the prevalence of the most common HPV type in OPSCC. The

other two assays depend on amplification of the HPV L1 region, which can potentially be lost as a result of viral integration, resulting in false negative data. HPV16 E6 PCR was used as an end-point assay. A result was called positive if a 161 bp band was observed on the gel as well as in the CasKi positive control, in conjunction with the absence of a band in the negative control.

GP5+/6+ EIA: A widely applied "in-house" method depending on amplification of the HPV L1 region and two separate EIA using a cocktail of probes detecting 14 high risk types and 6 low risk types. This allowed comparison of the data with previously generated data, and could potentially show whether positive results were due to changes to the sampling regime or greater assay sensitivity.

SPF10-DEIA: A commercially available assay that amplifies a small region (65 bp) of the HPV L1 gene. It is regarded as the most sensitive assay available and has been used in several large epidemiological studies (Tjalma et al. 2013). It also enabled the detection and typing of a wider range of HPV types (40 types).

3.3 Results

3.3.1 HPV typing

DNA extraction was performed on 524 cases. The β-globin fragment was successfully amplified in 511/524 cases (97.5%). Positivity was determined by the visual presence of a 209 bp amplicon following gel electrophoresis (Figure 3.1). Only β-globin positive samples were taken into account in further assays. HPV DNA was not identified in any samples using any of the assays (Table 3.1). All positive and negative controls gave the expected results. Positivity of samples was assessed by gel electrophoresis for HPV16 E6 PCR (Figure 3.2) and absorbance readings for GP5+/6+ and SPF10-DEIA as illustrated for SPF10-DEIA in Figure 3.3.

Table 3.1: Sample adequacy (β-globin) and HPV testing stratified by age

Age Group (years)	Samples (n)	β-globin	HPV detection Assay		
			HPV16 E6	GP5+/6+	SPF10-DEIA
24-29	167	162	0	0	0
30-34	168	162	0	0	0
45-49	114	113	0	0	0
50+	75	74	0	0	0
Total	524	511	0	0	0

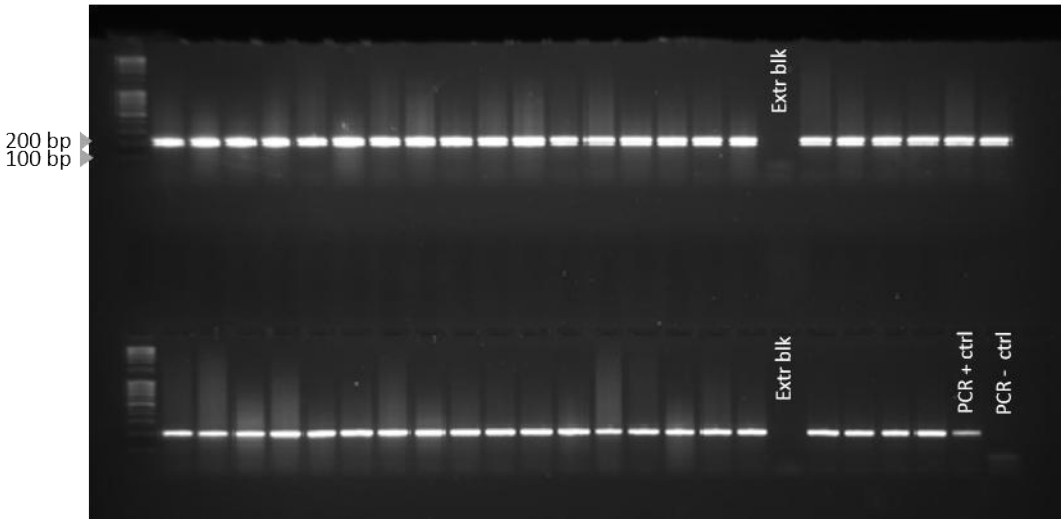


Figure 3.1: Example of β-globin PCR gel image

DNA extracted from all tonsil homogenates samples and blanks was assessed for suitability of PCR-based assays with a β-globin PCR. Extraction blanks (Extr blk) were included every 24 samples. PCR positive (CasKi) and negative (master mix only) controls were included in each PCR round. These confirmed the lack of technical issues and contamination during the PCR process. Samples for which a band of 209 bp was visible by gel electrophoresis were taken into account for HPV typing.

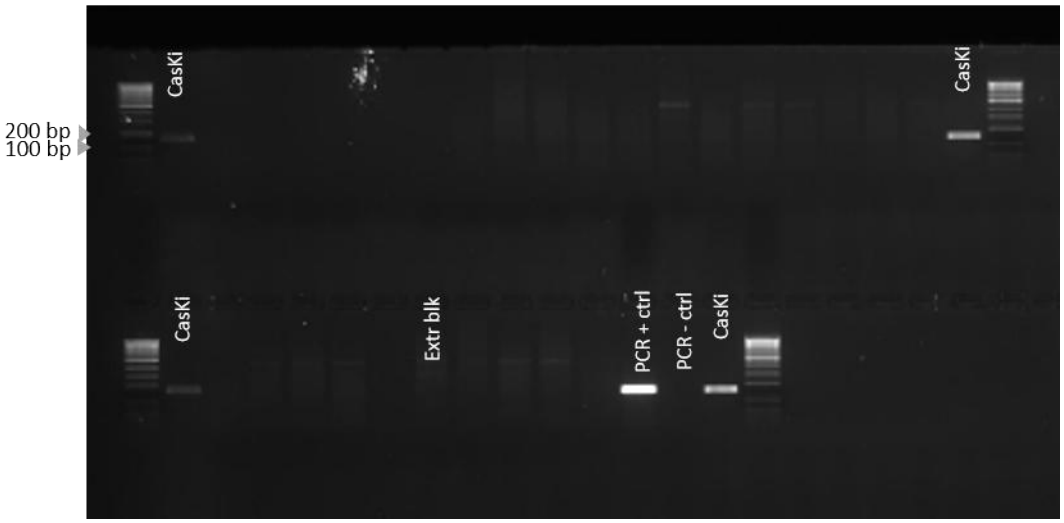


Figure 3.2: Example of E6 PCR gel image

A previously amplified fragment of the HPV16-positive CasKi cell line was used as a size marker (161 bp) next to the DNA ladder. In each PCR rounds internal PCR positive (CasKi) and negative (master mix only) controls confirmed the absence of technical issues and contamination. No band of 161 bp was visible in any of the samples. All the samples were HPV16 negative.

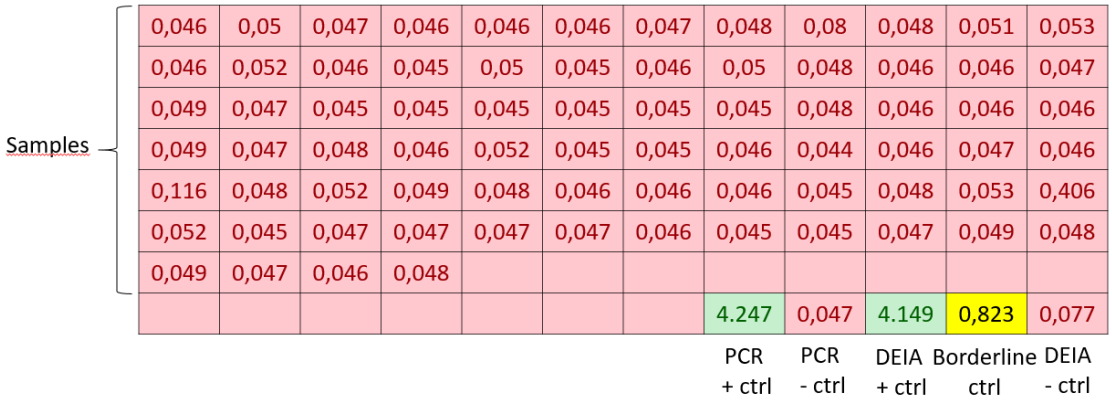


Figure 3.3: Example of SPF10-DEIA data: absorbance at 450 nm

Results from each plate were validated by the controls according to the manufacturer's protocol and only interpreted if all validation steps were passed. HPV-positive samples were highlighted in green and HPV-negative in red.

3.4 Discussion

No high risk or low risk HPV DNA was detected in any of the 511 non-malignant tonsil homogenates samples from men in the age groups most likely to harbour HPV infection. Amplification of the housekeeping gene β-globin in 97.5% of samples and inclusion of PCR controls showed that negative results were not due to inhibition during the PCR or low-quality DNA. This data appeared to validate and support the results of phase one of the study. Together, these findings suggested that HPV infection in non-malignant tonsil is a very rare event. The greatest possible population prevalence for HPV, consistent with the absence of HPV in the 3377 English and Scottish males and females included in the study, is 0.089% (95% CI 0–0.089%).

During the first part of the study, no HPV DNA was detected in 3377 FFPE samples. The material from which DNA was extracted comprised two 10 µM cuts through the middle of the tonsil (this method was chosen as it potentially would allow subsequent histopathological investigations in positive samples). A possible explanation for the absence of HPV was that small foci of infection in the tonsil could have been missed. The tonsil homogenates were used to address this issue as they contained the whole tonsil. The sample panel was modified to concentrate on groups with the highest susceptibility to HPV infection: men aged 25-34 and over 44 years old. These groups were designed based on published HPV prevalence data. A study using oral rinse samples reported an overall higher prevalence in men and a bimodal distribution with regard to age. The prevalence peaked in the 30-34 and 60-64 age groups (Gillison et al. 2012). In England in 2010 (year of study design), tonsil SCC were most common in men aged 45-75 years old (Office for National Statistics 2015).

The possible lack of detection sensitivity was also addressed. The sensitive nature of GP5+/6+ EIA and HPV16 E6 PCR had previously been assessed in the laboratory through the 2011 LabNet WHO study (Eklund et al. 2012). Its sensitivity was determined to be the equivalent of 500 HPV genome copies per PCR. The detection sensitivity of the HPV16 E6 PCR was assessed using dilutions of CasKi cells known to contain around 600 copies of HPV genome. It was estimated to be 500 genome copies per PCR. These two tests provided a robust detection of HPV16, most commonly associated with tonsil SCC, and detection of another 13 high risk and 6 low risk types. To increase the detection sensitivity, a third test was selected to use on the tonsil homogenates samples. SPF10-DEIA enabled the detection of 2

HPV genome copies per PCR and a wider range of HPV types. The combination of whole tonsils, high sensitivity assays and samples from men most likely to harbour HPV16 infection strengthened the results.

Variable HPV prevalence rates have previously been reported in tonsils. Several small studies focused on paediatric populations in which tonsil samples are more readily available. A study conducted in Finland on 106 samples from children aged 2-14 years old with tonsillar or adenoid hyperplasia found a prevalence of 8.5% not correlated with sex. HPV16 accounted for 66.7% of these (Mammas et al. 2006). Other studies in Brazil (n=100) and the USA (n=50) did not report any cases of high risk HPV infection, the only positive case identified was an HPV11 infection in the USA study (Moreira et al. 2006; Sisk et al. 2006). Small sample size and variation in detection methods, as well as geographical differences, could explain the range of prevalence detected. The origin of HPV infection in children may differ from adults, vertical and horizontal HPV transmission between mother and child have been described but might not be the single mode of HPV acquisition in children (Castellsagué et al. 2009). The consequences of HPV infection in paediatric populations in relation to the rise of tonsil SCC or the HPV prevalence in adults are currently unknown.

Only a small number of studies have reported HPV prevalence from adult tonsil samples. A 6.3% HPV16 prevalence was observed across 206 samples from people ranging from 1.5-72 years old in Finland. The highest incidence occurred in the 1.5-6 and 17-25 year-old groups (Chen et al. 2005). However, others reported findings similar of those presented in this chapter, i.e. very rare occurrence of HPV infection in non-malignant tonsils. No HPV16 or 18 were detected in a retrospective study of 212 FFPE samples from patients aged 21 years or older in the USA (Ernster et al. 2009). More recently, the results of the SPLIT study were published. In this study, samples were obtained from 200 patients aged 18 or over undergoing tonsillectomies for benign disease. Cells were collected from deep brushing of the tonsil and used for DNA extraction and liquid based cytology for each sample. Two women were HPV16-positive, one had no cytological lesion and the other could not be interpreted. Both of these studies are in concordance with the results presented in this chapter, indicating that HPV infections are a rare occurrence in non-malignant tonsils in the adult population (Franceschi et al. 2015).

HPV prevalence at genital sites varies between countries (De Vuyst et al. 2009; Hariri et al. 2011). Similar variations could occur in the tonsils, therefore, results from one population might not be relatable to another. The estimated fraction of OPSCC attributed to HPV varies across the world and within Europe (Castellsagué et al. 2016). Samples used in this study were collected between 2004 and 2008 in England and Scotland. During this period the age-standardised rate of tonsillar squamous cell carcinoma per 100000 people in England increased from 2.6 to 3.4 according to the Cancer registration statistics (Office for National Statistics 2015). The results are likely to be a good representation of the British population but might not apply to other geographical areas.

As opposed to the low prevalence reported in tonsils, HPV is more frequently detected in the oral cavity of cancer-free subjects through oral rinse or gargle samples (Kreimer et al. 2010; Gillison et al. 2012). Epithelial cells contained in these samples originate from the oral cavity. Oral HPV infection is common at diagnosis in patients with OPSCC but not always present (Rettig et al. 2015).

3.4.1 Strengths and Weaknesses

The power of this study lies in the high number of samples tested: 3377 in phase one, and the additional 511 presented in this chapter. So far, this is the largest sample cohort published (Palmer et al. 2014), in Appendix 3. The results are also strengthened by the use of multiple assays targeting several regions of the HPV genome, and by the high sensitivity of the SPF10 PCR DEIA. The use of whole tonsils presented some limitations even though it ensured that any potentially HPV-infected cells in the tonsil were taken into account. The high amount of lymphoid cells mixed with the squamous epithelial cells could render the detection of small foci of HPV infection harder.

3.4.2 Conclusion

The findings described in this chapter highlight the differences between the natural history of HPV-driven cervical and tonsillar SCC. In the cervix, high risk HPV infections are common and easily detectable, with a small proportion resulting in SCC. Furthermore, virus-induced neoplasias are well described and can be observed for years before leading to neoplasia in some cases. In the tonsils, HPV infection appears to be a very rare event despite

the possible exposure from virus present in the oral cavity. This suggests that when HPV does infect tonsil tissue, this could often result in tumour development.

These data have important implications for the prevention of OPSCC. Screening for and treatment of HPV-associated pre-cancers has been highly effective in reducing the incidence of cervical SCC. Screening for OPSCC or its precursors has been considered (Gillison et al. 2013) However, testing requires the identification of a treatable HPV-associated premalignant neoplasia, and this was not observed in the current study. Therefore, it appears that prophylactic vaccination, potentially including males, may be the most promising option for prevention of OPSCC.

Chapter 4 - Primary culture of oropharyngeal cells

4.1 Introduction

The study of head and neck cancer biology, and of HPV-positive disease in particular, is hampered by the lack of *in vitro* models. Relatively few HPV-positive head and neck cancer cell lines have been derived and described in the published literature. Five were described in publications in 2012, however only one of these originated from the oropharynx (Steenbergen et al. 1995; Balló et al. 1999; Lansford et al. 1999; Ferris et al. 2005; Tang et al. 2012). These lines have characteristics commonly found in HPV-positive tumours, such as wild-type p53 gene and p16 positivity, indicating that the disease could be HPV-driven (although one line, 93VU147T, differs from the others with described genetic alterations and mutated p53) (Steenbergen 1995). All 5 lines originate from patients who were current or ex-smokers at the time of diagnosis. Smoking has been identified as an adverse prognostic factor in both HPV-positive and HPV-negative tumours (Maxwell et al. 2010). The increased sensitivity to treatment and survival often observed in HPV-positive patients tends to be seen in non-smokers, whereas survival of HPV-positive smokers is reduced. Another group which encompasses HPV-positive smokers with positive lymph nodes and HPV-negative non-smokers with low disease stage has been described (Ang et al. 2010). It shows intermediate survival between “pure” HPV-positive and negative cases. Hence, the available cell lines may not provide the best possible representation of typical HPV-positive non-smoking patients with OPSCC. New cell lines, from typical HPV-positive patients, are needed to increase the understanding of treatment response of both HPV-positive smokers and non-smokers.

The aims of the work described in this chapter were to derive novel cell lines from oropharyngeal tumour biopsies, and to validate and characterise the successfully grown lines as models for HPV-positive oropharyngeal cancer. Lines were phenotypically described, including their morphology and growth characteristics. The cell lines were validated by confirming the type of cell cultured and by comparing the STR profiles of cell lines with the original biopsies.

4.2 Terminology

Biopsies were processed using an explant method. Each biopsy was cut into small tissue pieces and placed in culture. The term explant was only applied to biopsy pieces producing new cell growth. Throughout this chapter, the term primary culture designates developed cells up until the first passage. Once they have been passaged multiple times, they are called short-term cultures. Once cells sustained growth to a higher passage, and more importantly a greater number of population doublings than what would be expected from "normal" tonsil epithelium cells, a cell line can be deemed immortal. Such cells are likely to have overcome telomere mediated senescence and programmed cell death. This suggests that they are likely to derive from cancer cells present in the biopsy.

Because the study was named Primary Culture of Oropharyngeal Cells (PCOC), biopsies collected, and any cell line derived from a biopsy, were called PCOC#, with # corresponding to the chronological order of biopsy collection.

4.3 Protocol development and sample collection

4.3.1 Protocol development

The explant protocol had previously been used by other members of the group for the culture of gynaecological biopsies. It was first followed without modifications (Stanley 2002). However, in initial experiments, culture plates containing pieces of the biopsy were found to contain yeast on day 2 and discarded. Yeast infections (e.g. *Candida albicans*), can be common in the oral cavity and oropharynx. To address this, an antifungal agent (amphotericin B), was added to the initial washing step. Amphotericin B was also added in the media of the PCOC2 early cultures. However, the antifungal had an adverse effect on the 3T3 feeder cells. The length of time they remained attached to the plates was shortened. Hence, amphotericin B treatment was stopped after 6 days. Thereafter amphotericin B was only used in the washing step and as a short treatment if yeast contamination occurred.

4.3.2 Study samples

The PCOC study was sponsored by Cardiff University and approved by the Cardiff and Vale University Health Board R&D Office. Ethical approval was obtained from the Research

Ethics Committee for Wales and the study was adopted into the NISCHR portfolio. Patients with a suspected oropharyngeal primary tumour were potentially eligible for inclusion (see study protocol in Appendix 1 for full inclusion criteria). Patients undergoing treatment at the Cardiff and Vale NHS Trust were identified by their head and neck surgical consultants or consultant oncologists in the course of their clinical practice. Biopsies could be obtained at two points in the patient pathway. This was important as patients are treated differently according to the size and stage of the tumour. Biopsies could be collected during panendoscopy (a biopsy taken under general anaesthetic for histopathological analysis) or after surgical resection (surplus material was collected from the pathologist). In both cases, biopsies were stored in transport media to minimise tissue degradation. Twelve patients were consented between 27 Feb 2013 and 12 May 2014, and a total of 13 biopsies were collected (for the 4th patient recruited into the study, PCOC4, samples were received on two occasions: at the time of the diagnostic biopsy and during tumour resection, both were recorded as PCOC4). PCOC9 was obtained from a 73-year-old man who was later diagnosed with a lymphoma. As a result, PCOC9 was not included in further descriptions of study samples.

Approximately one-third of the patients were women, and the mean age was 52.66 years. Biopsies were taken from the tongue base (4) or the tonsil (6) prior to treatment. Characteristics of each biopsy regarding sex, age, smoking status, previous tumour history, site of biopsy, TNM stage, p16 status, treatment and outcome are described in Table 4.1.

Table 4.1: Description of the PCOC study sample

Study number	Sex (M/F)	Age (years)	Smoking status*	Site	TNM Stage	p16 IHC	Treatm
PCOC1	M	63	Current, 30 pack/yrs	Tongue base	T4aN2bM0	Not available	Right neck dissec chemoRT to
PCOC2	M	52	Never smoker	Left tonsil	T4aN2bM0	Positive	Chem
PCOC3	M	44	Ex-smoker, <10pack/yr	Right tonsil	T1N2aM0	Positive	Transoral las oropharynx and post-c
PCOC4	M	50	Never smoker	Left tonsil	T2N2bM0	Positive	Left neck dissec
PCOC5	F	62	Not available	Tongue base	T2N2bM0	Positive	Neck dissectio chem
PCOC6	F	59	Ex-smoker, 15pack/yr	Right tonsil	T2N2bM0	Positive	Neck dissectio chem
PCOC7	F	56	Ex-smoker, <10pack/yr	Tongue base	T2N2bM0	Positive	Neck dissectio chem
PCOC8	M	71	Never smoker	Right tonsil	T2N2bM0	Positive	Chem
PCOC10	M	41	Ex-smoker, 20pack/yr	Tongue base	T2N2bM0	Positive	Chem
PCOC11	M	50	Never smoker	Right tonsil	T3N2bM0	Positive	Neck dissectio chem
PCOC12	F	53	Current <10pack/yr	Right tonsil	T4aN2bM0	Positive	Chem

* Pack-years refers to the number of cigarette packs smoked per day multiplied by the number of years the perso
** Radiotherapy (RT), post-operative (post-op).

The biopsies varied in size and were composed of both tumour and parts of the surrounding tissue. This resulted in the growth of a variety of cells, this was apparent following 48 h of culture when the media was replaced. During the first 2-4 days, some unidentified non-adherent cells and red blood cells were observed in vascularised biopsies. Both were subsequently discarded with media changes. Explants were visible after 2-7 days of culture. Biopsy pieces that did not produce outgrowths 7 days after plating were regarded as unsuccessful. Some tissue fragments did not attach to the plates, additionally low adhesion of some explants resulted in loss/detachment following the first media change at day 2. Afterwards, all explants remained attached. Of the 13 patient biopsies, 8 produced visible new cell growth after 7 or fewer days (Table 4.2). Two of these, PCOC2 and PCOC3, were successfully grown over multiple passages.

Extended cell growth from explants of the other biopsies was unsuccessful for several reasons. PCOC1 cultures were discarded after two days due to fungal contamination. A yeast infection in PCOC6 developed amphotericin B resistance after 3 passages, resulting in the loss of cultures. A widespread bacterial contamination occurred when several biopsies (PCOC7-11) were in early stages of culture. The contamination spread to all biopsies and cell lines present in the incubator. The bacterial spread seemed independent of the operator and reagents used. After initial antibiotic treatment with gentamicin failed, the bacteria was typed to tailor the treatment to its sensitivity. It was identified as a *Stenotrophomonas* using MALDI-TOF. *Stenotrophomonas* are ubiquitous environmental bacteria that are naturally resistant to many broad-spectrum antibiotics, but which are often sensitive to co-trimoxazole. Remaining cultures were treated with co-trimoxazole. The concentration required to kill the bacteria was toxic for the cells and the cultures of PCOC7 passage (p) 1, PCOC8 p5, PCOC9 p1, PCOC10 p1 and PCOC11 p1 were lost. The origin of the contamination was not identified and resulted in the closure of cell culture for a month. Explants from the PCOC12 biopsy produced cell outgrowths which were successfully passaged twice. Cell growth slowed down after the second passage, and a change of morphology was observed (cells became bigger and heterogeneous in shape). Cultures were passaged a third time, but this was followed by senescence in all cells (Table 4.2).

Table 4.2: Growth status of PCOC biopsies

Study number	Explant growth	Outcome
PCOC1	Yes	Discarded Yeast contamination
PCOC2	Yes	Established line
PCOC3	Yes	Established line
PCOC4	No	No explants growth
PCOC5	No	No explants growth
PCOC6	Yes	Discarded Resistant yeast contamination
PCOC7	Yes	Discarded Stenotrophomonas contamination
PCOC8	Yes	Discarded (passage 5) Stenotrophomonas contamination
PCOC9	Yes	Discarded (passage 1) Stenotrophomonas contamination
PCOC10	Yes	Discarded (no passage) Stenotrophomonas contamination
PCOC11	Yes	Discarded (no passage) Stenotrophomonas contamination
PCOC12	Yes	Senescence (passage 3)

4.4 Growth and characterisation of the PCOC2 and PCOC3 cell lines

4.4.1 Histology

Sections of the original tumours from which the PCOC2 and PCOC3 cell lines were derived, were obtained to assess cellular morphology, and the presence of specific proliferation, cell cycle and differentiation markers. These were then stained with Haematoxylin and Eosin (HE) to observe general histology, Ki67 as a proliferation marker and p16 as a marker of HPV infection. Staining was performed by UCL-Advanced Diagnostics and University Hospital of Wales pathology department, the pathological description was kindly provided by Dr Adam Christian (UHW pathologist).

The PCOC2 tumour was initially described as a moderately differentiated squamous cell carcinoma. A mix of severe and lower grade dysplasia was observed (Figure 4.1 A and B). The severely dysplastic epithelia was characterised by the presence of mitotic figures in all layers, in addition to pleomorphic cells and a loss of maturation (Figure 4.1 C). Consistent with the features observed with HE staining, Ki67 staining was present in all layers of the severely dysplastic epithelia. p16 staining was strong and diffuse confirming likely HPV infection in the affected area (Figure 4.1 D and E).

PCOC3 was described as a moderately differentiated squamous cell carcinoma with a complex papillary architecture. A section through the resected tonsil showed a complex papillary architecture with evident severe dysplasia. This was illustrated by the presence of pleomorphic nuclei (i.e. nuclei showed diverse size and shapes) and mitotic figures in all layers of the epithelia as well as a loss of maturation in the top layers (Figure 4.2 A and C). A pocket of cells was visible invading into lymphoid tissue (Figure 4.2 B). The presence of neoplasia and disordered architecture was confirmed by the presence of strong Ki67 staining in all layers of the epithelia both in the dysplastic and invasive tissue (Figure 4.2 D). Strong and diffuse p16 staining showed that both the dysplastic and tumoral area were likely to be infected with HPV (Figure 4.2 E).

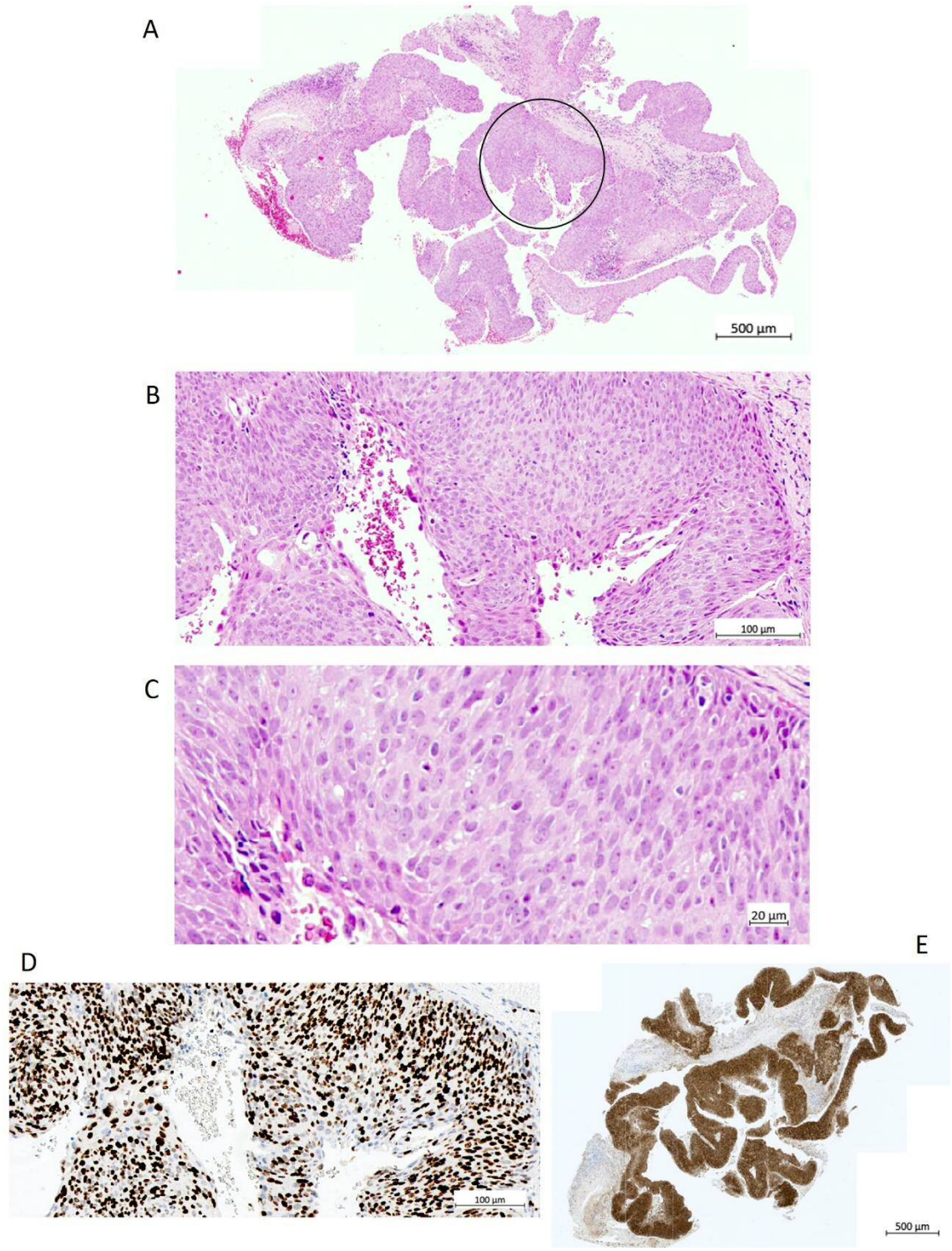


Figure 4.1: PCOC2 HE, Ki67 and p16 IHC staining

A: Low magnification view of a PCOC2 FFPE block stained with HE.
B: Higher magnification image of the zone circled in black in figure A: the epithelia shows signs of severe dysplasia.
C: Higher magnification image of severely dysplasic epithelia: mitotic figures in all layers, no maturation , disorganised tissue.
D: Ki67 staining of dysplasic epithelia showing proliferation throughout the epithelia.
E: p16 staining: Cells strongly stained for p16.

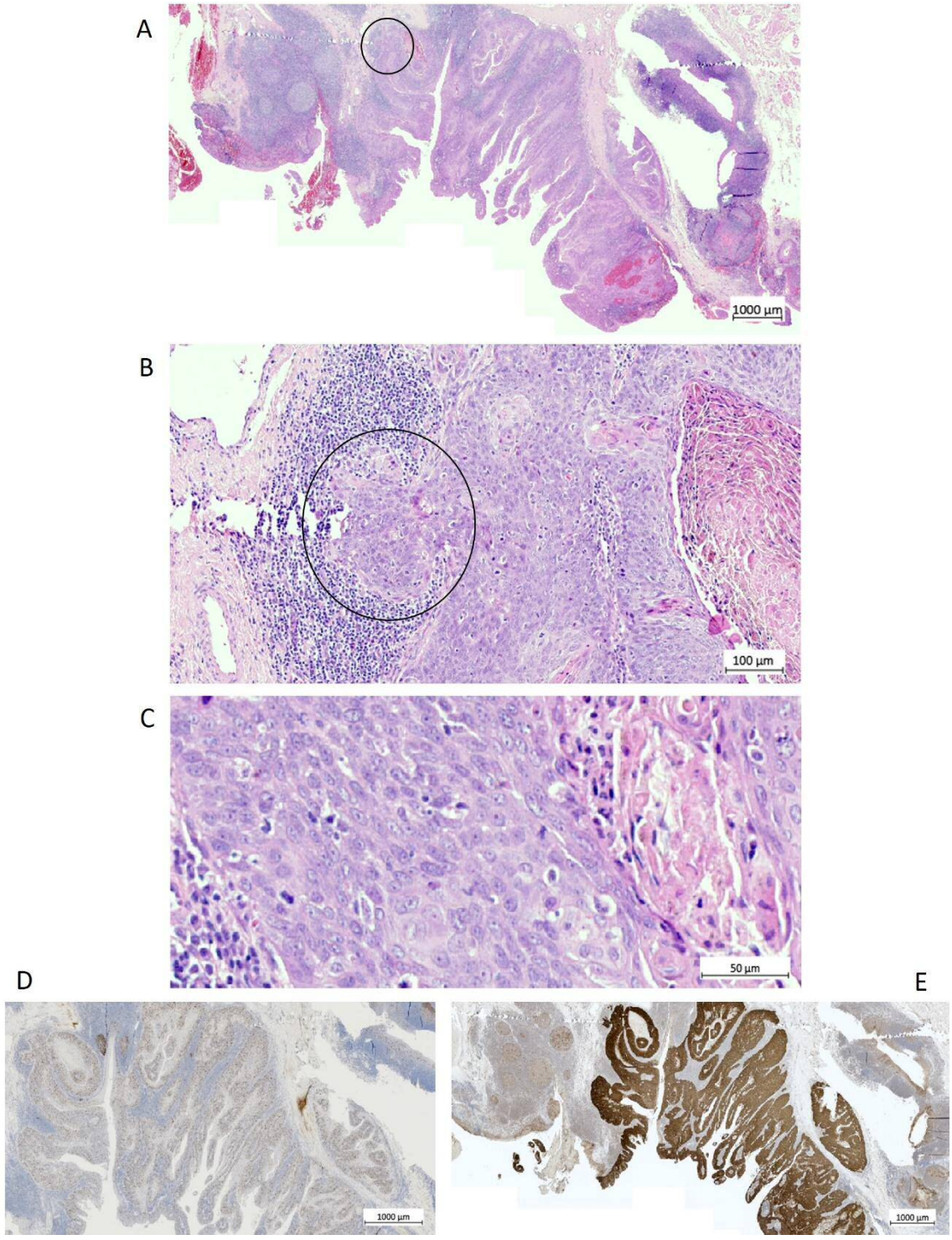


Figure 4.2: PCOC3 HE, Ki67 and p16 IHC staining

A: Low magnification view of a PCOC3 FFPE block stained with HE.

B: Higher magnification image of the zone circled in black in figure A: signs of invasion of epithelial cells in the underlying lymphocytes.

C: Higher magnification of dysplasic epithelia: mitotic figures in all layers, no maturation, disorganised tissue.

D: Ki67 staining: proliferation throughout the epithelia.

E: p16 staining: Cells strongly stained for p16.

4.4.2 Early stages of culture

Cells grown from PCOC2 and PCOC3 explants were successfully cultured for around 20 passages (24 and 19 respectively). Growth pattern, cell and colony morphology were assessed in early passages of cultures. The phenotypic characterisation provided informations related to the cell types present and the health of cells derived from explants. Cells were also checked for signs of apoptosis and senescence (e.g. blebbing or enlargement).

Tissue from the tumours of both patients was subjected to the explant culture procedure. In both cases, non-adherent cells were present at day 2 but were easily removed with a washing step prior to the media change. Cell outgrowths were visible from explanted tissue at day 4 and 6 respectively. Cells grew around the explanted tissue resulting in a “fried egg” colony shape, with no space visible between the cells (Figure 4.3). Colonies expanded as a result of the division of the peripheral cells, displacing the lowly adherent 3T3 feeder cells layer (Figure 4.4). The growth rate differed between PCOC2 and 3, with PCOC2 cells growing faster from the beginning. As a result, PCOC3 cultures developed over a longer period.

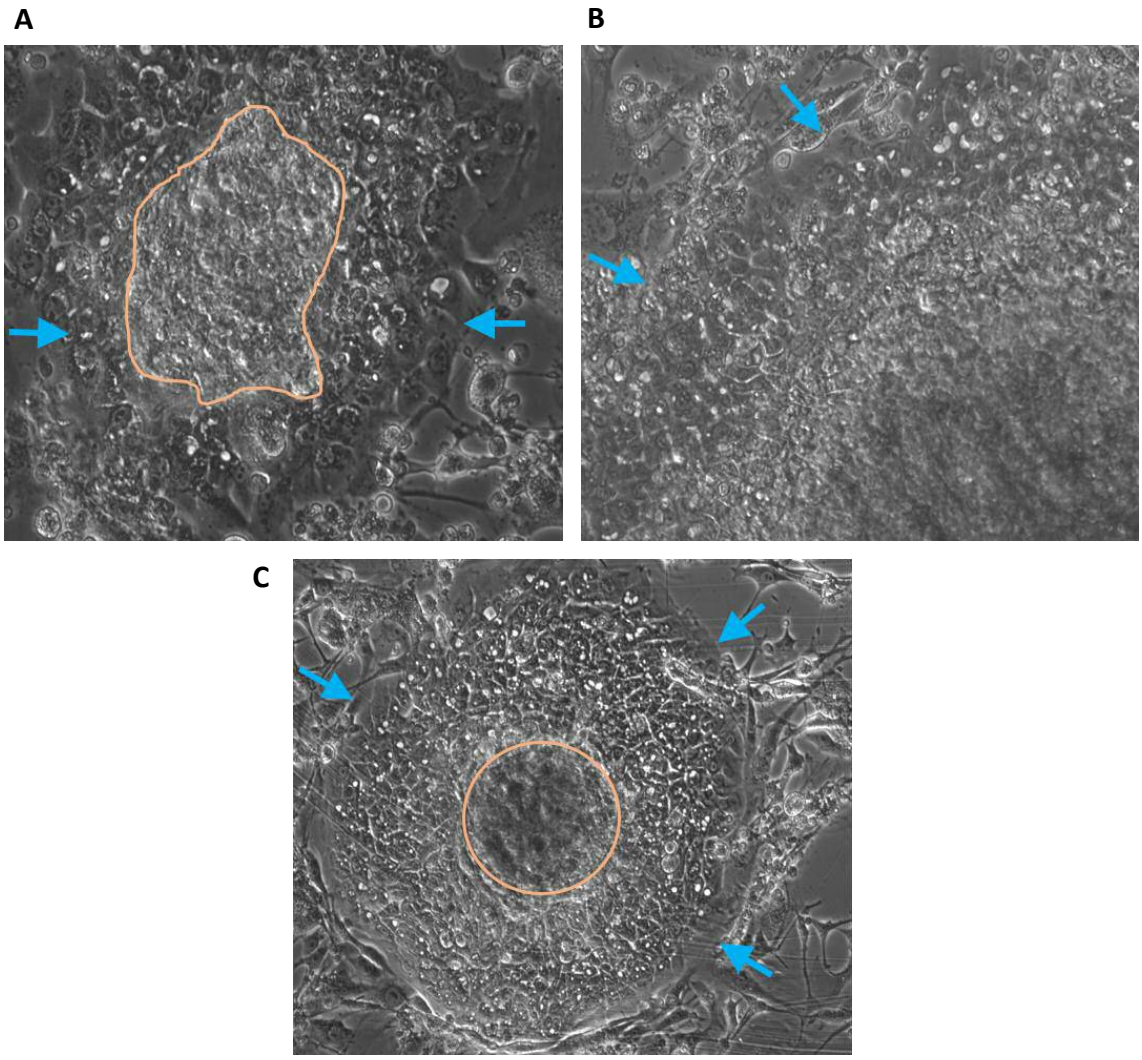


Figure 4.3: Morphology of PCOC2 and PCOC3 explants

A: PCOC2 explant growth. Explants (blue arrow) grew in a roughly circular pattern around the parental biopsy piece (circled in orange) (100X magnification).

B: PCOC2 cells of polygonal shape ressembling keratinocytes growing tightly (200X magnification).

C: PCOC3 cells, growing from the parental biopsy piece as polygonal cells forming tight colonies (100X magnification).

After 7 days, the cell outgrowths from PCOC2 had reached sufficient size to be passaged. Cells were trypsinised and plated as single cells, this provided growth space to the cells previously located in the centre closest to the tissue pieces (and hence avoiding cell cycle exit into G0). As the optimum length of trypsinisation was unknown, 5 minutes increments were used, removing the trypsin and collecting cells each time. New cells were polygonal, similar to keratinocytes, and grew as tight round colonies. Thereafter cells were split when 80% confluence was reached. In early stages of culture, PCOC2 plates were contaminated with fibroblasts colonies. Fibroblasts originated from the connective tissue surrounding the tonsil and were easily distinguished from the keratinocytes by their elongated shape. The fibroblast growth was reduced by maintaining a high density of 3T3 feeder cells. Differential trypsinisation was also used to remove the fibroblasts. Due to their low adhesion to the plate, fibroblasts detached sooner than keratinocytes following a short trypsinisation 1-2 min (Figure 4.5).

A yeast contamination was detected in PCOC3 plates within the first week. Cultures were treated with amphotericin B and the contamination cleared after 2 weeks. During the 15 days of treatment with amphotericin B, cells outgrowth slowly developed from the biopsy pieces. Amphotericin b reduced the time 3T3 feeder cells remained attached, and they were replaced every 2 days to maintain the cover. Cells were first passaged and re-plated as single cells after 13 days as part of a washing step to eliminate the contamination. PCOC3 cells were noticeably bigger than PCOC2 cells while also having polygonal shapes and growing tightly (Figure 4.4).

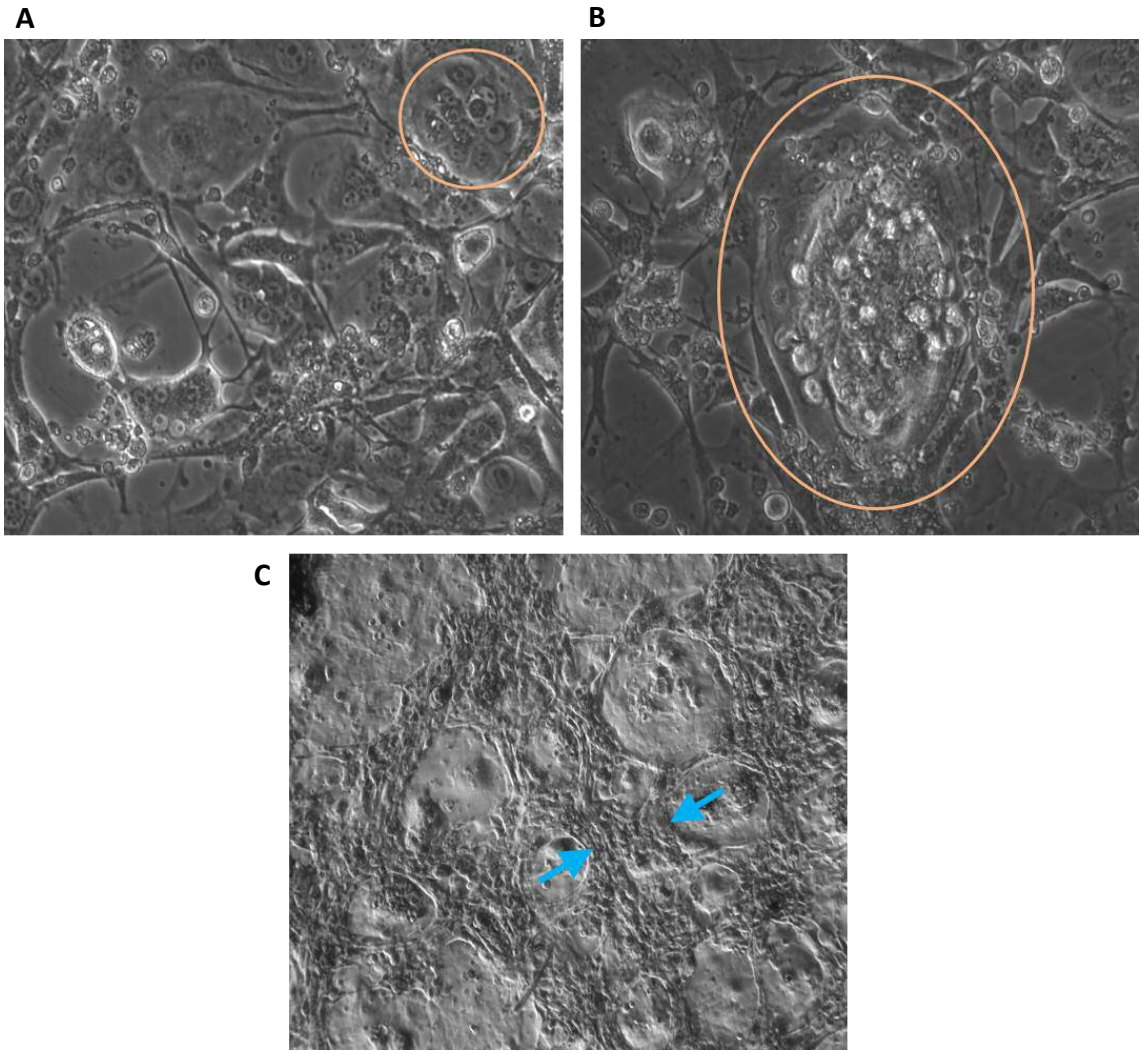


Figure 4.4: Morphology of PCOC2 and PCOC3 cells in culture and 3T3 feeder cells displacement

A**:** Small PCOC2 colony, formed of around 9 dividing cells. Polygonal cells grow without gaps between them hence forming tight colonies roughly circular in shape (200 x magnification).

B**:** PCOC3 colony, formed of two parts. The outer cells are polygonal and growing as a tight monolayer, the centre is composed of a multilayer of cells. Cells are rounding up at the top and detaching from the colony (200x magnification).

C**:** 3T3 feeder cells covering the plates being displaced by the growing colonies. As a result, 3T3 feeder cells are compacted between the growing PCOC colonies (blue arrows) (50x magnification).

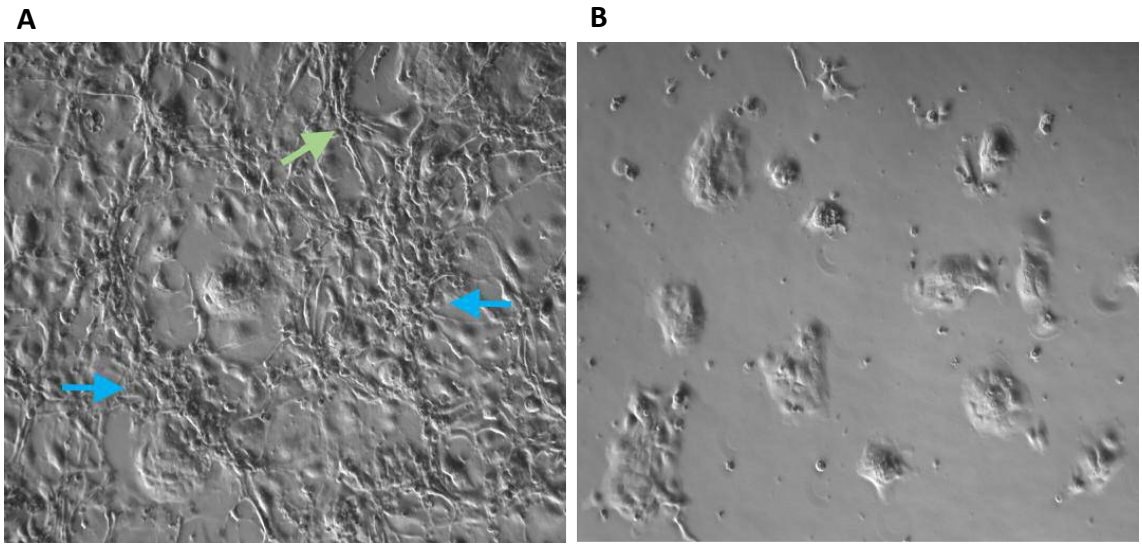


Figure 4.5: Human fibroblast contamination on PCOC2 p1 culture

A**:** 3T3 feeder cells which cover the plate (pointed by blue arrows) did not have a characteristic fibroblast morphology and did not divide in culture due to the 60 Gy radiation they received. A human fibroblast colony (pointed by a green arrow) is composed of dividing cells with a characteristic elongated shape.

B**:** Difference in adhesion to the dish between keratinocytes and fibroblasts was used to remove both 3T3 feeder cells and human fibroblasts with a short TE treatment. The image was taken immediately after a short TE treatment.

4.4.3 Heterogeneous colony morphology

Two types of colonies were observed in both PCOC2 and PCOC3. Some cells, deemed colony type 1, grew as a flat monolayer. These colonies had a tendency to be faster growing and cover larger surfaces. In the second type of colony (type 2), cells grew upwards, forming multi-layers. This resembled the growth pattern of a stratified epithelium, with cells detaching once they reached the top layer. These colonies were slower growing and occupied smaller surfaces. The proportion of the two types of colonies differed in PCOC2 and PCOC3 and varied over time.

At p1, the PCOC2 culture was composed equally of flat growing type 1 colonies and differentiating type 2. The proportion of type 1 colonies increased over time in PCOC2 cultures, and by p6 only flat growing colonies were present (Figure 4.6 A and B). PCOC3 cells grew as differentiating type 2 colonies only at p1. This resulted in slow colony development not reaching 80% confluence. Consequently, the cell line could not be frozen at early passages (due to a lack of cells). The colony morphology changed over time. By p3 all colonies were still differentiating but grew larger. These colonies were intermediate between type 1 and 2, composed of cells growing upward in the centre and laterally at the edges. By p7 most colonies grew flat and at p14 differentiating colonies were no longer present (Figure 4.6 C, D and E). The differences in colony morphology and growth speed greatly influenced the time needed to develop the lines: PCOC2 reached p7 after 2 ½ months, while PCOC3 took 5 months.

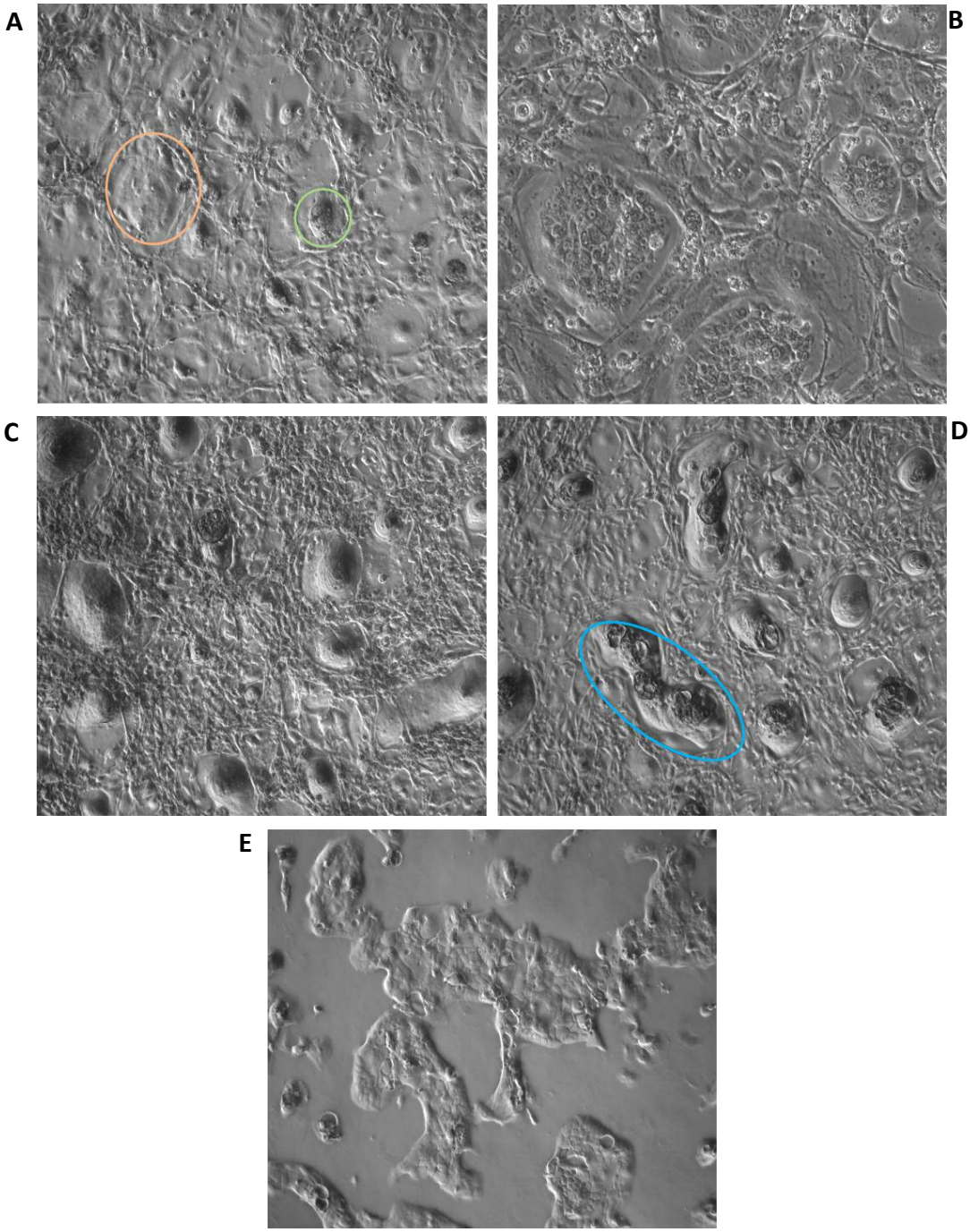


Figure 4.6: Variations in PCOC2 and PCOC3 colony morphology

A**:** PCOC2 p1. Two colony morphologies are present: 50% of flat colonies (circled in orange) and 50% of colonies growing as a multilayer (circled in green) (50x magnification).

B**:** PCOC2 p7. The amount of multilayered colonies decreased over time, by p7 only flat colonies were present (100x magnification).

C**:** PCOC3 p1. Only colonies growing as multilayers, similar to cells differentiating, were present at p1 (50x magnification).

D**:** PCOC3 p5. Over time, PCOC3 colony morphology changed as more cells grew as a monolayer resulting in colonies with a mix of cells growing laterally on the edges and upwards in the centre (an example is circled in blue) (50x magnification).

E**:** PCOC3 p7. Colonies composed of monolayers were faster growing than differentiating cells, by p7 all cells grew as monolayers, forming tight colonies.

4.4.4 Established cultures and amplification

To judge whether PCOC2 and PCOC3 were likely to be immortal lines, population doublings (PD) were estimated and compared with the number of PD expected from non-immortal cells. Untransformed neonatal foreskin keratinocytes (HEKn) cells, available from Life Technologies, can be grown in culture for approximately 30 PD before senescing. Commercially available human tonsil epithelial cells (HTEpiC) can be cultured for approximately 15 PD (ScienCell, Carslbad, USA; ref: www.sciencellonline.com/human-tonsil-epithelial-cells.html).

Based on culture records, PCOC2 plating efficiency was estimated to be 10.7%, and 30 PD were achieved by p9. Due to the growth pattern of PCOC3 accurate cell counts were not available for the first few passages and the number of PD can only be roughly estimated. With this caveat, and based on the calculated plating efficiency of 12.5%, 3 population doublings per passage were estimated. Counting from p3, 30 population doublings were reached by p13.

Stocks of PCOC2 and PCOC3 cell lines were expanded using the earliest available frozen aliquots to obtain a large number of cryopreserved samples of the same passage. At least 50 aliquots of PCOC2 p7 and PCOC3 p14 were frozen. The stored PCOC2 and PCOC3 cultures were likely to contain a heterogeneous population as they were expanded from early passage samples. In order to minimise the risk of experimental artefacts caused by selection of sub-populations in long term culture, these heterogeneous populations of p7 and 14 were used in all further experiments. An alternative approach would have been to use single cell cloning to isolate different clonal populations. However, heterogeneous cell populations were judged to be more representative of the original tumours. This was a dominant concern as the intended use of these cell lines was to study treatment response.

4.5 Model validation

It was important that the newly derived cell lines should be validated as appropriate models of HPV-positive OPSCC. This included confirmation of their genetic origin from the parent biopsy and that the cells were keratinocytes. The absence of mycoplasma contamination was also confirmed.

