**Division of Cancer and Genetics**

**School of Medicine**

**Cardiff University**


# Effect of Ionising Radiation on HPV-positive and HPV-negative Oropharyngeal Cancer Cell Lines

**By**

**Stefan Holzhauser (MSc.)**

A thesis submitted to the School of Medicine, Cardiff University in partial fulfilment of the requirements for the degree of Doctor of Philosophy

**October 2018**

## Acknowledgements

Firstly, I would like to thank my supervisors Dr. Stephen Man, Dr. Ned Powell and Dr. Mererid Evans for their generous support of my PhD study, for their immense knowledge, guidance and patience during the whole time of my research and writing of my thesis, in good and difficult moments.

Beside my supervisory team I would like to thank the Wales Gene Park team for their help during my research and the clinicians at the Heath Hospital for providing several biopsy samples for my study.

Many thanks to my fellow lab-mates at the Cancer and Genetics building and for all my other friends in Cardiff for all their help and fun during my PhD. I am very grateful for their support during all the difficult and emotional times at and off work. A special thank you to Joanne Jones for all the support during my studies, being a helpful hand if needed and for being a great friend.

I would also thank Cancer Research Wales for funding this project, as well as the Cardiff University School of Medicine.

Special thanks to my dear friend Dr. Javier Uceda who has sadly passed away, whom I was lucky to call a great friend during the 4 years of our PhD's. You have always been a good listener, motivator, an amazing flatmate, and for cheering me up in bad times with your happy spirit and smile. I will always remember the amazing time we have spent together. You will be truly missed, but I know you will keep an eye on your family and friends. Many thanks my dear friend and may you rest in peace.

Finally, I would like to thank my family: my parents and my brothers for supporting me spiritually throughout my PhD, especially the writing period. The great support of my family during the difficult times kept me strong and focused and I am very grateful for having you in my life.

## Summary

In recent decades, the incidence of human papillomavirus (HPV) associated oropharyngeal squamous cell carcinoma (OPSCC) has increased world-wide. Overall, HPV-positive OPSCC patients respond better to treatment (increased survival rate) compared to HPV-negative patients. This might be partially associated with a deficiency in repair of double-strand DNA breaks, and/or with residual p53 activity in HPV-positive tumours. However new studies, specific to HPV-positive OPSCC, are limited due to the low number of relevant OPSCC *in-vitro* models.

The general aims of this study were to develop new HPV-positive OPSCC cell lines and use them, together with established OPSCC cell lines, to investigate responses to ionising radiation (IR). These experiments were intended to test that HPV-positive OPSCC cell lines were more sensitive to IR than HPV-negative OPSCC cell lines. Two novel OPSCC cell lines, one HPV-positive and one HPV-negative, were derived and characterised.

The HPV-positive OPSCC cell lines demonstrated greater variation in radio-sensitivity compared to HPV-negative OPSCC cell lines. However, radio-sensitivity was not associated with p53 accumulation and/or cell cycle arrest. All HPV-positive OPSCC cell lines showed G2 arrest after IR, but so did several HPV-negative lines. The mRNA sequencing data confirmed expression of HPV oncogenes and integration of HPV DNA into the host genome, with an increase of integration sites after IR.

Comparison of irradiated HPV-positive and HPV-negative cell lines did not show consistent differences in gene expression associated with DNA repair. The transcription of DNA repair factors did not correlate with radio-sensitivity within the HPV-positive cell lines.

The study was successful in generating and characterising new OPSCC cell lines but did not find evidence that the better prognosis of HPV-positive tumours is associated with defects in DNA repair. This suggests that additional mechanisms may be responsible for the improved prognosis of HPV-positive OPSCC patients following treatment.