4.5.1 The PCOC cell lines were derived from the original tumour biopsies

STR profiling is based on the amplification of a series of polymorphic STR loci. The numbers of repeats of each allele at each site creates a unique profile for each individual (Nakamura et al. 1987). To allow future identification of the PCOC cell lines, and to confirm that they were derived from the original tumour biopsies, the STR profiles of the original biopsies were compared with those of the derived cells lines (Figure 4.7 and Figure 4.8). This confirmed that PCOC2 and PCOC3 cells were derived from the associated biopsy. It also demonstrated that PCOC2 and 3 cell lines differed from each other, and from established HPV-positive cell lines used in the laboratory (e.g. Hela, Siha, and Caski cells). Small variations were observed between the profiles obtained from the main biopsy and the cell lines. For example, the PCOC2 cell line profile appeared to have lost the 4 repeat alleles of THO1. These variations are consistent with loss of heterozygosity (such variations are common in immortal cell lines).

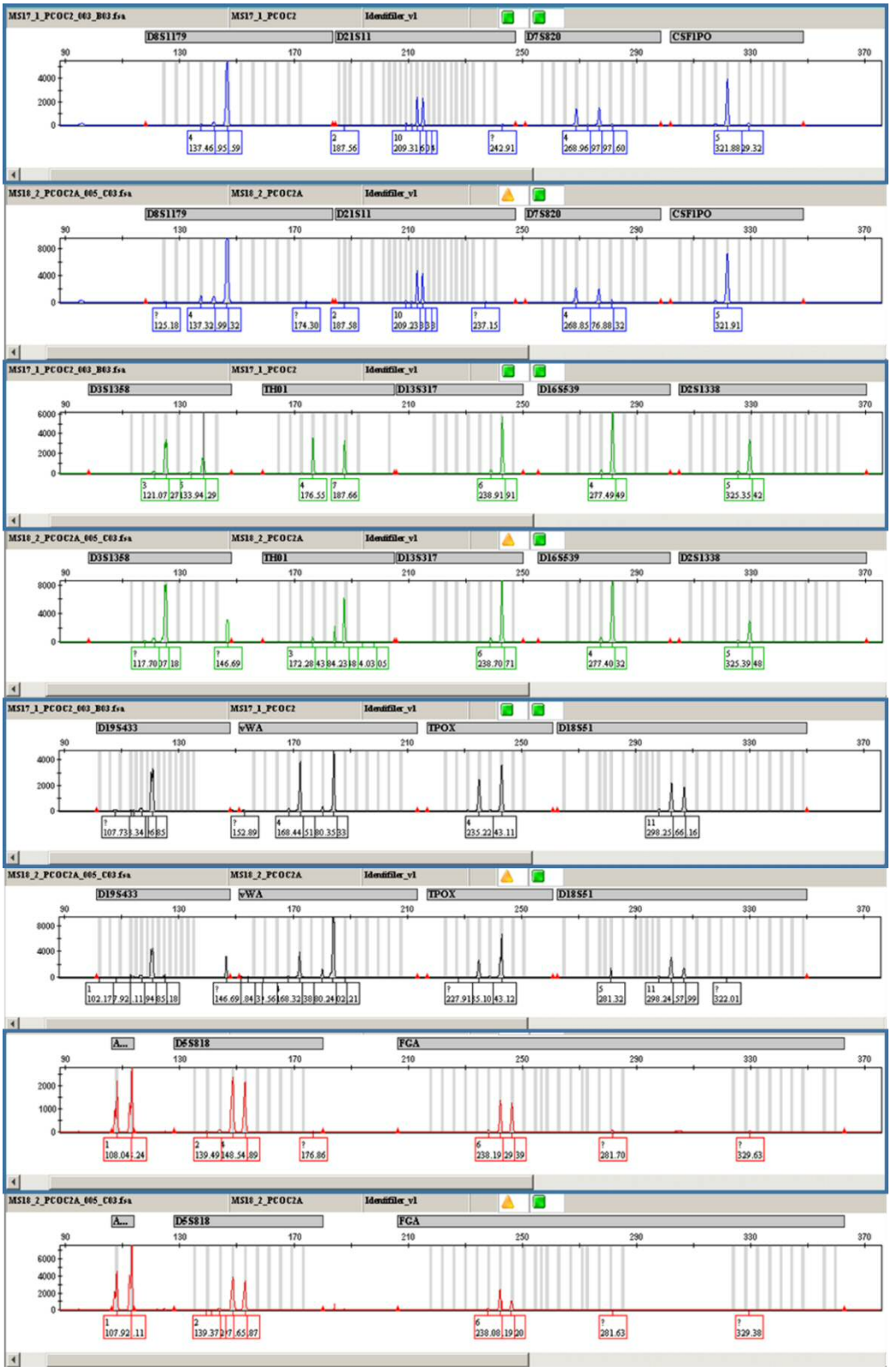


Figure 4.7: STR profile of PCOC2 biopsy and derived cell line

Short tandem repeats were amplified at 16 loci. Alleles are indicated in the grey bars immediately above the traces, and the number of repeats by the grey vertical lines. The biopsy profile is shown in the upper panel of each pair of images, and is outlined in blue. The cell line profile is shown in the lower panel of each pair.

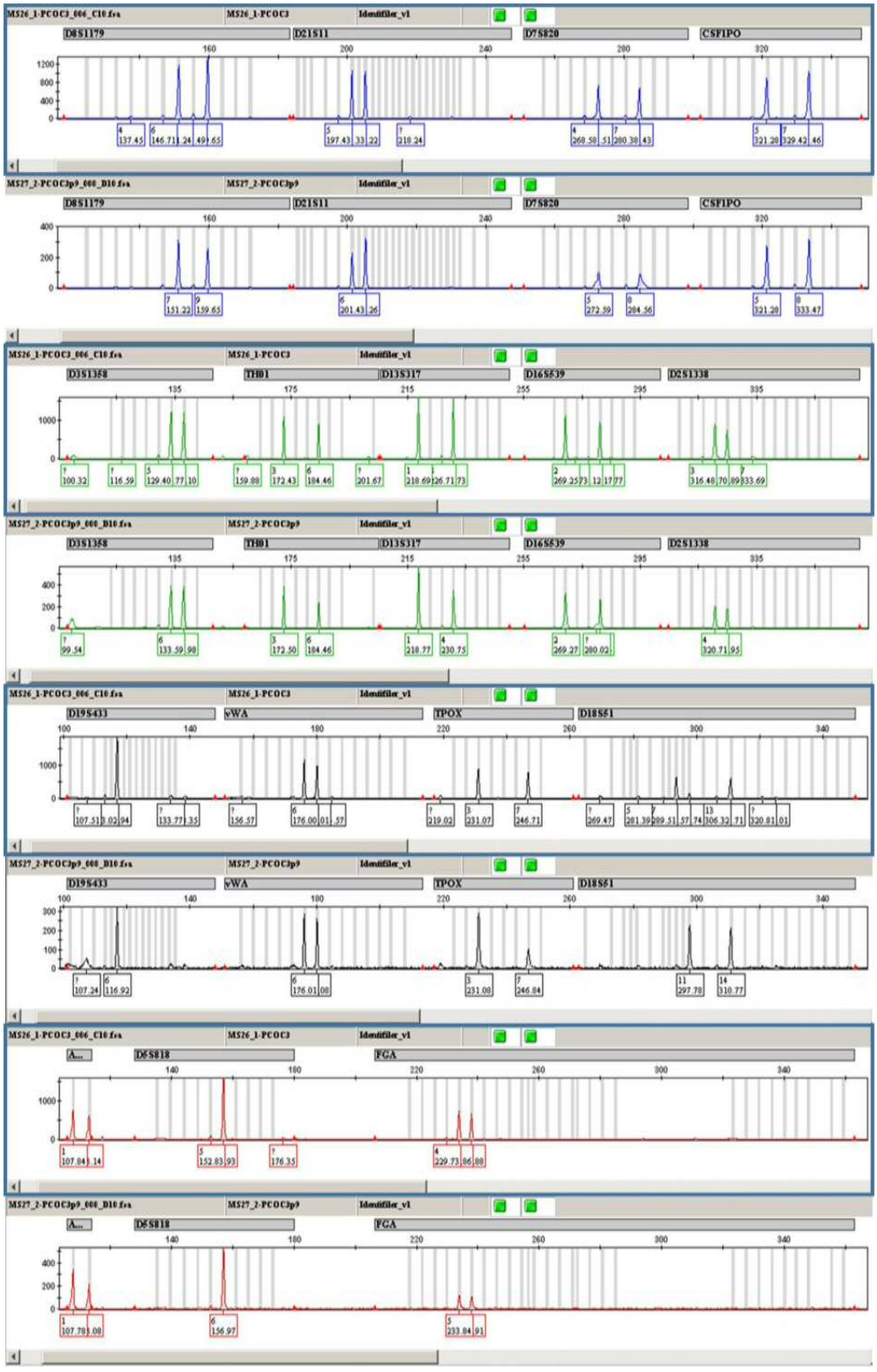


Figure 4.8: STR profile of PCOC3 biopsy and derived cell line

Short tandem repeats were amplified at 16 loci. Alleles are indicated in the grey bars immediately above the traces, and the number of repeats by the grey vertical lines. The biopsy profile is shown in the upper panel of each pair of images, and is outlined in blue. The cell line profile is shown in the lower panel of each pair.

4.5.2 PCOC cells are keratinocytes

Initial cell type characterisation was solely based on morphology. Cytokeratin (CK) expression is often used to define epithelial cells. Various cytokeratins, (e.g. 7, 8 and 19) are expressed in the tonsil (Clark et al. 2000). Their expression in PCOC2 and 3 cells and the original tumours was used to confirm the cell type on a molecular level. PCOC2 and PCOC3 cells stained strongly for all 3 cytokeratins. Cells in the original tumours were also stained by all 3 cytokeratins. The only difference was a weaker expression of CK8 in the PCOC2 tumour. Abnormal cytokeratin expression can occur in dysplastic and tumour cells. The CK pattern displayed by PCOC2 and PCOC3 was consistent with an epithelial origin. Overall, this confirmed a similar cytokeratin pattern in the tumours and derived cell lines, and strongly suggests that the cells are derived from tonsil epithelia and are keratinocytes (Figure 4.9).

4.5.3 Mycoplasma detection

Mycoplasma, due to their small size (± 100nm), are not detectable by light microscopy. Mycoplasma contamination in cell cultures causes biochemical and metabolic changes affecting a wide range of pathways which could alter experiment results (Miller et al. 2003). PCOC2 and PCOC3 cultures were tested for mycoplasma contamination by PCR. Both cell lines were confirmed to be free of mycoplasma after amplification and storage (data not shown).

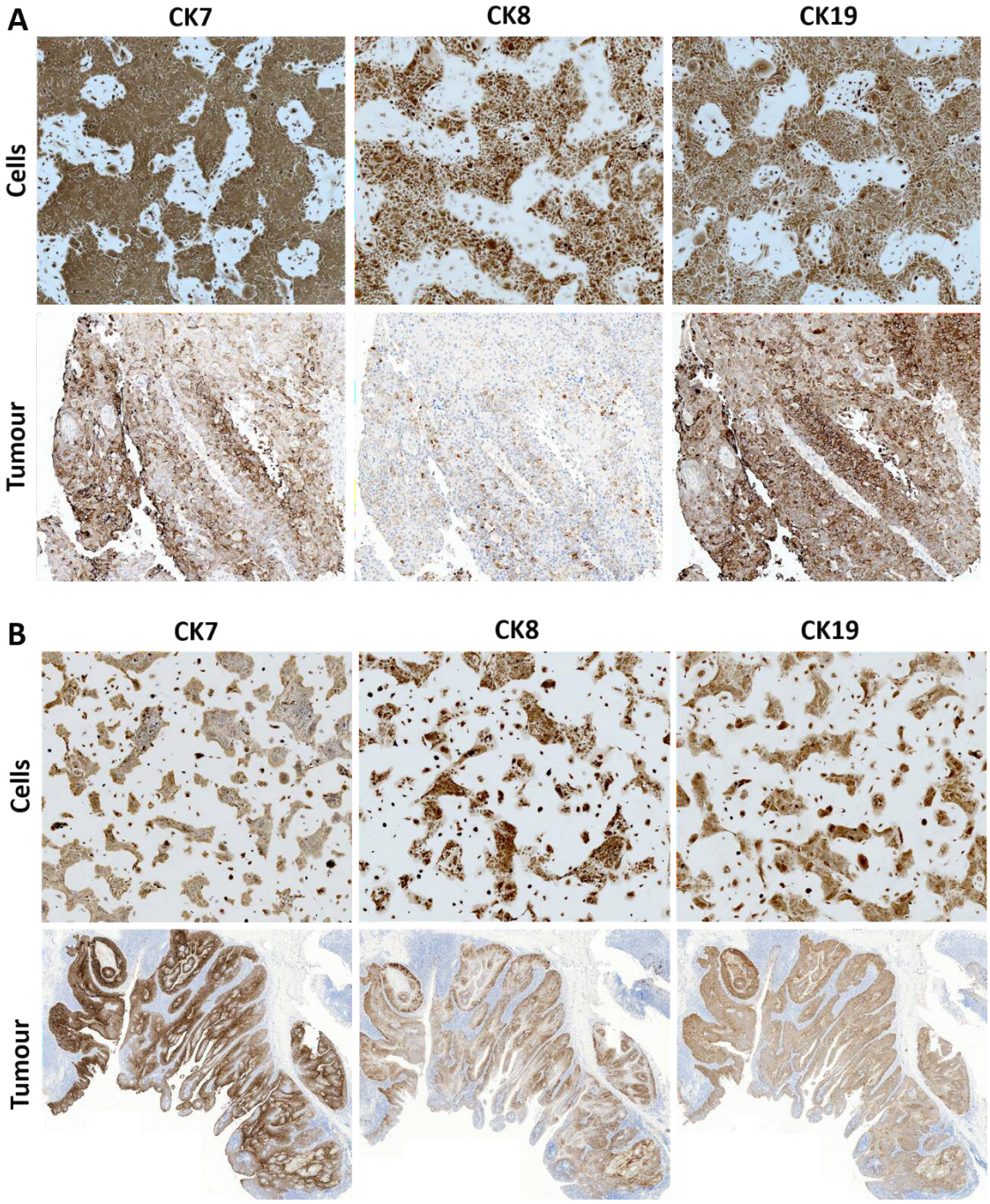


Figure 4.9: IHC staining for cytokeratins 7, 8 and 19 of PCOC2 and PCOC3 cells and original tumours

A: PCOC2 cells stained strongly for all 3 cytokeratins (CK) while cells in the tumours were strongly stained for CK 7 and 19. The tumour showed weak diffuse staining for CK8.

B: PCOC3 cultured cells and the original tumour cells were stained strongly for all 3 cytokeratins.

4.6 Assessment of HPV status

DNA was extracted from each biopsy and assessed using PCR targeting a 161 bp region of the HPV16 E6 gene. The adequacy of each sample for PCR-based analysis was confirmed by the amplification of a 209 bp fragment of the human β-globin gene. In accordance with the p16 IHC data obtained from the pathology department at the University Hospital of Wales, all oropharyngeal biopsies except PCOC9 were confirmed HPV16 positive (Table 4.1 and Figure 4.10). PCOC9 was shown to be HPV-negative, which is consistent with the patient being diagnosed with lymphoma. Explant cells from this biopsy were probably normal oropharyngeal cells, which were unlikely to lead to an established cell line.

In cervical SCC, HPV16 E1 and E2 genes are often disrupted as a result of integration of HPV DNA into the host genome (Pett and Coleman 2007). Tiling PCR assays amplifying overlapping fragments that span the E1 and E2 regions were used to detect HPV gene disruption (Figure 4.11). Disruption of the E1 and E2 regions was used as a surrogate marker for the presence of HPV integration. In several cases (PCOC1-5) the shorter fragments of E1 (1069 and 995 bp) were amplified, but the full length fragment (1954 bp) was not. This could indicate that DNA fragmentation during extraction or processing for PCR rendered the amplification of the larger fragment difficult. All E1 and E2 gene fragments were present in PCOC1-PCOC6 and PCOC10-PCOC12. Full length E1 and E2 genes were present in these biopsies, indicating presence of either episomal or mixed (episomal and integrated) infections. The second fragment of E2 and all E1 fragments were detected in PCOC7. This might indicate a disruption of E1 due to an integration event. Only the first fragment of E1 was detected in PCOC8. This was a sign of an integration event disrupting the second part of E1 and causing the loss of E2. No E1 or E2 fragments were detected in PCOC9 further confirming the absence of HPV16 infection (Figure 4.12).

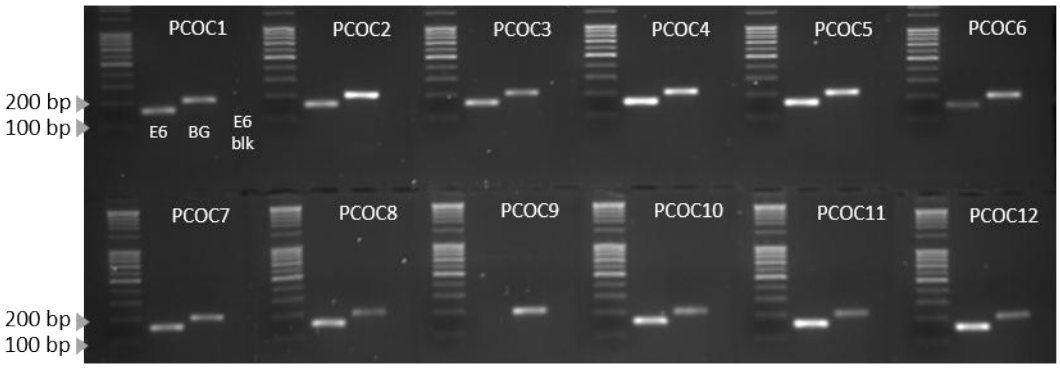


Figure 4.10: PCR amplification of β-globin and HPV16 E6 fragments in PCOC biopsy DNA
Three PCRs were performed for each samples: HPV16 E6 (E6 161 bp), β-globin (BG, 209 bp) and an E6 PCR blank as a negative control (E6 blk).

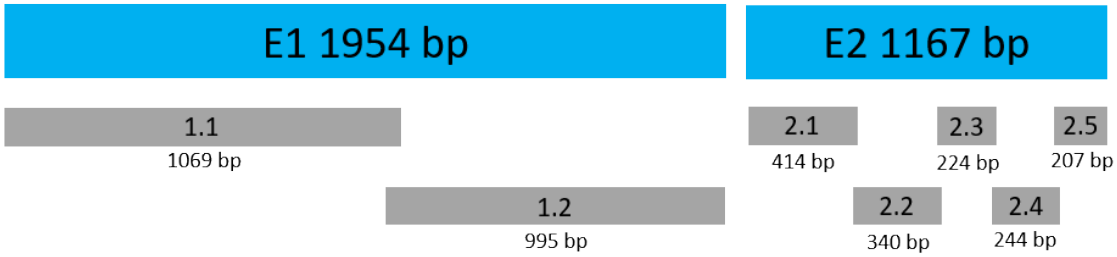


Figure 4.11: Schematic representation of HPV16 E1 and E2 gene tiling PCR assay

Whole genes and overlapping fragments of E1 and E2 genes were amplified by PCR. The absence of some fragments or of the whole gene indicates a disruption likely caused by an integration event (N.B: These are surrogate markers of integration as they do not enable to differentiate between fully integrated cases or a mix of episomes and integrated virus).

Study number	E1			E2					
	E1.1	E1.2	E1 full	E2.1	E2.2	E2.3	E2.4	E2.5	E2 full
PCOC1									
PCOC2									
PCOC3									
PCOC4									
PCOC5									
PCOC6									
PCOC7									
PCOC8									
PCOC9									
PCOC10									
PCOC11									
PCOC12									

Figure 4.12: PCR amplification of E1 and E2 HPV16 genes in PCOC biopsy DNA

Successfully amplified fragments or full genes are indicated in green and unsuccessful in red. E1 and E2 genes were present in PCOC1-PCOC6 and PCOC10-PCOC12, likely indicating either an episomal or mixed infection. HPV appeared to be integrated into the host genome in PCOC8 resulting in the loss of the second part of E1 and the whole E2. PCOC7 could not be interpreted due to an inconclusive E1 result. The HPV-negative status of PCOC9 was confirmed.

4.7 PCOC cell lines require post-mitotic 3T3 feeder cells to support optimal growth

The PCOC lines were derived with the support of post-mitotic 3T3 feeder cells. The use of 3T3 feeder cells was stipulated in the protocol originally used for gynaecological biopsies. 3T3 feeder cells support, combined with media additives, are a well described technique to enhance keratinocytes growth (Rheinwald and Green 1975). However for convenience, it would be simpler to culture them without 3T3 feeder cells. The effects of growth with and without 3T3 feeder cells were therefore investigated by generating growth curves. The optimum seeding densities were previously determined in 10 cm plates so that cells did not reach more than 80% confluence after one week (0.8 million cells and 2 millions 3T3 feeder cells). Seeding densities were scaled down to suit the culture area of 6 wells plates. Over a 9 days period, cell counts were performed each day by manual counting with a haemocytometer, following trypan blue staining. All experiments were performed in triplicates. As in standard culture conditions, media (EGF+) was replaced every other day, and 3T3 feeder cells were replaced as needed.

The resulting growth curves are shown in Figure 4.13. PCOC2 and PCOC3 cells were not able to grow efficiently without the support of 3T3 feeder cells. PCOC2 cells attached to the plate but failed to divide, apparently remaining in lag phase. Their number remained constant throughout the experiment. PCOC3 cells attached to the plate and completed two PD over 5 days. However, the cells did not reach a log phase of growth. After 5 days, the number of cells remained constant indicating a growth arrest.

In contrast, the growth curves with support of 3T3 feeder cells both showed significant growth. PCOC3 cells spent a longer time in lag phase compared to PCOC2 before entering the log growth phase. PCOC2 entered an exponential growth phase one day earlier than PCOC3 (day 4 versus day 5). In the log phase, the PD times were similar for both lines, at around 2 days. By day 10 PCOC2 had entered the plateau phase, this was not reached by PCOC3.

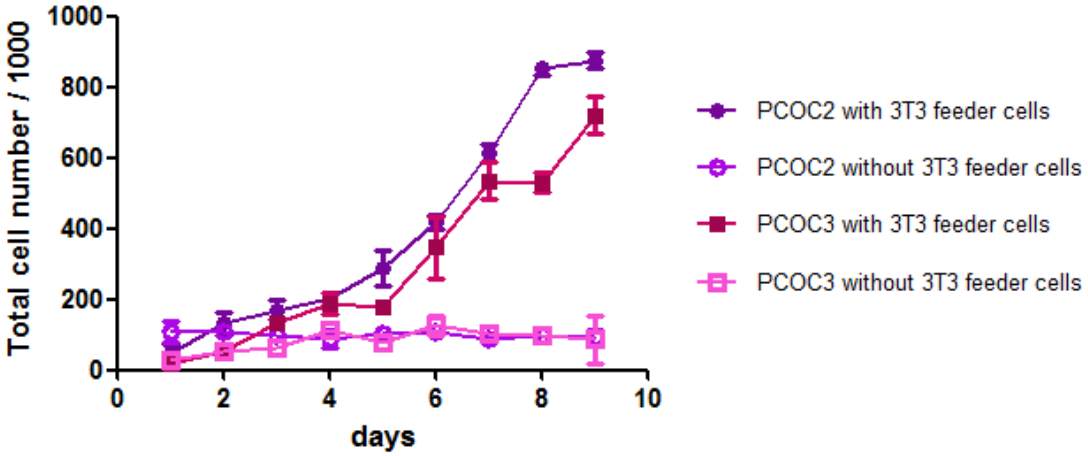


Figure 4.13: Growth curve of PCOC2 and PCOC3 cultured with and without 3T3 feeder cells
Triplicates of cells counts (divided by 1000) are represented over a 9 days period. Error bars indicate the standard deviation. In the absence of 3T3 feeder cells neither PCOC2 nor PCOC3 cells were able to sustain growth. In exponential growth, with the support of 3T3 feeder cells, the population doubling time of both cell line was approximately two days.

4.8 Discussion

Biopsies were cultured using protocols to support the growth of explants. Two out of 13 cultures produced cell lines that appear to be immortal. These cultures showed an ability to grow for over 30 PD, exceeding the growth span of normal keratinocytes. STR analysis confirmed these lines were derived from the original biopsies, and immunohistochemical analysis confirmed the cells were of the expected cell type (i.e. keratinocytes). Additional analyses demonstrated that these cells grew optimally when supported by post-mitotic 3T3 feeder cells.

The growth pattern of PCOC2 and 3 differed in the first passages, PCOC3 was slower to establish as a consistently growing cell culture. Such differences are likely to be due to individual genetic factors relating to the cells contained in the biopsy. This growth pattern changed as faster growing non-differentiating cells were selected. In the growth curves, the growth rate during the exponential phase was similar. Cells were seeded at low density in the growth curve experiment to reach confluence after 9 days or more. It is worth noting that the low density used likely contributed to the length of the lag phase. Culture without 3T3 feeder cells may be feasible if higher seeding densities were used.

Several methods have been described to establish cell lines from tissue samples including enzymatic/chemical dissociation, mechanical dissociation and explant. The explant method as described in Stanley 2002 was selected as it was previously successfully used to develop HPV-positive gynaecological cell lines in our laboratory (Stanley 2002; Bryant et al. 2014). The main drawback of this method was the lack of a single cell suspension, which reduced the number of cells able to grow in culture to the outside layer of the explant. To compensate for this, the biopsies were processed before culture to obtain fragments smaller than 2 mm. The advantage of the explant method was the reduced exposure to chemicals such as trypsin which could diminish cell viability. Even though mechanical disaggregation damaged some cells, the use of tissue fragments instead of single cells ensured some cells were still intact within each biopsy piece.

In this study in 76.9% of biopsies processed produced new cell growth. This showed the adequacy of the technique to sustain the growth of tonsil epithelial cells. However, the overall success rate at establishing long-term cell culture was low (15.4%). Although, it should be

recognised that contamination was the primary cause of failure, with a single event affecting 5 biopsies (38.4%).

The advantages and disadvantages of single cell cloning and heterogeneous cell populations were considered. Single cell cloning could have led to the isolation of different clones from the same tumour especially if performed during the first passages. Characterised clones can be a useful tool to determine if different levels of HPV gene expression or integration status influence treatment response. However, tumours contain heterogeneous cell populations. As the main intended use of the cells was to study the impact of a relatively new drug, the use of heterogeneous populations seemed appropriate. Early aliquots are still available if there is a need for single clone isolation in the future.

A drawback of long-term culture process is the selective pressure for the clones best suited to grow in tissue culture conditions. Using the lowest passage available reduced the impact of this process. Moreover, the use of the same stock in each experiment avoids clonal variations over time affecting the results. The selective pressure is likely to be greatest in the first few passages. Therefore the cell populations at p7 and p14 were probably already relatively homogenous. This could be assessed genetically using deep sequencing of multiple sub-clones, but this was not performed within the current study. However, the cultures were assessed microscopically, and alteration of colony morphology between p1 and 7 or 14 was an indication of the culture homogenisation (i.e. lateral growth instead of upwards). The peak-shaped colonies which disappeared over time were likely to contain normal cells which after a few passages underwent senescence due to telomere shortening.

In comparison to the already available HNSCC lines 93VU147T, UDSCC2, UPCISCC90, UMSCC47 and UMSCC104, the PCOC lines have several advantages. They originate from primary OPSCC tumours in treatment naive patients. One being a never-smoker and the other a ex-smoker, neither of them have a reported heavy drinking history. None of the other lines combine these characteristics. The main difference between the protocols used to develop PCOC2 and 3, and those used for the published lines, was the absence of 3T3 feeder cells to support the growth of the published lines. The type of media, washes and inversion of the plates to promote attachment also varied between protocols. A new method of establishing cell lines from head and neck tissue biopsy was recently described (Owen et al. 2016). Following enzymatic dissociation, cell surface markers including epithelial cell adhesion

molecules and fibroblast surface protein were used to sort keratinocytes and fibroblasts respectively using fluorescence associated cell sorting. This technique avoids the need for differential trypsinisation and washes to remove non-adherent cells. Fibroblast growth was not a major issue in the development of PCOC2 and PCOC3, possibly because the use of 3T3 feeder cells minimised their growth.

The low success rate in establishing new lines is consistent with the low number of HPV-positive head and neck, and more specifically OPSCC, cell lines already available. Other factors such as the relatively recent interest in HPV-positive head and neck SCC and the fact that not all developed lines are published also influences the amount of lines available.

All SCC biopsies collected were HPV-positive, this is higher than expected in view of the published HPV prevalence rate in oropharyngeal cancers of around 51.8% reported in the UK between 2002 and 2011 and, 55% in South Wales between 2001 and 2006 (Evans et al. 2013; Schache et al. 2016). In South Wales, the percentage of HPV-positive cases might be higher now than previously. Based on p16 data between 2010 and 2013 of a small number of patients treated at Velindre Hospital, 75% of tumours were HPV-positive (unpublished data provided by Dr Mererid Evans). Transcription of the CDK inhibitor p16 (CDKN2A) is commonly up-regulated in HPV-positive tumours as a consequence of binding of Rb by HPV E7 protein. Hence p16 expression is often used as a biomarker of HPV infection. The HPV positivity of the biopsies collected in the study could also be explained by a bias towards HPV-positive cases. Some patients were already known to be p16 positive before being consented for the study. Finally, the small cohort size could also have contributed an unrepresentative sample with regard to HPV prevalence.

4.8.1 Strengths and weaknesses

The culture method used had a high success of implants. The absence of explants in some biopsies suggested that only some tumour cells were able to grow in the culture environment provided. It also highlights the heterogeneity of the tumours, pieces containing a high number of necrotic cells are less likely to develop. The low take rate following explant growth was the main weakness. Contamination with environmental microbes contributed to the low success rate. Senescence observed after several passages indicated that the culture conditions were also suitable for normal cell growth.

Basic characterisation of the model was performed, including robust quality controls. STR profile combined with the HPV status and the keratinocytes staining provided substantial evidence of the cell type grown and confirmed their epithelial origin. However it should be noted that the cytokeratin panel used was not specific for tonsils or oropharyngeal tissue, and could be present in other types of epithelia. Loss of heterozygosity was identified in PCOC2 in comparison with the biopsy. Such events are known to occur in the development of tumour lines and to remain stable once cultures are established (Dracopoli and Fogh 1983). Genetic instability in culture has been described in several head and neck lines (Brenner et al. 2010). However this has not been tested in regards to PCOC2 and PCOC3. A strength of those models is the use of relatively low passages compared to what is often seen in immortalised lines, providing a closer representation to the original tumours than other available models.

The cells were not weaned off the feeders over time. Such a process would have required a gradual reduction of 3T3 feeders amount over 3-4 passages. Cells were cultured with a feeder layer before cryopreservation, the cultures were revived in the same conditions, and used straight away in experiments. This culture method was more labour intensive. However, culture conditions were developed in accordance with it and used throughout the project.

4.8.2 Conclusion

Two new HPV-positive cell lines have been established from fresh biopsies of oropharyngeal cancer. The novel lines are from patients who are typical of the demographic affected by HPV-driven oropharyngeal cancers (i.e. relatively young and fit, with limited co-morbidities associated with alcohol and tobacco use). These lines represent a valuable and scarce resource to support investigation of the biology of, and potential treatment of, HPV-positive oropharyngeal cancers.

Chapter 5 - messenger RNA sequencing

5.1 Introduction

Whole transcriptome sequencing (also known as RNA-seq) is a quantitative and sensitive technique that shows all transcripts present at a particular time point (Mortazavi et al. 2008). In recent years, it has emerged as a valuable technique for tumour characterisation, and has had a huge impact on fundamental understanding of the biology of cancer (Parfenov et al. 2014; Cancer Genome Atlas Network 2015). In the current study, mRNA-seq was used to investigate differences in gene expression between, and among, the HPV-positive and negative cell lines for following reasons:

- To characterise and/or confirm the status of the cell lines with regard to HPV gene expression, HPV integration and p53 mutation.
- To establish whether the HPV-positive cell lines are characteristic of typical HPV-driven tumours, and if our HPV-negative cell lines are typical of non-HPV-driven HNSCC.
- To compare expression of different genes, including those involved in DNA repair, between HPV-positive and negative lines. This was fundamentally important in relation to differential responses to PARPi (described in the next chapter).

5.2 Data and quality control

Preliminary bioinformatic quality control analyses were performed by Dr Peter Giles and Dr Kevin Ashelford from the Wales Gene Park. A random selection of 10% of reads was used for data quality analysis. All cell lines passed quality control measures and were used for downstream analysis. Reads were mapped against human reference sequence hg19 (genome build 19) and HPV16 reference sequence NC_001526 (Table 5.1).

Table 5.1: Mapped reads to human and HPV sequences

Cell line (HPV +/-)	Median insert size (bp)	Reference genome*	Mapped reads
UPCISC90 (+)	141	hg19	50339462
		HPV16	19831
PCOC2 (+)	137	hg19	34041846
		HPV16	5014
PCOC3 (+)	151	hg19	31352778
		HPV16	13087
UMSCC47 (+)	150	hg19	49239786
		HPV16	3969
UMSCC4 (-)	151	hg19	48897461
		HPV16	0
UMSCC6 (-)	151	hg19	47881266
		HPV16	0
UMSCC19 (-)	148	hg19	49508832
		HPV16	0
UMSCC74 (-)	154	hg19	51169790
		HPV16	12

*: hg19 denotes human genome build 19; HPV16 denotes HPV16 reference genome NC_001526.2.

5.3 Characterisation of novel and established cell lines

5.3.1 Cell line panel

Eight cell lines, 4 HPV-positives and 4 HPV-negatives, were selected to use as a study panel. The aim was to select lines representing the diversity of OPSCC. The main HPV-positive lines were the novel lines PCOC2 and PCOC3 (described in Chapter 4). Two commonly used lines, derived from heavy smokers, UMSCC47 and UPCISCC90 were also included. HPV-negative lines varied in p53 status. UMSCC4 and UMSCC19 have previously described mutations in p53, whilst UMSCC6 and UMSCC74a harbour wild-type p53.

UMSCC74a originated from a recurrent tumour, which was diagnosed after the patient had undergone chemoradiotherapy. This line is also described as having undergone epithelial-mesenchymal transition, giving the culture a fibroblastoid appearance. UMSCC4 was also derived post-treatment from a patient who had undergone 2 rounds of chemotherapy: bleomycin, oncovin, mitomycin C, and methotrexate.(Brenner et al. 2010).

5.3.2 HPV gene expression in four HPV-positive cell lines

Transcripts mapped to HPV16 sequence were visualised for each cell line using the Integrative Genomics Viewer (IGV Version 2.3.65, Broad Institute) as shown in Figure 5.1 (Robinson et al. 2011). Expression of HPV encoded genes was quantified (expressed as Reads per kilobase of transcripts per Million mapped reads (RPKM)).The HPV gene expression pattern varied between the lines (Figure 5.2). Late genes L1 and L2 were not expressed in any of the lines, indicating the presence of non-productive HPV infection. This corresponds to the gene expression pattern commonly observed in HPV-driven tumours, where deregulation of HPV gene expression prevents completion of the viral life cycle, and hence the L1 and L2 capsid genes are not expressed. The highest levels of E6/E7 expression were observed in UPCISCC90 and PCOC3. Proportionally higher expression of E4 was found in PCOC2 and PCOC3, and of E5 in UMSCC47. UPCISCC90, PCOC2, and PCOC3 showed expression of E2/E4 and E5.

5.3.3 Splicing of HPV encoded genes

The HPV genome is highly spliced. The two main oncogenes E6 and E7 are of particular interest as they are transcribed from the P97 promoter as an E6E7 transcript that undergoes alternative splicing. The resulting transcripts have been associated with different oncogenic potential. Splicing patterns were investigated using the Sashimi plot function of IGV as illustrated in Figure 5.3, and were compared with the most common HPV16 transcripts as described on the Papillomavirus Episteme reference website (https://pave.niaid.nih.gov/). To avoid confounding by sequencing artefacts, only splice sites with at least 10 reads mapped were considered for the analysis. Different frequencies of distinct splicing sites were present across the cell line panel. Quantified levels of each transcript and their position are shown in Table 5.2.

The most common oncogenic transcripts, E6*I and E6*II, were present in all lines. A splicing site across E1 and E4 was also found in all lines. This site is often present after the E6*I and E6*II transcripts. Less common splice sites located in E2 were present in UPCISCC90, PCOC2 and PCOC3, another less common site was found in E5 in PCOC2. UMSCC47 expressed a reduced variety of transcripts compared to the other lines. A splicing variant which has not been commonly described, across E6 and E7, was detected at a low level.

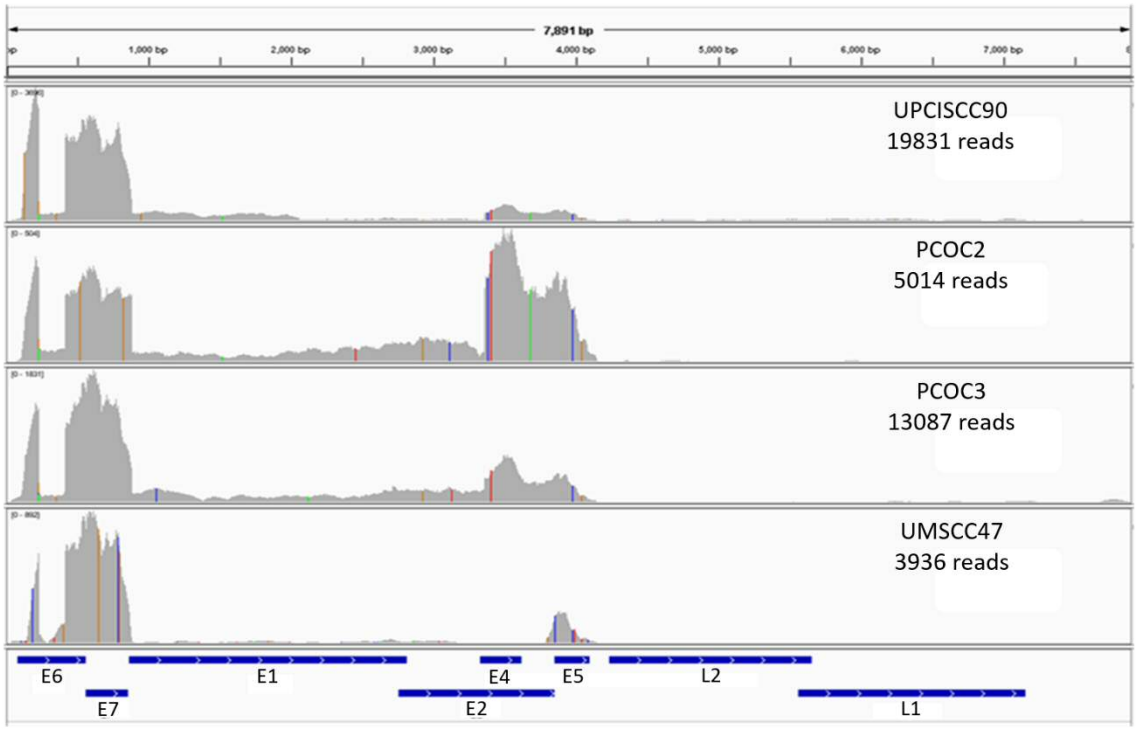


Figure 5.1: Mapping of reads against HPV16

The left axis represents the raw number of reads; the horizontal axis represents the HPV16 genome and the exons. For each cell line, the number of reads mapped to the HPV16 NC_001526.2 sequence is indicated. Vertical coloured lines indicate mismatches with the reference sequence.

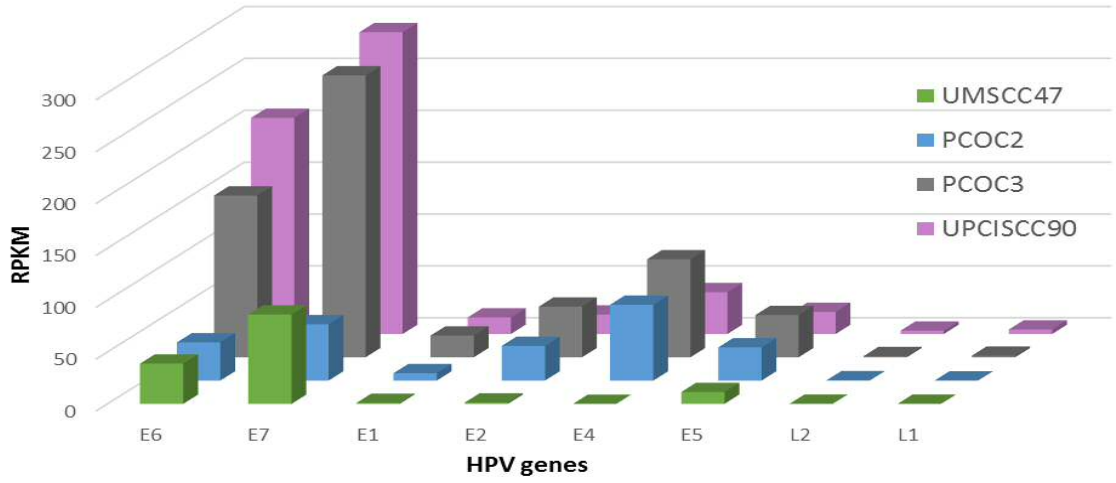


Figure 5.2: HPV gene expression in UMSCC47, PCOC2, PCOC3, UPCISCC90

Reads Per Kilobase of transcript per million Mapped reads (RPKM) are plotted for each HPV gene.

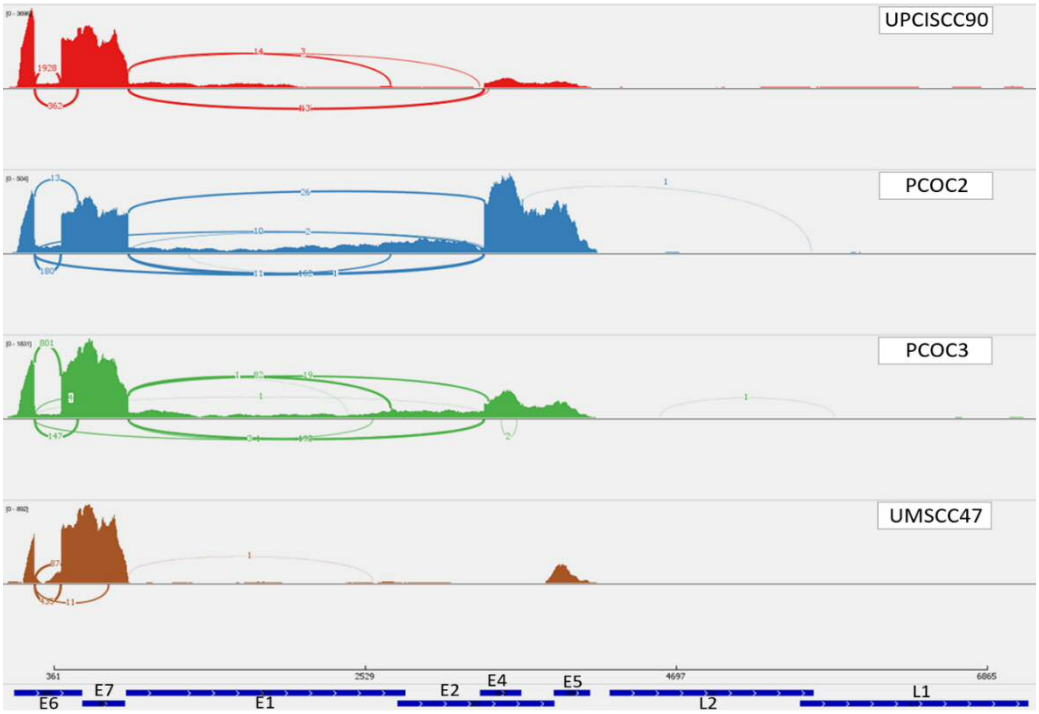


Figure 5.3: Splice junctions in UPCISCC90, PCOC2, PCOC3 and UMSCC47
Showing reads with ends mapping at different sites of the HPV genome, visualised using IGV.

Table 5.2: Description of HPV16 transcripts

Cell line	Start (nt)*	End (nt)	Count	Transcript
UPCISCC90	226	408	1928	E6*I
	226	525	362	E6*II
	880	3356	156	E1^E4
	880	2707	14	E2
	880	3359	12	E1^E4
PCOC2	226	408	180	E6*I
	226	525	13	E6*II
	226	3356	57	E5
	226	3359	10	E5
	880	2707	11	E2
	880	3356	162	E1^E4
	880	3359	26	E1^E4
PCOC3	226	408	801	E6*I
	226	525	147	E6*II
	880	2707	82	E2
	880	3356	132	E1^E4
	880	3389	19	E1^E4
UMSCC47	226	408	435	E6*I
	226	525	87	E6*II
	226	741	11	Novel

*: Nucleotide (nt)

5.3.4 Identification of HPV variants

Small differences in HPV16 sequence can be found in different geographic areas, these differences only involve a small number of nucleotides, and have been classified as HPV16 variants (Burk et al. 2013). The main variants described are European, African, Asian and American, with subtypes in each group. Nucleotide mismatches in the highly conserved E6 and LCR regions, which characterise each variant have been described (Cornet et al. 2012). Using IGV, the E6 nucleotide sequence of each cell line was compared with the mismatched sites described for each HPV16 variants, the results are presented in Table 5.3. UPCISCC90, PCOC2 and PCOC3 harbour infection by the main European variant of HPV16. This result was expected as all those lines originated from European patients. The HPV variant present in UMSCC47 was identified as African 2A (AF2A), this line was originally developed in the United States.

Table 5.3: Comparison of HPV16 E6 sequence in UPCISCC90, PCOC2, PCOC3 and UMSCC47 with known HPV16 variants

Variant or cell line	Position and nucleotide*												
	109	131	132	143	145	178	286	289	295	335	350	403	532
Reference**	T	A	G	C	G	T	T	A	T	C	T	A	A
EUR	/C	/G	-	-	-	/A	-	-	-	-	/G	-	-
EAS	-	-	-	-	-	G/C	-	-	-	-	-	-	-
AF1a	-	-	C	G	T	-	A	G	-	T	-	-	-
AF1b	-	G	-	G	T	-	A	G	G	T	G	-	-
AF2a	C	-	T	G	T	-	A	G	-	T	-	G	-
AF2b	-	-	-	G	T	-	A	G	-	T	-	-	-
NA	-	-	-	-	T	-	A	G	-	T	G	-	-
AA1	-	-	-	-	T	-	A	G	-	T	G	G	-
AA2	-	-	-	-	T	-	A	G	-	T	G	-	/G
UPCISCC90	ND	G	G	C	G	T	ND	A	T	C	G	A	A
PCOC2	ND	A	G	C	G	T	ND	ND	ND	C	T	A	A
PCOC3	ND	A	G	C	G	T	ND	ND	ND	C	G	ND	A
UMSCC47	ND	A	T	ND	T	T	ND	ND	T	T	T	G	A

* For each nucleotide position, the mismatch is indicated or a dash when no nucleotides were changed from the HPV16 reference sequence. Nucleotides separated by a slash are different nucleotides that can be in the given position for the given sublineage. The top half of the table shows known HPV16 variants described by Cornet et al (Cornet et al, 2012). The lower part represents the sequence from the 4 lines characterised. Diagnostic SNPs are shown in red.

** Abbreviations: EAS, European-Asian; EUR, European; AFR, African; NA, North American; AA, Asian-American; ND , no data - i.e. not transcribed at a sufficient level to allow detection.

5.3.5 HPV integration sites

Integration of HPV DNA into the host genome confers a proliferative advantage *in vitro,* and integrated virus is present in the majority of HPV-positive head and neck cancers (Jeon et al. 1995; Parfenov et al. 2014). Viral integration can potentially disrupt the pattern of HPV gene expression as well as alter the expression of some human genes, leading to genetic instability and disease progression. It was therefore important to determine the integration status of the lines. Integrated virus had previously been described for UMSCC47 and UPCISCC90. Hence it was possible to compare the mRNA-seq data with existing publications (Olthof et al. 2015), thereby potentially validating our methodology.

5.3.6 Fusion transcripts

HPV integration sites in the host genome were detected by analysing paired-end reads that had one end mapped to HPV16 sequence and the other to a human sequence (these are referred to as fusion transcripts). Circos plots were used to visualise fusion transcripts (Figure 5.4). Sites of insertions were further analysed using IGV (Table 5.4). The location of integration sites was determined to identify whether insertional mutagenesis was a likely factor in the development of the original tumour. Information on genes into which insertion was apparent, was obtained from the GeneCards Human Gene Database (http://www.genecards.org/).

Relatively high numbers of fusion transcripts were observed in the UPCISCC90 line, these were further analysed using IGV. One main integration site was detected in exon 4 of gene C9orf156 in chromosome 9, with reads mapping in E1, E6 and E7. Reads also mapped in intron 2 of the same gene from E1. Another site of integration was detected around 200 bp before FOXE1 from E7. The point of fusion in the HPV gene E6 in the two most common fusion transcripts coincided with the ends of E6*I and E6*II transcripts suggesting that transcription of these mRNAs was driven by an HPV promoter.

The analysis indicated integration upstream of, and production of a fusion transcript with, C9orf156 (Chromosome 9 Open Reading Frame 156). C9orf156 is a poorly characterised protein, also known as Nef Associated Protein 1. The suggested function of this protein, according to its GeneCards entry is a potential hydrolase, which hydrolyses acyl-CoA thioesters (*in vitro*). However, its physiological substrate is not known, and its only known disease association is with cleft lip. FOXE1 is a thyroid transcription factor involved in thyroid

morphogenesis. Several diseases have been associated with FOXE1 including cleft palate, thyroid associated conditions and thyroid cancers.

Four integration sites were detected in PCOC2. The main integration site was located in chromosome 10, with reads originating in E2 and one read from L1. All reads mapped within the gene YMEL1 at two sites, introns 6 and 7. The function of YME1L1 is described as the human orthologue of yeast mitochondrial AAA metalloprotease, Yme1p. The protein is localised in the mitochondria, and it is proposed that this gene plays a role in mitochondrial protein metabolism and could be involved in mitochondrial pathologies. The only confirmed disease association is with pulmonary valve insufficiency. The presence of the E6^E4 splice products suggests that if integration had taken place, it had not separated the E6 and E4 ORFs, implying that E2 sequences likely remain intact. This has been confirmed by PCR with full amplification of both E2 and E1 genes indicating that intact copies of both are present. Two other minor integration sites were detected, from E2 into exon 7 of SUX22 in chromosome 15, and from E7 into exons 8 and 9 of DAGLB in chromosome 7.

A number of integration sites were also observed in PCOC3. The main site of integration was located in chromosome 20. The majority of these fusion transcripts mapped to E7, with a minority mapping to E1. The site of integration was the gene CEBPB, which is formed of only 1 exon. The analysis showed several fusion transcripts that appear to link between the E7 splice donor region and exonic sequences in the CEBPB gene. CEBPB gene encodes a transcription factor that contains a basic leucine zipper domain. The encoded protein functions as a homodimer and is important in the regulation of genes involved in immune and inflammatory responses. Mutations in CEBPB are associated with several diseases including acute promyelocytic leukemia and glioma (Du et al. 2016).

Only two sites of integration were detected in UMSCC47. The main one was located in chromosome 3, with fusion transcripts originating from E7 and a few from E5. Within chromosome 3, the integration site was located in exon 11 of the TP63 gene. TP63 is a member of the p53 family of transcription factors. The encoded protein acts as a sequence-specific DNA binding transcriptional activator or repressor. Activity may be required in conjunction with TP73/p73 for initiation of TP53/p53 dependent apoptosis in response to genotoxic insults and the presence of activated oncogenes. It may also be involved in Notch signalling. Diseases associated with TP63 include ectrodactyly, ectodermal dysplasia, cleft

lip/palate syndrome and split-hand/foot malformation. The other integration event detected spans from E6 to chromosome 11. This read maps to Exon 9 of the BIRC2 gene. BIRC2 is a member of a family of proteins that inhibits apoptosis by binding to tumour necrosis factor receptor-associated factors TRAF1 and TRAF2. Mutation of BIRC2 may be associated with lung cancer and lymphoma. However, given that only a single read mapped to this location, further validation would be required to view this as a robust finding.

5.3.6.1 Consistency of HPV integration findings with published data

Integration events present in UPCISCC90 and UMSCC47 have been previously described (Ragin et al. 2004; Olthof et al. 2015). In UPCISCC90, the same integration sites were detected in C9orf156 exon 4 and intron 2 as previously observed by Olthof et al. The mRNA-seq also detected additional sites of integration within this gene, which had not been reported previously. An additional integration site 200 bp before FOXE1, observed with mRNA-seq, was also novel. The detection of these extra sites could be explained by the higher sensitivity of mRNA-seq. However, these data differ from the integration sites published by Ragin et al., who used fluorescence in situ hybridization and spectral karyotyping, and reported chromosome 9 at 9q31 as the main site of integration (Ragin et al. 2004). While located in the same chromosome, the location is not directly adjoining to the one found with mRNA-seq located at 9q22.33. These differences might be due to the techniques used, alternatively, this could be an example of evolution and selection of subclones in cell lines grown in different laboratories.

In UMSCC47, the mRNA-seq analysis identified one published integration event, previously detected using a rapid-amplification-of-cDNA-ends methodology, into exon 7 of TP63 (labelled exon 11 in the reference sequence used here). Using a ligation-mediated PCR method, another integration site, into intron 4 of TP63 was detected. This latter integration site was not observed in the mRNA-seq data. This could be explained by the proximity of the sites and also by differences in the techniques.

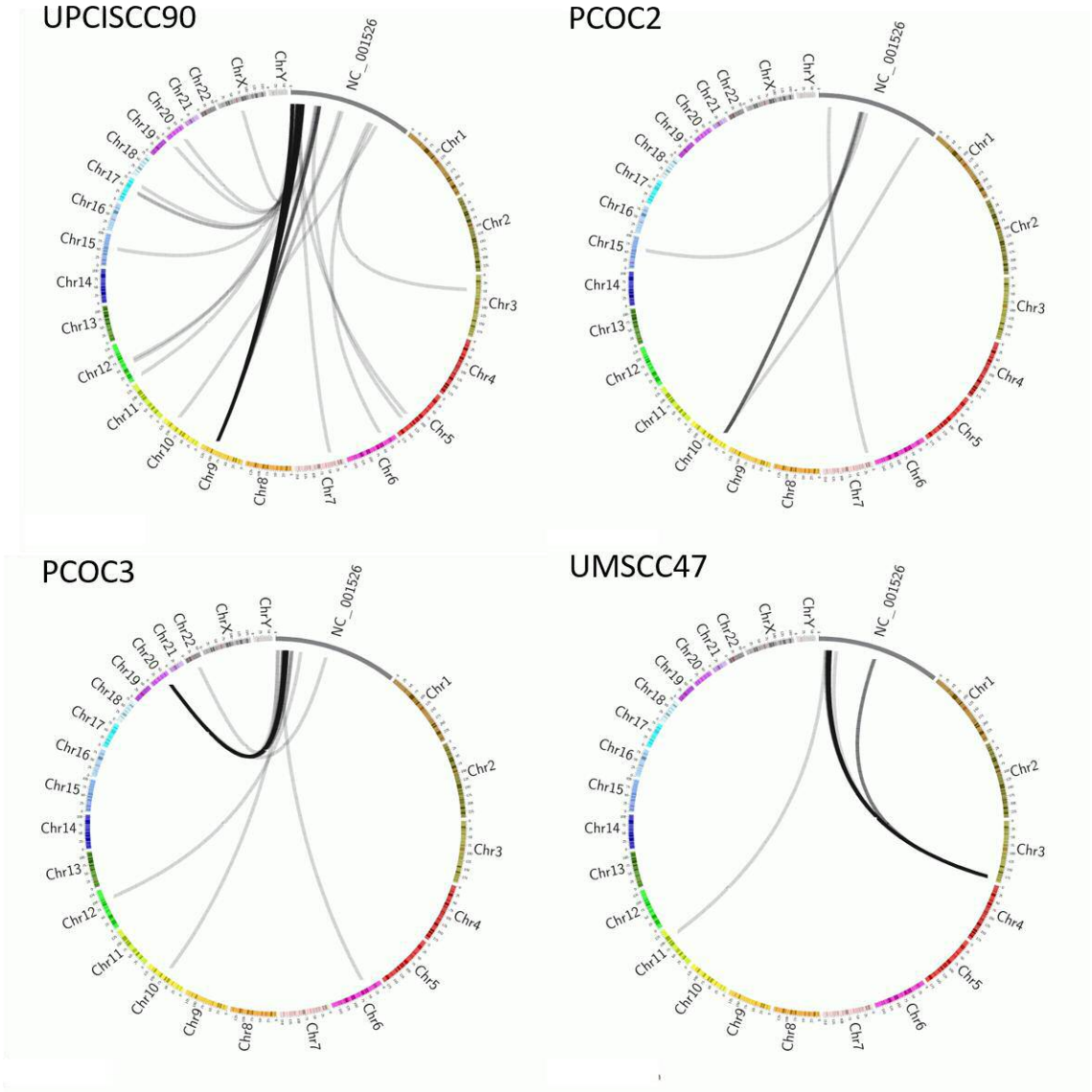


Figure 5.4: Circos plot showing fusion transcripts between HPV16 and human genome for UPCISCC90, PCOC2, PCOC3 and UMSCC47

The circle consists of the HPV16 sequence in grey and the human chromosomes as coloured bars. Fusion transcripts are represented as grey lines joining their point of insertion in both human and HPV sequence. Thicker/darker lines are the result of multiple reads detected.