# Abbreviations

| | |
|---|---|
| AIN | Anal Intraepithelial Neoplasia |
| ATM | Ataxia Telangiectasia Mutated |
| ATP | Adenosine Triphosphate |
| ATR | Ataxia Telangiectasia and RAD3 related |
| BER | Base Excision Repair |
| BP | Biological Processes |
| CC | Cellular Components |
| CI | Confidence Interval |
| CIN | Cervical Intraepithelial Neoplasia |
| CRT | Chemo-radiotherapy |
| Cs-137 | Cesium-137 |
| CU-OP | Cardiff University Oropharyngeal Cancer |
| DMEM | Dulbecco's Modified Eagle's Medium |
| dNTP | Deoxynucleotide |
| DSB | Double Strand Breaks |
| E | Early |
| E6-AP | E6-associated protein |
| EBV | Epstein-Barr virus |
| EDTA | Ethylenediaminetetraaecetic acid |
| ELISA | Enzyme Linked Immunosorbent Assay |
| ERK1/2 | Extracellular-signal-regulated kinase 1/2 |
| FBS | Fetal Bovine Serum |
| FDR | False Discovery Rate |
| FFPE | Formalin fixed and embedded in paraffin |
| GO | Gene Ontology |
| GO ORA | Gene Ontology Over-representative analysis |
| Gy | Gray |
| HBV | Hepatitis B virus |
| HCV | Hepatitis C virus |
| HEKn | Human Epidermal Keratinocytes (neonatal) |
| HKGS | Human Keratinocyte Growth Supplement |
| HHV8 | Human Herpes virus 8 |
| HNSCC | Head and Neck Squamous Cell Carcinomas |
| HPV | Human Papillomavirus |
| HR | Homologous Recombination |
| HR-HPV | High-risk HPV |
| HSPG | Heparin Sulphate Proteoglycans |
| hTERT | Human Telomerase Reverse Transcriptase |
| HTLV-1 | Human T-lymphotropic virus type 1 |
| ICTV | International Committee on the Taxonomy of Viruses |
| IGV | Integrative Genomics Viewer |
| IHC | Immunohistochemistry |

| | |
|---|---|
| IMRT | Intensity-Modulated Radiotherapy |
| IR | Ionising Radiation |
| L | Late |
| LCR | Long-control region |
| LOH | Loss of Heterozygosity |
| LR-HPV | Low-risk HPV |
| MAP | Mitogen-activated protein |
| Mb | Mega base |
| MCPyV | Merkel Cell Polyomavirus |
| MF | Molecular Functions |
| MMR | Mismatch Repair |
| NEAA | Non-essential Amino acids |
| NER | Nucleotide Excision Repair |
| NHEJ | Non-homologous End-joining |
| NISCHR | National Institute for Social Care and Health Research |
| OPSCC | Oropharyngeal Squamous Cell Carcinoma |
| ORF | Open Reading Frame |
| PBS | Phosphate Buffer Saline |
| PCOC | Primary Culture of Oropharyngeal Cells |
| PCR | Polymerase Chain Reaction |
| PD | Population Doublings |
| PDX | Patient Derived Xenografts |
| Pen Strep | Penicillin/Streptomycin |
| PHE | Public Health England |
| p-p53 | Phospho-p53 |
| pRb | Retinoblastoma protein |
| PV | Papillomavirus |
| R&D | Research and Development |
| RIN | RNA Integrity Number |
| RPKM | Reads Per Kilobase of transcripts per Million mapped reads |
| RT | Radiotherapy |
| SF | Surviving Fraction |
| SNP | Single Nucleotide Polymorphism |
| SSB | Single Strand Breaks |
| SSC | Squamous Cell Carcinoma |
| STR | Short Tandem Repeats |
| TBE | TRIS/Borate/EDTA |
| TE | Trypsin/EDTA |
| TILs | Tumour Infiltrating Lymphocytes |
| TNM | Tumour-Node-Metastasis |
| TORS | Trans-oral Robotic Surgery |
| UMSCC | University of Michigan Squamous Cell Carcinoma |
| UPCI-SCC | University of Pittsburgh Cancer Institute Squamous Cell Carcinoma |
| URR | Upstream Regulatory Region |

| | |
|---|---|
| UV | Ultraviolet Radiation |
| VAIN | Vaginal Intraepithelial Neoplasia |
| VIN | Vulval Intraepithelial Neoplasia |
| VLP | Virus-like particles |
| WHO | World Health Organisation |

# Table of Contents

# Table of Figures

# Table of Tables

# Chapter 1 : Introduction

## 1.1. Oropharyngeal Squamous Cell Carcinoma

### 1.1.1. Anatomy of the oropharynx

Head and Neck cancers are comprised of a wide variety of tumours from different anatomic structures including the craniofacial bones, soft tissue, salivary glands, skin and mucosal membranes (Pai and Westra, 2009). This includes cancers of the oral cavity, the hypopharynx, the larynx, the sinonasal tract, the nasopharynx and the oropharynx (Ramqvist and Dalianis, 2010). The focus of this study was the oropharynx, and specifically Oropharyngeal Squamous Cell Carcinoma (OPSCC).