Table 5.4: Main HPV integration sites in UPCISCC90, PCOC2, PCOC3 and UMSCC47*

Cell line	HPV gene	Human chromosome	Human gene
UPCISCC90	E1	9	C9orf156 Exon3
		9	C9orf156 Intron2
	E6	9	C9orf156 Exon3
	E7	9	C9orf156 Exon3
		9	Non coding +- 200 bp before FOXE1
PCOC2	E2	10	YMEL1 Intron 6 and 7
PCOC3	E7	20	CEPBP
UMSCC47	E7	3	TP63 Exon 11

*: Sites with at least 5 mapped reads are represented in the table.

5.3.7 Validation of p53 status

The p53 status of the established cell lines has been published: UPCISCC90, UMSCC6, 47 and 74a are described as wild-type for p53, while UMSCC4 and 19 are described as mutated (Somers et al. 1992; Bradford et al. 2003; Ferris et al. 2005; Mandic et al. 2005). The p53 status of PCOC2 and PCOC3 was also analysed using PCR amplification and Sanger sequencing (data not shown). Wild-type p53 was detected in both lines. This was consistent with the mRNA-seq data, where no deletions or mutations were observed relative to the wild-type sequence.

5.3.7.1 p53 SNPs and mutations

The mRNA-seq data set was used to assess SNPs and mutations in the p53 gene. IGV was used to identify mismatches with the reference sequence hg19. Mismatches were only counted if they were present in over 5 reads (this was to avoid confounding by counting isolated sequencing errors). The identified mismatches were searched against the NCBI dbSNP database to determine allele frequency and clinical significance. Numbering relates to the human genome hg19 reference sequence.

SNPs were identified in UPCISCC90, UMSCC47, UMSCC74a, PCOC2 and PCOC3. Comparing these with the NCBI dbSNP database showed no known clinically significant correlations. One pathogenic mutation in the DNA binding domain (a homozygous G→A mutation in exon 6), and two SNPs were identified in UMSCC4. A deletion was identified in UMSCC19 consisting of a gap of 22 nucleotides in exon 8. This deletion was also located in the DNA binding domain, resulting in an inactive form of p53.

One SNP, rs1042522, was detected in all lines except UMSCC6. This SNP has previously been associated with an increased risk of cervical cancer. However, its role is controversial as some large-scale studies found no association between this SNP and cancer (Klug et al. 2009). The role of rs1042522 has not been described in the context of head and neck tumours.

5.3.7.2 Consistency of p53 mutation findings with published data

Mutations and deletions were identified in UMSCC4 and UMSCC19. The localisation of the identified mutations in the DNA binding domain of UMSCC4 was consistent with previously published data (Mandic et al. 2005). A 10 bp deletion found in UMSCC19 has also been

described previously, this creates a frame shift (Somers et al. 1992). A non-pathogenic SNP has been reported by some papers in one allele of UMSCC47, this corresponds to what was observed in the current data set (Bradford et al. 2003). The PCR data previously obtained for PCOC2 and PCOC3 was also validated as the same SNPs were identified. Overall, the p53 status of all lines was confirmed.

5.4 Differential gene expression between HPV-positive and HPV-negative cell lines

5.4.1 Global unsupervised clustering

Differential gene expression between the 4 HPV-positive and 4 HPV-negative cell lines was investigated. For each cell line, gene expression was quantified as RPKM and differentially expressed genes were identified using an edgeR analysis as described in section 2.10.3.4 (Mortazavi et al. 2008; Robinson et al. 2010). Data was normalised and the P-values corrected for multiple testing and false discovery issues using the FDR method; a $p<0.05$ cut-off value was used (Benjamini and Hochberg 1995).

The results are visualised as heatmaps based on unsupervised hierarchical clustering (using average linkage and Pearson's correlation as the similarity metric) to group similar data together and reveal any patterns within the dataset. The resultant values are represented as colours, where green is under-expressed (compared to median), black is the median expression and red is overexpressed (Figure 5.5).

The heatmap demonstrates that the HPV-positive and negative samples form discrete groups, with further divisions visible within these groups. It also resulted in grouping of the PCOC lines relative to the other HPV-positive lines. To explore these differences further, a heatmap was generated to show differential gene expression between these groups (Figure 5.6). In these analyses, the number of genes differentially expressed was too high to sort through manually, so a Gene Ontology (GO) analysis was used to identify groups of genes that were differentially expressed as described in section 2.10.3.5 (Figure 5.6).

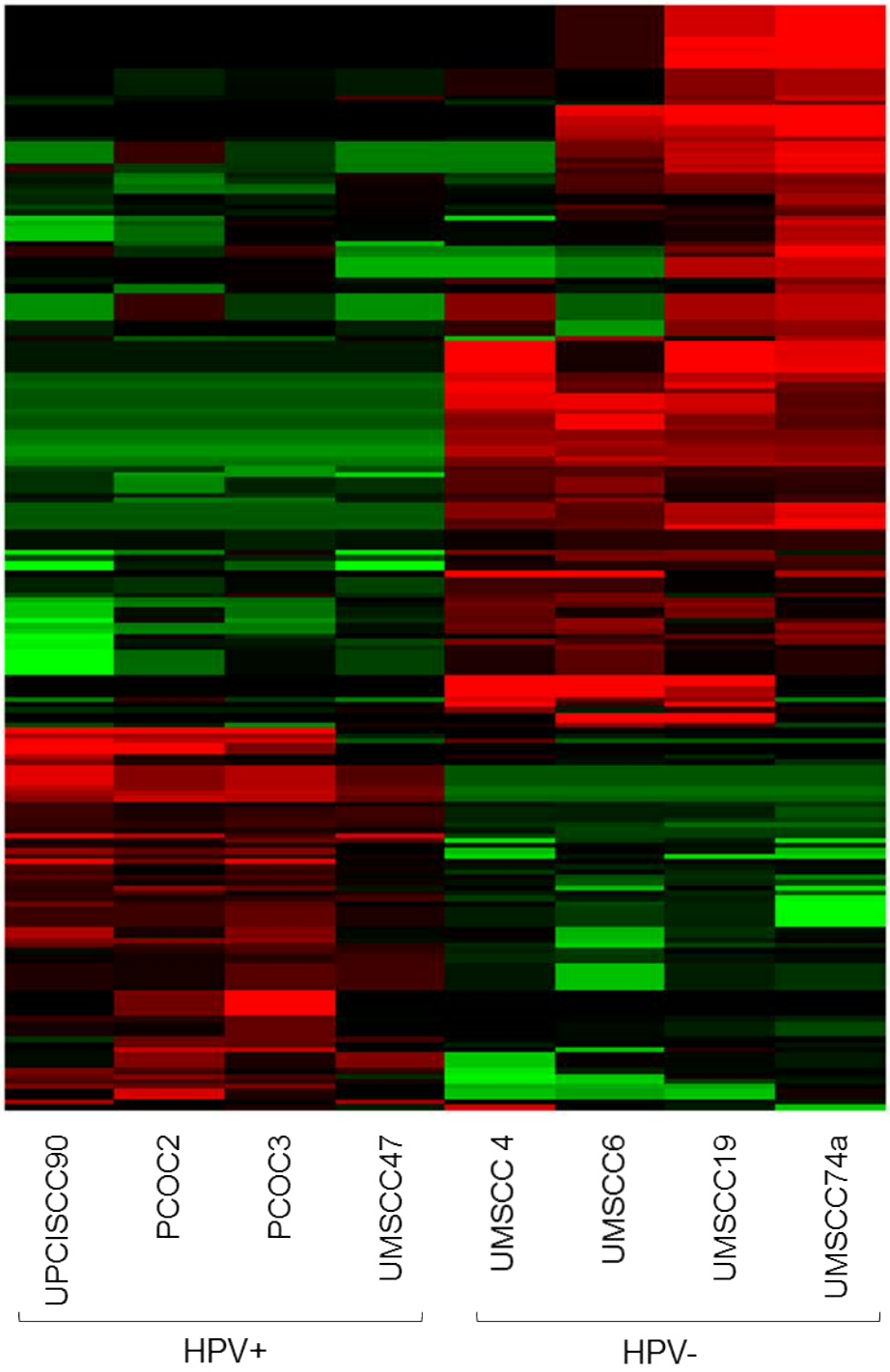


Figure 5.5: Heatmap representing differential gene expression between HPV-positive and negative cell lines

Gene expression was normalised across the lines and colour coded. Black represents the median, green a lower expression than the median value and red a higher expression. After correction for multiple testing, 211 transcripts were significantly differentially expressed between the two groups (FDR $p<0.05$).

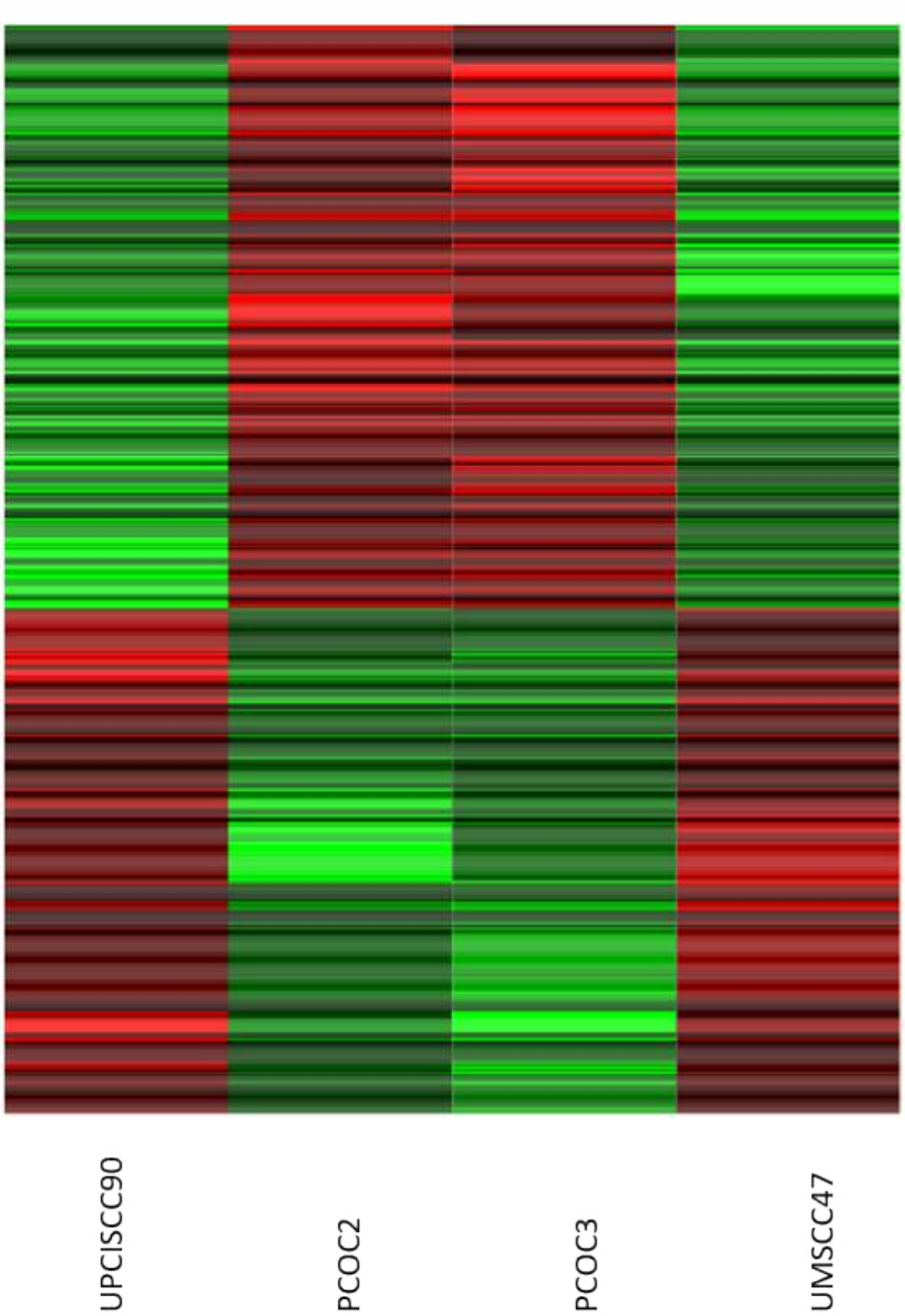


Figure 5.6: Heatmap representing differential gene expression between the PCOC cell lines compared to already established HPV-positive lines

Gene expression was normalised across the lines and colour coded. Black represents the median, green a lower expression than the median value and red a higher expression. After correction for multiple testing, 506 transcripts were significantly differentially expressed between the two groups (FDR $p<0.05$).

5.4.2 Differences in gene expression: Gene Ontology analyses

Genes that showed differences in transcript levels between groups were further investigated using a GO Over Representation Analysis (GO-ORA) (Young et al. 2010). The data was corrected for multiple testing and false discovery using the FDR method. GO analyses use a controlled vocabulary to classify genes under three headings of 1) biological processes (BP) 2) molecular functions (MF) and 3) cellular component (CC).

5.4.2.1 Comparison of GO subsets in HPV-positive and negative lines

The results for the top 500 differentially expressed transcripts are presented. The partition of significant transcripts in to the GO categories after p-value correction for multiple testing was: 51 BP, 5 MF, and 5 CC. The sensitivity of the gene expression analyses was limited due to the lack of biological repeats. Hence these analyses should be regarded as exploratory. The most significantly regulated ontologies are shown below for the BP, MF and CC classifications (Table 5.5, Table 5.6 and Table 5.7).

A variety of BP were differentially expressed between HPV-positive and negative cell lines. This included over-representation of pathways involved in several metabolic processes including glucoronidation, flavonoid and uronic acid metabolism. Several cellular processes including extracellular matrix (ECM) organisation and chromosome organisation were also significantly differentially regulated. Glucoronidation is the transfer of the glucuronic acid component of uridine diphosphate glucuronic acid to a substrate by UDP-glucuronosyltransferase. It is of interest, as it is used in the detoxification of many xenobiotics and bioactive compounds including many drugs and hormones. This is achieved by increasing the polarity of the substrate and increasing efflux via ABC class transporter proteins (Mazerska et al. 2016). This pathway appears to be significantly more active in HPV-negative cell lines, and is also highlighted in the analysis of MF ontologies. Genes involved in the organisation and composition of the ECM were represented in the BP and CC analyses.

Table 5.5: Significant Biological Process ontologies (HPV-positive vs negative cell lines)

GO ID	Term	p.fdr* value
GO:2001029	regulation of cellular glucuronidation	5.07044E-11
GO:0052696	flavonoid glucuronidation	1.21725E-08
GO:0052695	cellular glucuronidation	1.31571E-07
GO:0009812	flavonoid metabolic process	1.93747E-07
GO:0006063	uronic acid metabolic process	5.02586E-07
GO:0019585	glucuronate metabolic process	5.02586E-07
GO:0030198	extracellular matrix organization	1.33965E-06
GO:0043062	extracellular structure organization	1.33965E-06
GO:0045922	negative regulation of fatty acid metabolic process	3.86198E-06
GO:0051552	flavone metabolic process	4.55623E-06
GO:0009813	flavonoid biosynthetic process	4.55623E-06
GO:0045912	negative regulation of carbohydrate metabolic process	4.72082E-06
GO:0042573	retinoic acid metabolic process	1.21982E-05
GO:0010677	negative regulation of carbohydrate metabolic process	1.83022E-05
GO:0045833	negative regulation of lipid metabolic process	0.000105032
GO:0006805	xenobiotic metabolic process	0.000266518
GO:0071466	cellular response to xenobiotic stimulus	0.000335597
GO:0009410	response to xenobiotic stimulus	0.000434141
GO:0009888	tissue development	0.000928789
GO:0032787	monocarboxylic acid metabolic process	0.001857035
GO:0006631	fatty acid metabolic process	0.001857035
GO:0019217	regulation of fatty acid metabolic process	0.003262184
GO:0022617	extracellular matrix disassembly	0.003525732
GO:0001523	retinoid metabolic process	0.015858668
GO:0045143	homologous chromosome segregation	0.019387946
GO:0006109	regulation of carbohydrate metabolic process	0.019387946
GO:0016101	diterpenoid metabolic process	0.025809632
GO:0010511	regulation of phosphatidylinositol biosynthetic process	0.028288198
GO:0070192	chromosome organization involved in meiotic cell cycle	0.030297076
GO:0043436	oxoacid metabolic process	0.037869608

*: The p-value represents data corrected for multiple testing and false discovery. 51 GO were significantly different (p.fdr <0.05) between the two groups. The 30 most significant categories are shown.

Table 5.6: Significant Molecular Function ontologies (HPV-positive vs negative cell lines)

GO ID	Term	p.fdr* value
GO:0015020	glucuronosyltransferase activity	1.10672E-06
GO:0005501	retinoid binding	0.000106315
GO:0019840	isoprenoid binding	0.000166721
GO:0001972	retinoic acid binding	0.001408542
GO:0008194	UDP-glycosyltransferase activity	0.01020897

*: The p-value represents data corrected for multiple testing and false discovery. Five GO were significantly different (p.fdr <0.05) between the two groups.

Table 5.7: Significant Cellular Component ontologies (HPV-positive vs negative cell lines)

GO ID	Term	p.fdr* value
GO:0005578	proteinaceous extracellular matrix	0.001857035
GO:0031012	extracellular matrix	0.003534985
GO:0044421	extracellular region part	0.013170708
GO:0005576	extracellular region	0.017227438
GO:0000795	synaptonemal complex	0.045361421

*: The p-value represents data corrected for multiple testing and false discovery. Five GO were significantly different (p.fdr <0.05) between the two groups.

5.4.2.2 Differential Gene Expression and Gene Ontology analysis between the PCOC lines and previously established HPV-positive lines

The initial differential expression heatmap highlighted clusters of genes with reduced expression in the PCOC lines compared to the two other HPV-positive lines. To further investigate the gene expression profiles of these two groups, analyses of differential gene expression and GO were performed grouping PCOC2 and PCOC3 vs UPCISCC90 and UMSCC47.

Results showing transcripts differentially expressed ($p<0.05$) are presented as a heatmap (Figure 5.6) This highlighted two clear groups of genes, one with higher and one with lower expression in PCOC2 and PCOC3 compared to UPCISCC90 and UMSCC47. A GO analysis was used to look further at which groups of genes were differentially expressed (Table 5.8). The analysis was based on the top 500 significantly differently expressed transcripts. Pathways in the BP ontologies that showed differences between the groups, mainly related to cell differentiation and tissue development, including development of skin. Among the MF ontologies, several pathways relating to fatty acid metabolism were present, along with some involved with glycoprotein binding. In the CC ontologies, the primary differences were in pathways involved with composition of the extracellular space and composition of vesicles.

Table 5.8: PCOC cells versus established HPV-positive cells: Top 10 significant ontologies by GO category

Category*	GO ID	Term	p.fdr** value
BP	GO:0009888	tissue development	8.13507E-07
BP	GO:0051239	regulation of multicellular organismal process	0.000188443
BP	GO:0048856	anatomical structure development	0.000188443
BP	GO:0030154	cell differentiation	0.000328706
BP	GO:0032502	developmental process	0.000328706
BP	GO:0044707	single-multicellular organism process	0.00047516
BP	GO:0048513	animal organ development	0.000710766
BP	GO:0043588	skin development	0.001004473
BP	GO:0048869	cellular developmental process	0.00156408
BP	GO:0009605	response to external stimulus	0.001613676
MF	GO:0050544	arachidonic acid binding	0.003218797
MF	GO:0050543	icosatetraenoic acid binding	0.004844659
MF	GO:0050542	icosanoid binding	0.005199383
MF	GO:1901567	fatty acid derivative binding	0.005199383
MF	GO:0005539	glycosaminoglycan binding	0.005686227
MF	GO:0008201	heparin binding	0.010460524
MF	GO:0050786	RAGE receptor binding	0.017387801
MF	GO:0005102	receptor binding	0.033216283
MF	GO:0035662	Toll-like receptor 4 binding	0.034214762
MF	GO:0001948	glycoprotein binding	0.034263345
CC	GO:0005576	extracellular region	5.9795E-05
CC	GO:0044421	extracellular region part	0.000188443
CC	GO:0005615	extracellular space	0.0007416
CC	GO:0031982	vesicle	0.001613676
CC	GO:0031988	membrane-bounded vesicle	0.00228527
CC	GO:0065010	extracellular membrane-bounded organelle	0.002649475
CC	GO:0070062	extracellular exosome	0.002649475
CC	GO:0005587	collagen type IV trimer	0.002649475
CC	GO:0043230	extracellular organelle	0.002649475
CC	GO:1903561	extracellular vesicle	0.002649475

*: BP = biological process; MF = molecular function; CC = cellular component.

**: The p-value represents data corrected for multiple testing and false discovery (p.fdr <0.05). Ninety-three BP GO were significantly different between the two groups. Thirteen MF GO were significantly different between the two groups. Fifteen CC GO were significantly different between the two groups.

5.4.3 Relative expression of genes of interest

Based on the hypotheses under investigation, the mRNA levels of several genes and ontology groups were specifically assessed in the eight cell lines. These data are illustrated as histograms (with RPKM vs cell line) and heatmaps (where colour indicates fold-change between groups). Most of these genes and pathways of interest were not identified as significantly differentially expressed between the groups in previous unsupervised analyses. Therefore, this analysis is only exploratory.

5.4.3.1 p53 mRNA level

Considerably higher levels of p53 transcripts were observed in the HPV-positive cell lines than in HPV-negative lines (Figure 5.7). UMSCC74a showed higher levels of expression than other HPV-negative lines, which is consistent with the distinct pattern of gene expression compared to the other HPV-negative lines observed previously. In isolation, the difference between p53 transcript levels in the HPV-positive and negative lines is significant (T-test p=0.00567), but after correction for multiple tests the p-value dropped to 0.25.

5.4.3.2 p21 mRNA level

p21 (CDKN1A) is a well characterised p53 responsive gene. To gain insight into the likely levels of active p53 protein, mRNA levels of p21 were also assessed. The pattern of transcription of p21 was not similar to that observed for p53. There was no significant difference in expression between HPV-positive and negative lines (Figure 5.7). The highest levels of p21 transcription were present in UMSCC74a.

5.4.3.3 p16 mRNA level

mRNA for p16 was present in all the HPV-positive lines (Figure 5.7). However, the correlation between HPV status and p16 transcription was not perfect, as transcription of p16 was also evident in the HPV-negative line UMSCC19.

5.4.3.4 PARP1 mRNA level

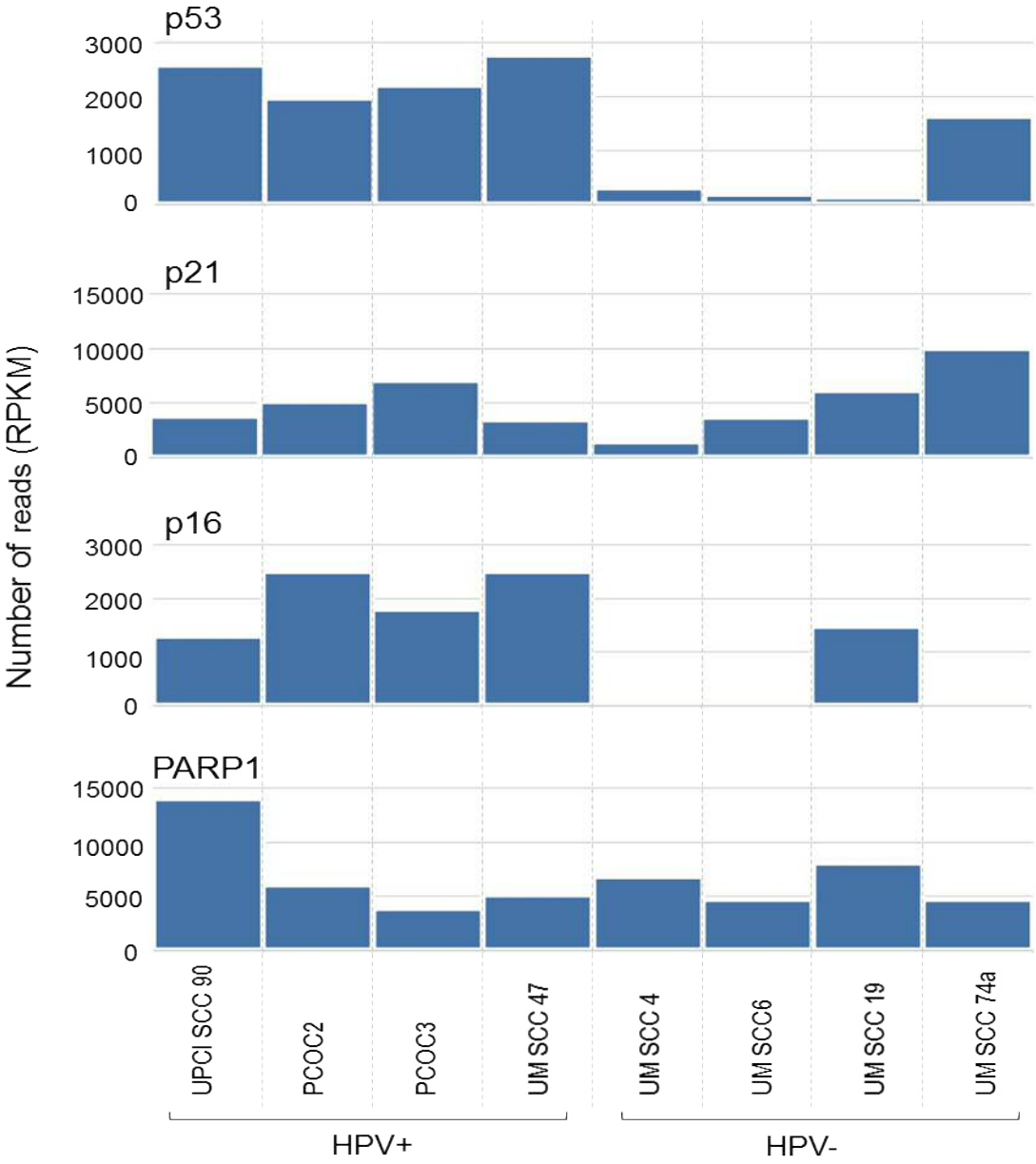
PARP1 is a gene of interest because the PARPi Olaparib is being investigated in clinical trials as a possible specific treatment for HNSCC (https://clinicaltrials.gov/ct2/show/

NCT02308072) and in the current pre-clinical *in vitro* study as a potential treatment for HPV-positive OPSCC cells. PARP1 transcripts were present in all cell lines tested. The highest levels of expression were observed in the HPV-positive cell line UPCISCC90, however, there was no obvious correlation between HPV status and PARP1 transcript levels (Figure 5.7).

5.4.4 Pathways of interest

The expression of genes in key DNA damage/repair pathways was investigated using GO directed analyses. The outputs relating to BER, Double Strand Break Repair (DSBR), and p53 pathways are shown as heatmaps (Figure 5.8, Figure 5.9 and Figure 5.10).

The BER heatmap (Figure 5.8) shows no difference in expression of genes involved in BER between the HPV-positive and negative cell lines. Similar trends were evident in the data for DSBR (Figure 5.9). Differential expression of genes involved in signalling via p53 in response to DNA damage was also assessed. This also showed no difference between the HPV-positive and negative cell lines in levels of mRNA for genes associated with p53 signalling (Figure 5.10).



Gene	Gene ID	Fold change	P-value	fdr p
TP53	NM_000546	-4.26	0.00567	0.25279
p21 (CDKN1A)	NM_000389	1.1	0.83265	1
p16 (CDKN2A)	NM_000077	-5.5	0.08457	0.98212
PARP1	NM_001618	-1.2	0.66105	1

Figure 5.7: mRNA levels of specific genes of interest

The number of reads (RPKM) for p53, p21, p16 and PARP1 are shown for the HPV-positive and HPV-negative lines. This is depicted as histograms in the upper panel and in tabulated form (showing fold change between the HPV-positive and negative groups, and signicance by T-test without and with fdr adjustment for multiple tests) in the lower panel. NB for comparison, a "highly expressed" gene eg. ACTB was transcribed at approximately 50,000 RPKM in these cell lines.

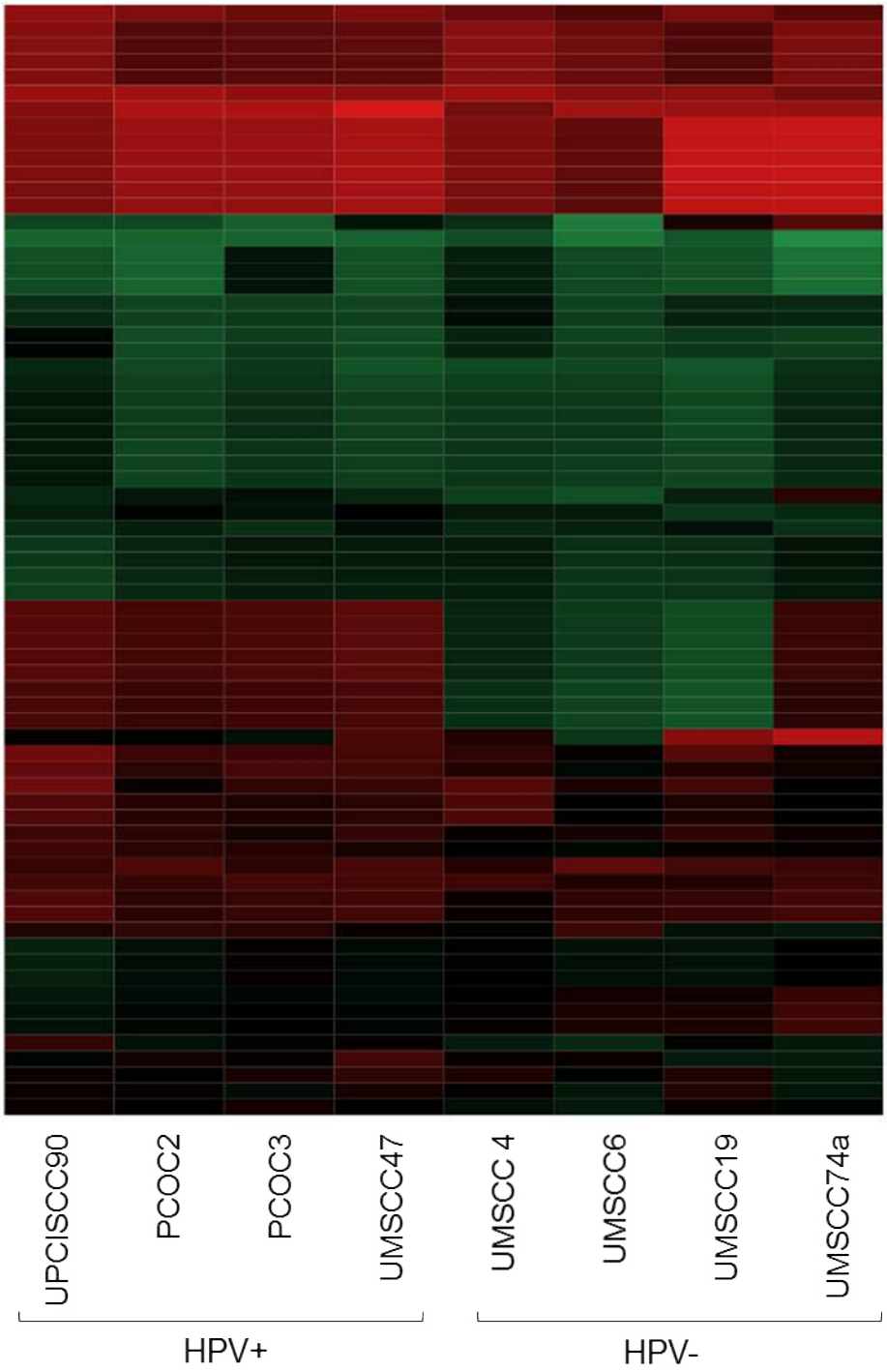


Figure 5.8: Differential gene expression in the BER pathway between HPV-positive and negative cells

Gene expression was normalised across the lines and colour coded. Black represents the median, green a lower expression than the median value and red a higher expression. The block of differentially expressed transcripts in the centre of the figure are mainly p53 transcripts. Heatmap based on analysis of 69 transcripts.

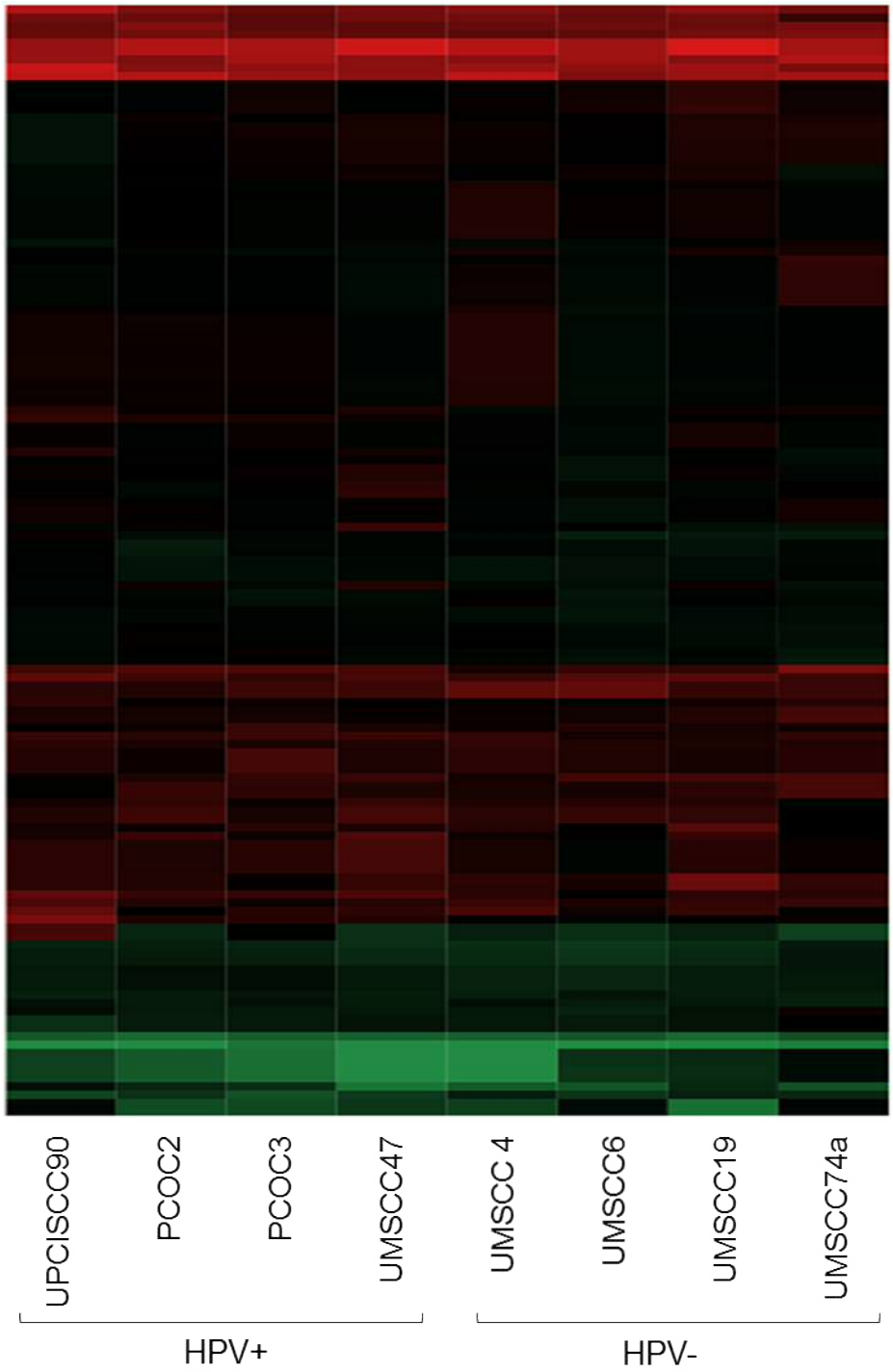


Figure 5.9: Differential gene expression in the DSB repair pathway between HPV-positive and negative cells

Gene expression was normalised across the lines and colour coded. Black represents the median, green a lower expression than the median value and red a higher expression. Heatmap based on analysis of 133 transcripts.

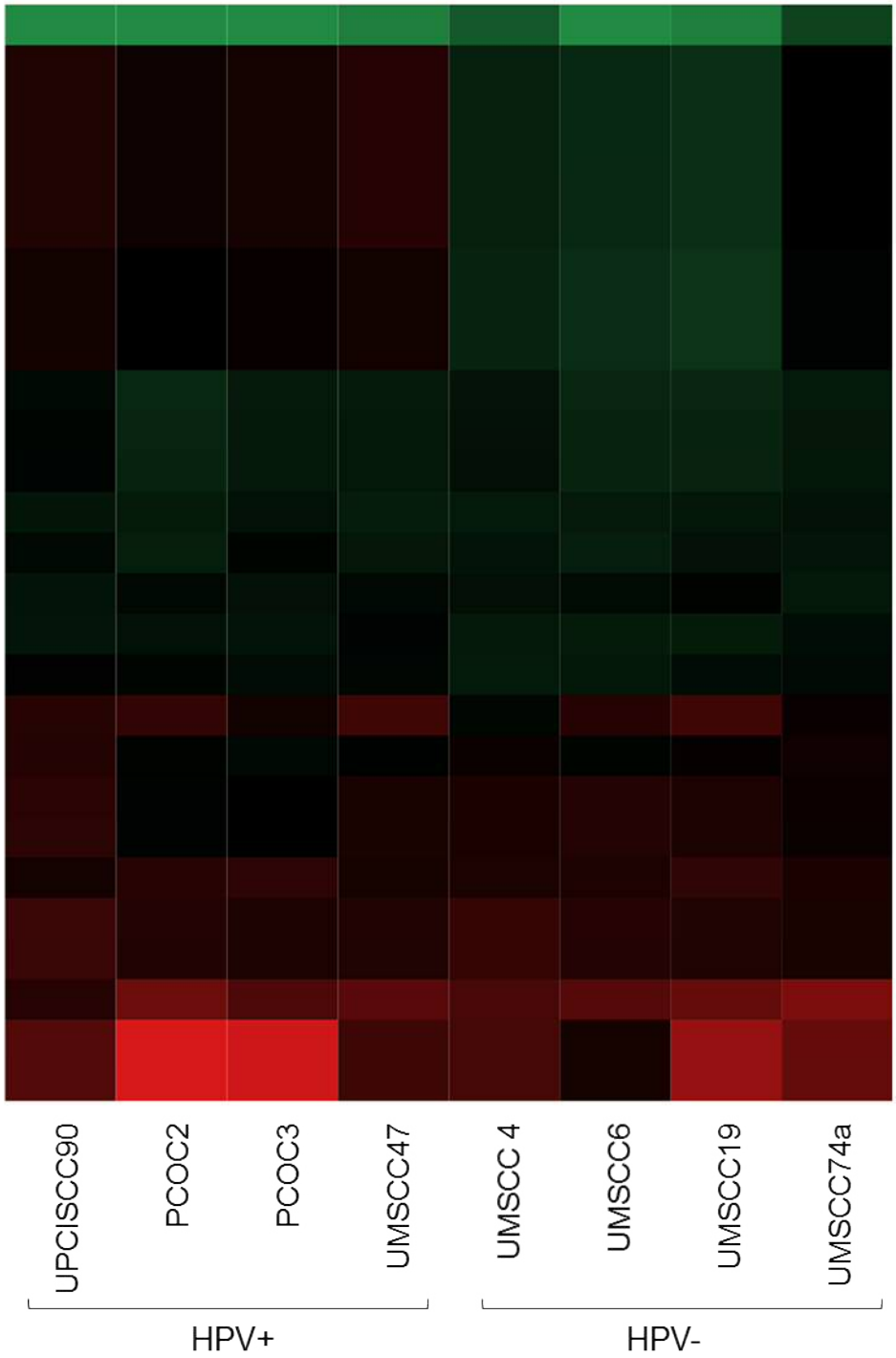


Figure 5.10: Differential gene expression in the "DNA damage response, signal transduction by p53 class mediator" pathway between HPV-positive and negative cells

Gene expression was normalised across the lines and colour coded. Black represents the median, green a lower expression than the median value and red a higher expression. Heatmap based on analysis of 27 transcripts. The botton two rows show transcript levels for NDRG1.

5.5 Discussion

5.5.1 Main Findings

The main findings of the mRNA-seq analyses are summarised below. The analyses:

- Confirmed that all "HPV-positive" cell lines showed active expression of HPV encoded oncogenes.
- Identified varying levels of expression of HPV encoded oncogenes (the highest levels were observed in UPCISCC90 and PCOC3).
- Confirmed the HPV16 variants present: UPCISCC90, PCOC2 and PCOC3 are European; UMSCC47 is an African 2A variant.
- Showed the presence of integrated HPV in the genomes of the HPV-positive lines. Some integrations occurred in sites with pathological associations. In PCOC3, integration occurred in the CEBPB gene, mutation of which is associated with acute promyelocytic leukaemia and glioma.
- Confirmed p53 status for all lines. For the established lines, this was consistent with published data. For the novel lines, the RNA-seq data was consistent with DNA sequencing data. In summary: no inactivating mutations were present in the HPV-positive lines, inactivating mutations were identified in two of the HPV-negative lines (UMSCC4 and UMSCC19).
- Using differential gene expression showed clustering of HPV-positive and HPV-negative cell lines. It also showed different patterns of gene expression between the novel and established HPV-positive cell lines.
- Using GO analysis between HPV-positive and HPV-negative cell lines primarily showed differences in: metabolic processes, including glucoronidation and in organisation and composition of the ECM.
- Using GO analysis between novel PCOC HPV-positive cell lines, and the established HPV-positive cell lines primarily showed differences in: cell and tissue differentiation and development; and composition of the extracellular space.
- Investigating expression levels of individual genes showed:
    - Higher transcript levels of p53 in HPV-positive lines, but no correlation between p21 and p53 transcript levels.

  - Transcription of p16 in all HPV-positive lines, but also p16 transcription in one HPV-negative line (UMSCC19).
  - PARP1 was transcribed in all the lines. The highest levels were observed in UPCISCC90, but there was no correlation between PARP1 levels and HPV status.
- Investigation of expression levels within specific pathways, based on GO analysis, showed no correlation between HPV status and transcription of genes involved in DNA repair pathways or p53 signalling.

The pattern of expression of HPV genes was investigated. This confirmed that there was active HPV gene expression in all cell lines. This, along with the HPV DNA typing and p16 complies with the "gold standard" definition of an HPV-positive tumour (Schache et al. 2011). The characteristics of the cell line panel were also validated, and matched previously published data regarding these lines. Splice site usage was investigated and confirmed the presence of the main oncogenic E6*I and E6*II transcripts. This supports the classification of the novel lines as deriving from true HPV-driven tumours and indicates that they are representative of "typical" HPV-positive OPSCC.

The novel lines PCOC2 and PCOC3 were shown to contain wild-type p53 and infection by a European variant of HPV16. Information regarding the physical state of the virus was obtained through the identification of fusion transcripts and one integration site in each line. PCOC2 expressed high levels of E4 and E5, the presence of these transcripts indicates either the integration of the full HPV16 genome or a mixed infection (presence of HPV both as an episome and in integrated form, this could be determined by Southern blot).

The sites of integration were assessed. Previous studies have suggested that integration of HPV into the host genome can result in insertional mutagenesis and loss of function of tumour suppressor genes (Schmitz et al. 2012). The sites of integration in the HPV-positive cell lines were therefore assessed. In PCOC3, insertion occurred in the CEBPB gene, which contributes to regulation of genes involved in immune and inflammatory responses. Mutations of CEBPB has been associated with acute promyelocytic leukaemia and glioma (Du et al. 2016). Hence, in this case, insertional mutagenesis may be relevant to disease aetiology. This may also be true for the insertion in TP63 in UMSCC47.

In HPV-positive head and neck cancer cells, p53 is degraded following the action of HPV E6 protein, the selective pressure for p53 mutation is therefore low in these cells, resulting in a wild-type p53 genotype in most cases (Stransky et al. 2011). In contrast, p53 mutations are often present in HPV-negative lines. p53 is the most commonly mutated gene in head and neck cancers, and clinically, mutation is significantly associated with shorter survival times and resistance to chemo and radiotherapy (Zhou et al. 2016). The mRNA-seq data confirmed the presence of published mutations in UMSCC4 and UMSCC19 (Somers et al. 1992; Mandic et al. 2005). No mutations were observed in the data for PCOC2 or PCOC3, again suggesting that they are representative of “typical” HPV-driven OPSCC.

The GO analysis comparing HPV-positive and negative cell lines identified primarily differences in levels of transcripts involved in metabolic pathways. This highlights the fundamental cellular differences associated with the differing aetiologies (HPV vs non-HPV). It was interesting that the glucoronidation pathway appeared more active in HPV-negative cells, as it could be speculated that greater detoxification of xenobiotics would be associated with poorer clinical outcomes (Mazerska et al. 2016). There was no significant up-regulation or down-regulation detected in major DNA damage response pathways or cell cycle regulation pathways. Such pathways have been previously identified as differentially expressed in HPV-positive tumours in large datasets but this was not evident in the current data (W. Zhang et al. 2016). This could be regarded as surprising, given the postulated deficiencies in DSBR associated with HPV-positive tumours (Kimple et al. 2013; Rieckmann et al. 2013). High levels of p16 have also been suggested to impair HR-mediated repair of DSB in HPV-positive HNSCC (Dok et al. 2014).

Induction of BER enzymes in HPV-positive cells might also have been expected. This hypothesis was based on data from The Cancer Genome Atlas Project that showed that the APOBEC mutation pattern is common in several HPV mediated cancers including cervical, and head and neck cancers (Roberts et al. 2013). In these tumours, APOBEC-mediated mutagenesis is pervasive and correlates with APOBEC mRNA levels. This is believed to be a consequence of induction of APOBEC activity by HPV encoded genes. For instance, APOBEC3A is a member of the “apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like" family of RNA editing proteins, and is up-regulated in response to expression of the E7 oncoprotein (Warren et al. 2015). As APOBEC mutations may be repaired via BER, it was surprising that the BER pathway

did not appear to be up-regulated in HPV-positive cells. This is however consistent with the lack of induction of APOBEC activities in the mRNA-seq dataset (data not shown).

The relative expression of several specific genes of interest was assessed. p53 transcript levels were consistently higher in HPV-positive cells. Regulation of p53 activity at the protein level is relatively well understood. In normal circumstances, p53 mRNA is rapidly transcribed and translated, but p53 protein is then ubiquitinated by Mdm2, and rapidly degraded. This results in low "steady-state" levels of p53 protein, but allows for rapid accumulation if proteolysis is blocked. Regulation of transcription of p53 itself, is less well studied and less well defined. Transcription appears to be regulated in a cell cycle dependent manner, and peaks in mid-S-phase (Reisman et al. 2012) but is also stimulated by serum factors (Reich and Levine 1984) and by c-Myc (Reisman et al. 1993). Interestingly, overexpression of Myc and HPV E7 has previously been shown to result in increased transcription of p53 (Liu et al. 2007) and numerous lines of enquiry suggest up-regulation of Myc in HPV infected cells and HPV associated tumours, hence it may be that these factors account for the increased p53 transcript levels in the HPV-positive lines (Strickland and Vande Pol 2016; Zhang et al. 2016).

An alternative explanation is that the difference between p53 mRNA levels in HPV-positive and negative cells is consistent with evidence that p53 protein stimulates transcription of p53 mRNA (Deffie et al. 1993; Kishore et al. 2007). The work by Deffie et al. also demonstrated that mutated p53 could not transactivate transcription from the p53 promotor. Hence, if there was some residual p53 activity in the HPV-positive cells, then this could stimulate p53 transcription, but this effect would not be present in the cells carrying inactivating mutations in p53 (UMSCC4 and UMSCC19). However this would not explain low levels of HPV transcription in UMSCC6 where p53 is apparently wild-type.

p21 (CDKN1A) is a relatively well characterised p53 responsive gene, and is often used a marker to indicate presence of active p53 (Riley et al. 2008; Abbas and Dutta 2009). The gene encodes a potent CDK inhibitor that binds to and inhibits the activity of CDK2 and CDK4 complexes, and thus functions as a regulator of cell cycle progression at G1. p21 expression is tightly controlled by p53, and mediates p53-dependent cell cycle G1-phase arrest in response to a variety of stress stimuli. There was no significant difference in expression between HPV-positive and negative lines, and no apparent correlation between transcript levels for p53 and p21. This suggests that although p53 was transcribed in the HPV-positive cell lines, functional

p53 protein appears not to be present. The highest levels of p21 transcription were present in UMSCC74a. This would be consistent with the presence of functional p53 in this line.

p16 over-expression arises from increased activity of E2F transcription factors secondary to binding of Rb by HPV E7 protein. p16 IHC is used diagnostically in conjunction with HPV DNA testing to identify HPV associated tumours (Schache et al. 2011). It was therefore expected that p16 would be transcribed at a higher level in HPV-positive cell lines. However, it was important to confirm this and show that the cell lines reflected the p16 IHC status of the original tumours. The high levels of p16 in the HPV-positive lines also confirm that proliferation in these cells, and in the original tumour, was driven by HPV, and these are hence appropriate models for investigation of HPV-driven disease. Transcription of p16 in the HPV-negative line UMSCC19 was not expected, but does demonstrate the acknowledged limited specificity of p16 testing to identify HPV-driven tumours (Evans et al. 2013).

Levels of PARP1 mRNA were assessed. This was relevant as PARP inhibition is one of the main themes of this thesis. PARP1 transcripts were present in all HPV-positive and negative cell lines. The lack of correlation with HPV status is consistent with the similar levels of other genes involved in BER between the two groups. UPCISCC90 showed very high levels of expression. This might suggest that this cell line may be more dependent upon ongoing PARP expression.

Several specific potentially relevant pathways were also investigated using GO directed analyses including BER, DSBR, and p53. There were no substantial differences in these pathways according to HPV status. Based on the suggestion that APOBEC3A induced in HPV-positive cells can induce DNA DSB, and activate the global DNA Damage Response (Landry et al. 2011), and high levels of reactive oxygen species in HPV mediated tumours (Marullo et al. 2015), it might be expected that HPV-positive cells would show constitutive activation of DNA repair pathways, but this is not supported by the mRNA-seq data.

Finally, mRNA levels for genes in the ontology “DNA damage response, signal transduction by p53 class mediator” were assessed. This was relevant to the question of whether functional p53 was present. The data appear consistent with the absence of p53 signalling in all the lines. This could be caused by a genuine absence of stress and/or DNA damage in the cells examined, or by the absence of functional p53 protein, or a combination of these factors.

Taken together, these data do not suggest differences in the induction and repair of DNA damage, between HPV-positive and negative cells.

5.5.2 Exclusion of UMSCC6

The UMSCC6 cell line was included in the mRNA-seq analyses but not in the subsequent experiments. This was because this cell line is exceptionally slow growing, with a doubling time of several weeks. This made it unfeasible to include it in clonogenic assays. The mRNA-seq data has been included however as it provides a useful comparison with the HPV-positive lines.

5.5.3 Strength and weaknesses

mRNA seq is a very sensitive technique, it provided in-depth analysis of the panel of cell lines and enabled the validation of their characteristics prior to use in other assays. This represents the first characterisation of these lines as well as the first in-depth characterisation of the novel HPV-positive PCOC lines.

The main weakness was the small sample size which limited the power to detect differences in gene expression between the groups. Similarly, the data was based on one sequencing reaction for each cell line, which again limits statistical power. Hence these results should be regarded as exploratory and not definitive. To confirm these differences would require replication in a larger data set.

It is difficult to be certain how well the gene expression patterns observed in cells grown in culture represent the gene expression patterns present in the original tumour. However, it is important to note that the growth conditions were standardised as far as possible (e.g. identical concentrations of FBS and culture conditions), and the differences observed hence appear likely to reflect the different molecular biology, including mutational profile and HPV status, of the original cells.

5.5.4 Conclusion

The main conclusions are that:

- The novel PCOC cell lines show characteristics of “typical” HPV-driven tumours (HPV genes expressed from integrated virus, p53 wild-type, p16 positive).
- There was no evidence for differences in induction or repair of DNA damage, between HPV-positive and negative cells.

Chapter 6 - Investigation of PARP inhibition in oropharyngeal cancer cell lines

6.1 Introduction

HPV-positive OPSCC patients tend to respond better than HPV-negative patients to radiotherapy (Dayyani et al. 2010). *In vitro,* some HPV-positive cell lines demonstrated an increased sensitivity to radiation compared to HPV-negative cells together with a G2 cell cycle arrest (Kimple et al. 2013; Rieckmann et al. 2013; Ziemann et al. 2015). This difference in response has been attributed to a deficiency in repair of DNA DSB in HPV-positive cells leading, to an accumulation of unrepaired breaks (Rieckmann et al. 2013). However, there are discrepancies in the literature with regards to the events following cell cycle arrest. One study reported observing apoptosis (Rieckmann et al. 2013) while others did not (Kimple et al. 2013; Ziemann et al. 2015).

The DSB repair deficiency in HPV-positive cells could potentially be targeted using a synthetic lethal therapeutic strategy. This synthetic lethal approach has already been used in BRCA deficient breast and ovarian cancers (Evers et al. 2008; Audeh et al. 2010). This provides a precedent for a similar approach in other DNA repair deficient tumours. We hypothesised that synthetic lethality could be achieved via the combination of SSB repair inhibition by the PARPi Olaparib and defective DSB repair in HPV-positive HNSCC cells.

In vitro studies in HNSCC cells lines have concentrated on the efficacy of combining PARPis with radiotherapy (Güster et al. 2014). Combination treatment is also being tested in an ongoing clinical trial regardless of HPV status (ORCA-2). However, there is still a lack of *in vitro* data for single treatment with Olaparib, especially with regard to a potential differential response between HPV-positive and negative cell lines.

An alternative explanation for the greater sensitivity of HPV-positive tumours to radiotherapy, might be that these tumours have low levels of functional p53 protein present due to incomplete degradation mediated by the HPV E6 protein, as opposed to entirely inactive p53 protein following p53 gene mutation (Kimple et al. 2013). Levels of wild-type p53 were therefore also assessed.

The work described in this chapter aimed to determine the sensitivity of a panel of HPV-positive and negative cell lines to a single dose of Olaparib and to investigate the mechanism of sensitivity.

It was hypothesised that:

1. Sensitivity to Olaparib would correlate with HPV status, and HPV-positive lines would be more sensitive to Olaparib than HPV-negative lines.
2. HPV-positive cells would accumulate higher levels of unrepaired DNA DSB than HPV-negative lines following treatment with Olaparib, as a consequence of the conversion of SSB to DSB during replication, coupled with defective repair of DSB in these lines.
3. The accumulation of DNA DSB would be associated with cell cycle arrest and possibly apoptosis.
4. In HPV-positive cells, cell cycle arrest and apoptosis may be linked to activation of p53.
5. High basal levels of PARP1 may correlate with sensitivity to Olaparib.

6.2 Optimisation of the Olaparib dosing protocol

A series of experiments were performed to optimise the various assays prior to investigating the main aims.

6.2.1 DMSO concentration

Olaparib is soluble in water below 0.002 mg/ml, which was lower than the concentrations needed for the experiments. Therefore the first dilution step was done in DMSO. However this led to residual DMSO in cell culture media and the possibility of DMSO mediated cytotoxicity irrespective of any drug action. The survival of PCOC2 cells cultured with 0.1%, 0.5% and 1% DMSO was measured to select a permissible final DMSO concentration which did not reduce survival. Cells were cultured in triplicates until 80% confluence was reached in the untreated control. The number of viable cells in each plate was then counted (Figure 6.1). A 1% concentration of DMSO reduced the number of viable cell by 43%, while 0.5% and 0.1% DMSO caused increases of 33% and 37% of viable cells respectively. A final concentration of 0.1% DMSO was selected for further experiments. DMSO concentrations were then standardised to 0.1% in all experiments to avoid variations in DMSO levels affecting the results.

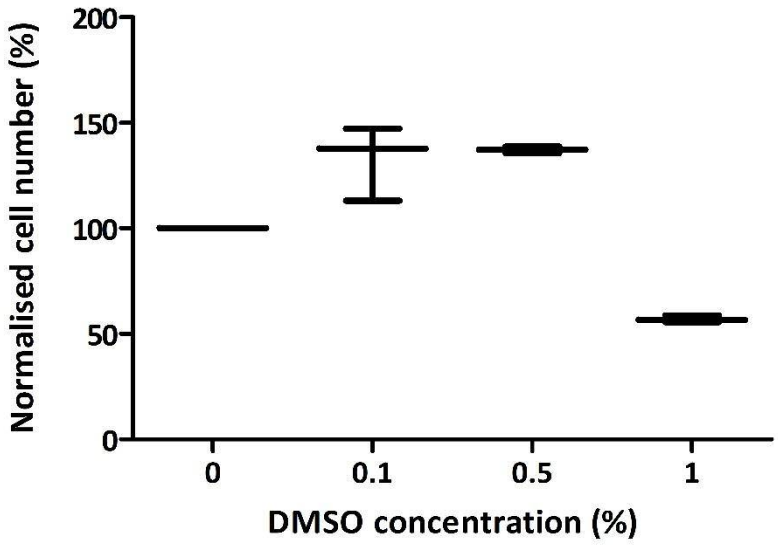


Figure 6.1: Effect of DMSO on PCOC2 cell culture

Viability counts of PCOC2 cells cultured with varying concentrations of DMSO were normalised to the untreated control. Data is represented as a box plot, indicating the median, minimum and maximum values. Cells were plated at 500000 cells per plate and incubated until the untreated plate reached approximately 80% confluence (c.7 days). 3T3 cells were removed by washing and viable cells were stained with trypan blue and counted manually using a haemocytometer. The median cell number was reduced by 43% in the presence of 1% DMSO. Concentrations of 0.1% and 0.5% DMSO caused a 37% and 33% increase of the median cell number respectively.

6.2.2 Seeding density

The seeding density (i.e. the number of cells initially plated) was optimised for each cell line as it represents a key parameter of clonogenic assays. A higher density than necessary could result in colonies not emerging from single cells, and could also mean control plates became confluent too soon. Alternatively, too low a seeding density would hamper the colony formation and increase the assay duration.

Some cell lines did not sustain growth at low density, as previously reported by Rieckmann et al. (Rieckmann et al. 2013). This issue was addressed through the use of 3T3 feeder cells that can promote growth by providing cell-cell contact. Clonogenic assays vary in duration (from 10 days to over 3 weeks), influenced by the plating efficiency and population doubling time of each cell line. A wide range of seeding densities (50 - 500000 cells per plate) were tested using UMSCC47 and UMSCC4 (data not shown). A selected narrower range of densities was then tested with each cell line in triplicates, with or without the presence of 3T3 feeder cells. The appropriate seeding density for PCOC2 and PCOC3 had already been determined in pilot experiments (4400 cells). Hence for these lines, only the 3T3 feeder cells requirements were tested. UMSCC4 was already known to grow very slowly at low densities without 3T3 feeder cells (Prof T. Carey, personal communication). Therefore, culture without 3T3 feeder cells was not attempted for this line. Colony morphology and density were observed 7 days after seeding and the media replaced. From day 10 onwards, plates were observed daily until the experiment was stopped, and the cells stained at day 15. Wide variations in optimal seeding densities were observed across the cell line panel (2500-12500 cells/plate). UMSCC47 and UPCISCC90 seeded at 20000 cells/plate were the only lines able to grow without 3T3 feeder cells. Overall cell growth at the lowest densities benefited from the support of 3T3 feeder cells (Figure 6.2). To homogenise conditions, it was decided to optimise all seeding densities with the presence of 3T3 feeder cells. Based on these results, seeding densities of 2500 cells for UMSCC47, 12500 cells for UPCISCC90, 7500 cells for UMSCC4, 7500 cells for UMSCC19 and 3000 cells for UMSCC74a were selected. A 10-15 days assay duration resulted in countable colonies for all cell lines.

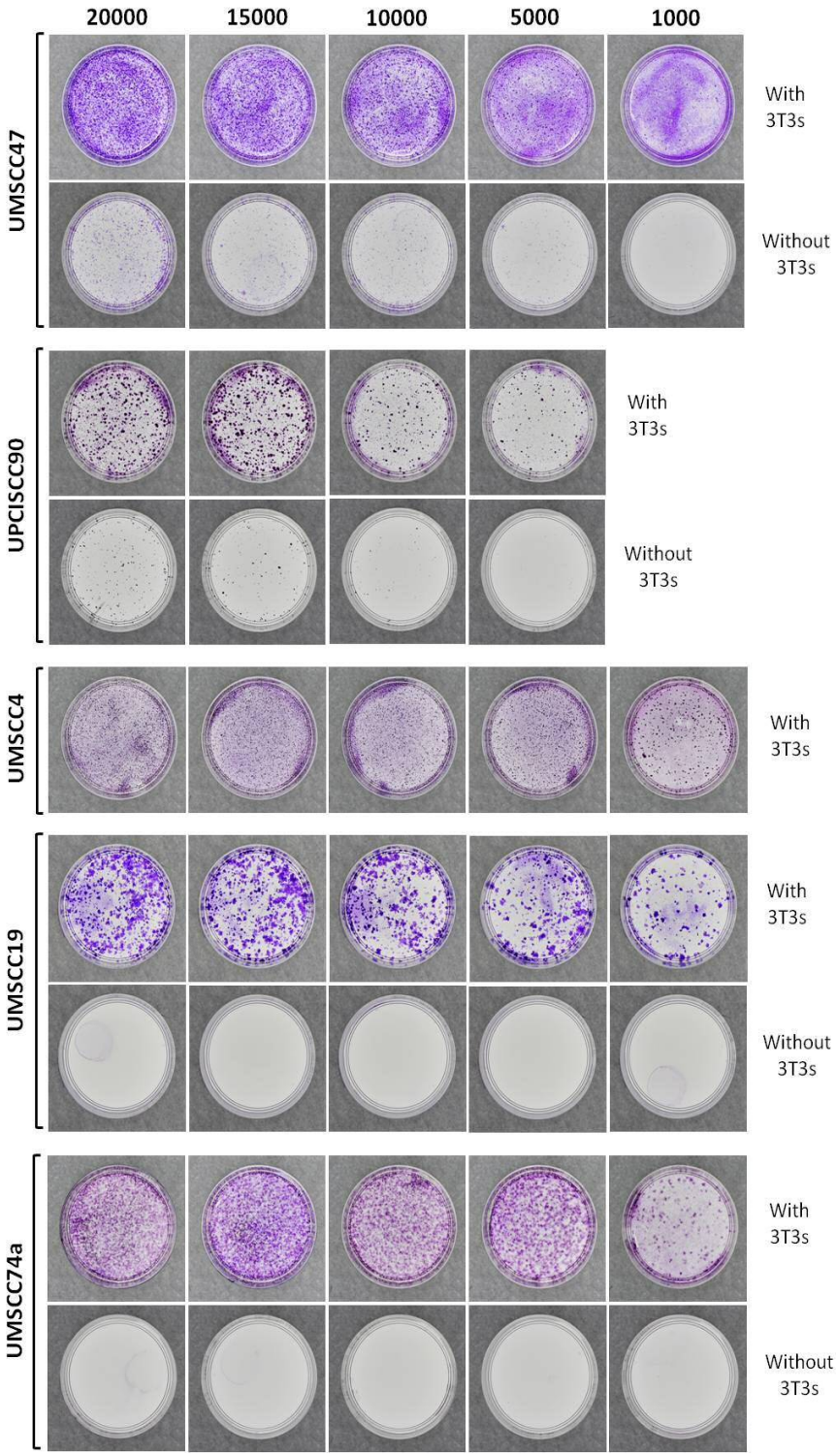


Figure 6.2: Effect of seeding density and presence of 3T3 feeder cells on colony formation with UMSCC47, UPCISCC90, UMSCC4, UMSCC19 and UMSCC74a

Cells were seeded at densities between 20000 and 1000 cells with and without the support of 3T3 feeder cells. Plates were incubated at 37°C for 10-15 days, then stained with crystal violet and photographed (images above). Colonies >50 cells were counted. UMSCC19 and UMSCC74a did not sustain growth without feeder cells. All lines had an enhanced capacity to form colonies with 3T3 feeder cells.

6.3 Colony formation capacity after a single dose of Olaparib

The effect of treatment with Olaparib was assessed using clonogenic assays in a panel of 4 HPV-positive and 3 HPV-negative cells lines (as described in Chapter 5). In addition, normal non-malignant keratinocyte cells (HEKn) were included as a control. One representative example of raw results using PCOC2 is shown in Figure 6.3. Examples for the remaining cell lines can be found in Appendix 2.

A wide range of responses to treatment was evident across the cell line panel (Figure 6.4). Within the HPV-positive lines, two groups with distinct colony formation capacity in response to Olaparib were present. PCOC2 and UPCISCC90 were the most sensitive lines. Their Surviving Fraction (SF) decreased in a dose-dependent manner with increasing exposure to Olaparib. A single treatment of 0.1 μM Olaparib reduced their mean SF by 30.85% and 45.8% respectively. The only significant difference between these lines was detected after 0.1 μM treatment (two way ANOVA $p<0.05$). PCOC2 and UPCISCC90 were also significantly more sensitive than the other HPV-positive lines at doses over 0.1 μM ($p<0.05$; two way ANOVA with Bonferroni post-test; Figure 6.4 A).

PCOC3 and UMSCC47 demonstrated lower sensitivity to Olaparib treatment compared to UPCISCC90 and PCOC2. These lines followed a similar dose response with no significant differences between them detected below 5 μM. The sensitivity increased at high doses with mean SF at 10 μM of 41.3% for UMSCC47 and 21.4% for PCOC3. UMSCC47 was only sensitive at 5 and 10 μM Olaparib treatment (one way ANOVA $p\leq0.01$). While PCOC3 was sensitive from 1 μM (one way ANOVA $p<0.05$).

HPV-negative lines also varied in their colony formation capacity following treatment with Olaparib (Figure 6.4 B). Two lines, UMSCC4 and UMSCC19, were not sensitive to treatment with doses below 5 μM. HEKn and UMSCC74a displayed higher Olaparib sensitivity compared to UMSCC4 and UMSCC19. They followed a similar dose response with no differences detected between the two curves (two way ANOVA $p<0.05$). Individually, UMSCC74a was significantly sensitive from 0.5 μM while HEKn displayed sensitivity at all doses (one way ANOVA $p<0.05$). This possibly suggested that UMSCC74a was an unusual case and did not behave like the other HPV-negative tumour cell lines included in the panel. The highest

doses of Olaparib tested (10 μM) caused a reduction of colony formation of nearly 100% in 5/8 lines.

In summary, three distinct categories of drug response were observed. These were defined according to sensitivity to Olaparib at 0.5-1 μM, at which the widest range of sensitivities were observed (Figure 6.5). Resistant cell lines (SF>80%) included the HPV-positive PCOC3 and UMSCC47 plus the HPV-negative lines UMSCC19 and UMSCC4, medium sensitivity cell lines (SF 40-80%) included HPV-negative HEKn and UMSCC74a, and sensitive cell lines (SF<40%) included HPV-positive cell lines PCOC2 and UPCISCC90.

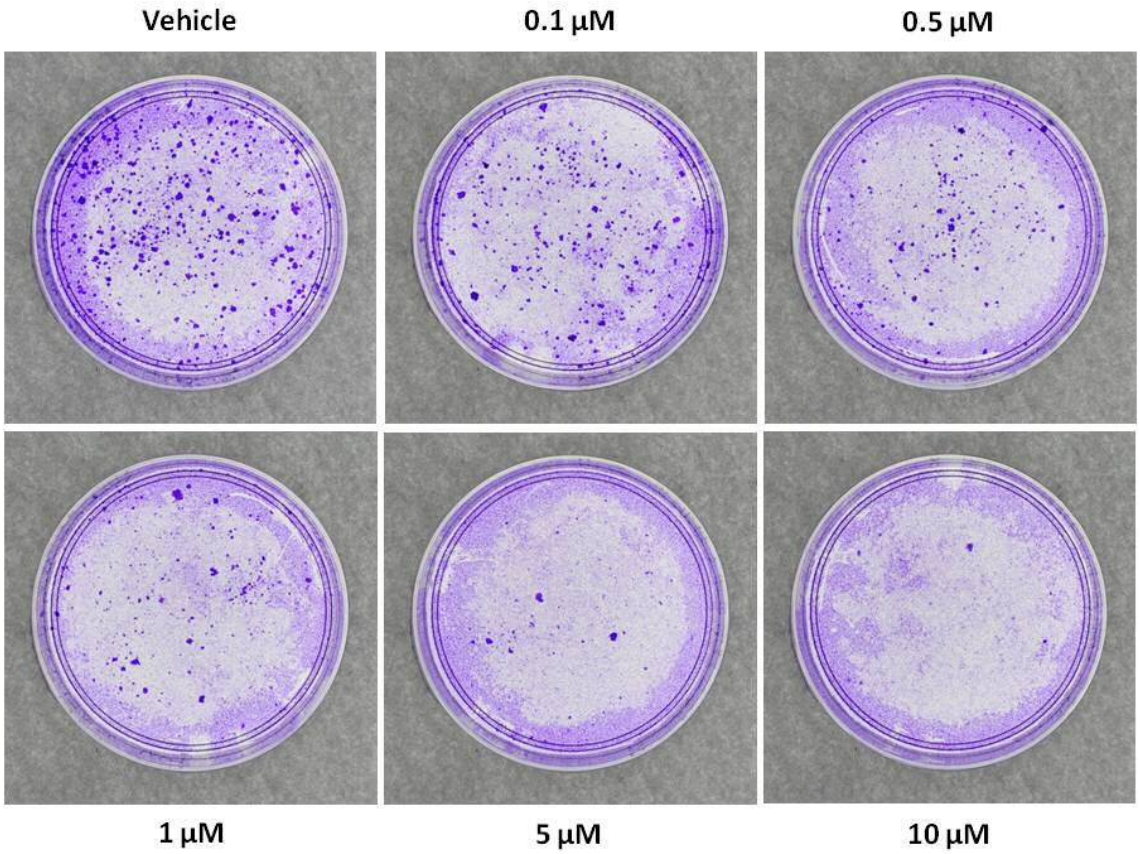


Figure 6.3: Example of PCOC2 clonogenic assay result

Cells were plated at low density, at 24 h media with/without Oparib was added, then plates were incubated for 10-15 days. Plates were stained using crystal violet and photographed (images above). Colonies of >50 cells were counted. The ability of PCOC2 to form colonies clearly decreased in a dose-dependent manner in response to increasing concentrations of Olaparib.

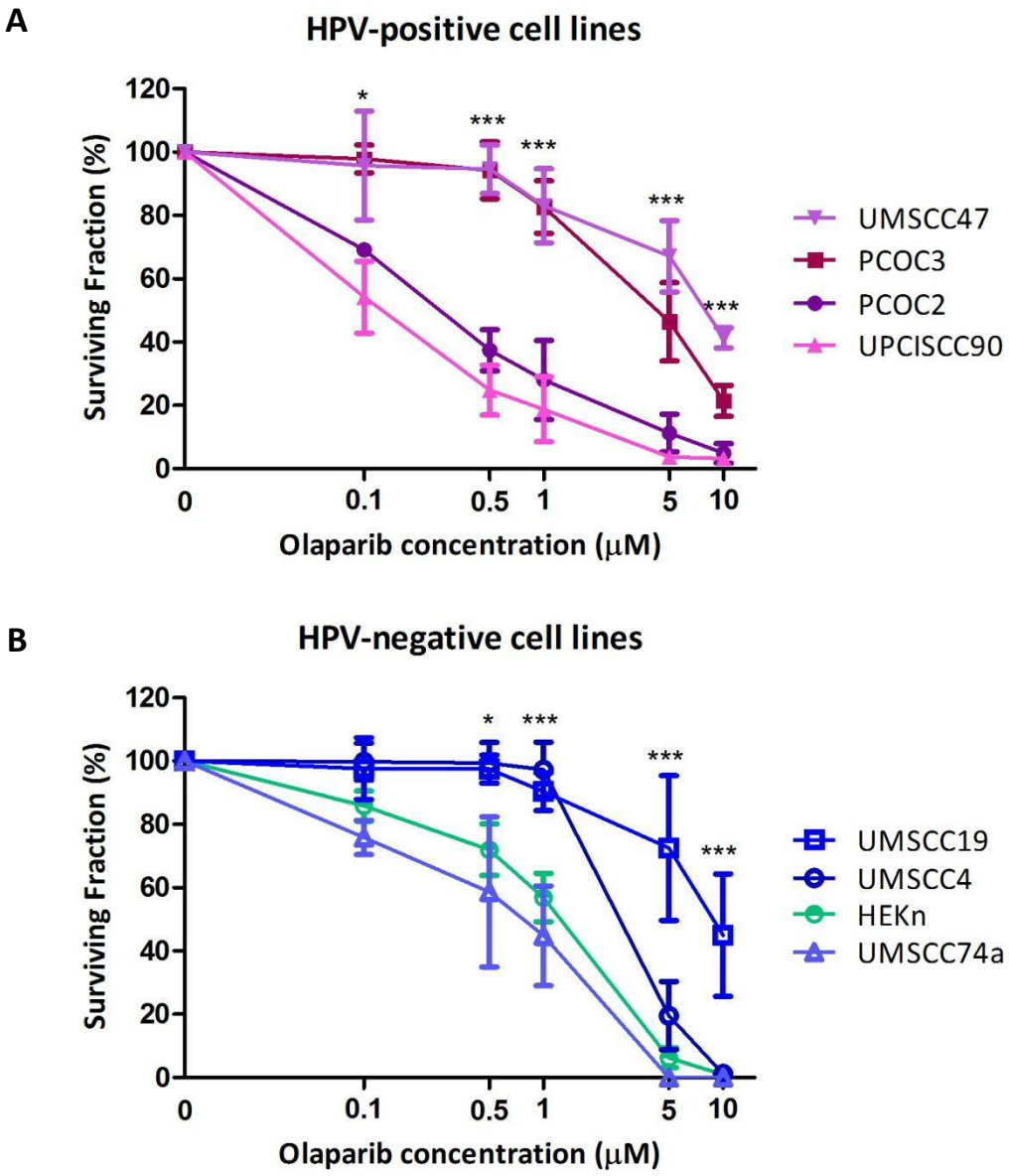


Figure 6.4: Surviving fractions of the cell line panel in response to Olaparib (by HPV status)
The results presented are representative of at least 3 repeats, including triplicates within each experiment (i.e. there are a minimum of 9 data points per condition). Statistical significance between the lines was assessed using two-way ANOVA with Bonferroni post-test. Error bars indicate SD.
Panel A (HPV-positive cell lines): The colony formation capacity of PCOC2 and UPCISCC90 was reduced compared to PCOC3 and UMSCC47 following a single dose exposure of Olaparib from 0.1 μM to 10 μM. As an example the result of the comparaison between UPCISCC90 and UMSCC47 is illustrated (*** $p \leq 0.001$, *$p < 0.05$).
Panel B (HPV-negative cell lines): UMSCC19 and UMSCC4 were least sensitive to Olaparib. Their colony formation capacity was reduced starting at 5 μM and 10 μM, whilst UMSCC74a and HEKn were more sensitive at lower doses. As an example the result of the comparison between UMSCC19 and UMSCC74a is illustrated (*** $p \leq 0.001$, *$p < 0.05$).

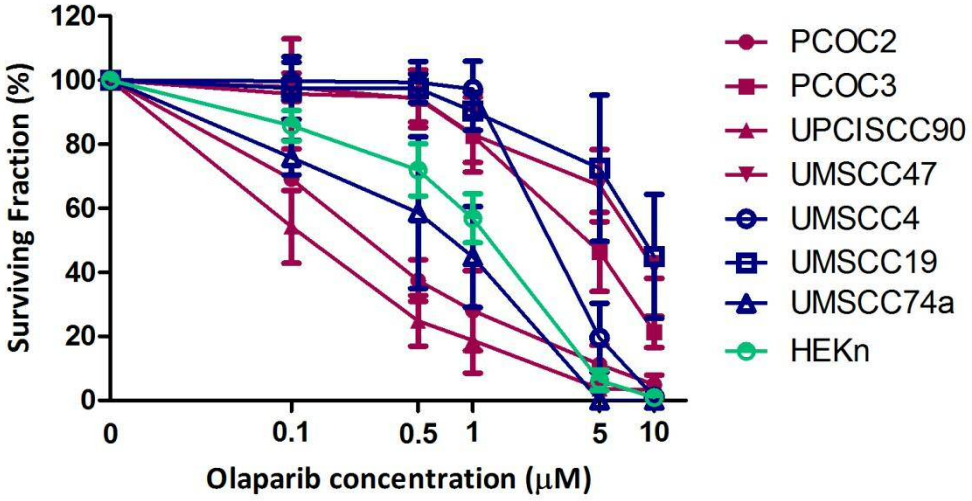


Figure 6.5: Surviving fractions of the cell line panel in response to Olaparib (all lines)
All lines were sensitive to high doses of Olaparib (10 μM), however at doses between 0.5-1μM, the surviving fractions differed significantly between lines (ANOVA $p<0.001$). The cell lines were divided into 3 categories based on sensitivity at 0.5-1 μM. The "sensitive" (SF<40%) cell lines were two HPV-positive lines (UPCISCC90 and PCOC2); "intermediate" (SF 40-80%) cell lines were an HPV-negative line and normal human epidermal keratinocytes (UMSCC74a and HEKn); and "resistant" (SF>80%) lines included 2 HPV-positive and 2 HPV-negative lines (UMSCC19, 4, 47 and PCOC3).

6.4 DSB detection protocol development

Measurement of phosphorylation of histone H2AX is a sensitive method for quantifying DNA DSB (Sharma et al. 2012). Following induction of DSB, serine 139 located on the C-terminal tail of the H2AX histone is quickly phosphorylated by kinases including ATM and ATR. This modified histone is referred to as γ-H2AX. The presence of γ-H2AX facilitates chromatin re-organisation to allow access for DNA damage repair proteins. De-phosphorylation occurs once the DSB has been repaired. γ-H2AX can hence be used as an indirect marker for the presence of DSB (Bonner et al. 2008). To allow quantification of γ-H2AX, cells require lysis to release the nuclei, followed by staining with an anti-serine 139 γ-H2AX antibody and/or a DNA stain before analysis by flow cytometry.

6.4.1 Lysis and staining specificity

Whole cells and nuclei populations were stained and analysed separately by flow cytometry to evaluate the lysis efficiency and staining specificity. HeLa cells were used for protocol development experiments as they were easily available. For assay optimisation, cells were treated with Ionising radiation (IR), rather than Olaparib, as IR is a well-characterised inducer of DNA damage, including DNA DSB. Cells were left untreated or exposed to 10 Gy IR and used after 30 min. Visualisation of data on a forward versus side scatter plot resulted in two distinct populations (Figure 6.6 A). The population comprised of whole cells was larger and more granular than the nuclei. Staining patterns also differed between whole cells and nuclei. The staining of whole cells using the γ-H2AX antibody was minimal, indicating the presence of mostly intact cell membranes (Figure 6.6 B). Staining for γ-H2AX, indicating the presence of DSB, was detected in the nuclei population, and increased following 10 Gy radiation (Figure 6.6 B).

In summary, these experiments demonstrated that the cell lysis procedure was effective in producing a nuclei population, and that staining with the γ-H2AX antibody appeared to correlate with the presence of DNA DSB.

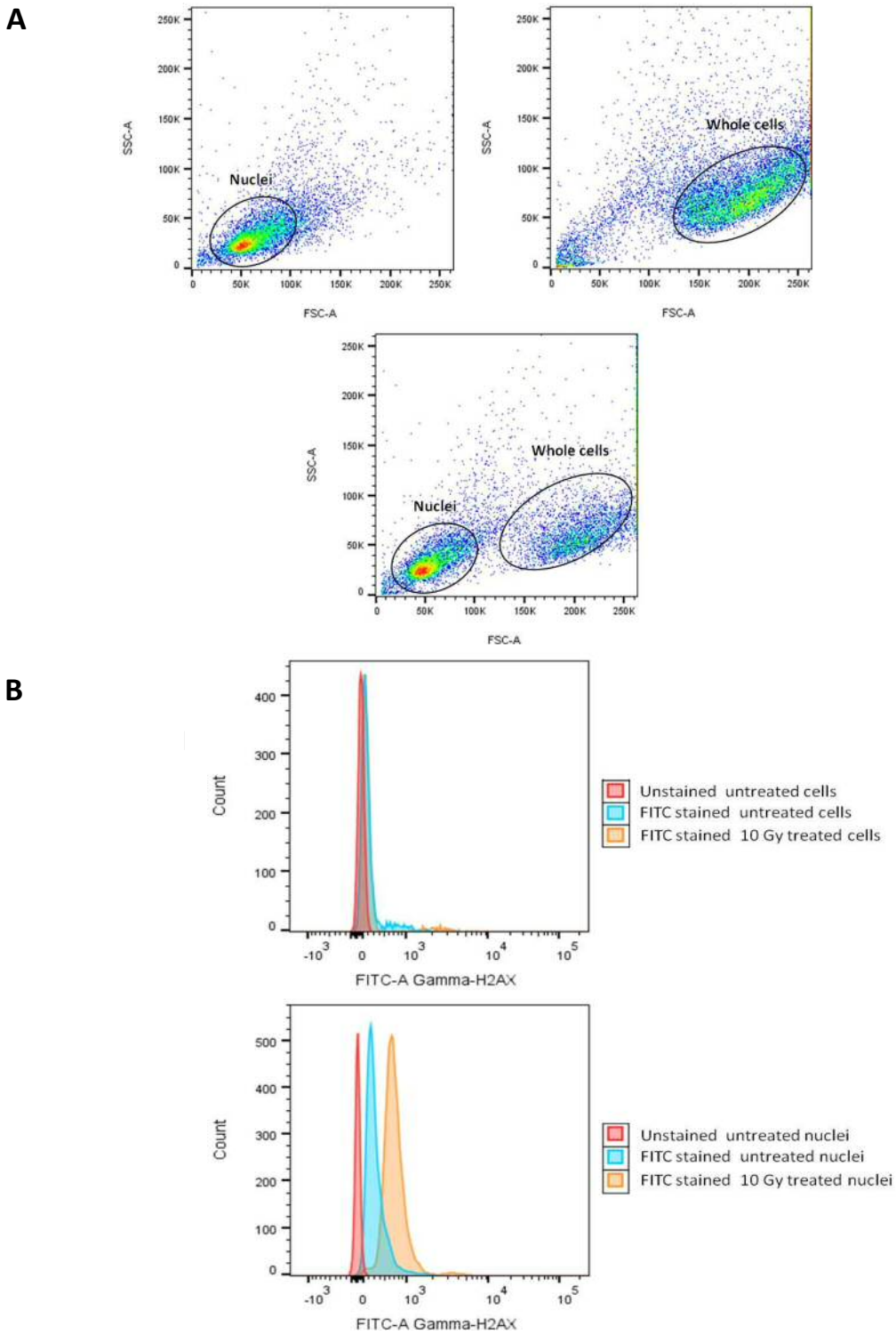


Figure 6.6: Visualisation of HeLa whole cells and nuclei population and quantification of DSB following irradiation

A: The gated nuclei population (first panel) and the whole cells population (second panel) and both populations (third panel). Cells and nuclei are distinguished by their size and granulosity.
B: Shows background and staining with γ-H2AX antibody in whole cells (first panel) and nuclei (second panel) untreated and after 10 Gy IR. Fluorecence is minimal in whole cells but detected in untreated in irradiated nuclei. This demonstrated increased antibody binding following induction of DSB.

6.4.2 Selection of experimental time-points

In order to observe unresolved DSB, present due to possible defects in DNA repair, the timeframe of the experiment was optimised. UPCISCC90 was selected for these experiments as it showed sensitivity to Olaparib in the clonogenic assay (HeLa would be inappropriate as its growth rate was much faster than the OPSCC lines). The cell cycle pattern and γ-H2AX response following treatment with 0.1, 1 and 10 μM Olaparib were measured after 12, 18, 24, and 48 h. Olaparib doses of 0.1, 1, and 10 μM were selected based on the clonogenic assay results as cell lines displayed a wide range of sensitivity at these doses.

The topoisomerase II inhibitor Etoposide was used as a positive control. Etoposide acts by stabilising DNA/topoisomerase II complexes, thereby inhibiting the ligation step in repair of topoisomerase induced DSB. The use of this drug specifically creates DSB, which can potentially be repaired using either HR or NHEJ. High doses (25 μM) of Etoposide were used to specifically induce DSB and increase γ-H2AX levels in the assay.

Variations in the amount of γ-H2AX were not detected following low dose treatment with Olaparib (0.1 and 1 μM) (Figure 6.7 A). However, treatment with 10 μM Olaparib resulted in 1.8 and 2.5 fold increases in γ-H2AX after 24 h and 48 h respectively (the most noticeable differences in the amount of γ-H2AX detected compared to the vehicle occurred after 24 h and 48 h). The positive control Etoposide was shown to be an appropriate control, as it induced an accumulation of γ-H2AX detectable after 12 h and this increased over time.

The effects of Olaparib on cell cycle distribution were also assessed. An increase in the proportion of cells in G2 and S-phase was observed after 24 h and 48 h for 10 μM Olaparib treatment whilst Etoposide caused a complete G2 arrest after 48 h (Figure 6.7 B). For further experiments, 24 h and 48 h were selected as appropriate time-points, as they showed the highest accumulation of γ-H2AX indicating the presence of unrepaired DSB.

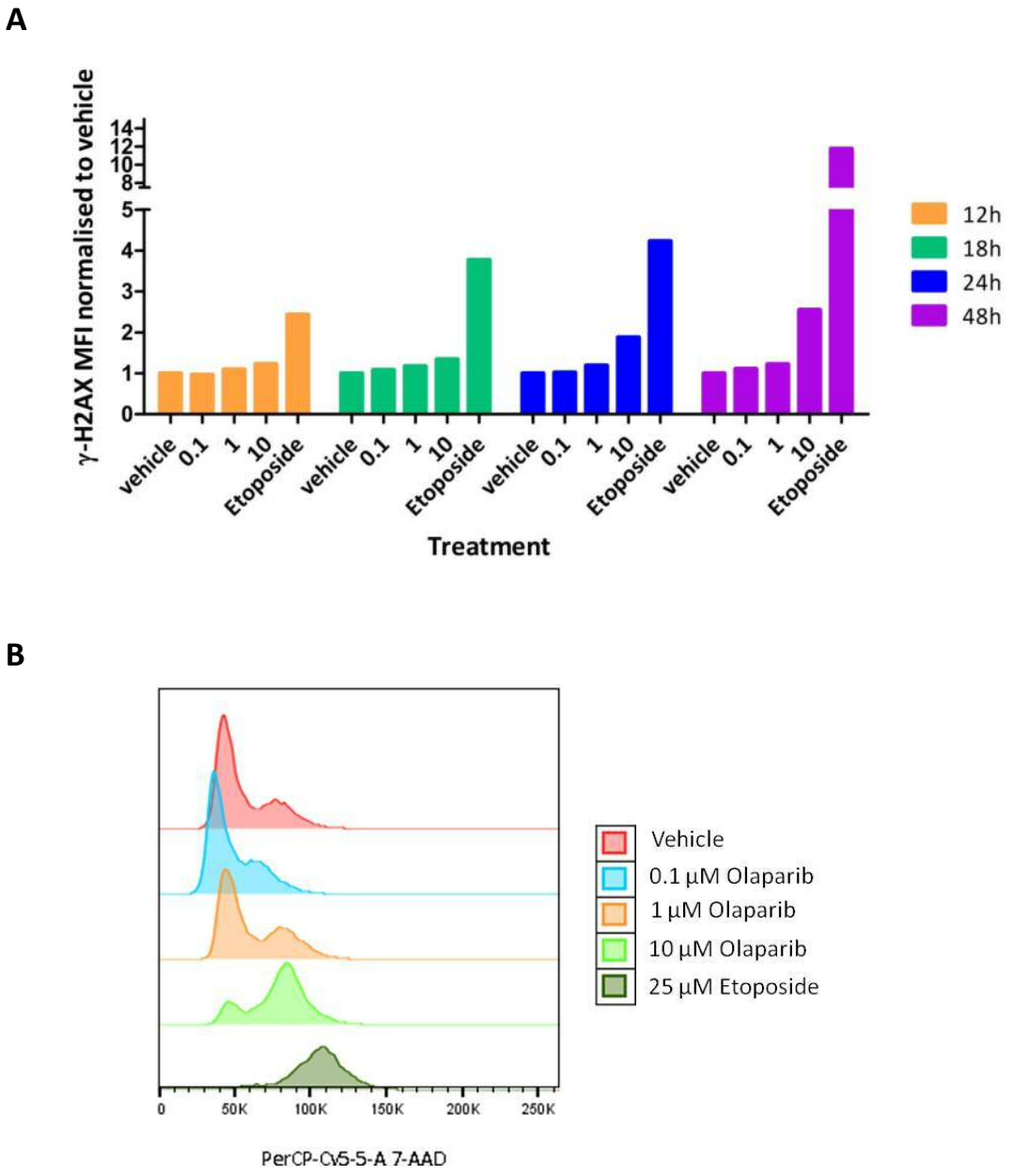


Figure 6.7: γ-H2AX and cell cycle following Olaparib treatment

A: The Mean Fluorescence Intensity (MFI) of the γ-H2AX FITC labelled antibody was measured 12, 18, 24 and 48 h after Olaparib (0-10 µM) or Etoposide (25 µM) treatment. Each value was normalised against the vehicle control. Increased fluorescence was observed after 24 and 48 h, indicating unrepaired DSB.

B: Fluorescence of 7-AAD versus cell count representing the cell cycle after 48 h of treatment with Olaparib or Etoposide. The proportion of cells in S and G2-phases increased after treatment with 1 and 10 µM Olaparib. Etoposide caused a G2 arrest.

6.4.3 Effect of 3T3 feeder cells in PCOC culture on γ-H2AX level

PCOC2 and 3 were cultured with the support of 3T3 feeder cells. These cells were rendered post-mitotic by administration of a 60 Gy dose of IR which caused sufficient DNA damage to prevent replication in culture. Most 3T3 feeder cells were removed using differential trypsinisation but a small number may have remained mixed with PCOC cells. Levels of unrepaired DSB in the contaminating 3T3 feeder cells could therefore potentially have influenced the results of the PCOC experiments. To assess whether this was a significant influence, an experiment was performed to investigate the γ-H2AX and cell cycle staining patterns of feeder cells. The general gating strategy used for all lines is shown in Figure 6.8.

The staining patterns were compared for nuclei populations obtained from: (1) PCOC2 cultured *with* 3T3 feeder cells; (2) PCOC2 cultured *without* 3T3 feeder cells; and (3) 3T3 feeder cells alone. The scatter plot illustrating the DNA content versus amount of γ-H2AX enabled discrimination between PCOC2 and 3T3 feeder cells. 3T3 feeder cells had a higher DNA content and due to irradiation displayed a higher level of γ-H2AX (Figure 6.9 A and B). The mouse genome is smaller than the human one (2.8×10^9 bp versus 3.3×10^9 bp). However, the 3T3 cell line is tetraploid which is consistent with the higher amount of DNA observed in the experiment (Leibiger et al. 2013). The more precise gating of the PCOC2 population enabled the elimination of the extra peaks in the cell cycle pattern caused by the 3T3 feeder cells (Figure 6.9 C). The γ-H2AX antibody used has been described as human-specific by the manufacturer, but with a disclaimer that due to close similarities there might be cross-species reactivity. This has clearly been observed in this case. It was demonstrated in previous experiments that PCOC cells do not grow well, especially at low plating densities, without 3T3 feeder cells. Hence removing them from the culture was therefore undesirable. The gating process shown here was used in all further PCOC experiments.

6.5 DNA double-strand break repair

6.5.1 Basal level of γ-H2AX and cell cycle distribution differs between cell lines

All lines had a baseline γ-H2AX expression which indicated the formation and repair of DSB as part of the normal cell processes. Different levels of expression were observed between the lines. UMSCC19 showed a significantly higher level of γ-H2AX than PCOC3, UMSCC47,

UPCISCC90 and UMSCC4 ($p<0.05$, result of ANOVA and Bonferroni's post-test). PCOC2, UMSCC74a and the normal keratinocyte line HEKn showed intermediate levels of γ-H2AX (Figure 6.10 A).

The cell cycle was not synchronised prior to these experiments and, perhaps because of this, variations in distribution were present between cell lines, with HEKn and PCOC3 showing the highest proportion of cells in G1 (Figure 6.10 B).

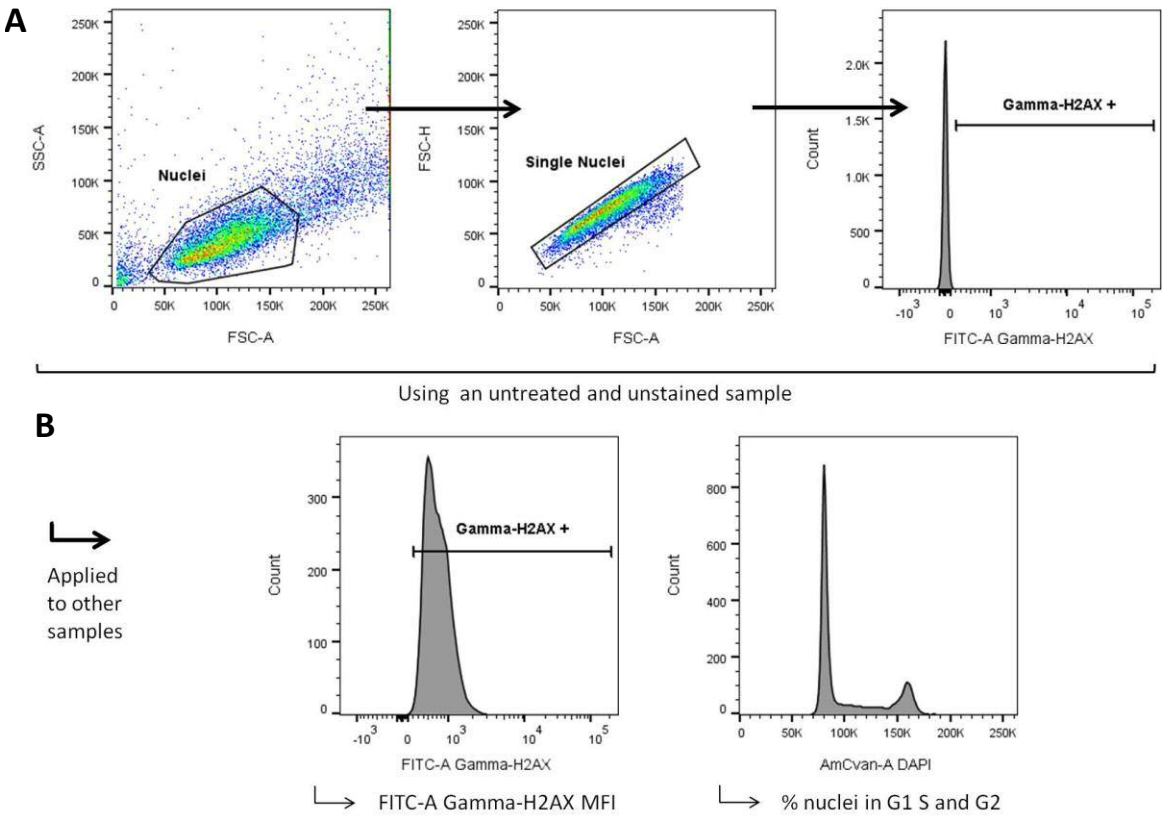


Figure 6.8: General gating strategy

The fluorescence in each channel was measured by flow cytometry (BD FACS Canto) and the data analysis carried out using FlowJo v10 software.

A: Within each experiment, the gates were fitted to the untreated and unstained control before being applied to the other samples. The nuclei population was distinguished from debris on the forward vs side scatter plot (Panel 1). Doublets were excluded based on the forward scatter area vs height plot (Panel 2). The Gamma-H2AX + gate was fitted by excluding the basal fluorescence in the FITC channel (Panel 3).

B: The MFI was calculated in the FITC-A channel and used as a representation of the Gamma-H2AX level (Panel 1). The cell cycle pattern was visualised in the single nuclei population using the AmCyan channel (panel 2).

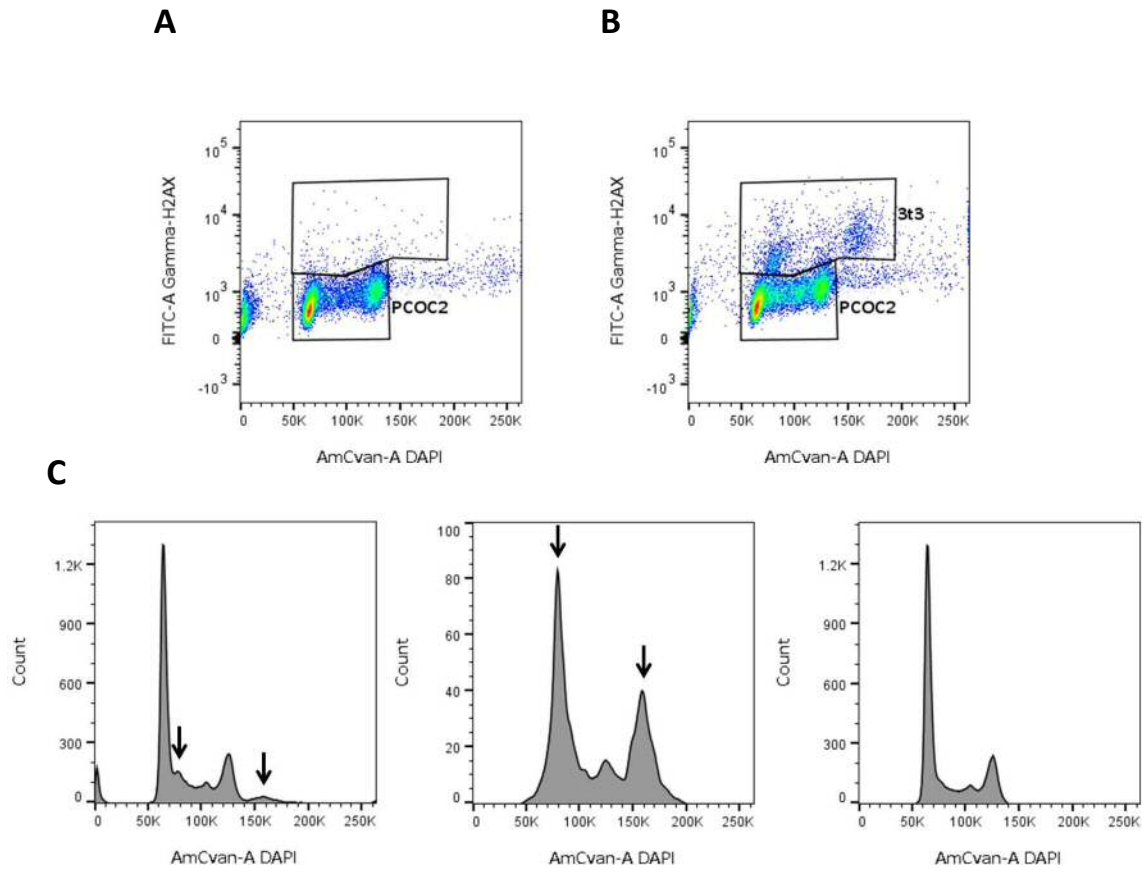


Figure 6.9: Gating strategy to reduce 3T3 feeder cells influence on measurement of the cell cycle

A: Gating of a population of nuclei isolated from PCOC2 cells cultured alone. DAPI represents the DNA content and the FITC labelled γ-H2AX antibody the presence of DSB.

B: Gating of a population of nuclei isolated from PCOC2 cells cultured with 3T3 feeder cells. Two populations can be distinguished; 3T3 feeder cells have a higher DNA content and elevated levels of DSB.

C: The first panel shows the cell cycle pattern of PCOC2 and 3T3 feeder cells analysed together. Two extra peaks (black arrows) are visible on the cell cycle pattern. The middle panel shows the profile of 3T3 feeder cells. The G1 and G2 peaks indicated by black arrows correspond to the extra peaks visible in the first panel. The third panel shows the PCOC2 gate analysed on its own, with the extra peaks no longer visible.

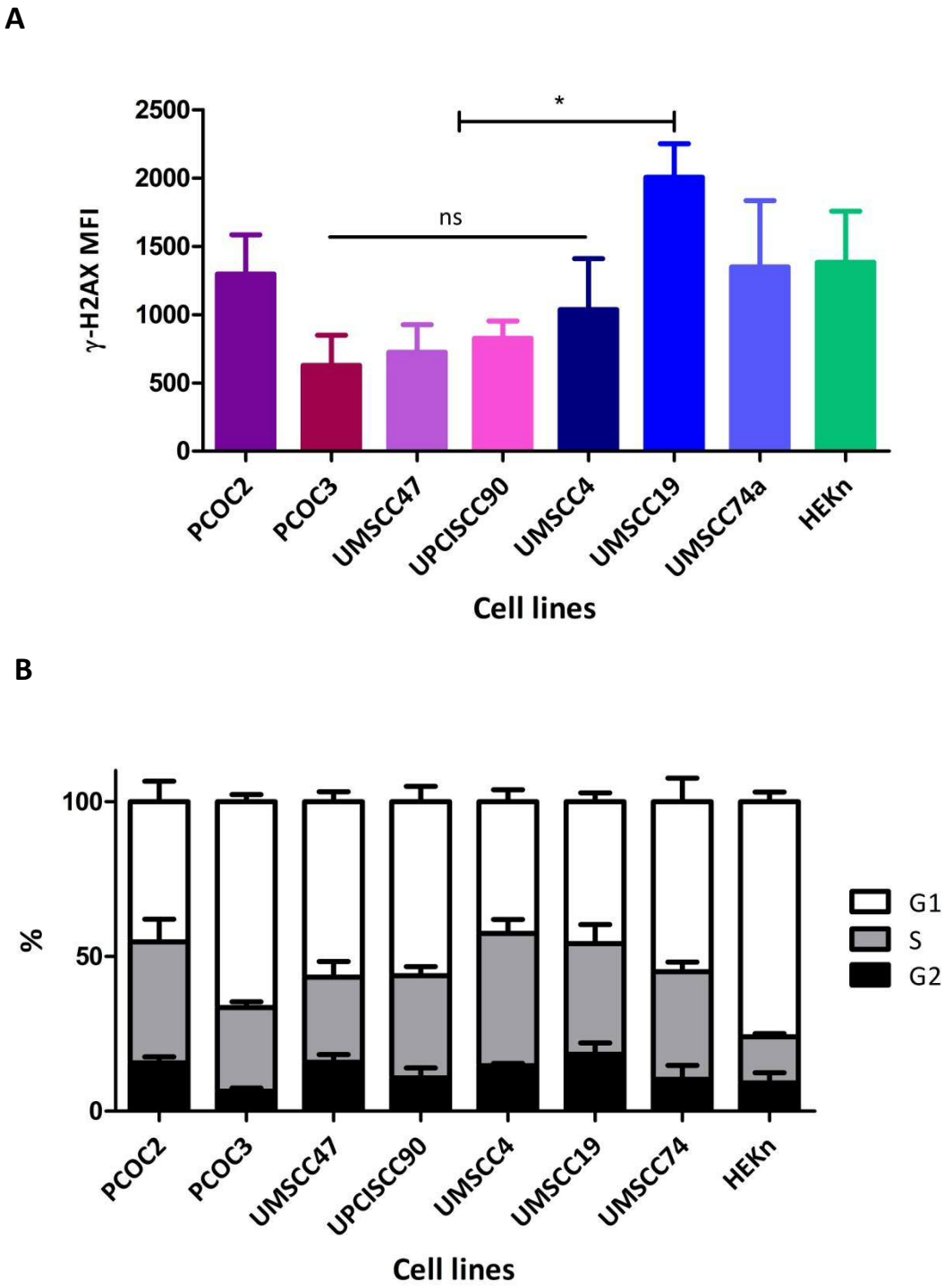


Figure 6.10: Basal level of γ-H2AX and cell cycle distribution differed between cell lines

A: γ-H2AX MFI representative of at least 3 experiments are plotted for each untreated cell line with the standard deviation. The basal γ-H2AX level differed between lines. UMSCC19 had an elevated level compared to PCOC3, UMSCC47, UPCISCC90 and UMSCC4 (ANOVA and Bonferroni's post test $p<0.05$).

B: Histogram showing cell cycle distribution for the cell line panel in untreated unsynchronised cells. Data is representative of at least 3 experiments.

6.5.2 High doses of Olaparib lead to unrepaired DSB

Low doses of Olaparib treatment (0.1 and 1 μM) did not lead to an increase in γ-H2AX in any of the cell lines, compared to the vehicle control. However increases in γ-H2AX after 10 μM Olaparib treatment were observed, and these varied between cell lines and over time. Representative experiments for each cell line are shown in Figure 6.11 to Figure 6.14 A. The data are summarised in Figure 6.15 and Figure 6.16. A significant increase in the level of γ-H2AX after 10 μM Olaparib compared to vehicle control was detected in the HPV-positive cell line UPCISCC90, and the HPV-negative cells UMSCC4, UMSCC19 and UMSCC74a both 24 h and 48 h after treatment. The highest increase was observed in UPCISCC90, in which the level of γ-H2AX was 1.5 and 2.3 fold higher than observed in the vehicle control after 24 h and 48 h respectively. Among the HPV-negative lines, the highest increase was seen in UMSCC74a with a 0.6 and 0.9 fold increase compared to vehicle after 24 h and 48 h.

Levels of γ-H2AX were elevated in PCOC2 and PCOC3 24 h after treatment (there was also a tendency for γ-H2AX to increase after 48 h in PCOC2). Normal keratinocytes (HEKn) cells were included to give an indication of the rate of DNA repair in "normal" cells. The HEKn displayed significant γ-H2AX increase after 24 h 10 μM Olaparib compared to vehicle and 1 μM Olaparib, but this was no longer visible after 48 h. This appears to indicate functioning repair of DSB in these 3 cell lines.

Following Etoposide treatment (25 μM), a significant increase in γ-H2AX was detected in all tumour cell lines (after both 24 h and 48 h). In HEKn γ-H2AX increased after 24 h but only in a non-significant tendency after 48 h. This confirmed the assay was functioning as expected, and that detected presence of unrepaired DSB in all lines.

In summary, the quantification of γ-H2AX level following Olaparib treatment showed HPV-positive cells did not accumulate higher levels of unrepaired DNA DSB than HPV-negative lines. Forty-eight hours after 10 μM Olaparib treatment, substantial levels of unrepaired DSB were evident in 2/4 HPV-positive lines and 3/3 HPV-negative lines.

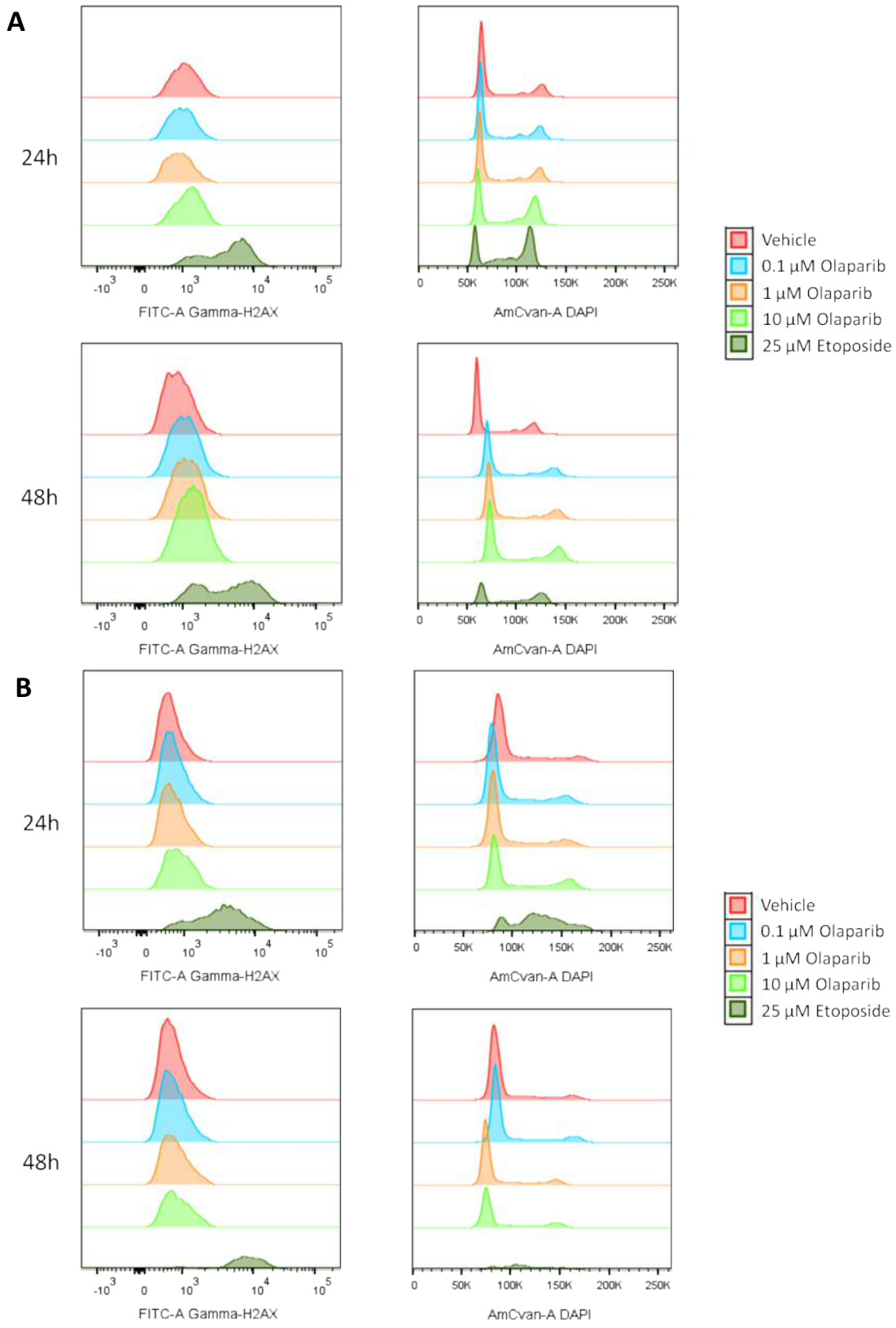


Figure 6.11: PCOC2 and PCOC3 γ-H2AX and cell cycle analysis

Histogram overlay representation of the γ-H2AX expression and cell cycle distribution 24 h and 48 h after treatment of A: PCOC2 and B: PCOC3.