The pharynx is classified into three sections. The nasopharynx is seen posterior to the nasal cavity, the hypopharynx lies behind the larynx, and the oropharynx is found posterior to the oral cavity (Hermans and Lenz, 1996). The oropharynx is hence located between the nasopharynx and larynx/hypopharynx. It is composed of four distinct parts: base of tongue (back 1/3 of the tongue), soft palate, tonsil and back site of the throat (posterior pharyngeal wall) (Figure 1-1) (Harrison et al., 2009).

The oropharynx includes a circular collection of submucosal lymphoid tissue (lingual and palatine tonsils) known as Waldeyer's ring, that is the first line of defence against airborne and ingested pathogens. The tonsil epithelium comprise part of Waldeyer's ring and consist of blind-ending crypts through the full thickness of the palatine tonsil. These patches of crypts are lined by a net-like squamous epithelium, known as reticulated epithelium, that is structured to enable transport of foreign antigens from the oropharynx to the tonsil lymphoid tissue (Pai and Westra, 2009).

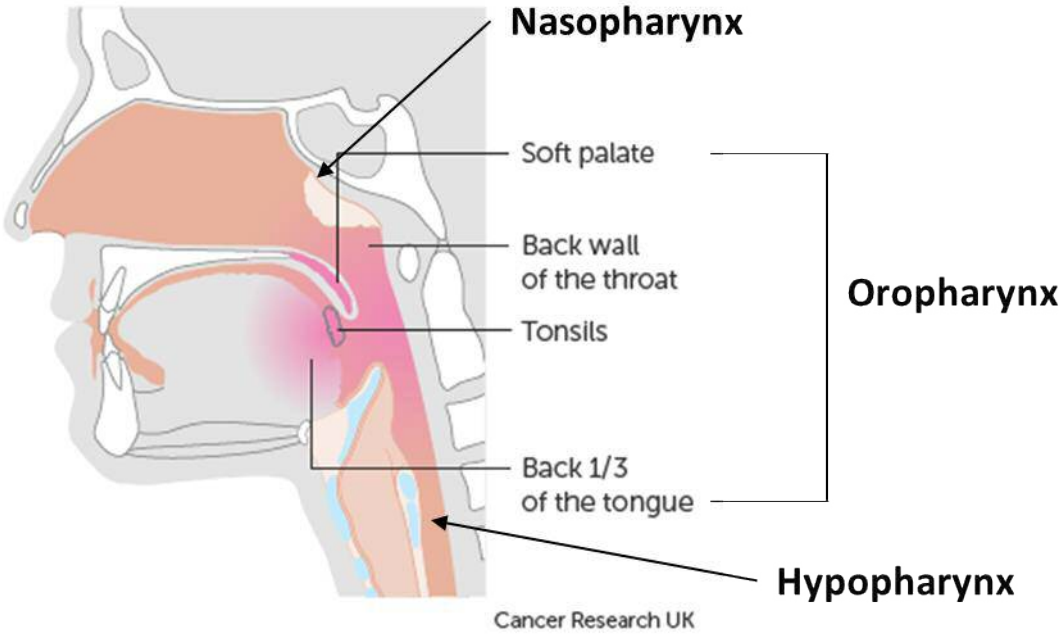



**Figure 1-1: Anatomy of the oropharynx**

The Oropharynx is located between the nasopharynx and larynx/hypopharynx and consists of the soft palate, back wall of the throat (, tonsils and back third of the tongue (Picture obtained and adapted from Cancer Research UK: www.cancerresearchuk.org; 07.08.2018).

### 1.1.2. Epidemiology of head and neck/oropharyngeal squamous cell carcinoma

Head and neck cancers are the sixth most common cancer types worldwide. Approximately 650,000 new patients are diagnosed with this type of cancer and around 350,000 deaths are reported per year (Argiris et al., 2008). Around 85% of head and neck cancers are head and neck squamous cell carcinomas (HNSCC). The incidence rate of HNSCC is known to be different according to the geographical region. A high incidence of HNSCC is found in Australia, France (Europe), Brazil (South America), Southern Africa and the Indian subcontinent (Argiris et al., 2008, Mehanna et al., 2010).