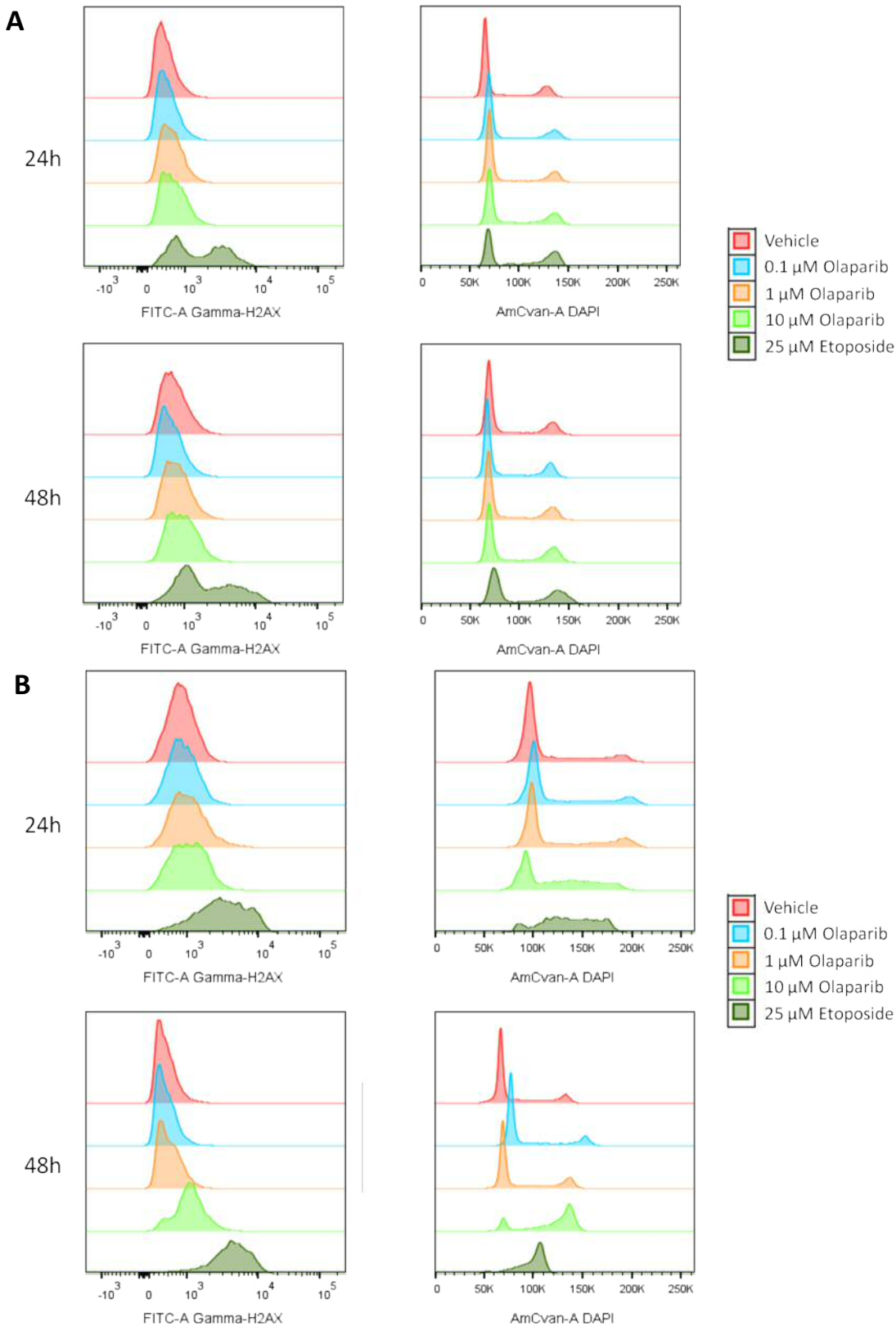


Figure 6.12: UMSCC47 and UPCISCC90 γ-H2AX and cell cycle analysis

Histogram overlay representation of the γ-H2AX expression and cell cycle distribution 24 h and 48 h after treatment of A: UMSCC47 and B: UPCISCC90.

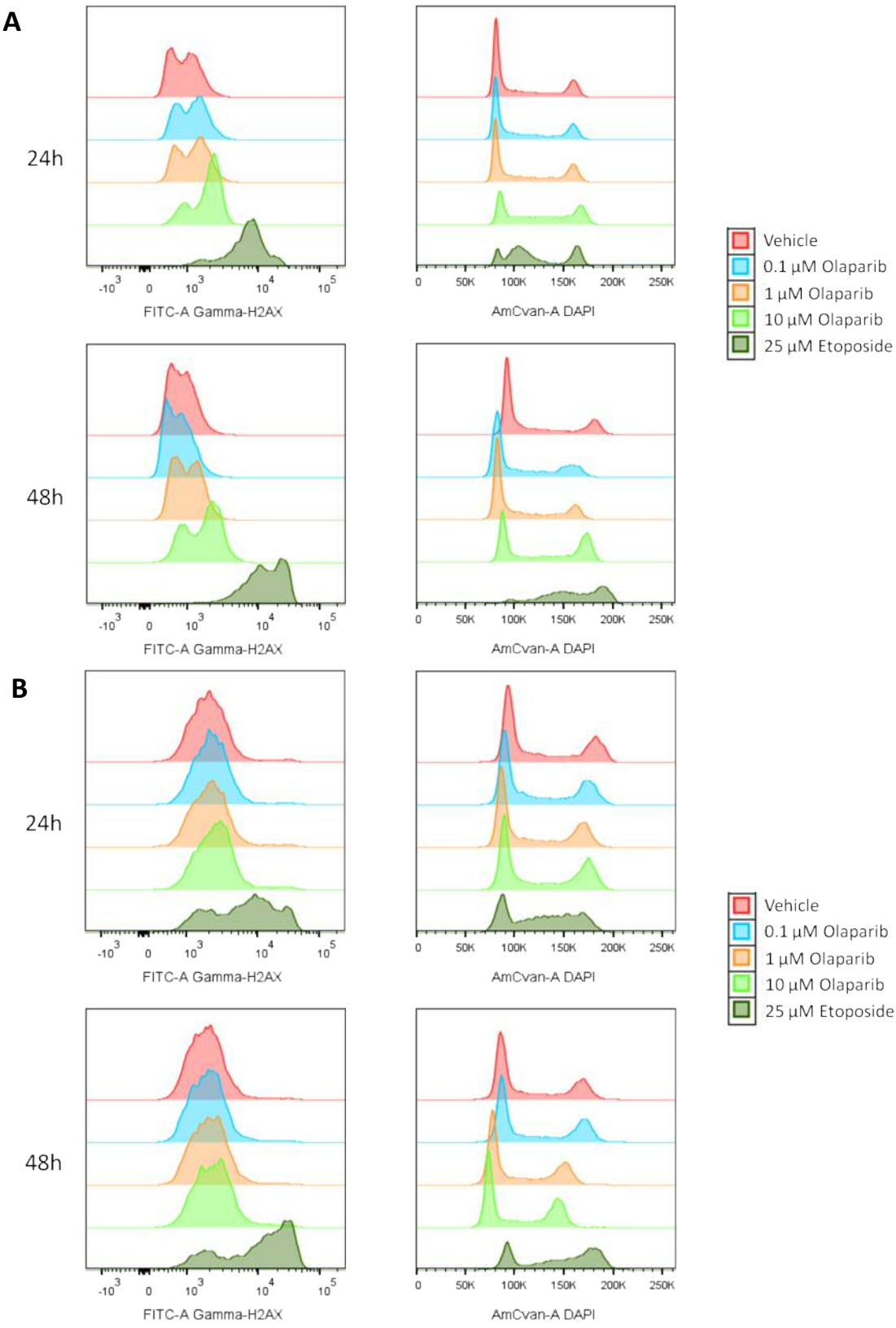


Figure 6.13: UMSCC4 and UMSCC19 γ-H2AX and cell cycle analysis

Histogram overlay representation of the γ-H2AX expression and cell cycle distribution 24 h and 48 h after treatment of A: UMSCC4 and B: UMSCC19.

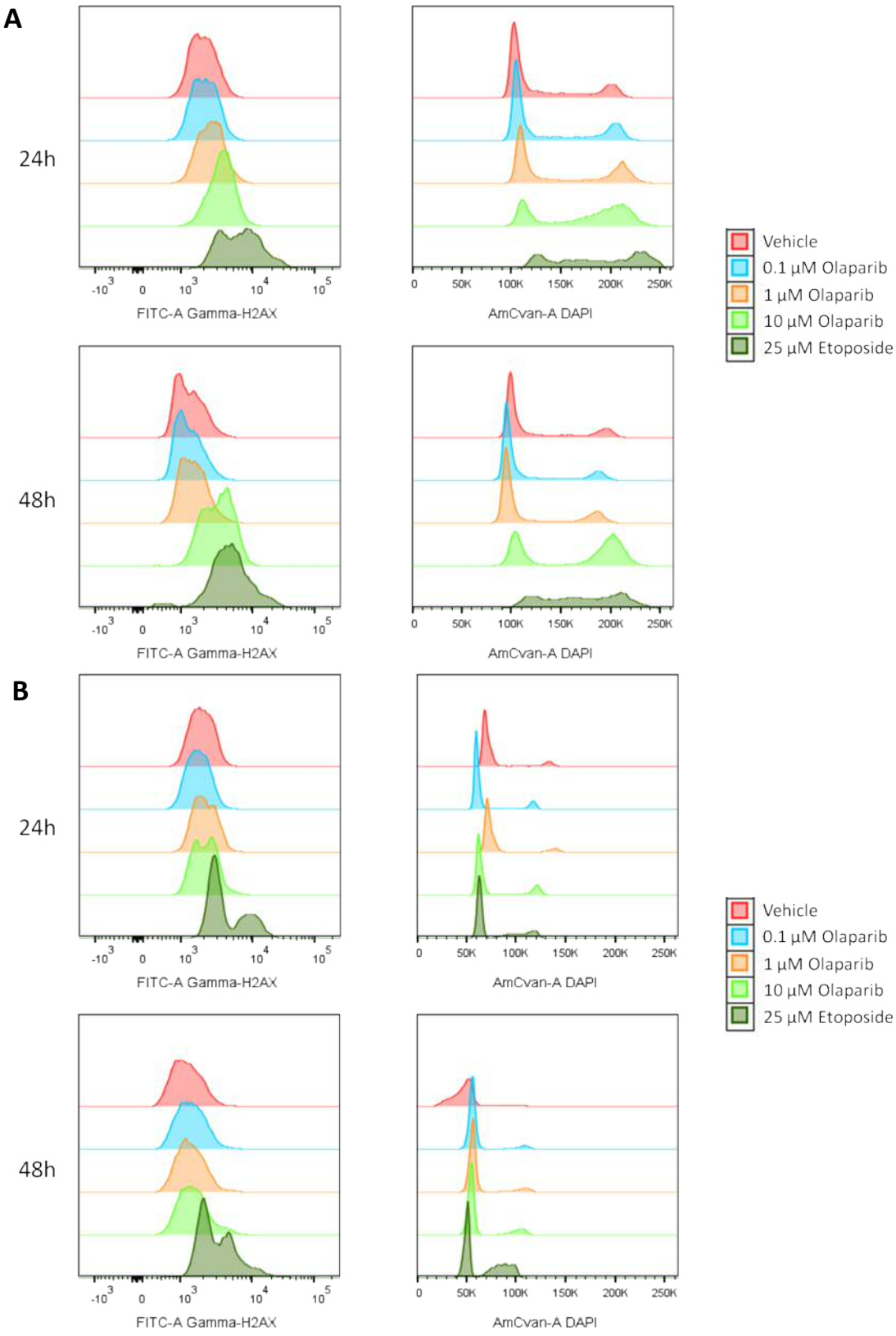


Figure 6.14: UMSCC74a and HEKn γ-H2AX and cell cycle analysis

Histogram overlay representation of the γ-H2AX expression and cell cycle distribution 24 h and 48 h after treatment of A: UMSCC74a and B: HEKn.

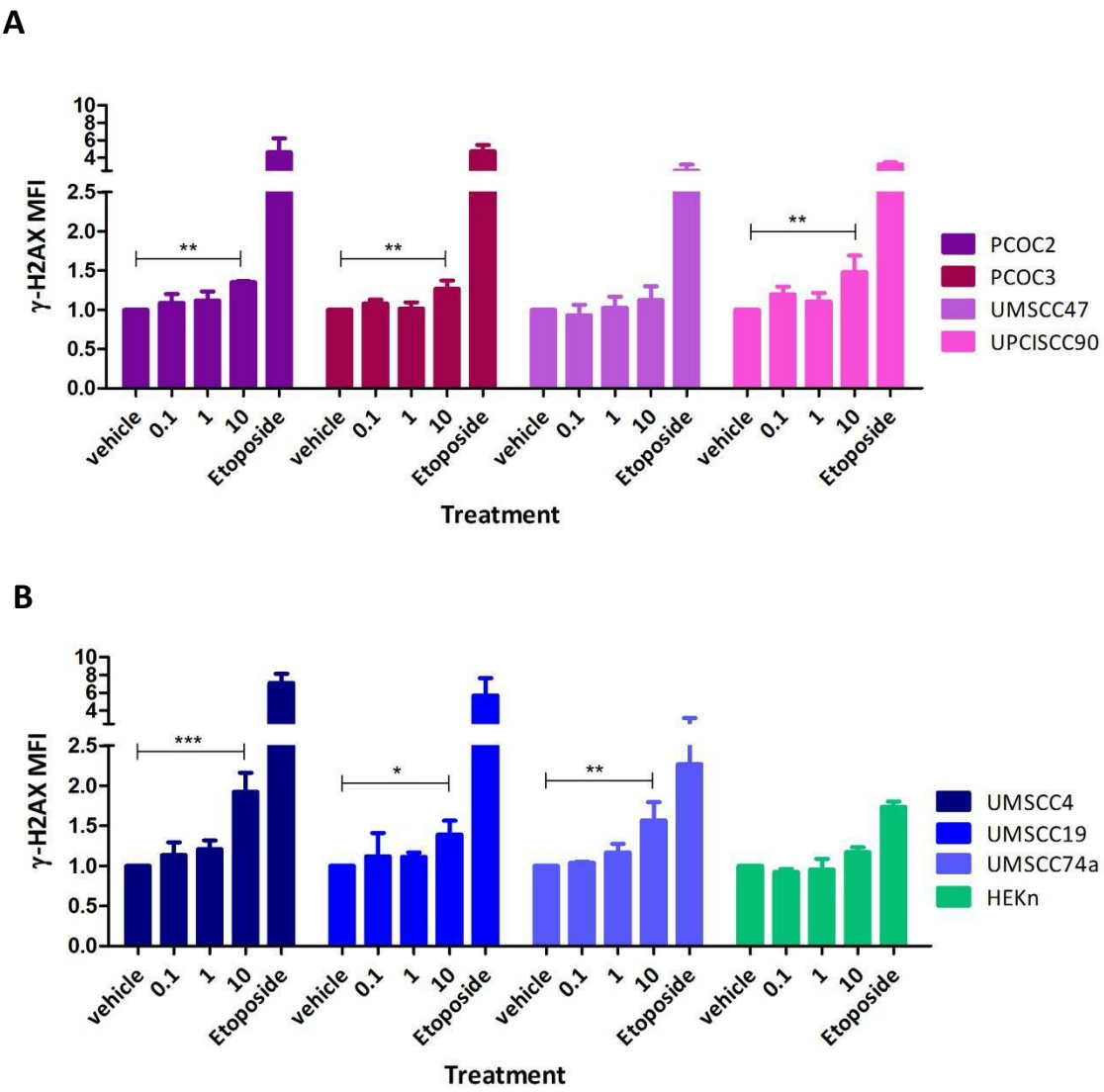


Figure 6.15: γ-H2AX measurement 24 h after Olaparib treatment

Data is representative of at least 3 independent experiments. The MFI normalised to vehicle is shown for each cell line after treatment with 0.1 µM 1 µM and 10 µM Olaparib. 25 µM Etoposide was used as a positive control. Etoposide treatment resulted in a significant γ-H2AX increase in all lines (ANOVA and Tukey post-test, * P < 0.05, ** P<0.01, *** P<0.001).

A: 10 µM Olaparib treatment caused a significant increase in the amount of γ-H2AX compared to the vehicle in the HPV-positive lines PCOC2, UPCISCC90 and PCOC3 but not in UMSCC47.

B: Following 10 µM Olaparib, the amount of γ-H2AX significantly increased compared to vehicle in the HPV-negative lines UMSCC4, UMSCC19 and UMSCC74a.

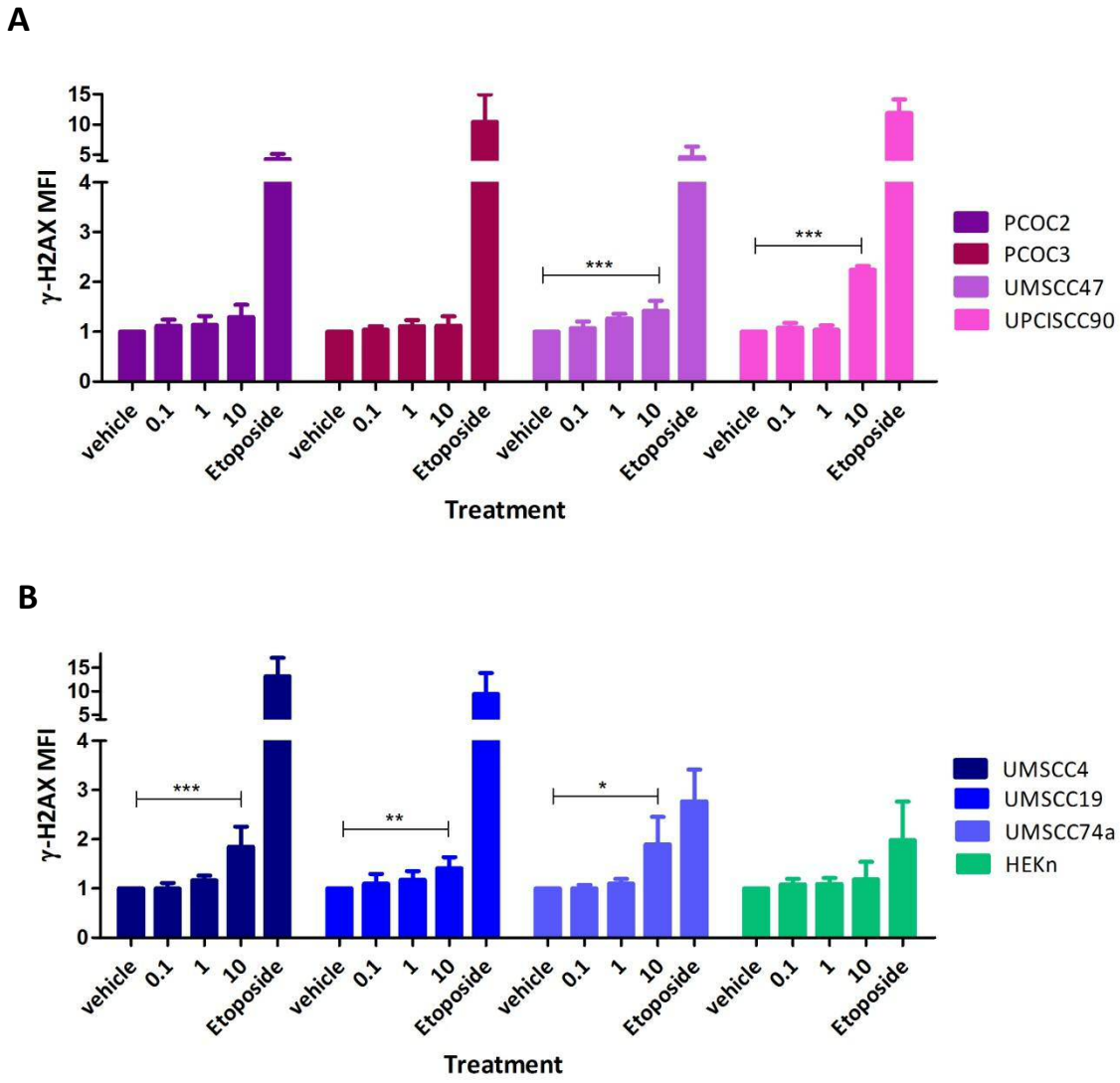


Figure 6.16: γ-H2AX measurement 48 h after Olaparib treatment

Data is representative of at least 3 independent experiments. The MFI normalised to vehicle is shown for each cell line after treatment with 0.1 µM 1 µM and 10 µM Olaparib. 25 µM Etoposide was used as a positive control. Etoposide treatment resulted in a significant γ-H2AX increase in all lines except HEKn (ANOVA and Tukey post-test, * $p<0.05$, ** $P<0.01$, *** $P<0.001$).

A: The highest increase in the amount of γ-H2AX detected compared to vehicle in HPV-positive lines occurred in UPCISCC90. A significant increase was observed in UMSCC47 after 10 µM compared to vehicle, which was not present after 24 h.

B: Within the HPV-negative lines, a significant increase in γ-H2AX at 10 µM compared to vehicle was observed in UMSCC4, UMSCC19 and UMSCC74a. This was not detected in HEKn.

6.5.3 High doses of Olaparib lead to cell cycle arrest

The effect of the DNA damage caused by Olaparib treatment on progression through the cell cycle was investigated using flow cytometry. One representative experiment for each cell line is shown in Figures 6.11-6.14 B, and summarised in Figure 6.17 and Figure 6.18.

The untreated samples showed some variations in cell cycle distribution among the cell lines. It was very noticeable that HEKn showed a much higher proportion of cells in G1 than the tumour lines, suggesting a lower proportion of actively cycling cells. Among the tumour lines, PCOC3 and UMSCC47 showed the highest G1 fractions, suggesting a lower rate of proliferation in these lines.

For all the cell lines, cell cycle distribution was not affected by treatment with 0.1 and 1 μM Olaparib for 24 h or 48 h. However, treatment with 10 μM Olaparib resulted in significant changes in cell cycle distribution in some cell lines. Olaparib treatment did not significantly affect the cell cycle distribution of PCOC3, or the HPV-negative lines UMSCC4, UMSCC19 and UMSCC74a. PCOC2 and HEKn cells accumulated in G2 24 h after treatment, however, the distribution was similar to the vehicle after 48 h. Significant effects on cell cycle distribution were detected in UPCISCC90, cells accumulated in S-phase after 24 h and G2 after 48 h. This pattern suggests cells slowing in the cycle to allow repair during S-phase, followed by G2 arrest due to the accumulation of unresolved DSB. UMSCC47 cells accumulated in G2 after 48 h treatment.

In summary, only the HPV-positive cell lines UPCISCC90 and UMSCC47 showed evidence of significant cell cycle arrest 48 h after treatment with Olaparib.

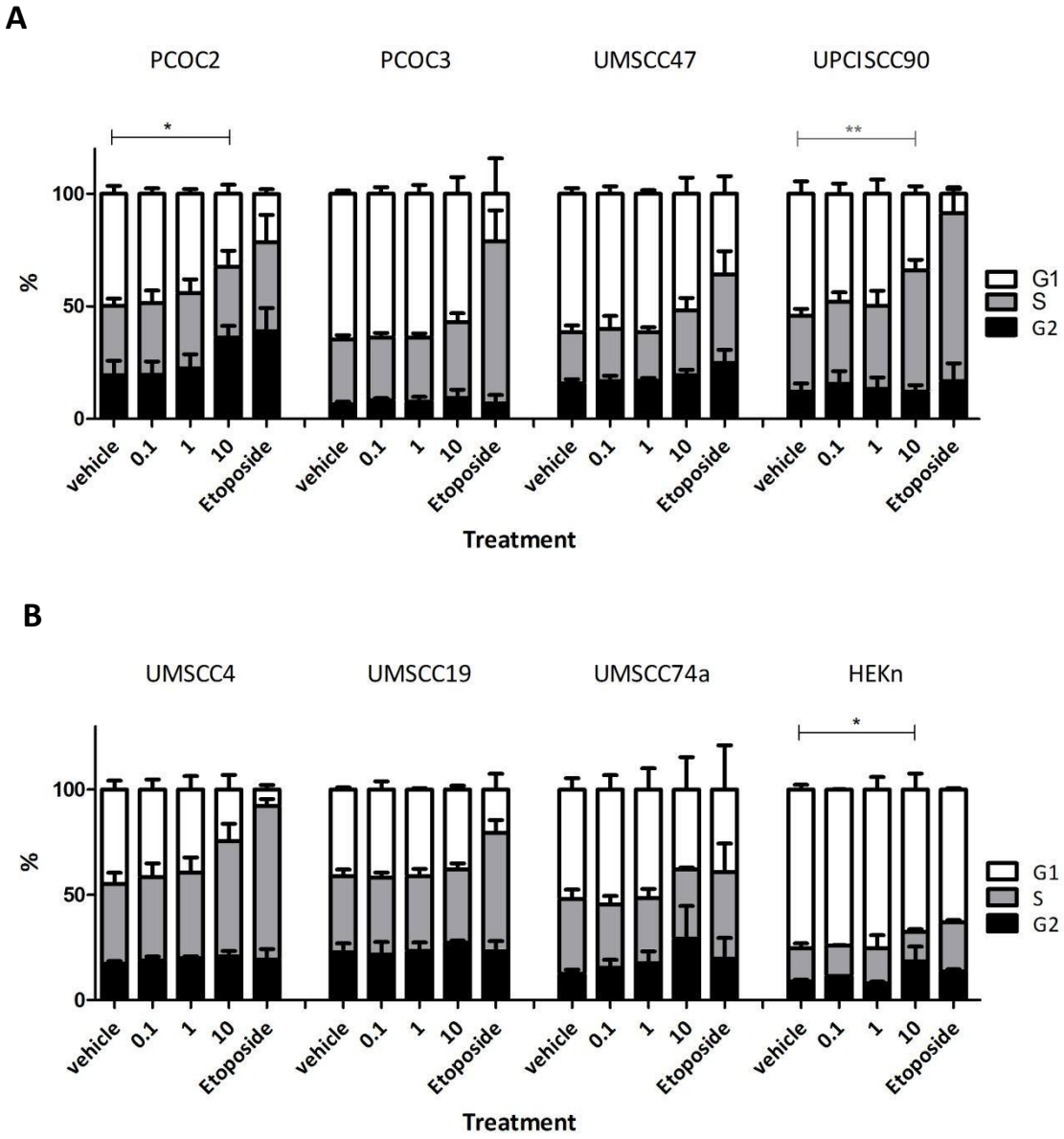


Figure 6.17: Cell cycle distribution after 24 h Olaparib treatment

Data is representative of at least 3 independent experiments. The percentage of cells in G1, S and G2-phases is shown for each cell line after treatment with 0.1 µM, 1 µM and 10 µM Olaparib. 25 µM Etoposide was used as a positive control.

A: Within HPV-positive lines, significant variations in cell cycle distribution were detected in PCOC2 and UPCISCC90 after 10 µM Olaparib treatment. This was reflected by an increased proportion of cells in G2 for PCOC2 and in S-phase for UPCISCC90.

B: Within the HPV-negative lines, a significant increase in cells in G2 was present in HEKn compared to vehicle . A tendency for cells to accumulate in S-phase was observed in UMSCC4 but this was not statistically significant.

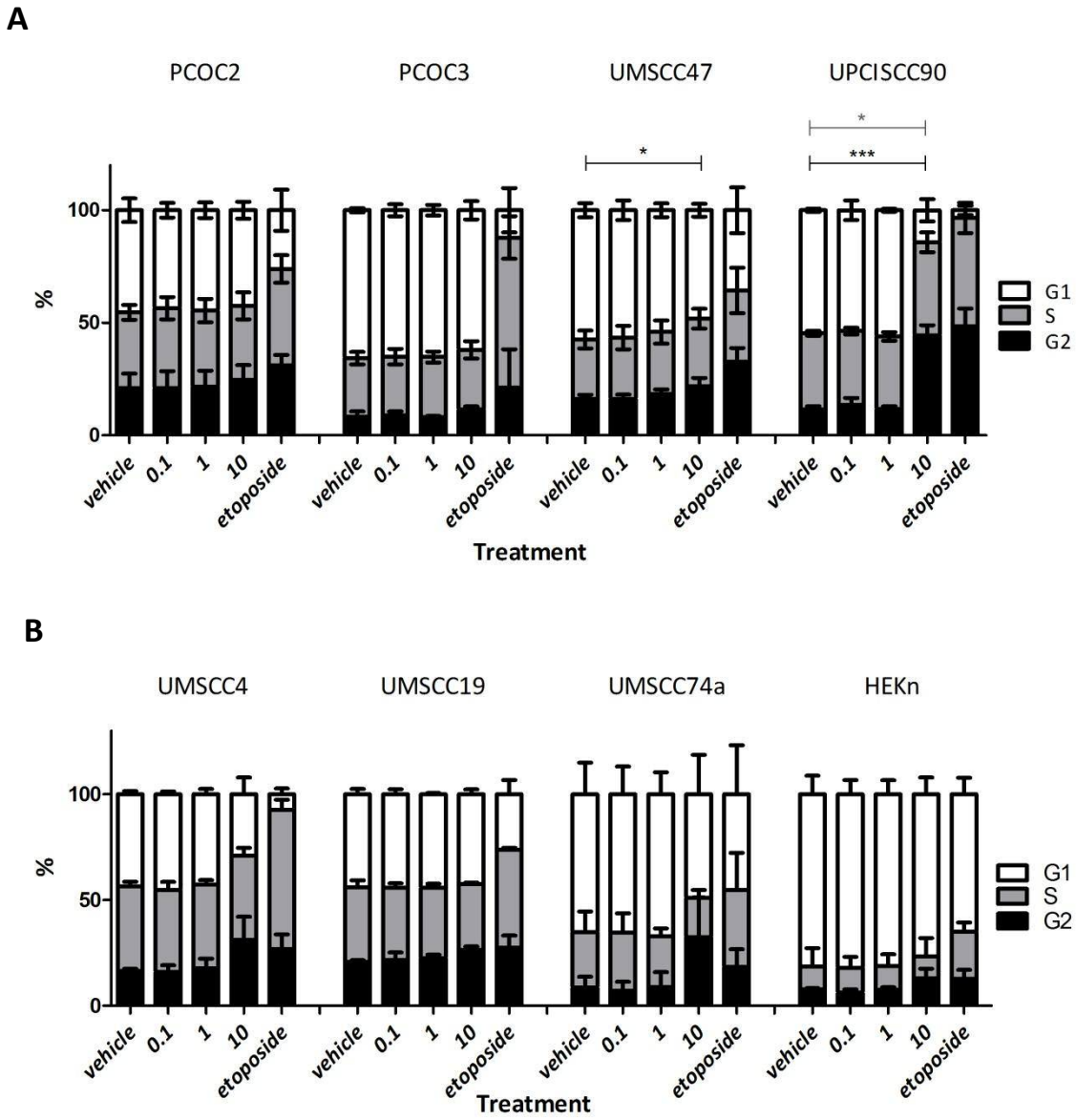


Figure 6.18: Cell cycle distribution after 48 h Olaparib treatment

Data is representative of at least 3 independent experiments. The percentage of cells in G1, S and G2-phases is shown for each cell line after treatment with 0.1 µM 1 µM and 10 µM Olaparib. 25 µM Etoposide was used as a positive control.

A: Within HPV-positive lines, significant changes in cell cycle distribution were detected in UPCISCC90 and UMSCC47 after 10 µM Olaparib treatment. This was reflected by an increased proportion of cells in both S and G2 in UPCISCC90 and of G2 in UMSCC47.

B: Within the HPV-negative lines, no significant cell cycle alteration were detetced between vehicle and 10 µM treatment. A tendency of cell accumulation in both S and G2 was observed in UMSCC4 and of G2 accumulation in UMSCC74a without being statistically significant.

6.6 Molecular effects of Olaparib treatment

Cleavage of PARP1 by caspases is considered to be a hallmark of apoptosis (Kaufmann et al. 1993). Cleavage of PARP1, mediated by caspase 3, yields two fragments of 24 and 89 kDa. Detection of the 89 kDa cleaved PARP1 fragment provided an opportunity to assess apoptosis in the cell line panel. Levels of p53 and phospho-p53 (p-p53) were assessed to determine whether any response to Olaparib treatment was mediated through p53 dependent or independent pathways.

6.6.1 Total and cleaved PARP1 expression in untreated cell lines

PARP1 expression was first examined in cell lysates from untreated cell lines. This analysis was performed to determine whether the basal level of PARP1 correlated with sensitivity to Olaparib. Total PARP1 western blot images and densitometry quantification plots are presented in Figure 6.19. There was wide variation in the levels of PARP1 between individual cell lines. It appeared that the basal level of PARP1 in untreated cells was highest in the HPV-positive line UPCISCC90 and in the HPV-negative line UMSCC19, the other lines showed approximately equal levels of expression. PARP1 levels did not show an obvious correlation with the HPV status of the cell lines.

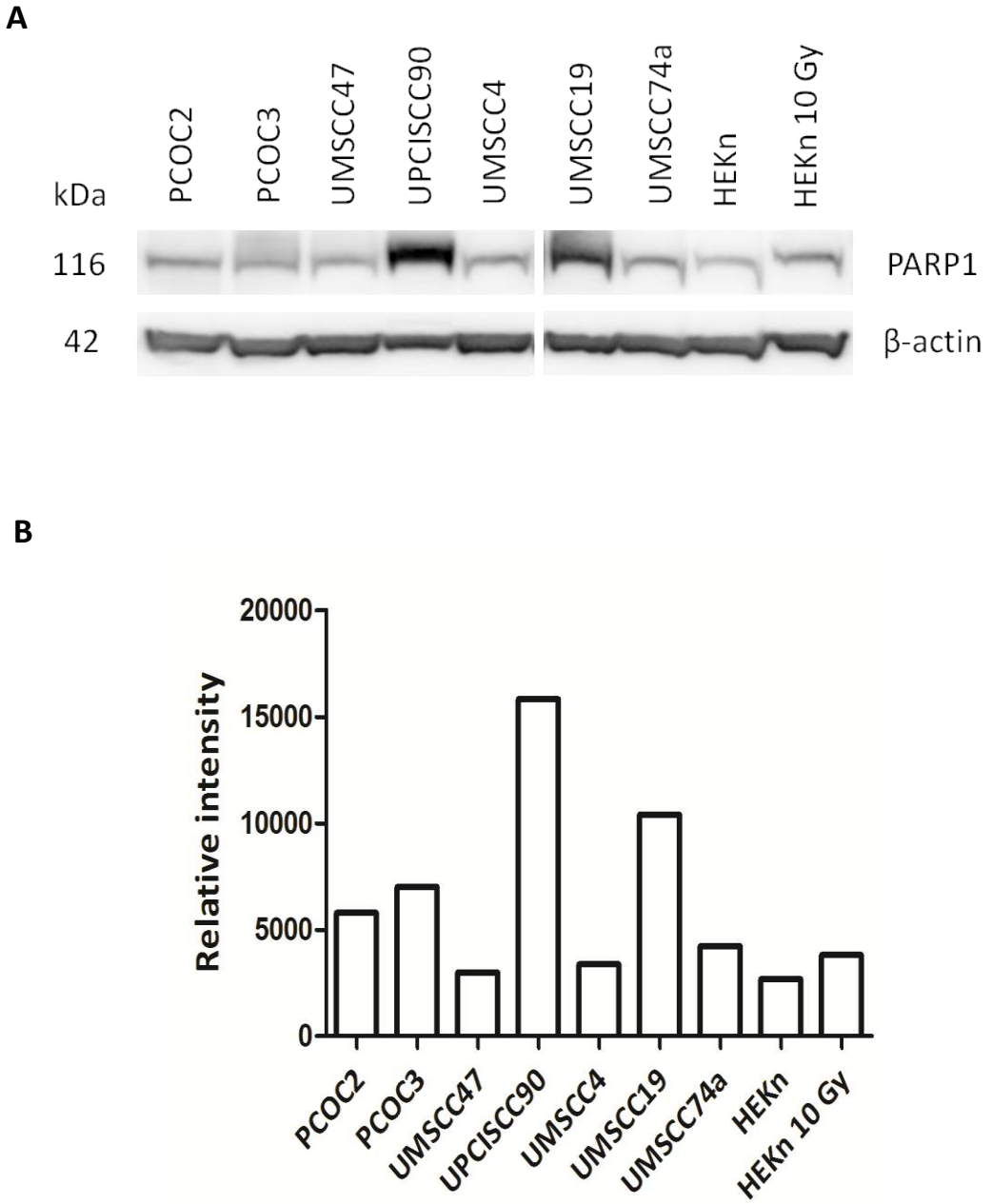


Figure 6.19: Basal PARP1 expression in untreated HPV-positive and negative cell lines

A: PARP1 western blot image of PCOC2, PCOC3, UMSCC47, UPCISCC90, UMSCC4, UMSCC19, UMSCC74a, HEKn untreated and HEKn 10 Gy. β-actin was used as a loading control and indicated even loading between wells.

B: The intensity of the western bloting bands as quantified in relation to β-actin level using Image studio Lite. The highest basal levels of PARP1 were seen in HPV-positive UPCISCC90 and HPV-negative UMSCC19 cells lines. 10 Gy treatment did not affect the HEKn PARP level.

6.6.2 Total and cleaved PARP1, p53 and p-p53 expression in Olaparib treated cell lines

A condensed summary of the Western blot images (48 h only) is shown in Figure 6.20, and the full dataset in Figures 6.21 to 6.28. In PCOC2, PARP1 expression showed no changes after 24 h treatment. However after 48 h, PARP1 levels increased in concert with Olaparib concentration. PARP1 also appeared significantly induced following 10 Gy IR. The 89 bp cleaved form of PARP1 was not detected in either untreated or Olaparib treated cells. However, cleaved PARP1 was clearly present after 10 Gy IR indicating the induction of apoptosis in irradiated but not Olaparib treated cells. Neither p53 nor p-p53 were detected in treated or untreated PCOC2 cells, but both were present in the HEKn 10 Gy positive control. This showed the absence of a p53 dependant response to Olaparib treatment at these timepoints.

No correlation between Olaparib dose and PARP1 expression was present in PCOC3. Cleaved PARP1 was not detected in this cell line under any condition (light unspecific bands were not taken into account). Neither p53 nor p-p53 expression were induced by Olaparib treatment or 10 Gy IR of PCOC3 but were present in the HEKn control. A similar pattern of expression was seen in UPCISCC90, indicating the absence of apoptosis or p53 dependent response in both of these lines after 24 h and 48 h Olaparib treatment.

In UMSCC47, an increase in PARP1 expression was seen after 10 µM Olaparib (48 h) and following 10 Gy IR. Cleaved PARP1, p53 and p-p53 were not present. This also showed an absence of apoptosis or p53 dependent response.

Among the HPV-negative cell lines, UMSCC4 and UMSCC19, there was no discernible dose-response relationship between Olaparib and PARP1. Cleaved PARP1 was not clearly detected. Some weak bands of variable sizes were visible around 89 kDa which are likely to be caused by the background and non-specific binding. Neither p53 nor p-p53 were detected. This indicated an absence of apoptosis or p53 dependent response in these cell lines.

PARP1 levels did not change after Olaparib treatment in UMSCC74a. However, both p53 and p-p53 appeared to be induced following 10 µM Olaparib (48 h) and 10 Gy IR, which suggests p53 mediated signalling was active in these cells but did not result in apoptosis.

In the HEKn cells, there was no dose-response relationship between Olaparib and PARP1, and PARP1 cleavage and p53 signalling were not induced by Olaparib. As previously observed, 10 Gy did result in accumulation of p53 and signalling via p-p53. This did not however result in apoptosis, as indicated by the absence of cleaved PARP1.

In summary, there was no evidence for induction of apoptosis (cleaved PARP1) in any of the cell lines following treatment with Olaparib. Furthermore, treatment with Olaparib was not associated with p53 reactivation (accumulation of p53 or presence of p-p53) in HPV-positive cell lines. However, p53 and p-p53 protein did appear to accumulate in UMSCC74a (HPV-negative) following treatment with 10 μM Olaparib.

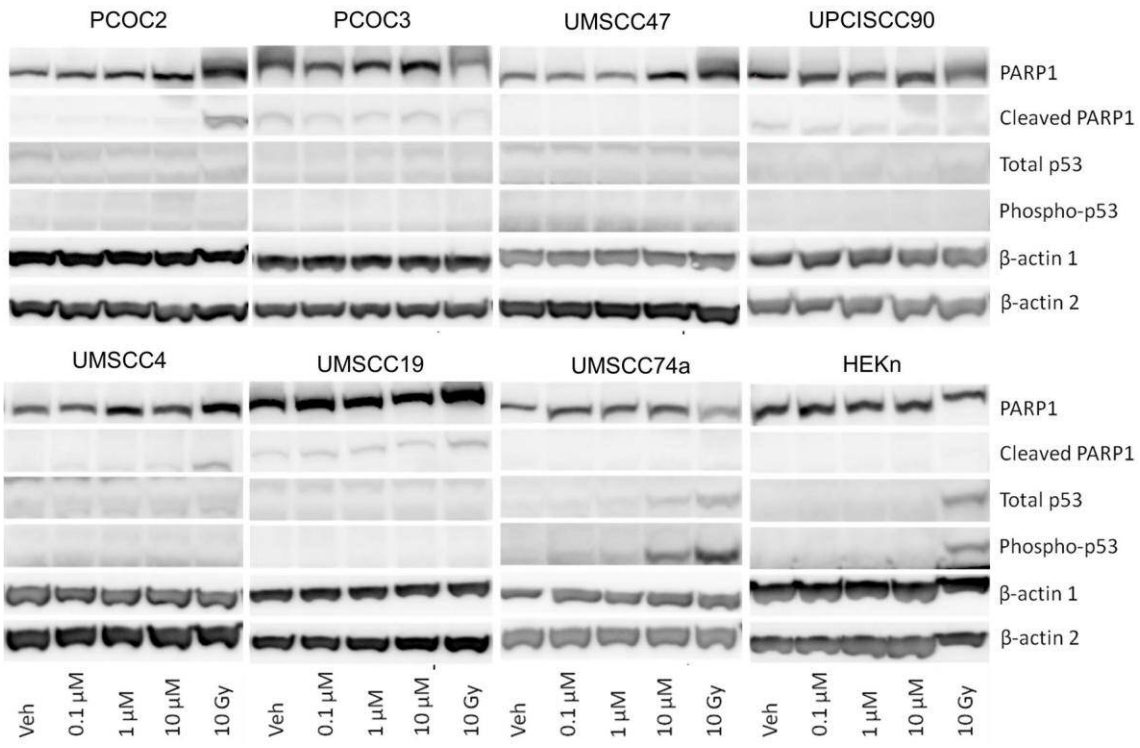


Figure 6.20: Protein levels of PARP1, cleaved PARP1, p53 and phospho-p53 48 h following Olaparib treatment

This figure shows PARP1, cleaved PARP1, p53 and phospho-p53 western blot images for all cell lines 48 h after Olaparib treatment. The cell lines treated with 10 Gy IR collected af 24 h following IR were included as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

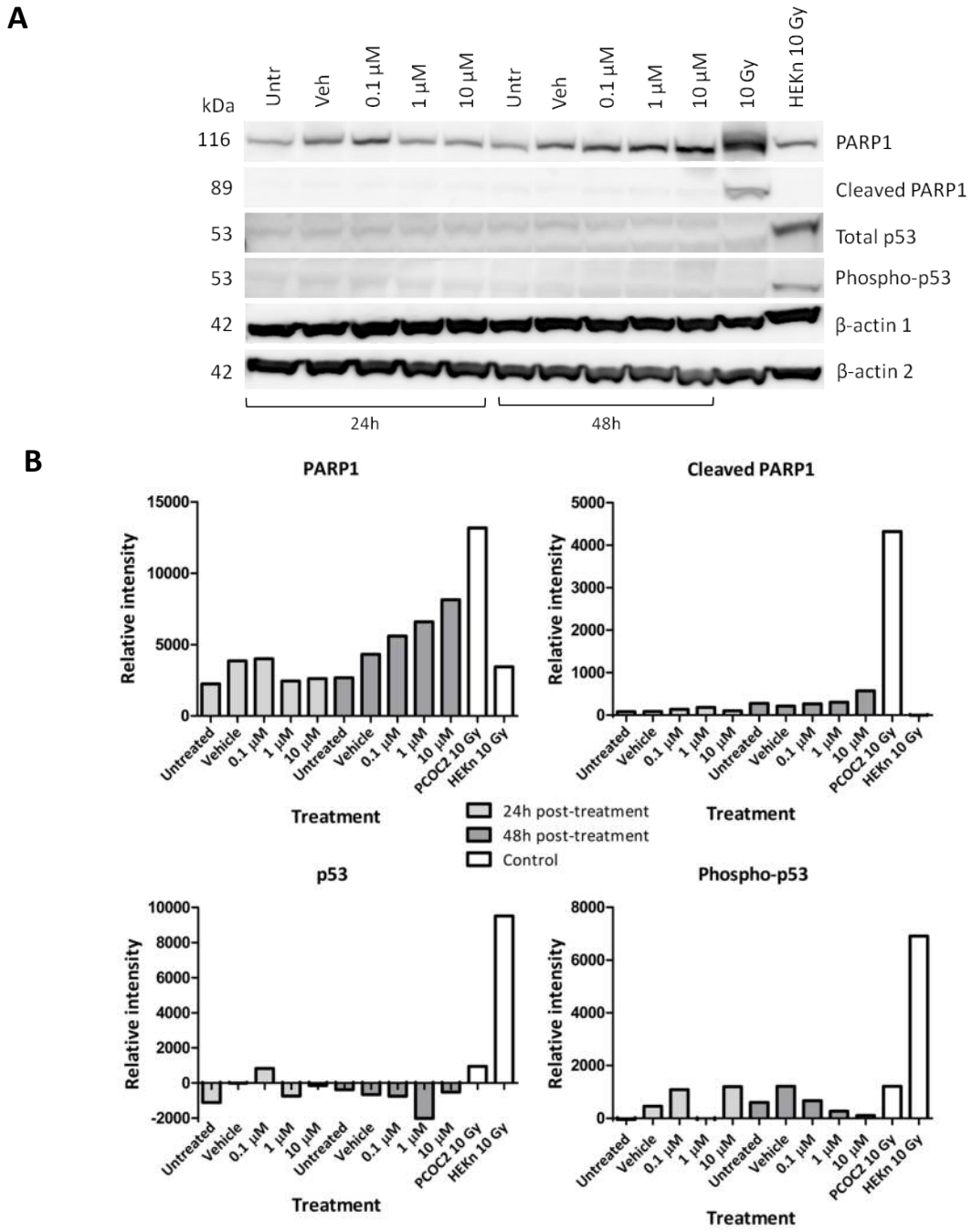


Figure 6.21: PCOC2 PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of PCOC2 after 24 h and 48 h Olaparib treatment. The cell lines treated with 10 Gy and HEKn treated with 10 Gy collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

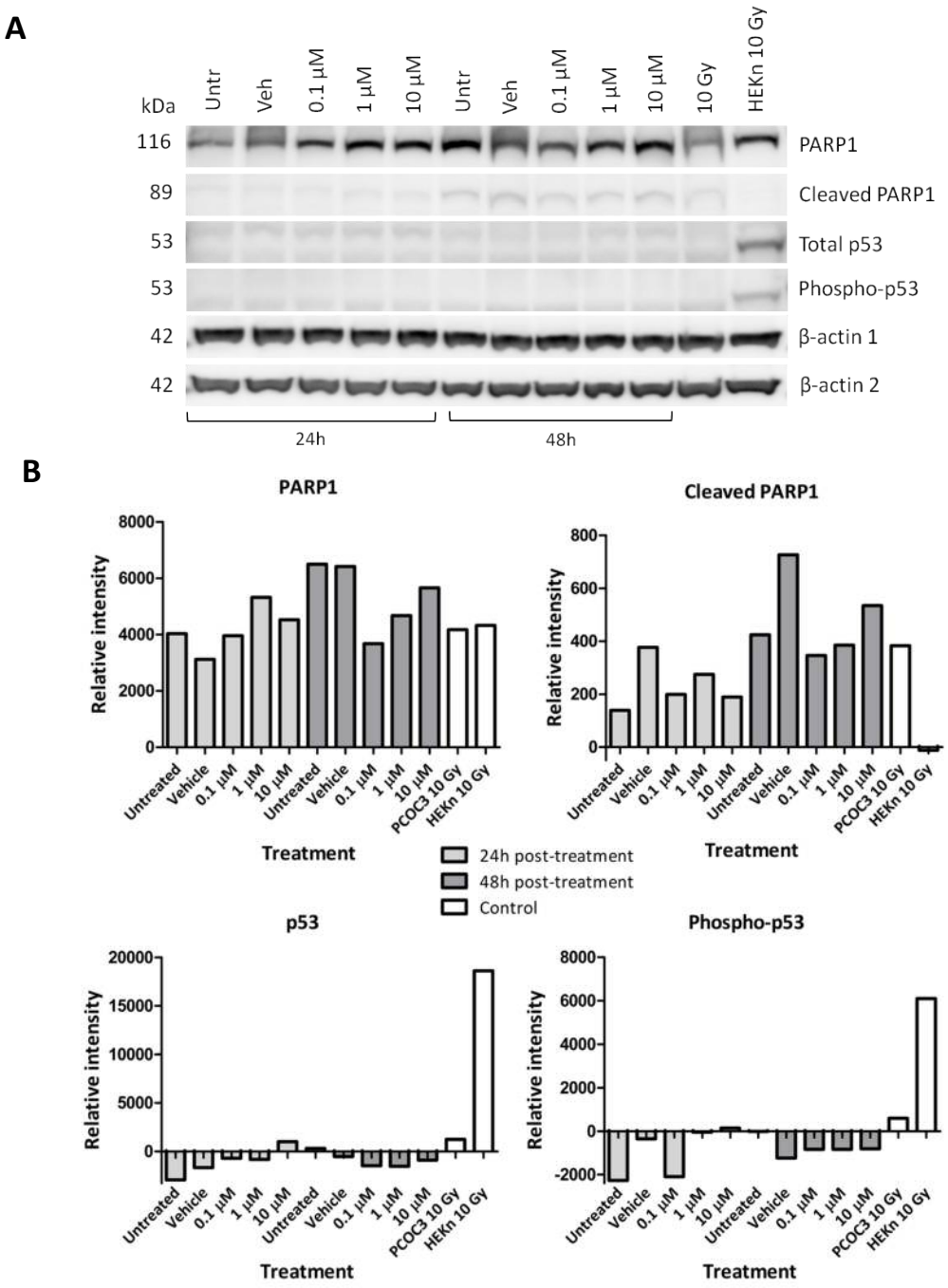


Figure 6.22: PCOC3 PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of PCOC3 after 24 h and 48 h Olaparib treatment. The cell line treated and HEKn treated with 10 Gy and collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

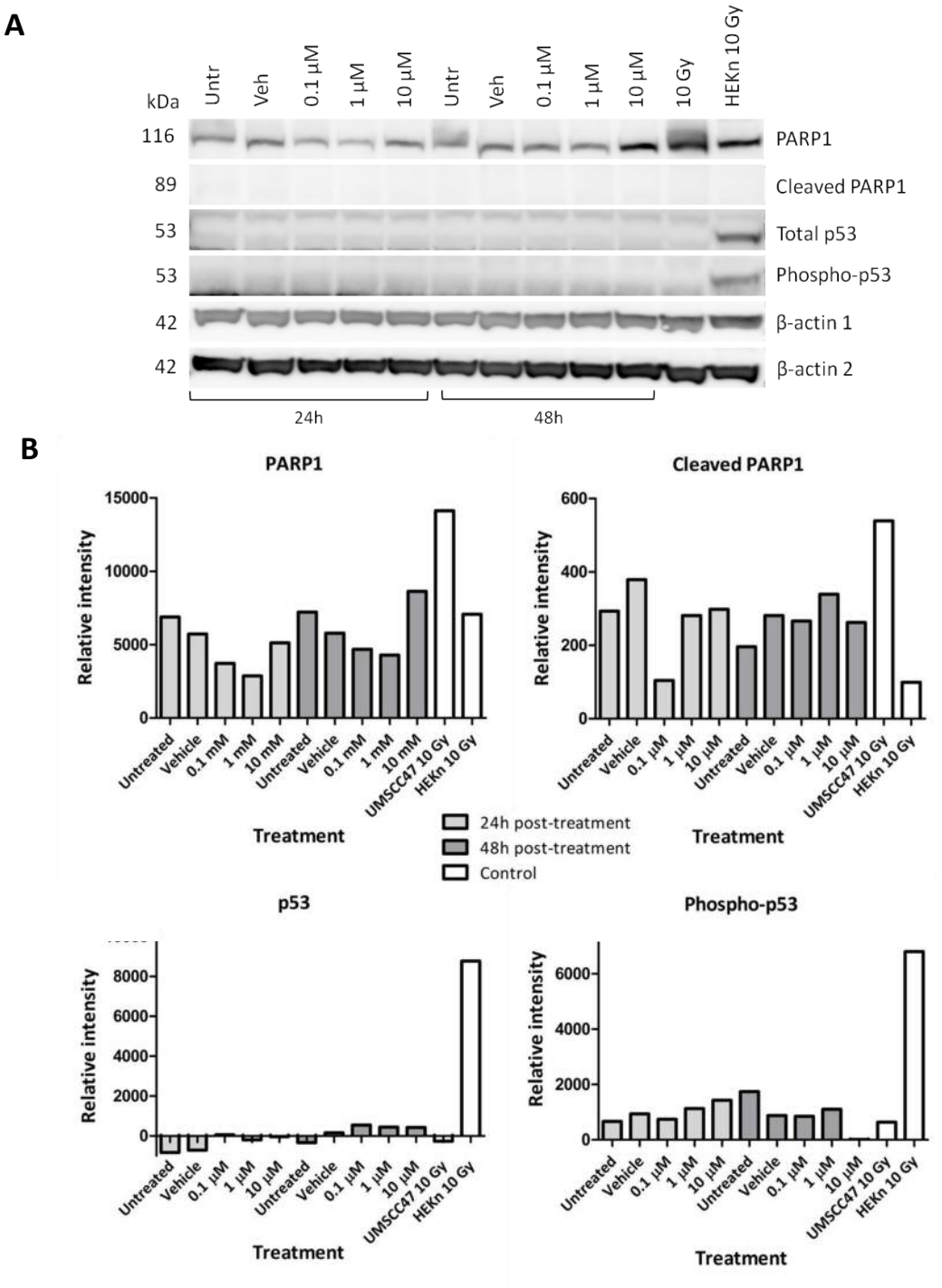


Figure 6.23: UMSCC47 PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of UMSCC47 after 24 h and 48 h Olaparib treatment. The cell line and HEKn treated with 10 Gy and collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