Over the decades, the incidence rates of HNSCC has changed in the US and the UK. The incidence rate for HNSCC overall has either decreased or remained stable. This might be attributed to anti-tobacco campaigns resulting in decreased consumption of tobacco. However, the incidence rates for tonsil and base of tongue has increased, particularly in young adults in the USA and European Countries (Argiris et al., 2008, Schache et al., 2016).

This increase of Human Papillomavirus (HPV)-positive related OPSCC's has been reported in several countries around the world (Stein et al., 2015). For example, in the Netherlands the prevalence of HPV in OPSCC rose from 5% to 29% (1990-2010), Australia an increase from 19% (1987-1990) to 47% (2001-2005) was observed. The United States reported an increase from 16.3% (1984-1989) to 72.7% (200-2004) (Stein et al., 2015).

Within the UK, increased incidence of HPV-positive and HPV-negative OPSCC cases has been documented with a 2-fold increase from 2002, with an age-standardised-rate of 2.1 per 100,000 (95% Confidence Interval (CI), 1.9-2.2) to 4.1 (95% CI, 4.0-4.3) in 2011 (Schache et al., 2016). The proportion of HPV-positive cases within the OPSCC patients has stayed stable at 50% (Schache et al., 2016). In a study cohort for South Wales, 55% of cases were HPV-positive in the time period of 2001-2006 (Evans et al., 2013). Compared to HPV-negative OPSCC, HPV-positive OPSCC show characteristic pathological, clinical and pathological features. Therefore, it can be characterised as a distinct disease entity (Gillison and Shah, 2001).

### 1.1.3. Risk Factors for head and neck cancer

Historically, tobacco products and alcohol consumption are the two main risk factors for head and neck cancers. The risk of head and neck cancer increases in proportion to the consumption of tobacco (Franceschi et al., 1990). This effect can also be indirect; increased risk is also associated with long duration of involuntary smoking exposure at home or at work. This effect was stronger for pharyngeal and laryngeal cancer than for other head and neck cancer types (Lee et al., 2008, Zhang et al., 2000). Polycyclic aromatic hydrocarbons (PAH), which are part of cigarette smoke, are known to have a genotoxic/carcinogenic effect. After they are metabolised, some of the resulting products can bind to DNA, developing mutagenic DNA adducts (Gelboin, 1980). Due to exposure of PAHs, several cytochrome P450 enzymes are induced (e.g CYP1A1 and CYP1B1) via the aryl hydrocarbon receptor (AhR). AhR then forms a complex with PAH's (with several other proteins, such as Hsp90, p23 and AhR interacting protein), which creates a heterodimer with an AhR nuclear translocator (ARNT) and binds to DNA via the xenobiotic response element (XRE). XRE is located in the promoter region of CYP1A and CYP1B genes, which then plays a significant role in tumorigenesis mediated by PAHs (Ewa and Danuta, 2017). One well known obtained PAH during is process is the activation of the carcinogen Benzo-(α)-pyrene (B(α)P) (Androutsopoulos et al., 2009, Ewa and Danuta, 2017). A study in 2008 has shown that xenobiotic metabolizing enzymes CYP1A1 and CYP1B1 were increased in the oral mucosa of smokers. This is of high importance as both CYP1A1 and CYP1B1 were found to convert a wide range of carcinogens to active metabolites that can form DNA adducts (Gumus et al., 2008).

A study in Brazil of 784 cases of mouth pharynx and larynx cancer and 1,578 non-cancer control cases, investigated the effect of different tobacco types. It was shown that the risk for all upper aerodigestive tract cancers and pharynx decreases for ex-smokers, but stays relatively higher for larynx. In general, relative risk remains higher than for non-smokers (Schlecht et al., 1999).