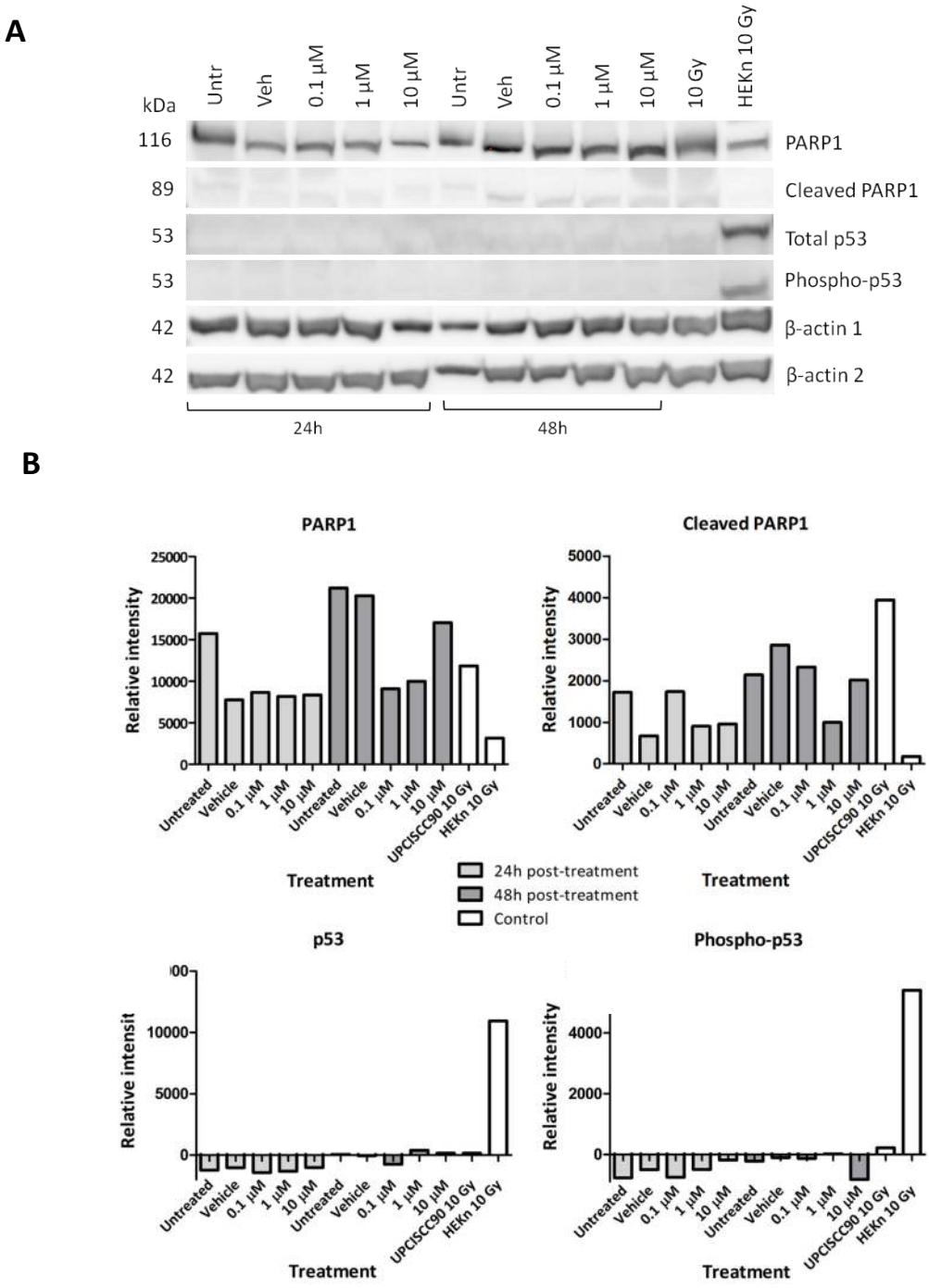


Figure 6.24: UPCISCC90 PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of UPCISCC90 after 24 h and 48 h Olaparib treatment. The cell line and HEKn treated with 10 Gy and collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

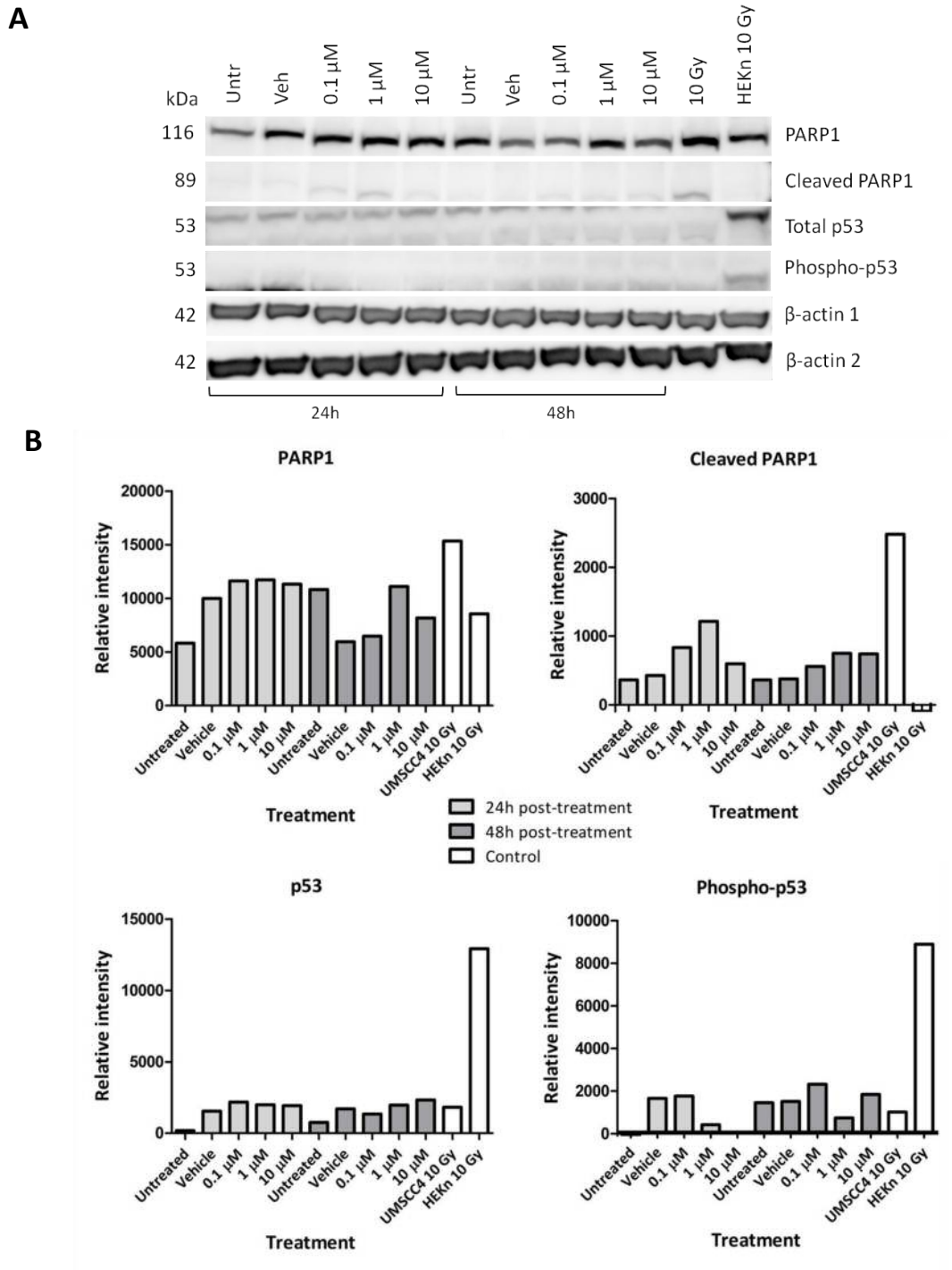


Figure 6.25: UMSCC4 PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of UMSCC4 after 24 h and 48 h Olaparib treatment. The cell line and HEKn treated with 10 Gy and collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

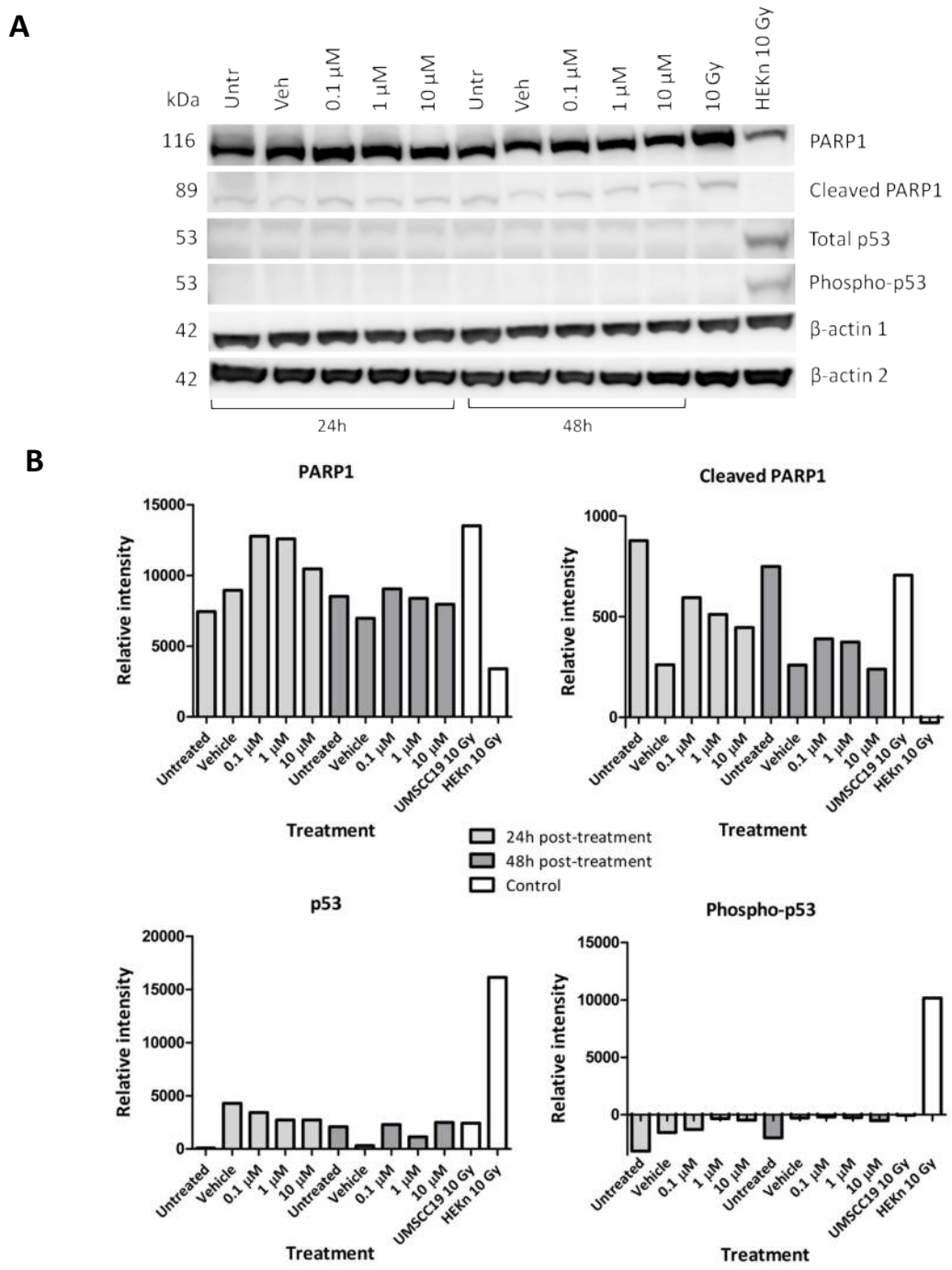


Figure 6.26: UMSCC19 PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of UMSCC19 after 24 h and 48 h Olaparib treatment. The cell line and HEKn treated with 10 Gy and collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

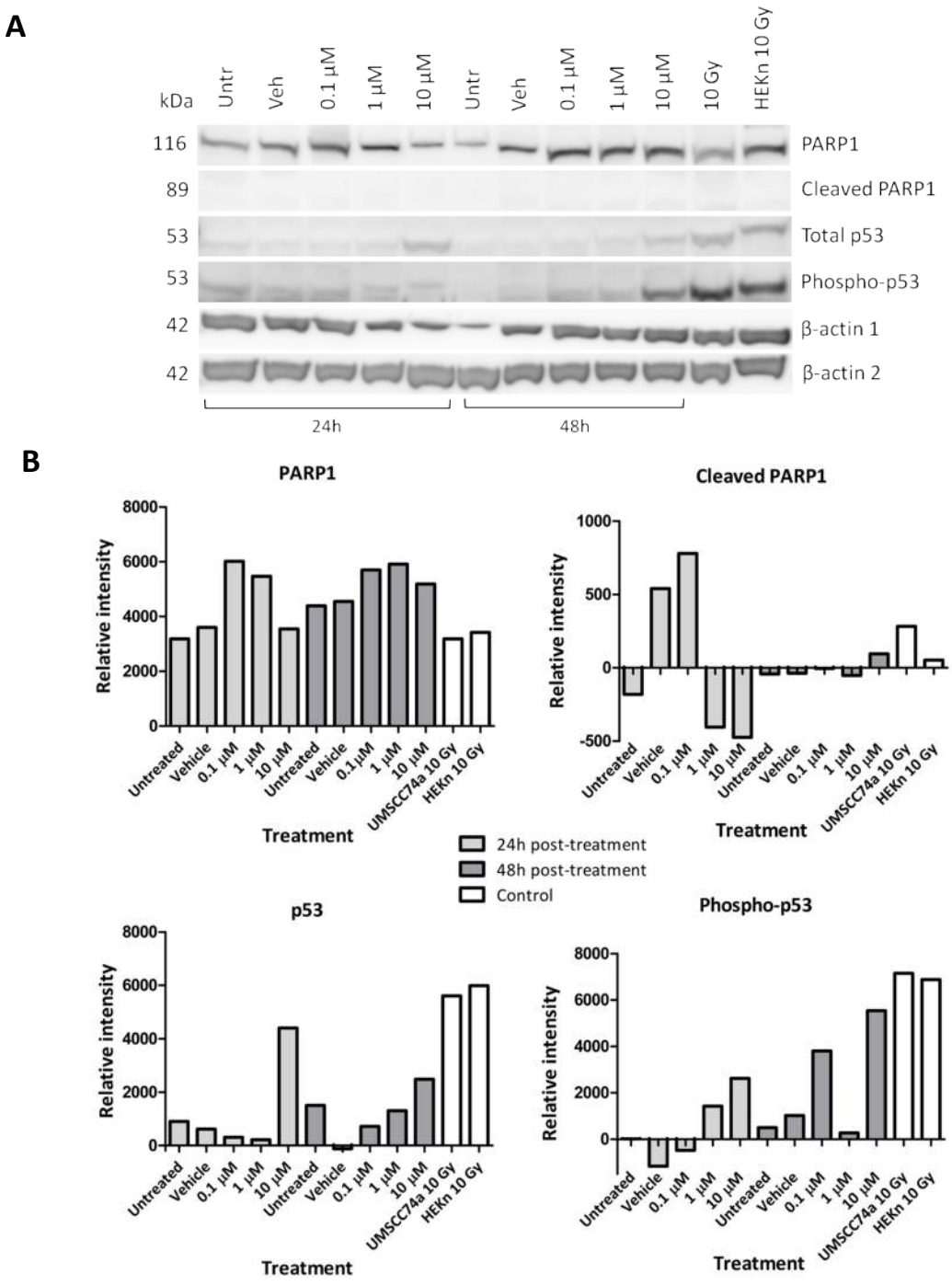


Figure 6.27: UMSCC74a PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of UMSCC74a after 24 h and 48 h Olaparib treatment. The cell line and HEKn treated with 10 Gy and collected after 24 h were used as controls. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

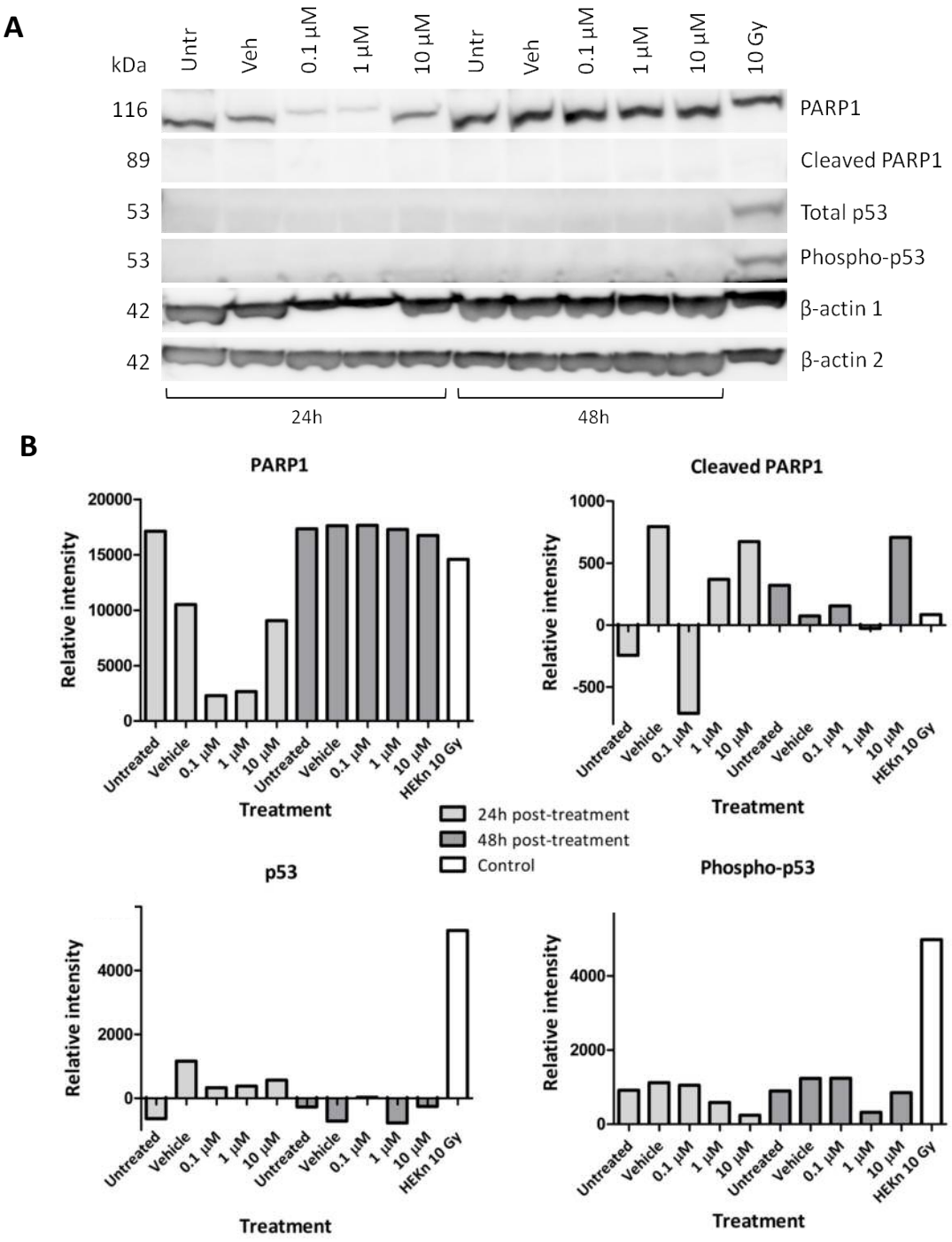


Figure 6.28: HEKn PARP1, cleaved PARP1, p53 and p-p53 western blot images and intensity plots after 24 h and 48 h Olaparib treatment

A: PARP1, cleaved PARP1, p53 and p-p53 western blot images of HEKn after 24 h and 48 h Olaparib treatment. The cell line treated 10 Gy and collected after 24 h was used as a control. Two membranes were used, one for PARP1, cleaved PARP1 (18 min exposure each), p-p53 (25 min exposure) and β-actin 1 (2 min exposure) and the other for p53 (25 min exposure) and β-Actin 2 (2 min exposure).

B: The intensity of the western blotting bands as quantified in relation to β-actin level using Image studio Lite.

6.7 Summary of results

The results of the assays described in this chapter are summarised in Table 6.1.

Table 6.1: Data summary

Cell line[a]	Clonogenic assay[1]		DSB[2]		S-phase[3]		G2-phase[4]		PARP1[5]	p53[6]		p-p53[7]		Cleaved PARP1[8]	
	0.5 µM[a]	10 µM[a]	10 µM[a]		10 µM[a]		10 µM[a]		0 µM[a]	10 µM[a]		10 µM[a]		10 µM[a]	
	10-15 days	10-15 days	24h	48h	24h	48h	24h	48h	0h	24h	48h	24h	48h	24h	48h
PCOC2															
PCOC3															
UMSCC47															
UPCISCC90															
UMSCC4															
UMSCC19															
UMSCC74a															
HEKn															

[a]Purple HPV-positive, Blue HPV-negative, Grey normal.

[1]Red indicates resistant cell lines (SF>80%); blue indicates medium sensitivity cell lines (SF 40-80%); green indicates sensitive cell lines (SF<40%). Colours are specific to indicated dose and timepoint.

[2]Green indicates that the relevant cell line accumulated DSB at the indicated dose and time-point.

[3]Green indicates that the relevant cell line showed significant changes to cell cycle distribution at the indicated dose and time-point.

[4]Green indicates that the relevant cell line showed significant changes to cell cycle distribution at the indicated dose and time-point.

[5]Green indicates that the relevant cell line showed an increase in PARP1 protein level at the indicated dose and time-point (+ indicates above average levels).

[6]Green indicates that the relevant cell line accumulated p53 protein at the indicated dose and time-point.

[7]Green indicates that the relevant cell line accumulated p-p53 protein at the indicated dose and time-point.

[8]Green indicates that the relevant cell line accumulated cleaved PARP1 protein at the indicated dose and time-point.

[a] Indicates dose of Olaparib.

6.8 Discussion

6.8.1 Main findings

With reference to the original hypotheses, these investigations established that:

- The cell lines displayed a wide range of Olaparib sensitivities:
    - Two HPV-positive lines (PCOC2 and UPCISCC90), the HPV-negative line UMSCC74a and the normal keratinocytes HEKn were sensitive to the lowest dose tested (0.1 μM).
    - All lines were sensitive to 10 μM treatment.
- Olaparib sensitivity did not correlate with HPV status, although the most sensitive lines were HPV-positive. Three groups were apparent after 0.5 μM and 1 μM treatment:
    - Sensitive: two HPV-positive lines.
    - Intermediate sensitivity: one HPV-negative line and normal keratinocytes.
    - Resistant: 2 HPV-positive and 2 HPV-negative lines.
- Following treatment with Olaparib, HPV-positive cells did not accumulate higher levels of unrepaired DNA DSB than HPV-negative lines. Forty-eight hours after 10 μM Olaparib treatment, substantial levels of unrepaired DSB were evident in 2/4 HPV-positive lines and 3/3 HPV-negative lines.
- Among the five cell lines that accumulated DSB, only the HPV-positive cell lines UPCISCC90 and UMSCC47 showed evidence of significant cell cycle arrest 48 h after treatment.
- There was a partial correlation between basal levels of PARP1 and sensitivity to Olaparib. UPCISCC90 showed the highest basal level and was the most sensitive cell line. However, UMSCC19 also showed relatively high basal level of PARP1 but was one of the least sensitive lines.
- There was no evidence of apoptosis induction (cleaved PARP1) in any of the cell lines following treatment with Olaparib.
- Treatment with Olaparib was not associated with p53 reactivation (accumulation of p53 or presence of p-p53) in HPV-positive cell lines. Although p53 and p-p53 protein did appear to accumulate in UMSCC74a (HPV-negative) following treatment with 10 μM Olaparib.

6.8.2 Clonogenic assay

The cell line panel showed a wide range in sensitivity to a single dose of Olaparib. Several of the HPV-positive lines demonstrated high sensitivity to low doses of Olaparib (0.1 μM) whilst all lines were sensitive to 10 μM. The widest range of responses was observed at 0.5 and 1 μM. These concentrations illustrated a clear distinction between the high sensitivity of some HPV-positive cells (PCOC2, UPCISCC90), the medium sensitivity of an HPV-negative line (UMSCC74a) and HEKn, and the low sensitivity of a mix of HPV-positive and negative lines (PCOC3, UMSCC47, UMSCC4 and UMSCC19). The dose range was chosen based upon studies testing the effects of Olaparib on BRCA2 mutant and proficient breast cancer cell lines. Doses up to 10 μM enabled the distinction between Olaparib sensitive (BRCA2 mutant) and non-sensitive (BRCA2 proficient) cells (Evers et al. 2008). The 0.5 μM and 1 μM Olaparib treatments were in the range reported to be selective for cells with BRCA2 mutation.

One unanticipated finding was that the normal keratinocytes (HEKn) were sensitive to relatively low doses of Olaparib. As these cells are presumed to be competent in DSB repair, Olaparib sensitivity was not expected after low dose treatment. Alternatively, this could be interpreted as indicating the presence of a functioning response to DNA damage, which may include arrest in G2-phase followed by effective DNA repair, that was absent in the majority of the cancer cell lines. This is consistent with expansion of the proportion of cells in G2 at 24 h following treatment but a return to a normal cell cycle distribution at 48 h.

An overlap in sensitivity between HPV-positive and negative OPSCC cell lines has also been reported following IR (Rieckmann et al. 2013). The absence of a clear difference in radio-sensitivity up to 6 Gy has also been confirmed in our laboratory using the same cell line panel as used in this study (Stefan Holzhauser, unpublished data). The main difference between Olaparib and radio-sensitivity was related to UMSCC47 which in the study by Rieckmann, was the most radio-sensitive of the HPV-positive lines but showed low sensitivity to Olaparib. Similarly, cisplatin treatment, as a single agent, did not highlight a clear difference between HPV-positive and negative HNSCC cell lines (Busch et al. 2016). It is noteworthy that several studies have observed correlations in cell line sensitivity between Olaparib and other DNA damaging chemotherapeutics including carboplatin, camptothecin and doxorubicin (testing undertaken by KuDOS Pharmaceuticals and reported in the European Medicine Agency (EMA)

report (Committee for Medicinal Products for Human Use 2014). Together, this suggests that responses to DNA damaging therapies may not show a clear correlation with HPV status.

Three other studies have examined the effects of PARP1 inhibition in HNSCC cell lines (Güster et al. 2014; Weaver et al. 2015; Nickson et al. 2017). The study by Weaver et al. compared the response to Veliparib in two HPV-positive and one HPV-negative cell lines (UMSCC47, UPCISCC154 and UMSCC1 respectively). It showed the HPV-negative line to be the least sensitive (c. 1.5 fold difference). The paper by Nickson et al. is of particular interest, as it used the same drug (Olaparib) and four of the same cell lines as used in the current study (UMSCC 47, 6, 74a and UPCISCC90). They observed that the HPV-positive cell lines were most sensitive to IR, but reported Olaparib (up to 1 μM) as having no effect on cell proliferation (determined using clonogenic assays) in any line.

The study by Guster et al. investigated the effects of Olaparib in a panel of 5 treatment-naïve HPV-positive HNSCC (with one line, UMSCC47, in common with the current study). This showed broad variation in response to Olaparib (between approximately 10-45% inhibition following 5 days growth with 1 μM Olaparib). However, it was notable that the two lines with highest baseline PARP1 levels showed the greatest sensitivity to Olaparib.

Taken together these reports show quite diverse findings. The reasons why other studies report very different results is unclear but may be related to differences in dose and the assay methodologies (e.g. differences in length of drug exposure) There is the suggestion that DSB repair is sub-optimal in some, but not all, HPV-positive cell lines, and that this may be associated with delayed resolution of DSB due to low levels of BRCA2, DNA-PK and associated proteins. There is also a suggestion that BER and SSB repair proteins may be up-regulated in HPV-positive cells, but levels of PARP1 do show variation. It also appears that high levels of basal PARP1 show some correlation with response to Olaparib.

With regard to the strengths and weaknesses of the current study, the clonogenic data is very robust. This was achieved by thorough protocol development to optimise culture conditions including assay length and seeding density for all cell lines. This ensured reliable plating efficiency calculations which were the main measure of a clonogenic assay. To minimise the effects of Olaparib treatment on cell adhesion, plates were treated 24 h after seeding. In addition, an untreated plate was present in each experiment to assess the effect of

DMSO alone (this indicated that no individual cell line was differentially sensitive to DMSO, data not shown). Manual counting was reduced to a minimum by the use of an image analysis software thereby reducing subjectivity in the counting process. The software was based on pixel shade which could be affected by picture lighting. Hence three pictures were taken and counted for each plate. If manual counting was required (due to the growth characteristics of the individual cell line), two people performed independent counts.

Several assays are available to measure drug sensitivity (e.g. MTT, cell count with dye exclusion, flow cytometric counting, etc.). Clonogenic assays were chosen primarily because the protocol was developed to test both radio and drug-sensitivity separately or in combination. Clonogenic assays are recognised as the assay of choice for radio-sensitivity as well as being adapted to determine the sensitivity of tumour cells to chemotherapy agents or other drugs (Rafehi et al. 2011). This assay measures the capacity of individual cells to divide following treatment and may detect a higher level of sensitivity than other assays (Bird et al. 1987). This effect has previously been observed in HNSCC cell lines which displayed sensitivity to chemotherapy agents at lower doses in clonogenic assays compared to MTT or counting assays (Henriksson et al. 2006). Similarly, no response to 0.1 μM or 1 μM treatment was detected by flow cytometry even in cell lines that were sensitive to these concentrations in the clonogenic assays. This could be an example of the greater potential sensitivity of the clonogenic assays.

Overall, the clonogenic assays demonstrated the efficacy of Olaparib against our panel of OPSCC cell lines, including two novel lines. The main findings were efficacy at potentially therapeutic levels in some cell lines, but a lack of correlation between HPV status and Olaparib sensitivity. Hence it appears that HPV status is unlikely to provide a convenient predictive biomarker to guide the use of Olaparib in the context of OPSCC. The same appears likely to be true of basal levels of PARP1. However, these predictions clearly require validation in the context of well controlled clinical trials. Clinical trials may also provide the bio-resources and data to assess correlations between treatment response and other markers of a “BRCA-like” phenotype that may be associated with response to Olaparib (Lim and Ngeow 2016).

6.8.3 γ-H2AX and cell cycle analysis

Levels of γ-H2AX, indicating the presence of DSB, and cell cycle distribution were assessed following treatment with Olaparib. Effects on these measures were only observed after 10 μM treatment. Lower doses of 0.1 μM and 1 μM did not lead to the formation of detectable DSB or cell cycle alterations. Cells were assessed at two timepoints (24 and 48 h), of these, 48 h is possibly a better indicator to the presence of delayed or suboptimal repair of DSB. Forty-eight hours after 10 μM treatment, substantial levels of unrepaired DSB were evident in 2/4 HPV-positive lines and 3/3 HPV-negative lines. Hence the data do not appear consistent with the hypothesis that HPV-positive cells would accumulate higher levels of unrepaired DNA DSB than HPV-negative lines due to defective DSB repair. Similarly, there was no correlation between HPV status and the degree of cell cycle perturbations in response to Olaparib (48 h after 10 μM Olaparib).

The presence of unrepaired DSB did not appear to correlate with effect on the cell cycle. Among the five cell lines that accumulated DSB, only UPCISCC90 and UMSCC47 showed evidence of cell cycle arrest 48 h after treatment. Conversely, all three of the HPV-negative OPSCC lines showed evidence of DSB 24 and 48 h after 10 μM treatment but did not show significant changes to cell cycle distribution at 24 or 48 h (although some increase in G2 fraction was apparent in UMSCC4 and 74a). This is consistent with delayed repair and impaired response to DNA damage in these lines. This possibility is supported by data from Kimple et al. that showed an absence of G2 arrest in several HPV-negative HNSCC cell lines following 4 Gy IR. However, they also reported significant G2 arrest in 4 HPV-positive lines (including UMSCC47 and UPCISCC90), which is not entirely consistent with our observations, where a clear distinction in response according to HPV status was not apparent.(Kimple et al. 2013).

Of all the lines, UPCISCC90 accumulated the highest level of DSB (both after 24 h and 48 h). After 24 h, this was associated with an increase in the proportion of cells in S-phase, which could be attributed to a pause in the cell cycle to allow DNA repair. After 48 h, a significant increase in the G2-phase was much more apparent, suggesting cell cycle arrest at this point, with cells unable to complete mitosis.

The γ-H2AX level in normal keratinocytes HEKn was unaffected by 10 μM Olaparib treatment, but there was a slight increase in G2 fraction after 48 h, suggesting DNA damage

occurred below detectable levels, and was responded to. Following treatment with Etoposide, it was notable that levels of DSB induced were generally 1-5 fold higher in the cancer cell lines than in the normal cells. This might suggest less effective repair of DSB in both HPV-positive and negative cells than in normal cells, however as this observation is based on a single cell line, it should be treated with caution.

The absence of significant differences between HPV-positive and negative HNSCC cell lines in basal level of γ-H2AX or cell cycle distribution has been observed in other studies. Nickson et al. found similar basal levels of γ-H2AX in 4 OPSCC cell lines, 3 of which were included in the panel used in this chapter (Nickson et al. 2017). They also reported faster repair of IR induced DSB (indicated by loss of γ-H2AX foci) in UMSCC47 compared to UMSCC6 and 74a. The same patterns (lower γ-H2AX levels in UMSCC47 compared to UMSCC6 and 74a) were apparent in the current study. Weaver et al. published a study based on 3 HNSCC cell lines and reported slower resolution of IR induced γ-H2AX foci in the two HPV-positive lines (including UMSCC47) (Weaver et al. 2015). The findings reported in this thesis are broadly consistent with previously published data but are based on a greater number of cell lines. The study by Weaver et al. is also notable as it proposed a specific mechanism for defective repair of DSB in HPV-positive OPSCC i.e. that HPV-positive cells show reduced repair by both NHEJ and HR as a consequence of deficiencies in protein expression and recruitment of DNA-PK and BRCA2. In a separate study, low levels of DNA-PK and BRCA2 were shown in UPCISCC90 but not in UMSCC47 (Nickson et al. 2017).

The strengths of this section included the thoroughly optimised and robust experimental methods. Experiments were performed in triplicate for each cell line using the same settings on a single flow cytometer. The main weakness was the absence of another technique to confirm the DNA damage caused by Olaparib. Such as Immunofluorescence to observe the formation and resolution of DSB, the main advantage of this technique is that the number of γ-H2AX foci has been correlated with the number of DSB (Willmore et al. 1998). However, flow cytometry based approaches allow simultaneous observation of DSB and cell cycle distribution. Furthermore, some exposures, e.g. cisplatin, can lead to pan-nuclear staining associated with stalling of multiple replication forks, in which case flow cytometric assessment may be more appropriate (Busch et al. 2016).

6.8.4 Western blotting

There was no apparent correlation between basal PARP1 level and HPV status. However, UPCISCC90 showed elevated PARP1 levels compared to the other lines. In the HPV-positive cell lines, following Olaparib treatment, there was no detectable p53, or p-p53 protein nor any induction of apoptosis. In the HPV-negative line UMSCC74a, p53 and p-p53 were detected following exposure to Olaparib. Cleaved PARP1, indicating apoptotic cells, was not detected in any of the cell lines in response to Olaparib.

Whether p53 accumulates and causes apoptosis in HPV-positive HNSCC lines is controversial. A study by Kimple et al. reported accumulation of p53, caspase activation and increased annexin V binding in HPV-positive cells following IR. Although basal p53 protein and weak p-p53 protein in irradiated UPCISCC90 were only detected using western blots with what they describe as “long exposure”, and they reported absence an of p53 and p-p53 in UMSCC47 (as observed in the current study) (Kimple et al. 2013). A similar study by Rieckmann et al. reported no accumulation of p53 and the absence of caspase activation and PARP1 cleavage (Rieckmann et al. 2013). It should be noted however that the 15 HNSCC lines assessed represented a very diverse group and only 3 were included in both studies. This highlights the pitfalls of drawing conclusions based on heterogeneous groups, and demonstrates the need for additional *in vitro* models.

Another study examined cisplatin sensitivity in HPV-positive and negative HNSCC cell lines and showed only very low levels of caspase activity in response to cisplatin (Busch et al. 2016). This study assessed only one cell line from the current panel (UMSCC47) in which no apoptosis was observed. The authors concluded that it was questionable whether apoptosis played a major role in response to cisplatin under the conditions investigated.

The elevated PARP1 protein in untreated UPCISCC90 cells, and lower expression in UMSCC74a, was consistent with a previous study (Nickson et al. 2017). Basal PARP1 levels are relevant as, when Olaparib was tested against a broad panel of cancer cell lines, it was observed that low expression of PARP1 correlated with low sensitivity to Olaparib (testing undertaken by KuDOS Pharmaceuticals and reported in a European Medicine Agency (EMA) report (Committee for Medicinal Products for Human Use 2014). This is consistent with the

highest sensitivity to Olaparib in the current study being observed in line with the highest expression of PARP1 (UPCISCC90).

There was no evidence to support the hypothesis that functional p53 was present in HPV-positive cells (i.e. that HPV conferred a less complete p53 phenotype than p53 mutation). It was interesting that accumulation and activation of p53 was not observed in UMSCC47, which supposedly expresses wild-type p53. This may suggest that either this line has an unidentified defect in p53 signalling, or that the stimuli used (Olaparib and IR) were not appropriate to stabilise p53. However, the latter seems unlikely as p53 was consistently induced by IR in HEKn cells.

The strengths of these analyses include the use of well-validated assays with appropriate controls. Potential weakness included the absence of a positive control for p53 specific for each line. However, this was compensated for by the presence of irradiated HEKn on all membrane with visible p53 and p-p53. There are also some questions regarding whether the use of a more sensitive γ-H2AX detection assay would have been helpful.

6.8.5 Further work

The findings of this chapter could be extended and further validated by repeating the investigation of induction of DSB using immunohistochemical staining. The mechanism by which PARP1 inhibits cell proliferation could be further assessed using an RNA-seq based investigation of treated vs untreated cells. The findings regarding apoptosis could also be validated by using additional markers (e.g. a direct measure of caspase activity). Perhaps the highest priority however is the derivation and analysis of a greater number of OPSCC cell lines. The vulnerability of studies using a very limited number of cell lines has recently been highlighted by Busch et al. who observed complex patterns of response among a panel of 11 cell lines (Busch et al. 2016). Using only a subset of these lines could easily lead to erroneously simplistic models.

6.8.6 Conclusion

There are several clear conclusions that can be drawn from the work described. The main one being that HPV status did not correlate with sensitivity to Olaparib. It was apparent that several lines were sensitive to Olaparib at doses that could be clinically relevant (by

analogy to the sensitivity of BRCA mutant breast cancer cell lines). However, it is difficult to discern any characteristic that would be adequate as a predictive marker. Regarding the mechanism of sensitivity, there was some evidence for an association with cell cycle arrest, but the correlation was not perfect. There was no evidence that sensitivity to Olaparib was associated with apoptosis.

The results described in this chapter do not support the often repeated suggestion that there is a specific defect in repair of DSB in HPV-positive cells (Weaver et al. 2015; Nickson et al. 2017). Or at least, one that is therapeutically exploitable via synthetic lethal approaches. However, it is important to note, that in the current study we have predominantly investigated DSB arising from treatment with Olaparib, rather than DSB directly induced by IR, and there may be significant differences in response to lesions inflicted by different methods. The absence of cell cycle arrest in response to accumulation of DSB in the HPV-negative lines may suggest impaired surveillance of and/or response to DNA damage in these lines.

Only the results for UPCISCC90 were consistent with the original hypothesis: that in HPV-positive cells treatment with Olaparib would cause accumulation of DSB, resulting in cell cycle delays, and this would be associated with increased sensitivity to Olaparib demonstrated by clonogenic assays. For all the other cell lines, the data are more complex and difficult to interpret. The diverse responses across the cell line panel appears likely to be a true reflection of the biological heterogeneity of OPSCC and suggests that caution should be exercised with regard to conclusions drawn in studies including only a small number of cell lines.

Chapter 7 - Final discussion

HPV-positive OPSCC is a relatively new disease entity, and the natural history of HPV infection in the oropharynx is poorly defined. The work described in Chapter 3 was the second phase of a natural history study, which aimed to determine whether HPV infection is common in non-malignant tonsil tissue, and identify premalignant lesions (analogous to cervical intraepithelial neoplasia). It was based on a sample of 511 non-malignant tonsils from patients representing the demographic groups most likely to harbour high-risk HPV infection. Three HPV detection techniques, including the most sensitive commercial test available at the time, failed to detect any HPV presence in the samples. This study showed that HPV infection is a very rare event in non-malignant tonsils. This finding has now been replicated in a similar study in France (Franceschi et al. 2015). The low prevalence of HPV infection, coupled with the absence of characterised premalignant lesions, and localisation of HPV within tonsillar crypts, renders early detection of OPSCC difficult. This suggests that for prevention of OPSCC, prophylactic vaccination, possibly including gender-neutral vaccination, may be a better investment of health resources. However, it will be approximately 30-40 years before vaccination results in reduced incidence of OPSCC. This is because, in the UK, vaccination of 12 year-old girls began in 2008, while diagnosis of OPSCC most commonly occurs in men aged 55-65 years. Hence there will be many OPSCC patients requiring treatment for many years to come. Patients are still likely to present with advanced lesions and suffer serious side effects and late toxicities from their treatment. This again highlights the need for improved therapies for this condition, and specifically treatments targeting the growing number of HPV-positive patients, who are likely to respond better to therapy, and live longer with the consequences of their treatment.

Understanding of HPV-positive OPSCC at the molecular level, and development of aetiology-specific treatments, is hampered by a lack of *in vitro* models. In the current literature, approximately 5-8 HPV-positive OPSCC cell lines have been described. The study presented in Chapter 4 aimed to develop novel *in vitro* models of OPSCC. The 61.5% rate of explant growth showed the method used was appropriate to sustain growth of epithelial cells derived from oropharyngeal tissue. However, the lower overall success rate (15.4%) for cell line development illustrates the technical challenges and time commitment required for primary and early cell culture. Ultimately, two apparently immortal lines were produced

(PCOC2 and PCOC3), and the analyses undertaken on these lines confirmed that they: originated from the corresponding biopsies, were keratinocytes, and were positive for HPV16 infection. Further characterisation using the mRNA-sequ data generated in Chapter 5, established that PCOC2 and PCOC3 contained wild-type p53 and that HPV had integrated into the host DNA (in chromosome 10 and 20 respectively). This validated the two cell lines as appropriate models to study HPV-positive OPSCC. The use of PCOC2 and PCOC3 for the rest of the investigations provided the advantage of low passage lines typical of the demographic affected by HPV-driven oropharyngeal cancers (i.e. relatively young and fit, with limited co-morbidities associated with alcohol and tobacco).

A panel composed of the two novel lines, plus two other HPV-positive lines and 4 HPV-negative lines was selected to reflect some of the heterogeneity found in OPSCC. This included variables such as wild type versus mutated p53, treatment naive versus recurrent, and smoking versus non-smoking. The mRNA-seq analysis, described in chapter 5, was used to confirm the molecular characteristics of the panel and compare gene expression between HPV-positive and negative lines. This confirmed the p53 status, and in UMSCC4 and UMSCC19 previously described mutations were identified (Somers et al. 1992; Bradford et al. 2003; Ferris et al. 2005). Integration events were identified in UMSCC47 and UPCISCC90 with an additional site in chromosome 9 detected in UPCISCC90. This descriptive data, together with the confirmed HPV status and STR typing, validated the cell line panel as appropriate for further experiments.

The mRNA-seq analysis allowed comparison of the molecular characteristics of the HPV-positive and negative groups. GO analysis highlighted several significantly differentially regulated ontologies, mostly relating to cellular metabolism. However none of these pathways had an obvious relation to viral oncogenesis or treatment response. The hypothesis that DDR pathways or expression of individual gene involved in DDR would be significantly different between HPV-positive and negative lines was not confirmed. This hypothesis was based on the suggestion that APOBEC3A, induced in HPV-positive cells (Landry et al. 2011), and high levels of reactive oxygen species in HPV-positive tumours (Marullo et al. 2015), would activate BER. Also, the utilisation of components of the DSB repair machinery for HPV replication, might result in up-regulation of DSB repair pathways. However, the mRNA-seq data did not provide evidence of differences in DNA damage response or repair pathways. This appears to contradict the suggestion that HPV-positive cell lines might be defective in DNA damage

response and/or repair (Kimple et al. 2013; Rieckmann et al. 2013; Weaver et al. 2015). Although it should be noted that these analyses were conducted using cells that had not received additional directly DNA damaging treatment.

Concurrent with the mRNA-seq study, the cell line panel was used to evaluate the potential of synthetic lethal therapy using the PARPi Olaparib. Chapter 6 described the results of this study. The sensitivity of the lines to a single dose of Olaparib was assessed. Olaparib sensitivity showed only partial correlation with HPV status. Of the four HPV-positive lines, two were classed as sensitive and two as resistant; while the HPV-negative lines were classed as resistant, or showing intermediate sensitivity. Further investigation of Olaparib treated cells was undertaken and this showed no evidence to suggest less effective repair of DSB in HPV-positive cells. In fact, the data seem to suggest the opposite trend. Three of the HPV-positive lines showed significantly increased levels of γ-H2AX 24 h after treatment with 10 μM Olaparib, but in two of these lines, levels had returned to baseline by 48 h. In contrast, all of the HPV-negative OPSCC lines showed significantly increased levels of γ-H2AX at 24 h that were not repaired at 48 h. However, the presence of unrepaired DSB 48 h after 10 μM treatment did not uniformly induce G2 arrest as had been hypothesised. Significant G2 arrest only occurred in UPCISCC90.

PARP1 expression, as measured by western blot and mRNA-seq, did not show obvious correlation with HPV status, or with drug sensitivity. Although it was notable that mRNA and protein levels of PARP1 were highest in UPCISCC90, and this line was also the most sensitive to Olaparib. This confirmed that high PARP1 level can be associated with Olaparib sensitivity, but only in some lines. The absence of p53 reactivation in HPV-positive cells seen in western blots was in accordance with the absence of differences in expression of “genes involved in DNA damage response, signal transduction by p53 class mediator” observed in the mRNA-seq data.

It had previously been suggested that in HPV-positive cells, accumulation of DSB might be caused by a deficiency in DNA repair resulting in delayed repair (Rieckmann et al. 2013; Weaver et al. 2015). Kimple et al. however, suggested that re-activation of p53 was involved in the HPV-positive lines response to radiation, but this did not seem to be the case in response to Olaparib (Kimple et al. 2013).

The results presented are based on a small number of cell lines. But it should be noted that the number of cell lines used is similar to or higher than what was used in key publications within the field (Kimple et al. 2013; Rieckmann et al. 2013; Weaver et al. 2015; Ziemann et al. 2015). Furthermore, the lines used in the current study may well represent better models (being exclusively OPSCC and including lines derived from HPV-positive non-smokers, rather than diverse HNSCC). The heterogeneity of the Olaparib sensitivity appears likely to be a reflection of the heterogeneity of the cell line panel. It may well also be a reasonable representation of the heterogeneity of OPSCC patients. Studies in HNSCC lines have previously reported wide variations in sensitivity to chemo and radio-therapy with overlap between HPV-positive and negative lines (Kimple et al. 2013; Busch et al. 2016).

Cell line studies, especially using novel and/or early passage cell lines may be limited in scale by the technical work necessary to maintain the lines in culture and their sometimes slow growth. However, perhaps the greatest limitations are the 2 dimensional nature of model and the absence of the microenvironment, especially with regard to the absence of cells of the immune system. Animal models could be seen as an intermediate level model. However, when conducted using immune compromised mice, they suffer from similar limitations.

The data were not consistent with the original hypothesis (i.e. HPV-positive tumours will show greater sensitivity to Olaparib), but did suggest that monotherapy with PARPi might be useful in some OPSCC patients irrespective of HPV status. Consistent with this suggestion, pre-clinical studies, described in the marketing authorisation of Lynparza, showed that several head and neck cell lines (HPV-negative, some non-oropharyngeal) were sensitive to Olaparib at similar concentrations as some ovarian lines.

To determine whether Olaparib will be useful in clinical practise will ultimately depend on the results of clinical trials. In this context, it is pragmatic to evaluate Olaparib in combination with radiotherapy rather than as monotherapy. Combination of Olaparib with radiotherapy is the subject of investigation of several ongoing clinical trials. Nonetheless, complementary *in vitro* studies still have the potential to increase our understanding of the molecular effects of these therapies.

None of the characteristics of the cell lines investigated in this study, such as HPV, p53 or PARP1, were indicators of sensitivity to Olaparib. Given that treatment with Olaparib may

be effective in treating a subset of OPSCC patients, it would be desirable to identify predictive biomarker to guide its use. An mRNA-seq approach using treated and untreated lines would provide an exploratory method of investigating the transcriptional profile associated with Olaparib response. However, given that clinical trials using Olaparib are underway, there may be bio-resources associated with clinical outcome data, which would allow identification of predictive markers. Such investigations should also be informed by the results of studies seeking predictive biomarkers for Olaparib (i.e. markers of BRCA-ness) in other tumours (Lim and Ngeow 2016).

There remains a pressing need for additional cell line models of HPV-positive OPSCC, which are required to increase understanding of this disease at the molecular level. It is also the case that understanding the natural history of HPV in the oropharynx is proving a challenging endeavour, and further well-designed epidemiological studies are required. It appears that the likely utility of Olaparib is best addressed empirically through appropriate clinical trials. In terms of aetiology-specific treatments for OPSCC, our data suggest that targeting DNA repair may not be the most fruitful avenue of investigation. Given the possible significance of the immune response in outcome of OPSCC, modulation of this response may be a more attractive option for ongoing investigation.

7.1 Conclusion

In conclusion, this investigation found little evidence to support the suggestion that HPV infection is associated with defects in DNA repair, and sensitivity to Olaparib. mRNA-seq did not show differences in basal transcription of genes involved in DNA repair (although these analyses were conducted in untreated cells). Similarly, there was no consistent trend in terms of sensitivity to Olaparib, and HPV-positive OPSCC cell lines did not appear to accumulate higher levels of DSB in response to the drug. Furthermore, there was no evidence for accumulation and activation of p53 in HPV-positive cells. This strongly suggests that improved survival in HPV-positive patients is likely to be due to alternative factors, most likely including the presence of a non-self immune target in HPV-positive tumours.

Bibliography

Abbas, T. and Dutta, A. 2009. p21 in cancer: intricate networks and multiple activities. *Nature Reviews. Cancer* 9(6), pp. 400–14.

von Ahlfen, S. et al. 2007. Determinants of RNA quality from FFPE samples. Fraser, P. ed. *PloS one* 2(12), p. e1261.

Alexandrov, L.B. et al. 2013. Signatures of mutational processes in human cancer. *Nature* 500(7463), pp. 415–21.

Ali, A.A.E. et al. 2012. The zinc-finger domains of PARP1 cooperate to recognize DNA strand breaks. *Nature Structural & Molecular Biology* 19(7), pp. 685–92.

Altmeyer, M. et al. 2009. Molecular mechanism of poly(ADP-ribosyl)ation by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. *Nucleic acids research* 37(11), pp. 3723–38.

Amé, J.-C. et al. 2009. Radiation-induced mitotic catastrophe in PARG-deficient cells. *Journal of cell science* 122(Pt 12), pp. 1990–2002.

Amé, J.C. et al. 1999. PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase. *Journal of Biological Chemistry* 274(25), pp. 17860–8.

Anant, S. and Davidson, N.O. 2000. An AU-rich sequence element (UUUN[A/U]U) downstream of the edited C in apolipoprotein B mRNA is a high-affinity binding site for Apobec-1: binding of Apobec-1 to this motif in the 3' untranslated region of c-myc increases mRNA stability. *Molecular and cellular biology* 20(6), pp. 1982–92.

Ang, K.K. et al. 2010. Human papillomavirus and survival of patients with oropharyngeal cancer. *New England Journal of Medicine* 363(1), pp. 24–35.

Ang, K.K. and Sturgis, E.M. 2012. Human Papillomavirus as a marker of the natural history and response to therapy of head and neck squamous cell carcinoma. *Seminars in Radiation Oncology* 22(2), pp. 128–42.

Audeh, M.W. et al. 2010. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. *The Lancet* 376(9737), pp. 245–51.

Baan, R. et al. 2007. Carcinogenicity of alcoholic beverages. *The Lancet Oncology* 8(4), pp. 292–3.

Baker, C.C. et al. 1987. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *Journal of Virology* 61(4), pp. 962–71.

Balermpas, P. et al. 2014. Tumour-infiltrating lymphocytes predict response to definitive chemoradiotherapy in head and neck cancer. *British Journal of Cancer* 110(2), pp. 501–9.

Balermpas, P. et al. 2016. CD8+ tumour-infiltrating lymphocytes in relation to HPV status and clinical outcome in patients with head and neck cancer after postoperative chemoradiotherapy: A multicentre study of the German cancer consortium radiation oncology group (DKTK-ROG). *International Journal of Cancer* 138(1), pp. 171–181.

Balermpas, P. et al. 2017. The PD-1/PD-L1 axis and human papilloma virus in patients with head and neck cancer after adjuvant chemoradiotherapy: A multicentre study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). *International Journal of Cancer* 141(3), pp. 594–603.

Balló, H. et al. 1999. Establishment and characterization of four cell lines derived from human head and neck squamous cell carcinomas for an autologous tumor-fibroblast in vitro model. *Anticancer Research* 19(5B), pp. 3827–36.

Begum, S. et al. 2005. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clinical Cancer Research* 11(16), pp. 5694–99.

Benjamini, Y. and Hochberg, Y. 1995a. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* 57(1), pp. 289–300.

Benjamini, Y. and Hochberg, Y. 1995b. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* 57(1), pp. 289–300.

Bird, M.C. et al. 1987. Comparison of in vitro drug sensitivity by the differential staining cytotoxicity (DiSC) and colony-forming assays. *British Journal of Cancer* 55(4), pp. 429–31.

Bonner, W.M. et al. 2008. γH2AX and cancer. *Nature Reviews Cancer* 8(12), pp. 957–67.

Bradford, C.R. et al. 2003. p53 mutation correlates with cisplatin sensitivity in head and neck squamous cell carcinoma lines. *Head and Neck* 25(8), pp. 654–61.

Branzei, D. and Foiani, M. 2010. Maintaining genome stability at the replication fork. *Nature Reviews Molecular Cell Biology* 11(3), pp. 208–19.

Brennan, J.A. et al. 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.

Brenner, C.J. et al. 2010. Genotyping of 73 UM-SCC head and neck squamous cell carcinoma cell lines. *Head and Neck* 32(4), pp. 417–26.

Bryant, D. et al. 2014. mRNA sequencing of novel cell lines from human papillomavirus type-16 related vulval intraepithelial neoplasia: Consequences of expression of HPV16 E4 and E5. *Journal of Medical Virology* 86(9), pp. 1534–41.

Bryant, H.E. et al. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* 434(7035), pp. 913–7.

Buck, C.B. et al. 2008. Arrangement of L2 within the papillomavirus capsid. *Journal of Virology* 82(11), pp. 5190–7.

Burk, R.D. et al. 2013. Human papillomavirus genome variants. *Virology* 445(0), pp. 232–43.

Burk, R.D. et al. 2017. Integrated genomic and molecular characterization of cervical cancer. *Nature* 543(7645), pp. 378–384.

Busch, C.-J. et al. 2016. Similar cisplatin sensitivity of HPV-positive and -negative HNSCC cell lines. *Oncotarget* 7(24), pp. 35832–42.

Cancer Genome Atlas Network 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 517(7536), pp. 576–82.

Castellsagué, X. et al. 2009. Human Papillomavirus (HPV) infection in pregnant women and mother-to-child transmission of genital HPV genotypes: a prospective study in Spain. *BMC Infectious Diseases* 9(74)

Castellsagué, X. et al. 2016. HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients. *Journal of the National Cancer Institute* 108(6), pp. 1–12.

Chen, R. et al. 2005. Presence of DNA of human papillomavirus 16 but no other types in tumor-free tonsillar tissue. *Journal of Clinical Microbiology* 43(3), pp. 1408–10.

Chiu, Y.-L. and Greene, W.C. 2008. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annual Review of Immunology* 26(1), pp. 317–53.

Choo, K.-B. et al. 1987. Integration of human papillomavirus type 16 into cellular DNA of cervical carcinoma: preferential deletion of the E2 gene and invariable retention of the long control region and the E6/E7 open reading frames. *Virology* 161(1), pp. 259–61.

Clark, M.A. et al. 2000. Differential cytokeratin and glycoconjugate expression by the surface and crypt epithelia of human palatine tonsils. *Histochemistry and Cell Biology* 114(4), pp. 311–21.

Clewley, J.P. et al. 2009. Prevalence of disease related prion protein in anonymous tonsil specimens in Britain: cross sectional opportunistic survey. *BMJ* 338, p. 1442.

Committee for Medicinal Products for Human Use 2014. *CHMP assessment report Lynparza*.

Cornet, I. et al. 2012. Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *Journal of Virology* 86(12), pp. 6855–61.

Cotner-Gohara, E. et al. 2008. Two DNA-binding and nick recognition modules in human DNA ligase III. *Journal of Biological Chemistry* 283(16), pp. 10764–72.

Culp, T.D. et al. 2006. Keratinocyte-secreted laminin 5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells. *Journal of Virology* 80(18), pp. 8940–50.

D'Amours, D. et al. 1999. Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *The Biochemical Journal* 342(Pt 2), pp. 249–68.

D'Souza, G. et al. 2007. Case–control study of human papillomavirus and oropharyngeal cancer. *New England Journal of Medicine* 356(19), pp. 1944–56.

D'Souza, G. et al. 2014. Differences in oral sexual behaviors by gender, age, and race explain observed differences in prevalence of oral human papillomavirus infection. *PloS one* 9(1), p. e86023.

David, S.S. et al. 2007. Base-excision repair of oxidative DNA damage. *Nature* 447(7147), pp. 941–50.

Day, P.M. et al. 2008. Heparan sulfate-independent cell binding and infection with furin-precleaved papillomavirus capsids. *Journal of Virology* 82(24), pp. 12565–8.

Dayyani, F. et al. 2010. Meta-analysis of the impact of human papillomavirus (HPV) on cancer risk and overall survival in head and neck squamous cell carcinomas (HNSCC). *Head & Neck Oncology* 2(15), pp. 1–11.

Deffie, A. et al. 1993. The tumor suppressor p53 regulates its own transcription. *Molecular and Cellular Biology* 13(6), pp. 3415–23.

Denis, F. et al. 2003. Late toxicity results of the GORTEC 94-01 randomized trial comparing radiotherapy with concomitant radiochemotherapy for advanced-stage oropharynx carcinoma: comparison of LENT/SOMA, RTOG/EORTC, and NCI-CTC scoring systems. *International Journal of Radiation Oncology, Biology, Physics* 55(1), pp. 93–8.

Desai, S. et al. 2011. Prevalence of human papillomavirus antibodies in males and females in England. *Sexually Transmitted Diseases* 38(7), pp. 622–9.

Dok, R. et al. 2014. p16INK4a impairs homologous recombination–mediated DNA repair in human papillomavirus–positive head and neck tumors. *Cancer Research* 74(6), pp. 1739–51.

Doorbar, J. 2005. The papillomavirus life cycle. *Journal of Clinical Virology* 32(SUPPL.), pp. 7–15.

Doorbar, J. et al. 2012. The biology and life-cycle of human papillomaviruses. *Vaccine* 30(S5), pp. F55-70.

Doorbar, J. et al. 2016. Human papillomavirus molecular biology and disease association. *Reviews in Medical Virology* 25, pp. 2–23.

Dracopoli, N.C. and Fogh, J. 1983. Loss of heterozygosity in cultured human tumor cell lines. *Journal of the National Cancer Institute* 70(1), pp. 83–7.

Du, C. et al. 2016. Microarray data analysis to identify crucial genes regulated by CEBPB in human SNB19 glioma cells. *World Journal of Surgical Oncology* 14(1), p. 258.

Duensing, S. et al. 2000. The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle. *Proceedings of the National Academy of Sciences of the United States of America* 97(18), pp. 10002–7.

Duensing, S. et al. 2001. Human papillomavirus type 16 E7 oncoprotein-induced abnormal centrosome synthesis is an early event in the evolving malignant phenotype. *Cancer Research* 61(6), pp. 2356–60.

Duensing, S. and Münger, K. 2004. Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *International Journal of Cancer* 109(2), pp. 157–62.

Dürst, M. et al. 1983. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proceedings of the National Academy of Sciences of the United States of America* 80(12), pp. 3812–5.

Dyson, N. 1998. The regulation of E2F by pRB-family proteins. *Genes & Development* 12(15), pp. 2245–62.

Egawa, N. et al. 2015. Human papillomaviruses; epithelial tropisms, and the development of neoplasia. *Viruses* 7(7), pp. 3863–90.

Eisbruch, A. et al. 2011. Chemo-IMRT of oropharyngeal cancer aiming to reduce dysphagia: swallowing organs late complication probabilities and dosimetric correlates. *International Journal of Radiation Oncology, Biology, Physics* 81(3), pp. e93-9.

Eklund, C. et al. 2012. The 2010 global proficiency study of human papillomavirus genotyping in vaccinology. *Journal of Clinical Microbiology* 50(7), pp. 2289–98.

El-Khamisy, S.F. et al. 2003. A requirement for PARP-1 for the assembly or stability of XRCC1 nuclear foci at sites of oxidative DNA damage. *Nucleic Acids Research* 31(19), pp. 5526–33.

Elenius, K. et al. 1991. Induced expression of syndecan in healing wounds. *The Journal of Cell Biology* 114(3), pp. 585–95.

Elsum, I. et al. 2012. The scribble–Dlg–Lgl polarity module in development and cancer: from flies to man. *Essays In Biochemistry* 53, pp. 141–68.

Ernster, J.A. et al. 2009. Prevalence of oncogenic human papillomavirus 16 and 18 in the palatine tonsils of the general adult population. *Archives of Otolaryngology Head and Neck Surgery* 135(6), pp. 554–7.

Evans, M. et al. 2013. Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population. *BMC cancer*

13(220)

Evans, M. and Jones, T.M. 2016. Transoral surgery or radiotherapy for oropharyngeal carcinoma – Is it either or…? *Clinical Oncology* 28(7), pp. 413–20.

Evers, B. et al. 2008. Selective inhibition of BRCA2-deficient mammary tumor cell growth by AZD2281 and cisplatin. *Clinical Cancer Research* 14(12), pp. 3916–25.

Eyfjord, J.E. and Bodvarsdottir, S.K. 2005. Genomic instability and cancer: networks involved in response to DNA damage. *Mutation Research* 592(1), pp. 18–28.

Faden, D.L. et al. 2017. Multi-modality analysis supports APOBEC as a major source of mutations in head and neck squamous cell carcinoma. *Oral Oncology* 74, pp. 8–14.

Fehrmann, F. et al. 2003. Human Papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation. *Journal of Virology* 77(5), pp. 2819–31.

Fenner, F. 1976. The classification and nomenclature of viruses. Summary of results of meetings of the International Committee on Taxonomy of Viruses in Madrid, September 1975. *Virology* 71(2), pp. 371–8.

Ferlay, J. et al. 2015. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* 136(5), pp. E359-86.

Ferris, R.L. et al. 2005. Human papillomavirus-16 associated squamous cell carcinoma of the head and neck (SCCHN): A natural disease model provides insights into viral carcinogenesis. *European Journal of Cancer* 41(5), pp. 807–15.

Flores, E.R. et al. 1999. Establishment of the Human Papillomavirus Type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line. *Virology* 262, pp. 344–54.

Fradet-Turcotte, A. et al. 2011. Nuclear accumulation of the papillomavirus E1 helicase blocks S-phase progression and triggers an ATM-dependent DNA damage response. *Journal of Virology* 85(17), pp. 8996–9012.

Franceschi, S. et al. 1990. Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in Northern Italy. *Cancer Research* 50(20), pp. 6502–7.

Franceschi, S. et al. 2015. Deep brush-based cytology in tonsils resected for benign diseases. *International Journal of Cancer* 137(12), pp. 2994–9.

Galande, S. and Kohwi-Shigematsu, T. 1999. Poly(ADP-ribose) polymerase and Ku autoantigen form a complex and synergistically bind to matrix attachment sequences. *The Journal of Biological Chemistry* 274(29), pp. 20521–8.

Gelboin, H. V 1980. Benzo[alpha]pyrene metabolism, activation and carcinogenesis: role and regulation of mixed-function oxidases and related enzymes. *Physiological Reviews* 60(4), pp. 1107–66.

Gillespie, K.A. et al. 2012. Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. *Journal of Virology* 86(17), pp. 9520–6.

Gillison, M. et al. 2012. Prevalence of Oral HPV Infection in the United States, 2009-2010. *JAMA* 307(7), p. 693.

Gillison, M.L. et al. 2000. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *Journal of the National Cancer Institute* 92(9), pp. 709–20.

Gillison, M.L. et al. 2012. Tobacco smoking and increased risk of death and progression for patients with p16-positive and p16-negative oropharyngeal cancer. *Journal of Clinical Oncology* 30(17), pp. 2102–11.

Gillison, M.L. et al. 2013. Eurogin roadmap: comparative epidemiology of HPV infection and associated cancers of the head and neck and cervix. *International Journal of Cancer* 134(3), pp. 497–507.

Gillison, M.L. et al. 2015. Epidemiology of human papillomavirus-positive head and neck squamous cell carcinoma. *Journal of Clinical Oncology* 33(29), pp. 3235–42.

Gillison, M.L. and Shah, K. V 2001. Human papillomavirus-associated head and neck squamous cell carcinoma: mounting evidence for an etiologic role for human papillomavirus in a subset of head and neck cancers. *Current Opinion in Oncology* 13(3), pp. 183–8.

Goldenberg, D. et al. 2008. Cystic lymph node metastasis in patients with head and neck cancer: an HPV-associated phenomenon. *Head & Neck* 30(7), pp. 898–903.

Gu, J. et al. 2007. XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO Journal* 26(4), pp. 1010–23.

Güster, J.D. et al. 2014. The inhibition of PARP but not EGFR results in the radiosensitization of HPV/p16-positive HNSCC cell lines. *Radiotherapy and Oncology* 113(3), pp. 345–51.

Haince, J.-F. et al. 2007. Ataxia telangiectasia mutated (ATM) signaling network is modulated by a novel poly(ADP-ribose)-dependent pathway in the early response to DNA-damaging agents. *The Journal of Biological Chemistry* 282(22), pp. 16441–53.

Haince, J.-F. et al. 2008. PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *The Journal of Biological Chemistry* 283(2), pp. 1197–208.

Hakenewerth, A.M. et al. 2011. Joint effects of alcohol consumption and polymorphisms in alcohol and oxidative stress metabolism genes on risk of head and neck cancer. *Cancer Epidemiology and Prevention Biomarkers* 20(11), pp. 2438–49.

Hariri, S. et al. 2011. Prevalence of genital human papillomavirus among females in the

United States, the National Health and Nutrition Examination Survey, 2003-2006. *Journal of Infectious Diseases* 204(4), pp. 566–73.

Hashim, D. et al. 2016. The role of oral hygiene in head and neck cancer: results from International Head and Neck Cancer Epidemiology (INHANCE) consortium. *Annals of Oncology* 27(8), pp. 1619–25.

zur Hausen, H. 2002. Papillomaviruses and cancer: from basic studies to clinical application. *Nature Reviews Cancer* 2(5), pp. 342–50.

Heck, J.E. et al. 2010. Sexual behaviours and the risk of head and neck cancers: a pooled analysis in the International Head and Neck Cancer Epidemiology (INHANCE) consortium. *International Journal of Epidemiology* 39(1), pp. 166–81.

Helleday, T. 2011. The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings. *Molecular Oncology* 5(4), pp. 387–93.

Henderson, S. et al. 2014. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. *Cell Reports* 7(6), pp. 1833–41.

Henriksson, E. et al. 2006. Differences in estimates of cisplatin-induced cell kill in vitro between colorimetric and cell count/colony assays. *In Vitro Cell & Developmental Biology. Animal.* 42(10), pp. 320–3.

Heusinkveld, M. et al. 2012. Systemic and local human papillomavirus 16-specific T-cell immunity in patients with head and neck cancer. *International Journal of Cancer* 131(2), pp. E74–E85.

Hoffmann, T.K. et al. 2006. T cells specific for HPV16 E7 epitopes in patients with squamous cell carcinoma of the oropharynx. *International Journal of Cancer* 118(8), pp. 1984–91.

Hong, S. et al. 2015. STAT-5 regulates transcription of the Topoisomerase IIβ-Binding Protein 1 (TopBP1) gene to activate the ATR pathway and promote human papillomavirus replication. *mBio* 6(6), pp. e02006-15.

IARC 2007. *IARC monographs on the evaluation of carcinogenic risks to human (vol 90 HPV)*. International Agency for Research on Cancer.

Israr, M. et al. 2016. Comparison of human papillomavirus type 16 replication in tonsil and foreskin epithelia. *Virology* 499, pp. 82–90.

Jackson, S.P. 2002. Sensing and repairing DNA double-strand breaks. *Carcinogenesis* 23(5), pp. 687–96.

Jacobs, M. V et al. 1997. A general primer GP5+/GP6+-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *Journal of Clinical Microbiology* 35(3), pp. 791–795.

Jagtap, P. and Szabó, C. 2005. Poly(ADP-ribose) polymerase and the therapeutic effects of its inhibitors. *Nature Reviews Drug Discovery* 4(5), pp. 421–40.

Jeon, S. et al. 1995. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *Journal of Virology* 69(5), pp. 2989–97.

Kaelin, W.G. 2005. The concept of synthetic lethality in the context of anticancer therapy. *Nature Reviews Cancer* 5(9), pp. 689–98.

Kameshita, I. et al. 1984. Poly (ADP-Ribose) synthetase. Separation and identification of three proteolytic fragments as the substrate-binding domain, the DNA-binding domain, and the automodification domain. *The Journal of Biological Chemistry* 259(8), pp. 4770–6.

Kammer, C. et al. 2000. Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *Journal of General Virology* 81, pp. 1975–81.

Kanginakudru, S. et al. 2015. Levels of the E2 interacting protein TopBP1 modulate papillomavirus maintenance stage replication. *Virology* 478, pp. 129–35.

Kaufmann, S.H. et al. 1993. Specific proteolytic cleavage of poly(ADP-ribose) polymerase: an early marker of chemotherapy-induced apoptosis. *Cancer Research* 53(17), pp. 3976–85.

Kim, H.S. et al. 2016. Association between PD-L1 and HPV status and the prognostic value of PD-L1 in oropharyngeal squamous cell carcinoma. *Cancer research and treatment* 48(2), pp. 527–36.

Kimple, R.J. et al. 2013. Enhanced radiation sensitivity in HPV-positive head and neck cancer. *Cancer Research* 73(15), pp. 4791–800.

Kishore, A.H. et al. 2007. p53 regulates its own activator: transcriptional co-activator PC4, a new p53-responsive gene. *The Biochemical Journal* 406(3), pp. 437–44.

Klingelhutz, A.J. et al. 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature* 380(6569), pp. 79–82.

Klug, S.J. et al. 2009. TP53 codon 72 polymorphism and cervical cancer: a pooled analysis of individual data from 49 studies. *The Lancet Oncology* 10(8), pp. 772–84.

Kolodner, R.D. et al. 2002. Maintenance of genome stability in saccharomyces cerevisiae. *Science* 297(5581), pp. 552–7.

Kowalski, M. et al. 2009. Genetic polymorphisms in DNA base excision repair gene XRCC1 and the risk of squamous cell carcinoma of the head and neck. *Journal of Experimental & Clinical Cancer Research* 28(37)

Kreimer, A.R. et al. 2010. Oral human papillomavirus in healthy individuals: a systematic

review of the literature. *Sexually Transmitted Diseases* 37(6), pp. 386–91.

Krishnakumar, R. and Kraus, W.L. 2010. The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Molecular Cell* 39(1), pp. 8–24.

Krokan, H.E. et al. 1997. DNA glycosylases in the base excision repair of DNA. *The Biochemical Journal* 325(1), pp. 1–16.

Krzywinski, M. et al. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research* 19(9), pp. 1639–45. doi: 10.1101/gr.092759.109.

Kubota, Y. et al. 1996. Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. *The EMBO journal* 15(23), pp. 6662–70.

Landry, S. et al. 2011. APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. *EMBO reports* 12(5), pp. 444–50.

Langelier, M.-F. et al. 2012. Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1. *Science* 336(6082), pp. 728–32.

Lansford, C.D. et al. 1999. Chapter 28 Head and Neck Cancers., pp. 185–255.

Lawrence, M. et al. 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 517(7536), pp. 576–82.

Leibiger, C. et al. 2013. First molecular cytogenetic high resolution characterization of the NIH 3T3 cell line by murine multicolor banding. *The Journal of Histochemistry and Cytochemistry* 61(4), pp. 306–12.

Liao, W. et al. 1999. APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochemical and Biophysical Research Communications* 260(2), pp. 398–404.

Lieber, M.R. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end joining pathway. *Annual Review of Biochemistry* 79, pp. 181–211.

Lim, D. and Ngeow, J. 2016. Evaluation of the methods to identify patients who may benefit from PARP inhibitor use. *Endocrine-Related Cancer* 23(6), pp. R267-85.

Lin, B.Y. et al. 2002. Chaperone proteins abrogate inhibition of the human papillomavirus (HPV) E1 replicative helicase by the HPV E2 protein. *Molecular and cellular biology* 22(18), pp. 6592–604.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature* 362(6422), pp. 709–15.

Lissowska, J. et al. 2003. Smoking, alcohol, diet, dentition and sexual practices in the epidemiology of oral cancer in Poland. *European Journal of Cancer Prevention* 12(1), pp. 25–33.

Liu, X. et al. 2007. Myc and human papillomavirus type 16 E7 genes cooperate to immortalize human keratinocytes. *Journal of Virology* 81(22), pp. 12689–95.

Longworth, M.S. and Laimins, L. a 2004. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiology and molecular biology reviews* 68(2), pp. 362–72.

Lyford-Pike, S. et al. 2013. Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma. *Cancer research* 73(6), pp. 1733–41.

Ma, Y. et al. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 108(6), pp. 781–94.

Ma, Y. et al. 2004. A biochemically defined system for mammalian nonhomologous DNA end joining. *Molecular Cell* 16(5), pp. 701–13.

Mammas, I.N. et al. 2006. Human papilloma virus in hyperplastic tonsillar and adenoid tissues in children. *The Pediatric Infectious Disease Journal* 25(12), pp. 1158–62.

Mandic, R. et al. 2005. Reduced cisplatin sensitivity of head and neck squamous cell carcinoma cell lines correlates with mutations affecting the COOH-terminal nuclear localization signal of p53. *Clinical Cancer Research* 11(19), pp. 6845–52.

Marullo, R. et al. 2015. HPV16 E6 and E7 proteins induce a chronic oxidative stress response via NOX2 that causes genomic instability and increased susceptibility to DNA damage in head and neck cancer cells. *Carcinogenesis* 36(11), pp. 1397–406.

Masson, M. et al. 1998. XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Molecular and Cellular Biology* 18(6), pp. 3563–71.

Masterson, L. et al. 2014. De-escalation treatment protocols for human papillomavirus-associated oropharyngeal squamous cell carcinoma: a systematic review and meta-analysis of current clinical trials. *European Journal of Cancer* 50(15), pp. 2636–48.

Maxwell, J.H. et al. 2010. Tobacco use in HPV-positive advanced oropharynx cancer patients related to increased risk of distant metastases and tumor recurrence. *Clinical Cancer Research* 16(4), p. 1226.

Mazerska, Z. et al. 2016. The role of glucuronidation in drug resistance. *Pharmacology & Therapeutics* 159, pp. 35–55.

Mehanna, H. et al. 2010. Head and neck cancer--Part 1: epidemiology, presentation, and prevention. *BMJ* 341(c4684), pp. 663–6.

Mehanna, H. et al. 2016. Geographic variation in human papillomavirus-related oropharyngeal cancer: data from 4 multinational randomized trials. *Head & neck* 38(Suppl 1), pp. E1863-9.

Menear, K.A. et al. 2008. 4-[3-(4-cyclopropanecarbonylpiperazine-1-carbonyl)-4-fluorobenzyl]-2 H-phthalazin-1-one: a novel bioavailable inhibitor of poly(ADP-ribose) polymerase-1. *Journal of Medicinal Chemistry* 51(20), pp. 6581–91.

Middleton, K. et al. 2003. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *Journal of Virology* 77(19), pp. 10186–201.

Miller, C.J. et al. 2003. Mycoplasma infection significantly alters microarray gene expression profiles. *BioTechniques* 35(4), pp. 812–4.

Mol, C.D. et al. 2000. Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means. *Mutation Research* 460(3–4), pp. 211–29.

Moody, C.A. and Laimins, L.A. 2009. Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. Galloway, D. ed. *PLoS Pathogens* 5(10), p. e1000605.

Moody, C.A. and Laimins, L.A. 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer* 10(8), pp. 550–60.

Moreira, K. et al. 2006. Detection of Human Papilloma Virus in the tonsils of children undergoing tonsillectomy. *The Brazilian Journal of Infectious Diseases* 10(3), pp. 165–8.

Mortazavi, A. et al. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5(7), pp. 621–8.

Mortusewicz, O. et al. 2007. Feedback-regulated poly(ADP-ribosyl)ation by PARP-1 is required for rapid response to DNA damage in living cells. *Nucleic acids research* 35(22), pp. 7665–75.

Moscicki, A.-B. et al. 2010. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. *Obstetrics and Gynecology* 116(6), pp. 1373–80.

Mourad, M. et al. 2017. Epidemiological trends of head and neck cancer in the United States: a SEER population study. *Journal of Oral and Maxillofacial Surgery* (June), pp. 1–11.

Münger, K. and Howley, P.M. 2002. Human papillomavirus immortalization and transformation functions. *Virus Research* 89(2), pp. 213–28.

Muñoz, N. et al. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine* 348(6), pp. 518–27.

Murai, J. et al. 2012. Trapping of PARP1 and PARP2 by clinical PARP inhibitors. *Cancer research* 72(21), pp. 5588–99.

Muramatsu, M. et al. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102(5), pp.

553–63.

Muwonge, R. et al. 2008. Role of tobacco smoking, chewing and alcohol drinking in the risk of oral cancer in Trivandrum, India: A nested case-control design using incident cancer cases. *Oral Oncology* 44(5), pp. 446–54.

Nakamura, Y. et al. 1987. Variable Number of Tandem Repeat (VNTR) markers for human gene mapping. *Science* 235(4796), pp. 1616–22.

Näsman, A. et al. 2012. Tumor infiltrating CD8 + and Foxp3 + lymphocytes correlate to clinical outcome and human papillomavirus (HPV) status in tonsillar cancer. *PLoS ONE* 7(6), pp. 1–8.

Nickson, C.M. et al. 2017. Misregulation of DNA damage repair pathways in HPV-positive head and neck squamous cell carcinoma contributes to cellular radiosensitivity. *Oncotarget* 8(18), pp. 29963–75.

Office for National Statistics 2015. Cancer Statistics Registrations, England - Office for National Statistics. Available at: http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2014-06-19#all-malignant-neoplasms-incidence-rates-by-sex-and-age-group-2012.

Olson, C. et al. 1965. Papilloma-like virus from bovine urinary bladder tumors. *Cancer Research* 25(6 Part 1), pp. 840–9.

Olthof, N.C. et al. 2015. Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. *International Journal of Cancer* 136(5), pp. E207-18.

Owen, J.H. et al. 2016. Novel method of cell line establishment utilizing fluorescence-activated cell sorting resulting in 6 new head and neck squamous cell carcinoma lines. *Head and Neck* 38(S1), pp. E459–E467.

Pai, S.I. and Westra, W.H. 2009. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annual Review of Pathology* 4, pp. 49–70.

Palmer, E. et al. 2014. Human papillomavirus infection is rare in nonmalignant tonsil tissue in the UK: Implications for tonsil cancer precursor lesions. *International Journal of Cancer* 135(10), pp. 2437–43.

Parfenov, M. et al. 2014. Characterization of HPV and host genome interactions in primary head and neck cancers. *Proceedings of the National Academy of Sciences of the United States of America* 111(43), pp. 15544–9.

Parsons, J.T. et al. 2002. Squamous cell carcinoma of the oropharynx. *Cancer* 94(11), pp. 2967–80.

Patel, S.G. and Shah, J.P. 2005. TNM staging of cancers of the head and neck: striving for uniformity among diversity. *CA: A Cancer Journal for Clinicians* 55(4), pp. 242–58.

Pattillo, R.A. et al. 1977. Tumor antigen and human chorionic gonadotropin in CaSki cells: a new epidermoid cervical cancer cell line. *Science* 196(4297), pp. 1456–8.

Perrone, F. et al. 2011. Isolating p16-positive/HPV-negative oropharyngeal cancer: an effort worth making. *The American Journal of Surgical Pathology* 35(5), pp. 774–8.

Pett, M. and Coleman, N. 2007. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *The Journal of Pathology* 212(4), pp. 356–67.

Pim, D. et al. 2012. Human papillomaviruses and the specificity of PDZ domain targeting. *FEBS Journal* 279(19), pp. 3530–7.

Powell, N.G. et al. 2015. Human papillomavirus-associated head and neck cancer: oncogenic mechanisms, epidemiology and clinical behaviour. *Diagnostic Histopathology* 21(2), pp. 49–64.

Prigge, E.-S. et al. 2017. Diagnostic accuracy of p16INK4A immunohistochemistry in oropharyngeal squamous cell carcinomas: a systematic review and meta-analysis. *International Journal of Cancer* 140(5), pp. 1186–98.

Rafehi, H. et al. 2011. Clonogenic assay: adherent cells. *Journal of Visualized Experiments* 49(2573), pp. 15–17.

Raff, A.B. et al. 2013. The evolving field of human papillomavirus receptor research: a review of binding and entry. *Journal of Virology* 87(11), pp. 6062–72.

Ragin, C.C.R. et al. 2004. Mapping and analysis of HPV16 integration sites in a head and neck cancer cell line. *International Journal of Cancer* 110(5), pp. 701–9.

Rass, U. et al. 2007. Actions of aprataxin in multiple DNA repair pathways. *The Journal of Biological Chemistry* 282(13), pp. 9469–74.

Reich, N.C. and Levine, A.J. 1984. Growth regulation of a cellular tumour antigen, p53, in nontransformed cells. *Nature* 308(5955), pp. 199–201.

Reinson, T. et al. 2013. Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification. *Journal of Virology* 87(2), pp. 951–64.

Reisman, D. et al. 1993. c-Myc trans-activates the p53 promoter through a required downstream CACGTG motif. *Cell Growth and Differenciation* 4, pp. 57–65.

Reisman, D. et al. 2012. Transcriptional regulation of the p53 tumor suppressor gene in S-phase of the cell-cycle and the cellular response to DNA damage. *Biochemistry Research International* 2012, pp. 1–5.

Rettig, E.M. et al. 2015. Prognostic implication of persistent human papillomavirus Type 16 DNA detection in oral rinses for human papillomavirus–related oropharyngeal carcinoma. *JAMA Oncology* 1(7), pp. 907–15.

Rheinwald, J. and Green, H. 1975. Serial cultivation of strains of human epidemal keratinocytes: the formation of keratinizing colonies from single cells. *Cell* 6, pp. 331–44.

Rieckmann, T. et al. 2013. HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. *Radiotherapy and Oncology* 107(2), pp. 242–6.

Riley, T. et al. 2008. Transcriptional control of human p53-regulated genes. *Nature Reviews Molecular Cell Biology* 9(5), pp. 402–12.

Roberts, S.A. et al. 2013. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nature Genetics* 45(9), pp. 970–6.

Robinson, J.T. et al. 2011. Integrative Genome Viewer. *Nature Biotechnology* 29(1), pp. 24–6.

Robinson, M. et al. 2012. HPV specific testing: a requirement for oropharyngeal squamous cell carcinoma patients. *Head and Neck Pathology* 6(S1), pp. 83–90.

Robinson, M.D. et al. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1), pp. 139–40.

de Roda Husman, A.M. et al. 1995. Processing of long-stored archival cervical smears for human papillomavirus detection by the polymerase chain reaction. *British Journal of Cancer* 72(2), pp. 412–7.

Rosenthal, F. et al. 2013. Macrodomain-containing proteins are new mono-ADP-ribosylhydrolases. *Nature Structural & Molecular Biology* 20(4), pp. 502–7.

Rothkamm, K. et al. 2003. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Molecular and Cellular Biology* 23(16), pp. 5706–15.

Russo, A.J. et al. 2006. E2F-1 overexpression in U2OS cells increases cyclin B1 levels and cdc2 kinase activity and sensitizes cells to antimitotic agents. *Cancer Research* 66(14), pp. 7253–60.

Safaeian, M. et al. 2010. Epidemiological study of anti-HPV16/18 seropositivity and subsequent risk of HPV16 and -18 infections. *Journal of the National Cancer Institute* 102(21), pp. 1653–62.

Saiki, R.K. et al. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839), pp. 487–491.

Saleh-Gohari, N. and Helleday, T. 2004. Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Research* 32(12), pp. 3683–8.

Salter, J.D. et al. 2016. The APOBEC protein family: united by structure, divergent in function. *Trends in biochemical sciences* 41(7), pp. 578–94.

Sancar, A. et al. 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual Review of Biochemistry* 73(1), pp. 39–85.

Schache, A.G. et al. 2011. Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous cell carcinoma: sensitivity, specificity, and prognostic discrimination. *Clinical Cancer Research* 17(19), pp. 6262–71.

Schache, A.G. et al. 2016. HPV-related oropharynx cancer in the United Kingdom: an evolution in the understanding of disease etiology. *Cancer Research* 76(22), pp. 6598–606.

Schelhaas, M. et al. 2012. Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis. *PLoS pathogens* 8(4), p. e1002657.

Scherer, W.F. et al. 1953. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *The Journal of Experimental Medicine* 97(5), pp. 695–710.

Schlecht, N.F. et al. 1999. Effect of smoking cessation and tobacco type on the risk of cancers of the upper aero-digestive tract in Brazil. *Epidemiology* 10(4), pp. 412–8.

Schmitz, M. et al. 2012. Loss of gene function as a consequence of human papillomavirus DNA integration. *International Journal of Cancer* 131(5), pp. E593–E602.

Schreiber, V. et al. 2002. Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *The Journal of Biological Chemistry* 277(25), pp. 23028–36.

Schwartz M. et al. 1998. Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection. *Journal of the National Cancer Institute* 90(21), pp. 1626–36.

Sedman, J. and Stenlund, A. 1998. The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *Journal of Virology* 72(8), pp. 6893–7.

Shafti-keramat, S. et al. 2003. Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. *Journal of Virology* 77(24), pp. 13125–35.

Sharma, A. et al. 2012. Histone H2AX phosphorylation: a marker for DNA damage. In: *Methods in Molecular Biology*., pp. 613–26.

Sisk, J. et al. 2006. Presence of human papillomavirus DNA in tonsillectomy specimens. *The Laryngoscope* 116(8), pp. 1372–4.

Slade, D. et al. 2011. The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* 477(7366), pp. 616–20.

Smeets, S.J. et al. 2007. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *International Journal of Cancer* 121(11),

pp. 2465–72.

Smith, E.M. et al. 1998. Human papillomavirus and risk of oral cancer. *The Laryngoscope* 108(7), pp. 1098–103.

Sobol, R.W. et al. 1996. Requirement of mammalian DNA polymerase-β in base-excision repair. *Nature* 379(6561), pp. 183–6.

Somers, K.D. et al. 1992. Frequent p53 mutations in head and neck cancer. *Cancer Research* 52(21), pp. 5997–6000.

Spardy, N. et al. 2009. HPV-16 E7 oncoprotein attenuates DNA damage checkpoint control by increasing the proteolytic turnover of claspin. *Cancer Research* 69(17), pp. 7022–9.

Spriggs, C.C. and Laimins, L.A. 2017. FANCD2 binds human papillomavirus genomes and associates with a distinct set of DNA repair proteins to regulate viral replication. *mBio* 8(1), p. e02340.

Stanley, M.A. 2002. *Culture of human cervical epithelial Cells*.

Steenbergen, R.D.M. et al. 1995. Integrated human papillomavirus Type 16 and loss of heterozygosity at 11q22 and 18q21 in an oral carcinoma and its derivative cell line. *Cancer Research* 55(22), pp. 5465–71.

Steger, G. and Corbach, S. 1997. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *Journal of Virology* 71(1), pp. 50–8.

Stenglein, M.D. et al. 2010. APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nature structural & molecular biology* 17(2), pp. 222–9.

Stransky, N. et al. 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science* 333(6046), pp. 1157–60.

Strickland, S.W. and Vande Pol, S. 2016. The human papillomavirus 16 E7 oncoprotein attenuates AKT signaling To promote internal ribosome entry site-dependent translation and expression of c-MYC. *Journal of Virology* 90(12), pp. 5611–21.

Sulli, G. et al. 2012. Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. *Nature Reviews Cancer* 12(10), pp. 709–20.

Tang, A.L. et al. 2012. UM-SCC-104: a new human papillomavirus-16 containing head and neck squamous cell carcinoma cell line. *Head & Neck* 34(10), pp. 1480–91.

Thomas, M. et al. 2008. Analysis of specificity determinants in the interactions of different HPV E6 proteins with their PDZ domain-containing substrates. *Virology* 376(2), pp. 371–8.

Thorland, E.C. et al. 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene* 22(8), pp. 1225–37.

Van Tine, B.A. et al. 2004. Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning. *Proceedings of the National Academy of Sciences of the United States of America* 101(12), pp. 4030–5.

Tjalma, W.A. et al. 2013. Differences in human papillomavirus type distribution in high-grade cervical intraepithelial neoplasia and invasive cervical cancer in Europe. *International Journal of Cancer* 132(4), pp. 854–67.

Todaro, G.J. and Green, H. 1963. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *The Journal of Cell Biology* 17(2), pp. 299–313.

Topalian, S.L. et al. 2016. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nature Reviews Cancer* 16(5), pp. 275–87.

Vesely, M.D. et al. 2011. Natural innate and adaptive immunity to cancer. *Annual Review of Immunology* 29(1), pp. 235–71.

Vieira, V.C. et al. 2014. Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio* 5(6), p. e02234.

de Villiers, E.M. et al. 2004. Classification of papillomaviruses. *Virology* 324(1), pp. 17–27.

De Vuyst, H. et al. 2009. HPV infection in Europe. *European Journal of Cancer* 45(15), pp. 2632–9.

Vyas, S. et al. 2014. Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nature communications* 5(4426)

Walboomers, J.M.M. et al. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of Pathology* 189(1), pp. 12–9.

Wallace, N.A. and Galloway, D.A. 2014. Manipulation of cellular DNA damage repair machinery facilitates propagation of human papillomaviruses. *Seminars in Cancer Biology* , pp. 30–42.

Wang, Z. et al. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10(1), pp. 57–63.

Warren, C.J. et al. 2015. APOBEC3A functions as a restriction factor of human papillomavirus. Imperiale, M. J. ed. *Journal of Virology* 89(1), pp. 688–702.

Weaver, A.N. et al. 2015. DNA double strand break repair defect and sensitivity to poly ADP-ribose polymerase (PARP) inhibition in human papillomavirus 16-positive head and neck squamous cell carcinoma. *Oncotarget* 6(29), pp. 26995–7007.

Weinberg, R.A. 2013. *The biology of cancer*. Second. Garland science.

Wiest, T. et al. 2002. Involvement of intact HPV16 E6/E7 gene expression in head and neck cancers with unaltered p53 status and perturbed pRb cell cycle control. *Oncogene* 21(10), pp. 1510–7.

Willmore, E. et al. 1998. Etoposide targets topoisomerase IIalpha and IIbeta in leukemic cells: isoform-specific cleavable complexes visualized and quantified in situ by a novel immunofluorescence technique. *Molecular Pharmacology* 54(1), pp. 78–85.

Wilson, R. et al. 2005. Role of the E1 E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31. *Journal of Virology* 79(11), pp. 6732–40.

Woodman, C.B.J. et al. 2007. The natural history of cervical HPV infection: unresolved issues. *Nature Reviews Cancer* 7(1), pp. 11–22.

Young, M.D. et al. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* 11(2), p. R14.

Yuan, Z. et al. 2013. Genetic variants at 10q23 are associated with risk of head and neck cancer in a Chinese population. *Oral Oncology* 49(4), pp. 332–5.

Zhang, R. et al. 2016. Dysregulation of host cellular genes targeted by human papillomavirus (HPV) integration contributes to HPV-related cervical carcinogenesis. *International Journal of Cancer* 138(5), pp. 1163–74.

Zhang, W. et al. 2016. Integrative genomics and transcriptomics analysis reveals potential mechanisms for favorable prognosis of patients with HPV-positive head and neck carcinomas. *Scientific Reports* 6(24927)

Zhang, Z.-F. et al. 2000. Environmental tobacco smoking, mutagen sensitivity, and head and neck squamous cell carcinoma. *Cancer Epidemiology and Prevention Biomarkers* 9(10), pp. 1043–9.

Zhou, G. et al. 2016. TP53 mutations in head and neck squamous cell carcinoma and their impact on disease progression and treatment response. *Journal of Cellular Biochemistry* 117(12), pp. 2682–92.

Ziemann, F. et al. 2015. Increased sensitivity of HPV-positive head and neck cancer cell lines to x-irradiation ± Cisplatin due to decreased expression of E6 and E7 oncoproteins and enhanced apoptosis. *American Journal of Cancer Research* 5(3), pp. 1017–31.

Appendix 1: PCOC study protocol

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

PCOC

Primary Culture of Oropharyngeal Cells

A project to establish the feasibility and utility of growing primary cell cultures from Oropharyngeal Cancer biopsies

Version 2.0
15th May 2014

MAIN SPONSOR:	Cardiff University
FUNDERS:	Cancer Research Wales
STUDY COORDINATION CENTRE:	HPV Research Group School of Medicine Cardiff University Heath Park CF14 4XN

NRES reference: 13/WA/0002

Protocol authorised by:

Name & Role	Date	Signature
Dr Mererid Evans (Clinical-PI)		
Dr Ned Powell (Laboratory-PI)		

Study Management Group

Clinical PI:

Dr Mererid Evans
Consultant Oncologist (Velindre Cancer Centre)
Senior Research Fellow (Cardiff University)
Velindre Cancer Centre,
Velindre Road,
Whitchurch,
Cardiff, UK
CF14 2TL
Tel: 029 20196160

Scientific PI:

Dr Ned Powell
Senior Lecturer
HPV Research Group
School of Medicine
Cardiff University
Heath Park
Cardiff, UK
CF14 4XN
Tel: 02920 744742

Clinical Collaborators:

Dr Adam Christian
Consultant Histopathologist
University Hospital of Wales
Heath Park
Cardiff, UK
CF14 4XN
Tel: 02920744278

Dr Kenneth May
Consultant Histopathologist
University Hospital of Wales
Heath Park
Cardiff, UK
CF14 4XN
Tel: 02920744278

Mr David Owens
Consultant Surgeon
University Hospital of Wales
Heath Park
Cardiff, UK
CF14 4XN
Tel: 02920744174

Prof Alison Fiander
Chair of Obstetrics and Gynaecology
School of Medicine
Cardiff University
Heath Park
Cardiff, UK
CF14 4XN
Tel: 02920 743234

Study Coordination

General queries should be directed to Dr Ned Powell:

Address: HPV Research Group
School of Medicine
Cardiff University
Heath Park
CF14 4XN

Tel: 02920 744742
E-mail: powellng@cf.ac.uk

Clinical Queries

Clinical queries should be directed to Dr Mererid Evans:

Address: Velindre Cancer Centre,
Velindre Road,
Whitchurch,
Cardiff, UK
CF14 2TL

Tel: 029 20196160
Email: mererid.evans@wales.nhs.uk

Sponsor

Cardiff University is the main research sponsor for this study. For further information regarding the sponsorship conditions, please contact the Research Governance Manager at:

Mr Chris Shaw
Research Governance Manager
Research and Commercial Division
30-36 Newport Rd
Cardiff
CF24 0DE

Email: shawc3@cf.ac.uk
Telephone : 02920879130
Fax : 02920874189

Funder

Cancer Research Wales
Registered office:
Velindre Hospital
Whitchurch,
Cardiff CF14 2TL
Wales, UK

This protocol describes the PCOC study and provides information about procedures for entering participants. Every care was taken in its drafting, but corrections or amendments may be necessary. Problems relating to this study should be referred, in the first instance, to the Chief Investigator.

This study will adhere to the principles outlined in the NHS Research Governance Framework for Health and Social Care (2nd edition). It will be conducted in compliance with the protocol, the Data Protection Act and other regulatory requirements as appropriate.

Table of Contents

STUDY SUMMARY

TITLE Primary Culture of Oropharyngeal Cells (PCOC)

DESIGN Pilot project to determine feasibility and utility of growing cells from oropharyngeal cancers

OUTCOME MEASURES The primary outcome will be determination of the feasibility of primary culture of oropharyngeal cancers. Secondary outcome measures include: characterisation of biopsies in terms of HPV infection, possible use of primary cultures as a model to assess potential novel therapies.

POPULATION Patients undergoing treatment for, or investigations to confirm, oropharyngeal cancer at Cardiff and Vale NHS Trust

ELIGIBILITY Patients diagnosed with squamous cell carcinoma of the head and neck, with the primary tumour suspected oropharyngeal.
Able to give informed consent

DURATION 3 years

1. INTRODUCTION

1.1 BACKGROUND

HPV-associated oropharyngeal cancer in Wales

Human Papillomavirus (HPV) has been long been recognized as the cause of cervical cancer as well as other cancers of the vulva, vagina, penis and anus. More recently, it has been established as a cause of Head and Neck (H&N) cancers involving the tonsils and tongue base (oropharyngeal cancers). The incidence of oropharyngeal cancers is increasing in Europe and the USA as a result of HPV; in Scotland, it is the fastest increasing cancer and in Sweden, the incidence almost doubled from 1970 to 2006 with a fourfold increase in the proportion of HPV-positive cases over this time period, prompting reports of an 'epidemic of a virus-induced carcinoma' [1].

HPV-associated H&N cancer is also relevant to the population of Wales. Whereas the incidence of most H&N cancers has remained static, the incidence of oropharyngeal cancer has doubled over the last 20 years, particularly in men aged between 45-65 years (Figure 1). In a retrospective analysis of 147 cases of oropharyngeal cancer presenting in South Wales 2001-2006, 55% were HPV-positive (M Evans, N Powell, data submitted for publication). Clinicians treating H&N cancer in Wales recognize the changing demographic of patients presenting to their clinics, typically young men who are fit and active with locally advanced disease of the oropharynx – the 'epidemic' described in Sweden also appears to be happening here.

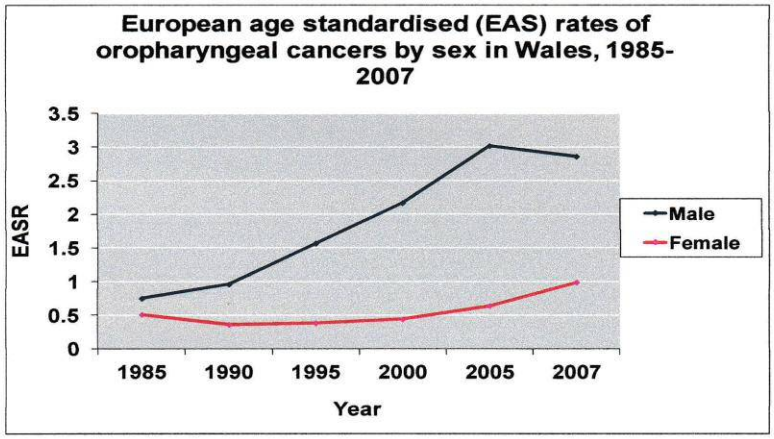


Fig 1: Incidence of oropharyngeal cancer in Wales in men and women 1987-2007 (WICSU data).

Molecular basis of disease

The natural history of HPV infection in the H&N is poorly understood and it is not known why the oropharynx, particularly the tonsils, are more susceptible to HPV-induced transformation than other subsites of the H&N. Our understanding of the processes underlying HPV-induced malignant transformation is mainly derived from studies in cervical carcinoma, but there appear to be important differences between the cervix and oropharynx. For example, 95% of HPV-positive oropharyngeal cancers are related to HPV type 16 infection, whereas HPV16 accounts for only 60% of all cervical cancers and other high-risk HPVs (HR-HPV) are responsible for the rest.

HPV-induced transformation is primarily mediated by the oncogenes E6 and E7, the protein products of which inactivate the cellular tumour suppressor proteins p53 and pRb respectively, resulting in loss of cell cycle control, loss of apoptosis ('programmed cell death' regulated by p53) and uncontrolled cellular proliferation. Accumulation of genetic damage in the resulting unstable population of cells can ultimately lead to malignant transformation and progression to invasive cancer. De-regulation of viral gene expression, at least in cervical cancer, is commonly the result of integration of viral DNA into the host genome and continued expression of viral E6 and E7 oncoproteins is necessary to maintain the transformed phenotype. Although integration of HPV DNA into genomic DNA is a common event in cervical carcinoma, the situation in H&N cancers is less clear and episomal HPV DNA has been found at high frequency in tonsillar cancers, albeit in small studies [2].

Unlike the situation in cervical cancer, where HPV is a 'necessary' cause of malignant transformation, not all oropharyngeal cancers are caused by HPV. The remainder (approximately 50% overall) typically arise in patients who smoke tobacco and drink alcohol, in common with cancers arising elsewhere within the H&N. Long-term exposure to the carcinogenic effects of tobacco and alcohol may produce mutations or epigenetic inactivations in a number of critical tumour suppressor genes, including p16, p53 and Rb, resulting in multistep progression from normal cell to dysplasia to frank carcinoma. The tumour suppressor gene p53, one of the most commonly mutated genes in all human malignancy, is mutated in 50-80% of (HPV negative) H&N cancers, usually at an early stage of cancer progression.

Diagnosing a HPV-associated H&N cancer

Because oropharyngeal cancers can arise by HPV-dependent and HPV-independent means, the presence of HPV DNA (detected by PCR or In Situ Hybridization, ISH) in a tumour *per se* is not sufficient to prove that HPV has induced malignant transformation. Evidence of a causal association is provided if HPV oncoproteins are shown to be expressed in transformed cells. E7-induced degradation of Rb results in the upregulation of p16 expression through loss of feedback inhibition and over-expression of p16 is considered a reliable surrogate marker of active HPV infection. Oropharyngeal cancers may be divided into 3 classes on the basis of p16 expression and presence of HPV DNA: class I - HPV negative, p16 low; class II - HPV positive, p16 low; class III - HPV positive, p16 high. Class II tumours may be multifactorial in origin, formed when tobacco/alcohol-related tumours are infected by high risk HPV [3].

Treatment response and prognosis

HPV-associated oropharyngeal cancer has a better response to radiotherapy [4], chemotherapy [5] and chemoradiotherapy [6] compared to HPV-negative oropharyngeal cancer. In a prospective phase II trial [6], patients with HPV-positive tumours showed a significantly higher response rate to induction chemotherapy (82% vs 55%, p=0.01) and definitive chemoradiotherapy (84% vs 57%, p=0.007) compared to HPV-negative tumours, as well as improved overall survival at 2 years (95% vs 62%, p=0.005).

The different molecular biological processes underlying malignant transformation in HPV-positive and negative cancers may underlie the different responses to treatment. The effect of HPV theoretically obviates the need for mutational inactivation of the p53, Rb and p16 genes in HPV-driven cancers and it is possible that 'wild-type' p53 and Rb can become reactivated during treatment, thereby restoring cellular functions like apoptosis. There is some evidence that HPV-associated H&N cancers are less likely to harbour p53 mutations than HPV-negative cancers [7]. Furthermore, HPV positive oropharyngeal cancers with 'wild type' p53 had better survival outcomes in a study of 90 patients treated with surgery and post-operative radiotherapy than patients with HPV positive or negative tumours containing mutated p53, but this study only included 17 patients with HPV-associated disease [8]. There is therefore a suggestion, but not conclusive proof, that p53 mutation status is related to HPV status in oropharyngeal cancer and may influence outcome.

Prognosis

HPV is a strong and independent prognostic factor for survival in patients with oropharyngeal cancer [6, 8, 9]. In a study of over 300 patients with locally advanced oropharyngeal cancer treated with

radical chemoradiotherapy [9], 3 year overall survival was 82.4% in HPV positive patients compared to 57.1% in HPV negative patients (p<0.001). However, survival was also affected by other factors, particularly tobacco smoking and patients could be stratified into three categories with respect to risk of death: low risk (3-year overall survival 93.0%) patients included HPV positive non-smokers and some HPV positive smokers (with low nodal stage); intermediate risk (3-year survival 70.8%) patients included most HPV positive smokers and HPV negative non-smokers; high risk patients (3-year survival 46.2%) consisted of HPV negative patients only. Thus smoking reduced 3 year survival in HPV positive patients with advanced nodal stage by 23%. This negative effect of smoking on survival from oropharyngeal cancer has been reported in other studies; furthermore, smoking cessation may not reduce the negative impact of smoking on outcome and disease recurrence rates are reportedly higher in HPV-positive ex-smokers than in patients who have never smoked.

Because of the excellent survival rates in HPV-positive (low risk) patients, there is a move to reduce their treatment intensity to spare them from the devastating long-term side-effects of current treatment protocols, including difficulties with speech and swallowing (13% of patients will still be dependent on gastrostomy tube feeding 2 years after chemoradiotherapy), facial disfigurement and body image issues which may affect patients for the remainder of their lives. A Phase III UK RCT, DeEscalate-HPV (Dr M Evans is a co-Investigator), funded by CR-UK, is due to open later in 2012 and will randomize patients with low-risk HPV-positive oropharyngeal cancer to receive chemoradiotherapy (standard arm) or radiotherapy with Cetuximab (experimental arm), which is a monoclonal antibody targeting the Epidermal Growth Factor Receptor (EGFR). However, HPV-positive patients with 'intermediate-risk' disease are excluded from this study. Data from our (CRW-funded) analysis of oropharyngeal cancers across South Wales 2001-2006 showed that 70% of patients were current or previous smokers and >50% of patients had advanced nodal disease (Dr M Evans, Dr N. Powell, data submitted for publication) meaning that a significant proportion of patients with HPV-positive oropharyngeal cancer in South Wales would not be eligible for upcoming clinical trials because of their smoking history.

We believe that better means of stratifying patients into risk groups must be developed, based on disease biomarkers rather than smoking habits. Furthermore, there is an urgent need to develop new treatment strategies, which target the mechanisms underlying HPV-induced transformation, to use instead of, or in conjunction with, current treatments to improve outcome for all patients with HPV-associated oropharyngeal cancer. In order to develop these targeted therapies, preclinical work is required to better understand the biology of HPV-mediated malignant transformation in the H&N and to investigate the response of HPV-infected and uninfected cells to chemotherapy and radiotherapy.

Aim of Study:

To determine the feasibility of establishing primary cultures of HPV infected and uninfected oropharyngeal carcinomas in the laboratory from surgical biopsies.

Once established, the primary cultures will be characterized with respect to presence of HPV infection, integration state of HPV and p53/Rb expression and function. Potentially, they will also be utilized to study the cellular effects of current treatments and to test novel strategies of drug-induced reversal of HPV-mediated p53 silencing.

1.2 RATIONALE FOR CURRENT STUDY

[To include: research question and hypothesis]

Immortalized cell lines, rather than primary cultures derived from fresh tissue biopsies, have been used in the past to study HPV-associated cancers. Whereas immortalised cell lines have many advantages related to ease of use, they are a poor model for *in vivo* cellular behaviour. In reality, tumours are far more diverse than the limited number of immortal lines available. Studies should ideally utilise more biologically relevant models, and this has led to increasing interest in the use of primary cultures to investigate tumour behaviour. This is particularly important when assessing the effects of potential therapeutic agents, as many established cell lines and even late passage primary

cultures, have developed resistance to such agents over time. Hence one of the main priorities for this project is to develop a primary culture model for oropharyngeal cancer, based on laboratory culture of surgical biopsies.

Following ethical approval and informed consent, a 4mm punch biopsy will be taken from a patients' tumour which will be surplus to that required for diagnostic purposes or pathology review/reporting. Currently, around 40 patients per annum present to the H&N clinic at C&V with a diagnosis of oropharyngeal cancer - not all of these cancers will be amenable to biopsy, thus we aim to collect biopsies from 1 patient per month. Over 12 months this should allow samples from a minimum of 12 patients to be collected. Based on previous experience in the HPV research group of growing primary cultures from biopsies of Vulval Intraepithelial Neoplasia (VIN), we anticipate that successful primary cultures will be obtained in 50% of these samples.

Methods for establishing primary cultures from vulval intraepithelial neoplasia (VIN) are already established in the laboratory, and the same methods will be used to establish primary cultures from H&N cancer biopsies. Samples will be placed immediately into transport medium (DMEM supplemented with Foetal Calf Serum (FCS), Gentamicin sulphate and Amphotericin B) and transferred to the HPV Research Laboratory, they will then be coarsely minced, exposed to FCS and cultured. Explanted epithelial cells will be subcultured at 7-10 days onto deactivated 3T3 feeder layers. Primary cultures are expected to divide for 3 to 6 months before senescing. This method of culture has been chosen due to the likelihood of producing large numbers of clonal cells with long *in vitro* life.

Cultures will be characterised in terms of HPV type, oncogene expression, integration state, p53 expression and p53 mutation state. HPV type will be determined by GP5/6 PCR ELISA; viral oncogene expression by QRTPCR for the E2, E6 and E7 genes; and viral integration by DIPS. These analyses are all well established in our laboratory. Levels of p53 will be quantified in protein extracted from cell lines using western blotting and p53 mutation status by complete sequencing of the p53 coding sequence (exons 4-9). The sequence analysis entails amplification with 7 separate primer sets followed by purification of the amplicon and sequencing using BigDye (ABI) chain termination.

2. STUDY OBJECTIVES

Primary objective: to determine the feasibility of establishing primary cultures of oropharyngeal cancer from fresh tissue biopsies.

Secondary objective: to characterize the primary cultures for presence and type of HPV, oncogene expression, integration state, p53 expression and p53 mutation state.

Other study objectives: once characterized, the cultures will be used as an *in vitro* system to test the effect of novel therapeutic agents.

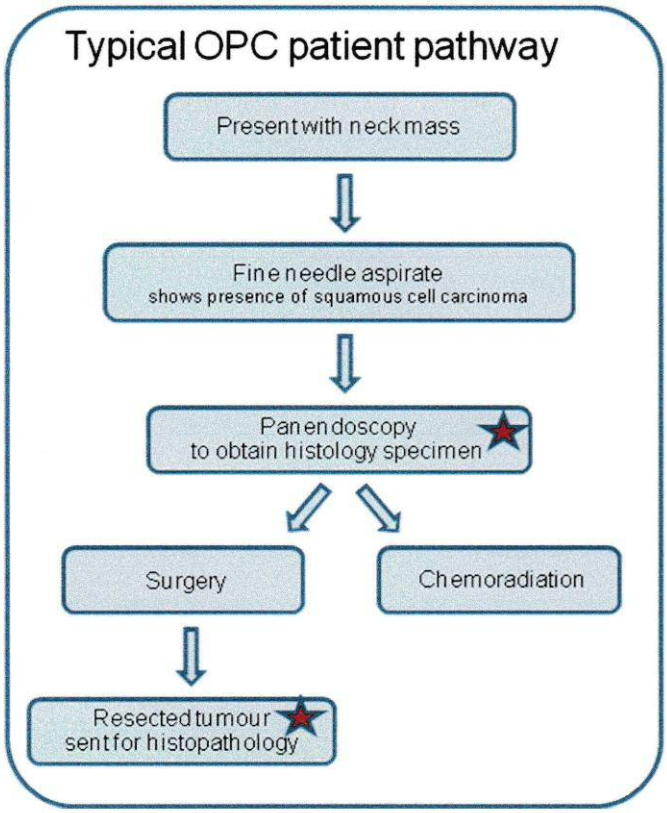
3. STUDY DESIGN

3.1 PATIENT IDENTIFICATION

Patients will be identified by their H&N surgical consultant (Surgical lead: Mr Dave Owens) and/or consultant oncologist (Dr Mererid Evans) in the course of their routine clinical practice.

3.2 PATIENT SAMPLES

Tissue samples will be obtained from patients at two possible points in the treatment pathway (shown in figure 1 below – red stars indicate points at which samples may be obtained).



Samples will be obtained at two points in the patient pathway, according to the clinical scenario.

Either at:

1. Panendoscopy biopsy. Panendoscopy is performed once a diagnosis of squamous cell carcinoma has been made from an FNAC and/or biopsy of a metastatic cervical lymph node. This is an examination conducted under general anaesthetic, during which biopsies will be taken for histopathological analysis. Patients will be consented for an additional research biopsy to be taken at the same time as their diagnostic biopsy.

2. Post surgery resection biopsy – for patients having surgery to resect their primary tumour. Once the resected tumour has been sent for histological analysis, a biopsy will be taken (by the reporting pathologist) for research use. This biopsy will be surplus to diagnostic/pathology requirements.

It is necessary to obtain biopsies from two points in the pathway as patients will be treated differently according to the site and stage of their tumour. In general, for tumours treated surgically it will be possible to obtain a tissue biopsy once the tumour has been resected and sent for pathological analysis. However, tumours involving the tongue base or soft palate and/or that are larger than 4cm (stage 3 or 4) are likely to be treated by chemoradiation, hence it would not be possible to obtain a

post treatment sample. In these patients, consent will be sought to obtain a biopsy specifically for research, at the time of their diagnostic panendoscopy.

3.3 PLANNED ANALYSIS / USE OF SAMPLES

Biopsy samples will be transferred immediately from theatre (for endoscopy biopsies) or pathology laboratory (for post surgery biopsies) to the HPV Research Laboratory. On arrival in the laboratory, they will be bisected and half the sample will be used for primary culture, and the other half placed in DNA/RNA preservation medium.

3.31 PRIMARY CULTURE OF OPC BIOPSIES

Biopsies will be initially processed according to established protocols for culture of HPV infected cervical tissue [10]. Samples will be coarsely minced, exposed to FCS and cultured in supplemented DMEM. Standard tissue culture procedures will be used to prevent microbial contamination. Explanted epithelial cells will be subcultured at 7-10 days onto deactivated 3T3 feeder layers in modified Glasgow media. Primary cultures are expected to divide for 3 to 6 months with a doubling time of 1 – 5 days before senescing. This method of culture has been chosen due to the likelihood of producing large numbers of clonal cells with long in vitro life. This procedure has been used successfully in our laboratory for culture of HPV infected vulval cells. Based on previous experience, we expect around 40% of biopsies to produce viable cultures.

A detailed record of the in vitro history of each line will be compiled including number of population doublings and plating efficiency at each passage. The cells of the oral cavity are often infected with *Mycoplasma orale,* and primary cultures are susceptible to mycoplasma infection [11], hence established cultures will be checked for mycoplasma contamination. At second passage cultures will be examined to confirm cellular lineage, based primarily on cell morphology. At third passage, DNA and RNA will be extracted and cultures will be characterised in terms of HPV infection (type and integration state). Over 130 types of HPV have been identified; up to 18 of these may have oncogenic potential [12]. HPV type will be determined according to the method of Walboomers et al. [13] which is used routinely in our laboratory [14, 15].

Integration of HPV into the host genome is linked to upregulation of the E6 and E7 genes due to disruption of the E2 regulatory gene. Disruption of E2 is also significant as it is a multifunctional protein with pro-apoptopic functions [16]. HPV integration status will be determined by Analysis of Papillomavirus Oncogene Transcripts (APOT). Cultures that appear to grow well will be analysed by short tandem repeat analysis to confirm that they have not been contaminated by established immortal lines.

The reason for characterisation of the individual cultures is to assess the degree of heterogeneity among cultures and to determine how accurately they reflect clinical disease. Ultimately the cultures will be used to investigate the utility of novel treatments for OPC, hence detailed characterisation is also necessary for identification of biomarkers that predict response to treatment.

3.32 CHARACTERISATION OF OPC BIOPSIES

Half of the biopsy will be used for primary culture. The other half will have DNA and RNA extracted. DNA and RNA will be used determine whether the cells are infected with HPV, and whether HPV oncogenes are actively transcribed. RNA will also be used to investigate HPV integration state in HPV positive cultures. Where cell cultures are produced from the other half of the biopsy, this material will be used to assess whether the cultured cells are representative of the starting material.

3.33 FORMALIN FIXED PARAFFIN EMBEDDED BIOPIES

Patients will also be requested to grant access to Formalin Fixed Paraffin Embedded (FFPE) tissue from their resected tumour. This will enable investigation of protein expression and tissue distribution of molecular markers identified in the primary cultures and biopsies.

3.4 STUDY OUTCOME MEASURES

The primary outcome will be to establish the feasibility of primary culture of OPC biopsies.
The secondary outcomes relate to the characterisation of any lines that do grow.

3.5 SAMPLE SIZE

In the region of 40 patients per year are treated for OPC at Cardiff and Vale NHS Trust. Around half of these patients will be treated surgically and half by chemoradiation. We aim to collect in the region of 12 biopsies per year. This will allow assessment of approximately 36 biopsies over the full term of the project. This will allow thorough assessment of the feasibility of this technique.

4. PARTICIPANT ENTRY

4.1 PATIENT IDENTIFICATION

Patients will be identified by Mr Owens and Dr Evans in the course of their routine clinical practice.

4.2 INCLUSION CRITERIA

The inclusion criteria are:

1. Presentation at the Head and Neck Oncology clinic at Cardiff and Vale NHS Trust
2. Fine needle aspirate and/or biopsy of a mass in the H&N showing presence of squamous cell carcinoma.
3. Ability to give written informed consent

4.3 EXCLUSION CRITERIA

1. Patients undergoing panendoscopy in whom there is no confirmation of malignant disease will be excluded. This will ensure that patients will only be approached once they have received a diagnosis of malignancy.
2. Malignancy in the H&N in a non-oropharyngeal site.
3. Prior radiotherapy to the biopsy site.

4.4 WITHDRAWAL CRITERIA

Patients can withdraw from the study at any time. Data and samples from patients choosing to withdraw will be destroyed.

5. ADVERSE EVENTS

Enrolment in the study will not affect the clinical management of patients. Hence there is no potential for adverse events related to the study.

7. STATISTICS AND DATA ANALYSIS

This is a pilot study to establish feasibility of a procedure. The data generated will not be subject to statistical analysis.

Data and all appropriate documentation will be stored for a minimum of 5 years after the completion of the study.

8. REGULATORY ISSUES

8.1 ETHICS APPROVAL

The study will be conducted in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th World Medical Assembly, Helsinki 1964 and later revisions.

The main ethical issue is the requirement for informed consent. Patients will be fully informed about the study and it will be made clear that they will not benefit directly from the results of the research. Patients with oropharyngeal cancer presenting to the H&N MDT at Cardiff and Vale NHS Trust are now routinely tested for a surrogate marker of HPV infection, p16, by immunohistochemical analysis of their diagnostic biopsy. The results of p16 staining are communicated to patients who are made aware of the viral aetiology of their cancer. HPV testing is also a requirement of a number of clinical trials that are opening for C&V patients in the next 1-2 months (De-escalate HPV, REALISTIC and TITAN), As a result, HPV testing of primary cultures in this study will not raise ethical dilemmas about whether to inform patients of the HPV status of their tumours (a consideration for previous studies) as that will have already been done.

8.2 CONSENT

Consent to enter the study will be sought from each participant only after a full explanation has been given and adequate time allowed for the patient to digest this information. A patient information leaflet will be offered and time allowed for consideration. Signed participant consent will be obtained. The right of the participant to refuse to participate without giving reasons will be respected. All participants are free to withdraw at any time from the study without giving reasons and without prejudicing further treatment.

8.3 CONFIDENTIALITY OF PATIENT DATA

On arrival in the laboratory patient samples will be allocated a unique study number. This number will be used to identify the samples throughout the course of the study. Dr Evans will have access to the patients' clinical data as part of her clinical practice, Dr Powell will collate consent forms. Laboratory staff will not have access to any patient identifiable data.

8.4 INDEMNITY

Cardiff University holds negligent harm and non-negligent harm insurance policies which apply to this study.

8.5 SPONSOR

Cardiff University will act as the main sponsor for this study. Delegated responsibilities will be assigned to the NHS trusts taking part in this study.

8.6 FUNDING

Cancer Research Wales are funding this study.

8.7 AUDITS

The study may be subject to inspection and audit by Cardiff University under their remit as sponsor and other regulatory bodies to ensure adherence to GCP and the NHS Research Governance Framework for Health and Social Care (2nd edition).

9. STUDY MANAGEMENT

The day-to-day management of the clinical aspects of the study will be co-ordinated by Dr Evans. Day-to-day management of the laboratory aspects will be undertaken by Dr Powell.

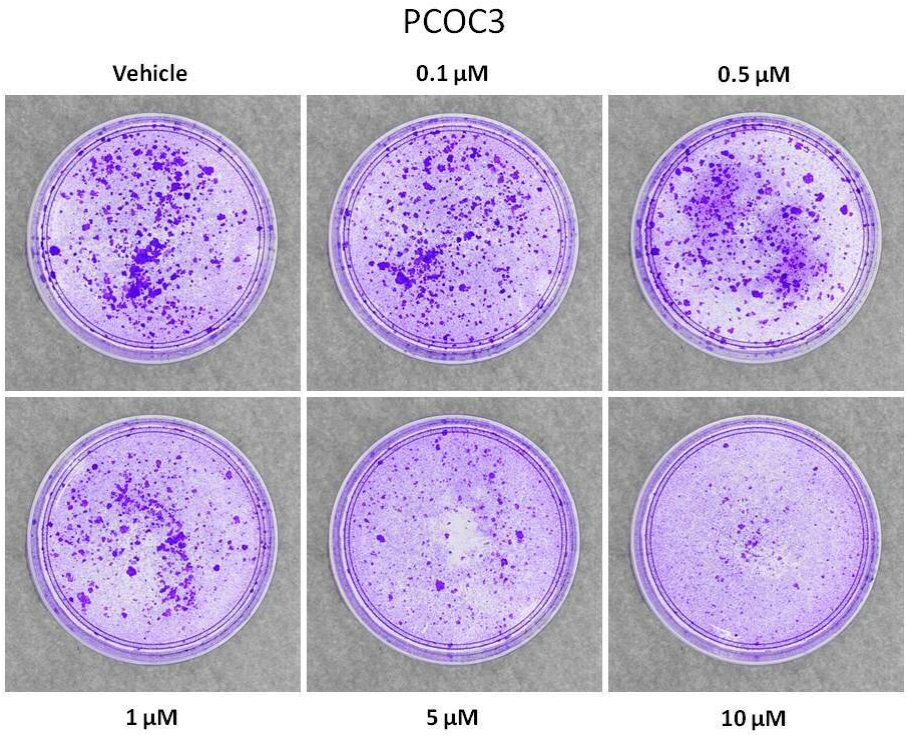
10. PUBLICATION POLICY

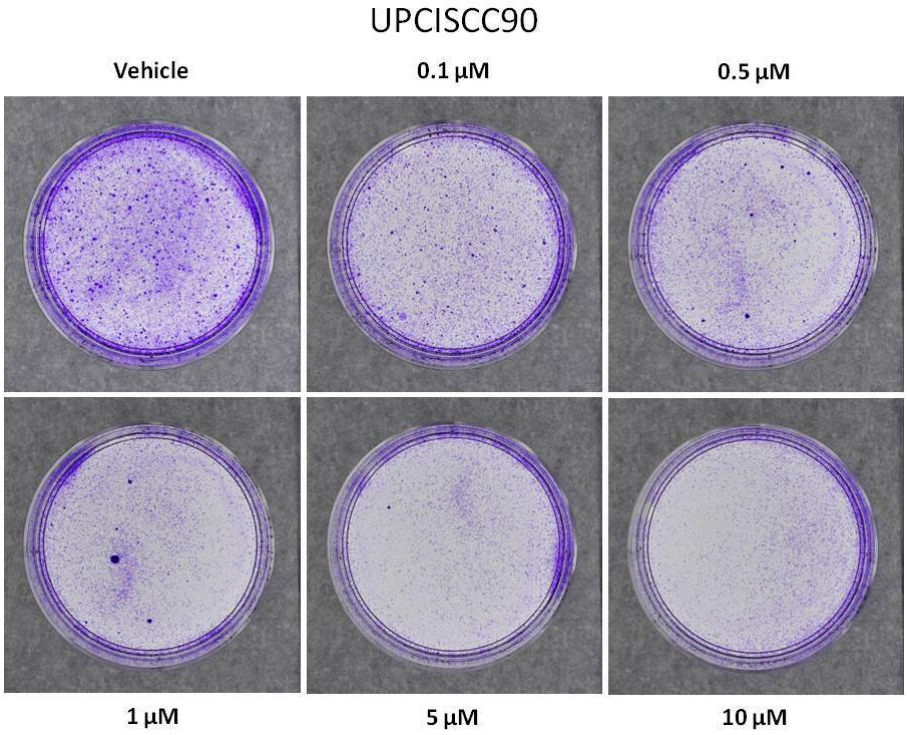
The results of these investigations will be published in the peer-reviewed scientific literature. No patient identifiable data will be published, but the contribution of the patients and of the funding body will be acknowledged.

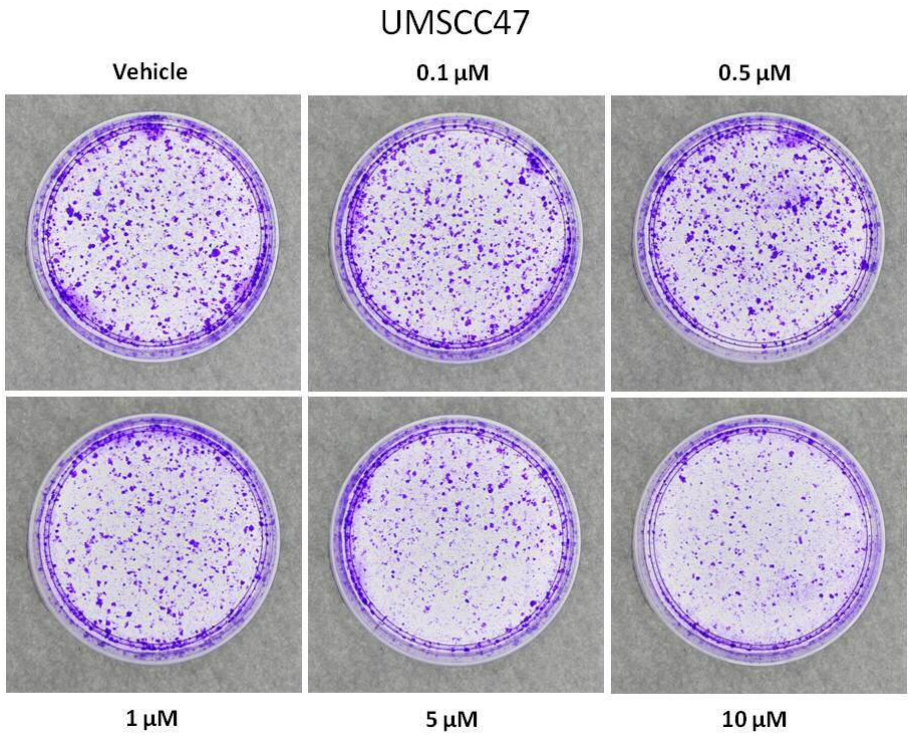
11. REFERENCES

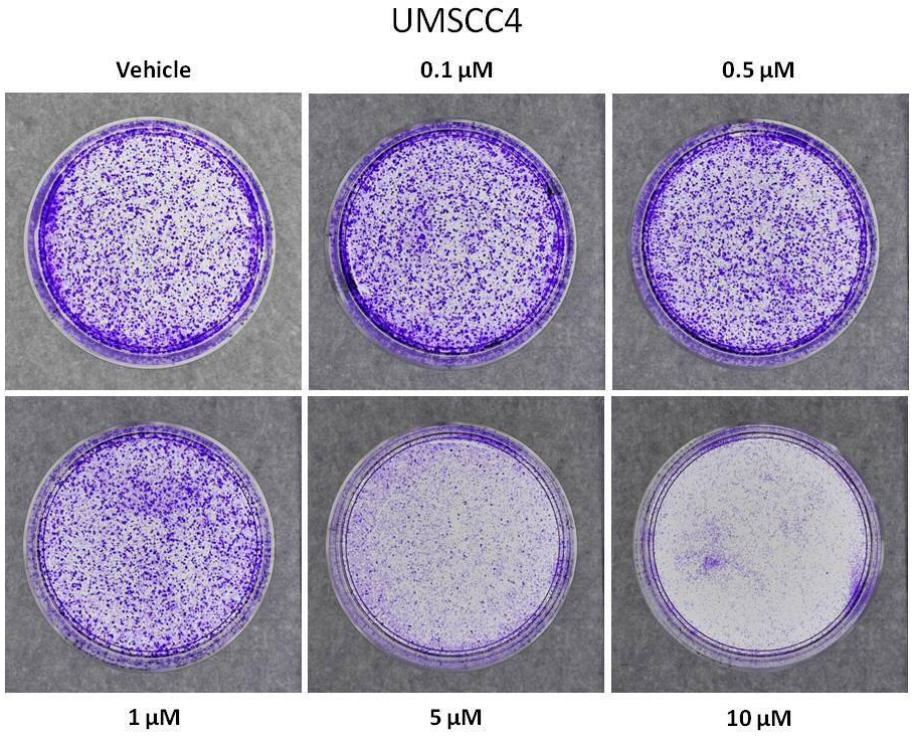
1. Nasman, A., et al., *Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma?* Int J Cancer, 2009. **125**(2): p. 362-6.
2. Mellin, H., et al., *Human papillomavirus type 16 is episomal and a high viral load may be correlated to better prognosis in tonsillar cancer.* International journal of cancer. Journal international du cancer, 2002. **102**(2): p. 152-8.
3. Weinberger, P.M., et al., *Molecular classification identifies a subset of human papillomavirus--associated oropharyngeal cancers with favorable prognosis.* J Clin Oncol, 2006. **24**(5): p. 736-47.
4. Lassen, P., et al., *Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck.* J Clin Oncol, 2009. **27**(12): p. 1992-8.
5. Worden, F.P., et al., *Chemoselection as a strategy for organ preservation in advanced oropharynx cancer: response and survival positively associated with HPV16 copy number.* J Clin Oncol, 2008. **26**(19): p. 3138-46.
6. Fakhry, C., et al., *Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial.* J Natl Cancer Inst, 2008. **100**(4): p. 261-9.
7. Dai, M., et al., *Human papillomavirus type 16 and TP53 mutation in oral cancer: matched analysis of the IARC multicenter study.* Cancer research, 2004. **64**(2): p. 468-71.
8. Licitra, L., et al., *High-risk human papillomavirus affects prognosis in patients with surgically treated oropharyngeal squamous cell carcinoma.* J Clin Oncol, 2006. **24**(36): p. 5630-6.
9. Ang, K.K., et al., *Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer.* N Engl J Med, 2010.
10. Stanley, M.A., *Establishing HPV-containing keratinocyte cell lines from tissue biopsies.* Methods Mol Med, 2005. **119**: p. 129-39.
11. Stanley, M.A., *Culture of Human Cervical Epithelial Cells*, in *Culture of Epithelial Cells*, R.I. Freshney and M.G. Freshney, Editors. 2002, Wiley: New York. p. 137-169.
12. Munoz, N., et al., *Epidemiologic classification of human papillomavirus types associated with cervical cancer.* N Engl J Med, 2003. **348**(6): p. 518-27.
13. Jacobs, M.V., et al., *A general primer GP5+/GP6(+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings.* J Clin Microbiol, 1997. **35**(3): p. 791-5.
14. Hibbitts, S., et al., *Human Papillomavirus Infection – an Anonymous Prevalence Study in South Wales, UK.* British Journal of Cancer, 2006. **95**(2): p. 226-32.
15. Wall, S.R., et al., *Cervical human papillomavirus infection and squamous intraepithelial lesions in rural Gambia, West Africa: viral sequence analysis and epidemiology.* Br J Cancer, 2005. **93**(9): p. 1068-76.
16. Parish, J.L., et al., *E2 proteins from high- and low-risk human papillomavirus types differ in their ability to bind p53 and induce apoptotic cell death.* J Virol, 2006. **80**(9): p. 4580-90.

Appendix 2: Clonogenic assay raw data example

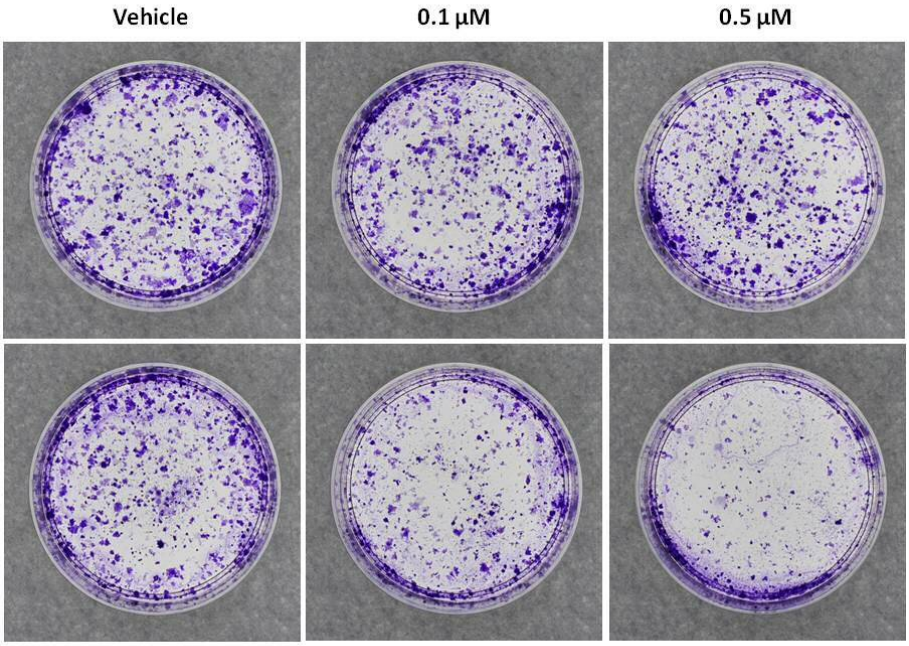




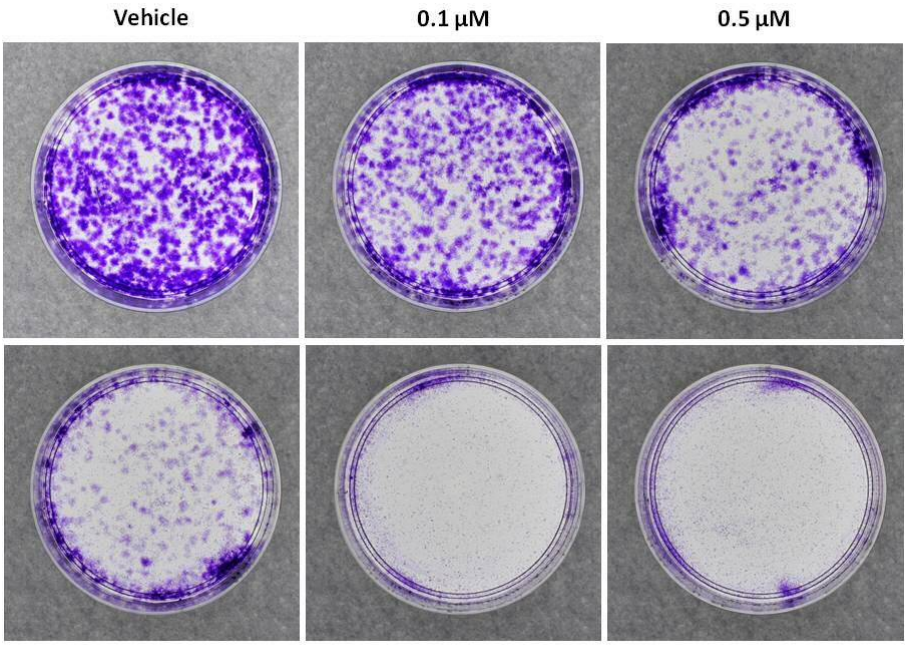

UMSCC47
Vehicle
0.1 µM
0.5 µM
1 µM
5 µM
10 µM


UMSCC4
Vehicle
0.1 µM
0.5 µM
1 µM
5 µM
10 µM

UMSCC19



UMSCC74a



Appendix 3: Paper

IJC
International Journal of Cancer

Human papillomavirus infection is rare in nonmalignant tonsil tissue in the UK: Implications for tonsil cancer precursor lesions

Elizabeth Palmer[1], Robert G. Newcombe[2], Adele C. Green[3,4], Carole Kelly[5], O. Noel Gill[5], Gillian Hall[6], Alison N. Fiander[1,7], Evelyne Pirotte[1], Sam J. Hibbitts[1], Jarrod Homer[8,9] and Ned G. Powell[1]

[1] HPV Research Group, Institute of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, United Kingdom
[2] Department of Primary Care and Public Health, School of Medicine, Cardiff University, Cardiff, United Kingdom
[3] Institute of Inflammation and Repair, University of Manchester, Manchester Academic Health Sciences Centre, Manchester, United Kingdom
[4] Cancer and Population Studies Group, Queensland Institute of Medical Research, Brisbane, Australia
[5] Public Health England Centre for Infectious Disease Surveillance and Control, Colindale, London, United Kingdom
[6] Department of Pathology, Manchester Royal Infirmary, Central Manchester University Hospitals Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom
[7] Department of Obstetrics and Gynaecology, School of Medicine, Cardiff University, Cardiff, United Kingdom
[8] Department of Otolaryngology-Head and Neck Surgery, Manchester Royal Infirmary, Central Manchester University Hospitals Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom
[9] Institute of Cancer Sciences, University of Manchester, Manchester, United Kingdom

The incidence of human papillomavirus (HPV)-associated tonsil cancer is increasing but the prevalence of HPV, and of premalignant precursors, in tonsil tissue is unknown. We aimed to assess prevalence of HPV infection in nonmalignant tonsillar crypt epithelia and to histopathologically characterise positive samples. Formalin-fixed paraffin-embedded (FFPE) tonsil tissue specimens were obtained from an age- and sex-stratified random sample of patients aged 0–69 years whose paired tonsils were archived following elective tonsillectomy at hospitals throughout England and Southern Scotland from 2004 to 2008. Homogenised fresh-frozen tonsil tissue was also obtained from archive for two random subsets of males aged 25–34 and over 44. HPV status was assessed in all samples for 20 mucosal HPV types by GP5+/6+ polymerase chain reaction (PCR) enzyme immunoassay and by HPV16 type-specific PCR targeting the E6 gene. In the homogenised material, HPV status was also assessed for 44 HPV types by SPF10-PCR enzyme immunoassay. Of 4,095 randomly sampled FFPE specimens, amplifiable DNA was extracted from 3,377 (82.5%) and from 511 of 524 (97.5%) homogenised tonsils. HPV DNA was identified in 0 of 3,377 (0%, 95% CI 0–0.089%) fixed samples and 0 of 511 (0%, 95% CI 0–0.58%) homogenised samples. This suggests HPV infection may be rare in tonsil reticulated crypt epithelia. Furthermore, we found no evidence of HPV-associated premalignant neoplasia. These data suggest that if HPV-associated premalignant lesions do occur, they are likely to be rare and may have a high risk of progression to carcinoma.

Epidemiology

Oropharyngeal squamous cell carcinoma (OPSCC) arises predominantly from the tonsils and base of the tongue. Tobacco and alcohol are known risk factors for OPSCC, but a growing proportion of cases are attributable to oncogenic human papillomavirus (HPV).[1–3] A meta-analysis of contemporary studies estimated the proportion of OPSCC associated with HPV to be 70% in North America and 73% in Europe,[4] and recent decades have seen sharp increases in the overall incidence of HPV-associated OPSCC.[5–7] In the USA, between 1988 and 2004, incidence rates for HPV-associated OPSCC

Key words: human papillomavirus, HPV, oropharyngeal, tonsil, cancer
Conflicts of interest: R.N. has received consultancy fees from Glaxo SmithKline for work on unrelated projects, mainly dental. A.F. has served on advisory boards for companies that produce HPV vaccines (GlaxoSmithKline Biologicals SA and Sanofi Pasteur MSD) and received research grant funding and support to attend HPV-related conferences from both companies. S.H. has received research grant funding and support to attend HPV-related conferences from companies that supply HPV tests (Roche). N.P. has received research funding and honoraria to speak/advise on HPV from companies that produce HPV vaccines (GlaxoSmithKline Biologicals SA and Sanofi Pasteur MSD)
Grant sponsor: Cancer Research UK (Population Research Committee); Grant number: C24566/A12523
DOI: 10.1002/ijc.28886
History: Received 9 Dec 2013; Accepted 11 Mar 2014; Online 11 Apr 2014
Correspondence to: Ned Powell, HPV Research Group, Institute of Cancer and Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, United Kingdom, Tel.: +44-2920-744742, Fax: +44-2920-744655, E-mail: powellng@cf.ac.uk

Int. J. Cancer: 135, 2437–2443 (2014) © 2014 UICC

What's new?

HPV, notorious for causing cervical cancer, also causes cancer of the tonsils. HPV-related tonsil cancers are on the rise in Britain, primarily in men. This study asked how common it is to find HPV in apparently healthy tonsils. The authors tested thousands of tonsils collected from elective tonsillectomies, but not a single one contained HPV, suggesting that HPV infects the tonsils only rarely. The authors also sought to uncover a pre-malignant lesion of the tonsils, analogous to the cervical intraepithelial neoplasia that afflicts the HPV-infected cervix, but found no such lesions in the tonsils.

rose from 0.8/100,000 to 2.6/100,000.[3] The incidence of tumours at specific subsites of the oropharynx may be underestimated owing to broad classification of many cases as ICD C10 "Malignant neoplasm of the oropharynx," but in England a minimum of 55% of OPSCC develop from the tonsils[8] and between 1995 and 2010 the age-standardised incidence rate (ASR) for ICD C09 "Malignant neoplasm of tonsil" increased threefold from 1.1 to 3.5/100,000.[8,9] In Sweden between 1970 and 2006, the incidence of HPV-associated tonsil carcinoma increased from 0.18/100,000 to 1.27/100,000.[7] Unlike other head and neck cancers, patients with HPV-positive OPSCC tend to be younger (usually <60 years) and often do not have a history of heavy alcohol and tobacco use.[5] Historically, HPV has been predominantly linked to cervical cancers in women; however, tonsil SCC predominantly affects men; in England in 2010, the ratio of male:female cases was 2.6:1.[8] A recent analysis suggests that in the USA by 2020, the number of HPV-positive OPSCC will exceed the number of cervical cancers.[3]

Despite now being recognised as the cause of the majority of OPSCC, the pathobiology and natural history of oropharyngeal HPV infection are not well understood. Data on the prevalence of HPV infection in the oral cavity, assessed by oral rinse, are accumulating,[10,11] and although such studies are essential for a full understanding of the epidemiology and aetiology of HPV-associated disease, they cannot be used to infer infection rates in the tonsillar reticulated crypt epithelia from which tonsillar carcinomas originate.[12] Determining infection rates in the tonsillar crypts requires studies based on invasively sampled tissue; however, to date the two largest studies in adults included only 212 and 229 cases and reported prevalence of 0 and 6.3%, respectively.[13,14] Among HPV-positive OPSCC, HPV16 is the predominant genotype, accounting for ~95% of cases[15]; however, it is not clear whether other HPV types also infect the tonsils but are not associated with malignancy. It is also unclear whether productive infection, resulting in release of infectious virions, can be established in the tonsils. The HPV life cycle is intimately linked to the differentiation of stratified squamous epithelium[16]; however, the squamous epithelium of the tonsillar crypts is reticulated (netlike) rather than stratified, and may not be compatible with completion of the viral life cycle. A final unanswered, and clinically very important question, is whether HPV infection is associated with a premalignant tonsillar lesion. In cervical epithelia, persistent viral infection leads to cervical intraepithelial neoplasia (CIN) followed by the accumulation of mutations that may lead to invasion and metastasis.[17] The premalignant CIN phase may last for months or decades, but whether an equivalent premalignant lesion occurs in the tonsils is unknown.

Our study aimed to address these questions by investigating HPV type-specific prevalence in a large nationally representative series of resected, clinically nonmalignant tonsils, followed by histopathological characterisation of HPV-positive cases, to determine whether productive infection and/or premalignant neoplasia were present. We also hypothesised that determination of HPV prevalence in noncancerous tonsils from males and females of different ages would allow inferences to be drawn regarding the age groups at risk of infection and potential routes of transmission.

Material and Methods

Study design

This was a cross-sectional study. Tonsil tissue was obtained from the National Anonymous Tonsil Archive. This archive was established to determine the prevalence of abnormal prion protein following the epidemic of bovine spongiform encephalopathy in cattle and related concerns regarding occurrence of variant Creutzfeldt–Jakob disease in humans. More than 100,000 paired tonsils were collected from children and adults following elective tonsillectomies performed at 134 hospitals throughout England and Southern Scotland between January 2004 and September 2008.[18] All tonsils were obtained with informed consent. One tonsil of each pair was collected as fresh tissue (chilled on wet ice during transit), then dissected, homogenised and stored at −80°C; the other tonsil was collected and transported in formalin, then embedded in paraffin at the study centre. Paired tonsils arrived at the study centre an average of 65 hr after operation.

For our study, 4,095 cases stratified for sex and geographical distribution were randomly sampled from the archive. As far as possible, these randomly sampled cases were equally distributed across the following age bands: 0–4, 5–9, 10–12, 13–15, 16–18, 19–21, 22–24, 25–29, 30–34, 35–39, 40–44, 45–49 and 50+ years (Table 1). The study had greater than 99% power to detect a difference between a prevalence rate of 1% and a prevalence rate of 5% between the two sexes and over 80% power to detect a difference between a prevalence rate of 1% and 5% between individual age bands. Research Ethics Committee approval was obtained before study commencement.

To control for the possibility that HPV present in very small foci of infection might be missed, two subsets of men

Table 1. Sample adequacy (beta-globin) and HPV testing for samples stratified by material type, age and gender

				Assay			
Material	**Age (years)**	**Sex (M/F)**	**Samples (*n*)**	**Beta-globin**	**GP5+/6+**	**HPV16 E6**	**SPF10 DEIA**
FFPE tonsils	0–4	M	180	154	0	0	–
		F	175	159	0	0	–
	5–9	M	165	136	0	0	–
		F	164	146	0	0	–
	10–12	M	164	143	0	0	–
		F	165	146	0	0	–
	13–15	M	164	148	0	0	–
		F	165	138	0	0	–
	16–18	M	163	149	0	0	–
		F	165	151	0	0	–
	19–21	M	165	142	0	0	–
		F	165	132	0	0	–
	22–24	M	165	135	0	0	–
		F	165	122	0	0	–
	25–29	M	168	131	0	0	–
		F	167	118	0	0	–
	30–34	M	168	129	0	0	–
		F	165	127	0	0	–
	35–39	M	169	138	0	0	–
		F	167	130	0	0	–
	40–44	M	165	133	0	0	–
		F	168	131	0	0	–
	45–49	M	116	101	0	0	–
		F	136	107	0	0	–
	50+	M	75	54	0	0	–
		F	101	77	0	0	–
	Total	M + F	4,095	3,377	0	0	–
Fresh homogenised tonsils	25–29	M	167	162	0	0	0
	30–34	M	168	162	0	0	0
	45–49	M	114	113	0	0	0
	50+	M	75	74	0	0	0
	Total	M	524	511	0	0	0

"–" denotes that the specified assay was not performed on this material.

from the original sample were selected for more intensive assessment of their tonsil tissue specimens. The results of oral rinse studies suggest that men in the age group 25–34 years are most likely to carry oral HPV infection followed by men over 44[10]; hence, archived homogenised fresh tonsil specimens from 335 and 189 men randomly sampled from these age groups were assayed for HPV DNA.

Laboratory analyses

For the main study sample HPV infection was initially assessed in sections cut from formalin-fixed paraffin-embedded (FFPE) blocks. This method was chosen as it allows further histopathological investigations in positive samples. FFPE blocks were sectioned with appropriate precautions to prevent block-to-block contamination (thorough cleaning of the microtome between samples, use of a fresh blade for every sample, use of Microsol wipes, *etc.*). Sections from blank (paraffin only) blocks were cut after every 20 samples and processed in parallel with cases. Sections from an FFPE HPV16-positive tonsil cancer were included as a positive control. DNA was extracted from 2 × 10 μm sections, taken through the centre of the tonsils, by overnight

digestion in 500 µl 50 mM Tris, 1 mM EDTA, 0.5% Tween with 1 mg/ml Proteinase K followed by heat inactivation (100°C, 5 min) and centrifugation. The supernatant was used as template in polymerase chain reactions (PCRs). Positive (HPV16-positive Caski cells) and negative controls were included in every extraction.

For the fresh tissue from the two subsets of adult males, DNA was extracted from homogenised tonsils by proteinase K (3 mg/ml) digestion overnight followed by phenol–chloroform extraction and ethanol precipitation. DNA purity and yield were determined by Nanodrop spectrophotometry. Positive (HPV16-positive Caski cells) and negative (water) controls were included for each extraction run. The adequacy of DNA for PCR analysis was assessed by amplification of a 209-bp region of the human beta-globin gene (PC03/PC05 primer set).[19] HPV typing was performed by PCR using GP5+/6+ primers (amplimer 142 bp) followed by enzyme immunoassay,[20] as previously described,[21] which detects 14 high-risk and six low-risk HPV types. Samples with absorbance greater than three times the mean of the negative controls were defined as positive[20]; in our laboratory this corresponded to 500 copies of the HPV16 genome per PCR (determined using WHO Labnet standards[22]). Positive (HPV16-positive Caski DNA) and negative (water) controls were included for each PCR run. Type-specific PCR to amplify a 161-bp fragment of the HPV16 E6 gene was performed using primers and conditions as previously described.[23] The HPV16 E6 type-specific PCR was an end-point assay and was assessed by visual examination of agarose gels including appropriate standards. The lower limit of sensitivity was identified as 500 genome equivalents using a dilution series of DNA extracted from a known number of Caski cells (each containing ~600 copies of HPV16). HPV testing was also performed using the SPF10-PCR Direct Enzyme Immuno Assay (DEIA) system (Labo Biomedical Products, Rijswijk, The Netherlands, based on licensed Innogenetics Technology),[24,25] as recently described (amplimer 65 bp).[26] Hundred nanograms of phenol–chloroform-extracted DNA was used as a template. The SPF10 DEIA is a licensed *in vitro* diagnostic assay and is validated for detection of 44 mucosal HPV types, including all 18 potentially carcinogenic HPV types (HPV16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82) plus a further 26 low-risk types.[27,28] The kit includes a "borderline" standard and absorbance readings above borderline are interpreted as positive; the manufacturer's kit insert specifies the limit of detection as two HPV16 genomes per PCR.

Statistical analyses

HPV prevalence rates were estimated with 95% confidence intervals (CI) determined by the method of Lancaster.[29]

Results

FFPE tissue specimens for 4,095 patients, randomly sampled across age and sex categories, were retrievable from archive and included in the analysis (Table 1). Amplifiable DNA, assessed by PCR for the human beta-globin gene, was extracted from 3,377 of 4,095 (82.5%) tonsil samples from 1,693 males of mean age 23.3 years and 1,684 females with mean age 23.6 years. The DNA extracted from these FFPE samples was tested for HPV DNA using the GP5+/6+ EIA and HPV16 E6 PCR assays. None of the 3,377 samples tested positive for HPV by either assay, corresponding to a prevalence of HPV infection in nonmalignant tonsil of 0 of 3,377 (0%, 95% CI 0–0.089%) samples. The extraction and PCR-positive and -negative controls gave appropriate results in both assays. In samples that gave OD readings approaching the positive cut-off value, the assay was repeated to determine whether these samples represented low-level positives, but on repetition OD readings were close to the mean value observed for negative controls. To confirm that the reticulated crypt epithelia, from which tonsil SCC arises, had been sampled, sections adjacent to those used for DNA extraction were taken from 20 randomly selected blocks and subjected to H&E staining and histopathological examination. This confirmed that, as is normal for this tissue, around 20% of the cells on each slide comprised reticulated crypt epithelia and 80% were lymphoid stroma.

HPV prevalence was then determined in DNA extracted from homogenised fresh tonsils from individuals likely to have the highest prevalence of oral HPV infection (*i.e.*, men 25–34 years and >44 years).[10] This material was representative of the whole tonsil (in contrast to sections through centre of embedded material). The human beta-globin gene was successfully amplified in 511 of 524 cases (97.5%), 324 were aged 25–34, 187 were aged 45–69 and the overall mean age was 36.9 years. HPV infection was assessed using the GP5+/6+ EIA and HPV16 E6 PCR. All extraction controls and PCR controls produced the appropriate results, but all the 511 samples tested negative for HPV DNA. The analytical sensitivity of the GP5+/6+ EIA assay was determined using a standardized sample set obtained from the WHO HPV Labnet project,[22] which demonstrated that the GP5+/6+ PCR EIA was sensitive to 500 genome equivalents for HPV16. The sensitivity of the E6 HPV16 type-specific PCR was also 500 genome equivalents.

Finally, a highly sensitive commercial assay, the SPF10-PCR DEIA, was applied to DNA extracted from the fresh homogenised tonsils only. The SPF10-PCR DEIA has an analytical sensitivity of two HPV16 genome equivalents per PCR, and has been widely applied in studies of archival tissue.[26] All controls gave the appropriate results and all 511 fresh tonsil tissue samples tested negative, confirming 0% HPV prevalence (95% CI 0–0.58%).

Discussion

Despite rising incidence rates of HPV-associated tonsil cancer in the UK population, no HPV DNA was identified in benign tonsil tissue specimens from 3,377 English and Scottish males and females aged 0–69 years (average age 23). There are

three possible explanations for this finding. First, the sampling regimen could have been inadequate. To address this question, HPV infection was also assessed in homogenised whole tonsils from two subsets of men in the age groups at highest risk of HPV infection, *i.e.*, 25–34 and over 44. The increased risk of HPV infection in these groups was predicated on two observations: first, that HPV infection, detected in oral rinse samples, has been shown to have a bimodal distribution with highest prevalence among individuals aged 30–34 years, with men showing a significantly higher prevalence than women,[10] and second, that HPV-associated tonsil carcinomas are most common in men aged 45–75.[8] Given the absence of HPV in tissue homogenates, it did not appear that the negative results were due to inadequate sampling.

The second possible explanation was use of inappropriate or insensitive assays, but this was not the case here because the GP5/6 and SPF10 DEIA systems amplify 20 and 44 HPV types, with analytical sensitivities of 500 and two HPV16 genome equivalents, respectively. Furthermore, successful amplification of the human beta-globin gene in more than 80% of samples showed that PCR inhibition and excessive degradation of DNA were not significant factors, and inclusion of appropriate extraction and PCR-positive controls confirmed that the assays were performed correctly.

Therefore, we propose that the third explanation is the most likely, namely that HPV infection in nonmalignant tonsil is a very rare event. Given our large sample size, the greatest population prevalence for HPV consistent with our entirely negative result is 0.089%.

One small study of tonsillectomy samples from adults (USA, 1979–2001, $n = 212$) has previously reported an HPV prevalence of 0%.[13,14] This is highly consistent with our data, but owing to the small sample size, the USA study gave an upper confidence limit of 3.5% compared to 0.089% in our study. Smaller studies of predominantly juvenile tonsillectomy samples (Finland, 2001–2002, $n = 229$; Greece, 1995–2000, $n = 102$; Brazil, before 2005, $n = 100$; USA, before 2006, $n = 50$) have reported prevalence rates of 6.3, 8.5, 0 and 4%, respectively.[30–32] Methodological and quality control issues might underlie these differences; however, just as genital HPV prevalence varies between populations,[33,34] some differences between HPV prevalence in tonsillectomy studies might reflect genuine variation between populations. Hence, it is not clear how generalisable our findings may be outside the UK.

In relation to previous studies, it is important to distinguish between investigations of oral HPV infection ascertained from oral rinse samples containing cells primarily derived from the stratified squamous epithelium of the oral cavity and, studies such as ours, of resected tonsil tissue containing reticulated crypt epithelium from which OPSCC develop. This is relevant, as there is no evidence that HPV infection rates in the oropharynx can be inferred from oral HPV prevalence. In contrast to studies of benign tonsil tissue, where data are sparse, there is an emerging consensus regarding HPV infection in the oral cavity with substantial studies reporting significant levels of infection, especially in the USA where a prevalence of 6.9% (sampled 2009–2010) has been reported.[10] A recent meta-analysis estimated a pooled prevalence of 4.5% in oral samples collected in 11 countries between 1994 and 2007.[11]

The strengths of this study include the examination of nonmalignant tonsils from an unselected series of individuals during a recent period (2004–2008), when rates of tonsillar carcinoma had become relatively high (ASR 2.6/100,000) in the source population.[8] The large sample meant the study was highly powered (>99%) to detect prevalent HPV. The availability of homogenised whole tonsils and the use of multiple highly sensitive methods for HPV testing, targeting separate regions of the HPV genome, were additional strengths. The tonsillectomy specimens came from unselected people in regions across England and Southern Scotland[18] and these results should be broadly representative of the UK population. The primary weakness was that we could not rule out HPV infection present below the limits of detection of the assays used; hence, small foci of HPV-infected cells may have been missed. The predominance of lymphoid cells in tonsil tissue may also hinder the detection of small foci of HPV-infected epithelial cells. Our data are, however, certainly inconsistent with the presence of diffuse or productive HPV infections.

These findings have significant implications for understanding of the natural history of HPV in the tonsils, and suggest it differs substantially from the natural history of HPV in cervical tissue. In the cervix, HPV infection is common but progression is not; for tonsils, our data are more consistent with infection being rare, but the proportion of infections that progress being higher. Lower infection rates in tonsil tissue would be consistent with greater immunological exposure in tonsil epithelia relative to cervical epithelia, and with the recently reported low prevalence of HPV E6-directed antibodies in sera of people without HPV-associated disease.[35] Although detection of E6 antibodies >10 years before diagnosis of OPSCC suggests that a long duration of infection before diagnosis may be usual for oropharyngeal as well as for cervical cancers.[35] The lower prevalence of HPV in the tonsils might also be explained by cells of the reticulated crypt epithelium, from which most HPV-associated tonsil SCC develop,[12] being semireceptive to viral infection and semipermissive to expression of viral proteins. This could arise from a receptor required for viral entry into the cell being rare in the population, or from infrequent missrecognition of a common receptor. A semipermissive model of viral gene expression is potentially relevant as virus-induced cancers often arise at sites where productive viral infection cannot be properly supported.[16] The reticulated epithelial crypts comprise specialised epithelia with immune and secretory functions, not stratified epithelia,[36] hence they may not support the full HPV life cycle but could allow expression of the viral early genes including oncogenic E6 and E7.

One of the original aims of our study was to identify and characterise HPV-associated premalignant lesions in tonsil tissue, but the absence of any HPV-positive tonsils prevented this. The existence of such lesions was hypothesised largely by analogy to cervical neoplasia, where prolonged persistent HPV infection is often associated with CIN.[16] The absence of HPV-associated neoplasia in this study does not prove that it did not occur, but suggests that it may be uncommon without subsequent development of OPSCC, although HPV-associated field change has not been observed in tissue adjacent to OPSCC.[12] These data are highly relevant in the context of prevention of OPSCC. Because of the increasing incidence of OPSCC and the success of screening in preventing HPV-associated cervical cancer,[37] screening for OPSCC or its precursors has been considered.[38] However, effective prevention would depend on identification of a treatable HPV-associated premalignant neoplasia, which was not observed in our study. Similarly, the potential of tonsillar cytology appears limited by difficulty in sampling the correct mucosa.[39,40] Therefore, it appears that prophylactic vaccination or serological screening for E6 antibodies may be more promising options for long-term prevention of tonsil cancer.[35,41]

Acknowledgements

The authors are grateful for the assistance of the staff of the HPA, especially Jon Clewley, Chris Shaw, Reza Dabaghian and Kate Soldan. They are also grateful to Dr. Annika Antonsson for critical reading of the manuscript.

References

1. D'Souza G, Kreimer AR, Viscidi R, et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med* 2007;356:1944–56.
2. Gillison ML, D'Souza G, Westra W, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst* 2008;100:407–20.
3. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol* 2011;29:4294–301.
4. Mehanna H, Beech T, Nicholson T, et al. Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer-systematic review and meta-analysis of trends by time and region. *Head Neck* 2012;35:747–55.
5. Chaturvedi AK, Engels EA, Anderson WF, et al. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol* 2008;26:612–19.
6. Hocking JS, Stein A, Conway EL, et al. Head and neck cancer in Australia between 1982 and 2005 show increasing incidence of potentially HPV-associated oropharyngeal cancers. *Br J Cancer* 2011;104:886–91.
7. Nasman A, Attner P, Hammarstedt L, et al. Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma? *Int J Cancer* 2009;125:362–6.
8. Anon. Cancer Statistics Registrations, England (Series MB1), No. 41, 2010. Available at: http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Cancer+Registrations#tab-data-tables (Accessed 08/07/2013).
9. Anon. Cancer Statistics Registrations, England (Series MB1), No. 28, 1995–1997. Available at: http://www.ons.gov.uk/ons/rel/vsob1/cancer-statistics-registrations-england-series-mb1-/no-28-1995-1997/index.html (Accessed 08/07/2013).
10. Gillison ML, Broutian T, Pickard RK, et al. Prevalence of oral HPV infection in the United States, 2009–2010. *JAMA* 2012;307:693–703.
11. Kreimer AR, Bhatia RK, Messeguer AL, et al. Oral human papillomavirus in healthy individuals: a systematic review of the literature. *Sex Tranm Dis* 2010;37:386–91.
12. Begum S, Cao D, Gillison M, et al. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clin Cancer Res* 2005;11:5694–9.
13. Ernster JA, Sciotto CG, O'Brien MM, et al. Prevalence of oncogenic human papillomavirus 16 and 18 in the palatine tonsils of the general adult population. *Arch Otolaryngol Head Neck Surg* 2009;135:554–7.
14. Chen R, Sehr P, Waterboer T, et al. Presence of DNA of human papillomavirus 16 but no other types in tumor-free tonsillar tissue. *J Clin Microbiol* 2005;43:1408–10.
15. Dayyani F, Etzel CJ, Liu M, et al. Meta-analysis of the impact of human papillomavirus (HPV) on cancer risk and overall survival in head and neck squamous cell carcinomas (HNSCC). *Head Neck Oncol* 2010;2:15.
16. Doorbar J. Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci* 2006;110:525–41.
17. Snijders PJ, Steenbergen RD, Heideman DA, et al. HPV-mediated cervical carcinogenesis: concepts and clinical implications. *J Pathol* 2006;208:152–64.
18. Clewley JP, Kelly CM, Andrews N, et al. Prevalence of disease related prion protein in anonymous tonsil specimens in Britain: cross sectional opportunistic survey. *BMJ* 2009;338:b1442.
19. de Roda Husman AM, Snijders PJ, Stel HV, et al. Processing of long-stored archival cervical smears for human papillomavirus detection by the polymerase chain reaction. *Br J Cancer* 1995;72:412–17.
20. Jacobs MV, Snijders PJ, van den Brule AJ, et al. A general primer GP5+/GP6(+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *J Clin Microbiol* 1997;35:791–5.
21. Powell NG, Boyde AM, Tristram AJ, et al. The potential impact of human papillomavirus vaccination in contemporary cytologically screened populations may be underestimated: an observational retrospective analysis of invasive cervical cancers. *Int J Cancer* 2009;125:2425–7.
22. Eklund C, Forslund O, Wallin KL, et al. The 2010 global proficiency study of human papillomavirus genotyping in vaccinology. *J Clin Microbiol* 2012;50:2289–98.
23. Collins SI, Constandinou-Williams C, Wen K, et al. Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study. *Cancer Res* 2009;69:3828–32.
24. Kleter B, van Doorn LJ, Schrauwen L, et al. Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *J Clin Microbiol* 1999;37:2508–17.
25. van Doorn LJ, Molijn A, Kleter B, et al. Highly effective detection of human papillomavirus 16 and 18 DNA by a testing algorithm combining broad-spectrum and type-specific PCR. *J Clin Microbiol* 2006;44:3292–8.
26. Tjalma WA, Fiander A, Reich O, et al. Differences in human papillomavirus type distribution in high-grade cervical intraepithelial neoplasia and invasive cervical cancer in Europe. *Int J Cancer* 2013;132:854–67.
27. de Villiers EM, Fauquet C, Broker TR, et al. Classification of papillomaviruses. *Virology* 2004;324:17–27.
28. Munoz N, Bosch FX, de Sanjose S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003;348:518–27.
29. Lancaster HO. The combination of probabilities arising from data in discrete distributions. *Biometrika* 1949;36:370–82.
30. Mammas IN, Sourvinos G, Michael C, et al. Human papilloma virus in hyperplastic tonsillar and adenoid tissues in children. *Pediatr Infect Dis J* 2006;25:1158–62.
31. Ribeiro KM, Alvez JM, Pignatari SS, et al. Detection of human papilloma virus in the tonsils of children undergoing tonsillectomy. *Braz J Infect Dis* 2006;10:165–8.
32. Sisk J, Schweinfurth JM, Wang XT, et al. Presence of human papillomavirus DNA in tonsillectomy specimens. *Laryngoscope* 2006;116:1372–4.
33. De Vuyst H, Clifford G, Li N, et al. HPV infection in Europe. *Eur J Cancer* 2009;45:2632–9.
34. Hariri S, Unger ER, Sternberg M, et al. Prevalence of genital human papillomavirus among females in the United States, the National Health And Nutrition Examination Survey, 2003–2006. *J Infect Dis* 2011;204:566–73.
35. Kreimer AR, Johansson M, Waterboer T, et al. Evaluation of human papillomavirus antibodies

and risk of subsequent head and neck cancer. *J Clin Oncol* 2013;31:2708–15.

36. Perry ME. The specialised structure of crypt epithelium in the human palatine tonsil and its functional significance. *J Anat* 1994;185 (Part 1): 111–27.
37. Peto J, Gilham C, Fletcher O, et al. The cervical cancer epidemic that screening has prevented in the UK. *Lancet* 2004;364:249–56.
38. Gillison ML, Castellsague X, Chaturvedi A, et al. Comparative epidemiology of HPV infection and associated cancers of the head and neck and cervix. *Int J Cancer* 2014;134:497–507.
39. Fakhry C, Rosenthal BT, Clark DP, et al. Associations between oral HPV16 infection and cytopathology: evaluation of an oropharyngeal "pap-test equivalent" in high-risk populations. *Cancer Prev Res* 2011;4:1378–84.
40. Lingen MW. Brush-based cytology screening in the tonsils and cervix: there is a difference! *Cancer Prev Res (Phila)* 2011;4: 1350–2.
41. Herrero R, Quint W, Hildesheim A, et al. Reduced prevalence of oral human papillomavirus (HPV) 4 years after bivalent HPV vaccination in a randomized clinical trial in Costa Rica. *PloS One* 2013;8:e68329.