The second main risk factor is chronic alcohol consumption. The carcinogenic effect of alcohol results from metabolism of ethanol into acetaldehyde, with direct DNA damage induction (Baan et al., 2007). The metabolism of ethanol occurs in two main steps. First alcohol dehydrogenase (ADH) converts ethanol to acetaldehyde, then acetaldehyde gets further converted to acetate by aldehyde dehydrogenase (ALDH). Although the major acetaldehyde production in humans is found in the liver and gut, the highest level after

alcohol consumption is found in the saliva, which increases the risk for cancer in the oral cavity and other sites of head and neck (Tsai et al., 2014). Alcohol seems to synergistically enhance the effect of tobacco metabolites on the cell, due to its characteristics as a solvent potentially increasing tissue exposure to the metabolites (Pai and Westra, 2009). There is one gene in particular, ALDH2, in which mutations or polymorphisms are known to increase the risk to develop HNSCC (Hakenewerth et al., 2011). Furthermore, in a study by Tsai et al. (2014), the interaction between oral hygiene and different genetic polymorphisms of ADH1B and ALDH2 was investigated. ADH1B is one of the major ADH enzymes for conversion of ethanol to acetaldehyde. Three different alleles are known of ADH1B which are formed by variations at codons 48 and 370. On the other hand, ALDH2 is well known for the conversion of acetaldehyde to acetate (well-known SNP rs671: allele *1 = glutamate at codon 504 and allele *2 = lysine at codon 504). Persons with *2/*2 genotype are unable to oxidize acetaldehyde to acetate. In this study it was stated that faster or slower activity of these enzymes (ADH1B or ALDH2) in correlation with oral hygiene has an effect on decrease/increase of head and neck cancer. The study concluded that both fast and slow ADH1B genotypes increase the risk for head and neck cancer. In persons with slow ADH1B, poor oral hygiene seemed to play a major role in risk for head and neck cancer. Furthermore, the risk was increased in people with fast ADH1B genotype and slow/non-functional ALDH2 genotypes (Tsai et al., 2014).

Besides alcohol and tobacco there are several other possible risk factors for HNSCC. Other known risk factors associated with development of HNSCC are the above mentioned poor oral hygiene, dental plaque formation, signs of chronic irritation of the mouth and/or family history can also play a critical role in HNSCC development (Hashim et al., 2016, Lissowska et al., 2003). Another main risk factor for HNSCC is caused by an oncovirus, known as HPV. The chapters below will give an overview about this virus.

## 1.2. Papillomaviruses

Papillomaviruses are found in a wide variety of animal species including humans. These viruses are likely to have evolved over the last 350 million years. Their origins appear to be connected to changes in the epithelium of their ancestral host, with viruses evolving to be specific to cutaneous and mucosal epithelia (Doorbar et al., 2015, Egawa et al., 2015) . Currently, over 200 different papillomavirus have been identified and completely sequenced, including around 170 types connected to humans (Doorbar et al., 2015, de Villiers, 2013).

### 1.2.1. Discovery of oncogenic viruses

In 1909 Francis Peyton Rous undertook his famous experiment on viral transmission of cancer. His experiment on sarcomas caused by a virus in chickens ("Rous sarcoma virus"), built the first milestone in virus-cancer research, which was later important for the discovery of human oncogenic viruses. In his experiment sarcomas were ground and filtered to eliminate intact cancer cells. The filtrate was then injected into a healthy chicken and tumour growth observed (Moore and Chang, 2010, Rous, 1910, Rous, 1911). Oncogenic viruses discovered over the last 54 years include hepatitis B virus (HBV), hepatitis C (HCV), Epstein-Barr virus (EBV), human herpes virus 8 (HHV8), Merkel cell polyomavirus (MCPyV), human T-lymphotropic virus type 1 (HTLV-1) and human papillomavirus (HPV) (Moore and Chang, 2010, Chang et al., 2017).

### 1.2.2. HPV and cancer

Before the discovery was made that linked HPV and cervical cancer, an initial link was proposed between herpes simplex virus type 2 and cervical cancer. The connection between HPV and cervical cancer was first postulated in 1972 (zur Hausen, 1975, zur Hausen, 1977, zur Hausen, 2009). In the 1980s the first HPV-16 sequence was published, and the HPV-16 DNA was successfully isolated from cervical cancer by southern blot hybridization (Durst et al., 1983, Gissmann et al., 1982a, Gissmann et al., 1982b, zur Hausen, 2009). In 2008, Harald zur Hausen was awarded a Nobel prize for establishing the causal link between HPV infection and cervical cancer.

In the early 1980s the first connection between HPV and oral and laryngeal squamous cell carcinogenesis was made (Syrjanen et al., 1983, Syrjanen et al., 1982, Syrjanen, 2010b). An increase of incidence of head and neck cancers (HNSCC) caused by HPV has been reported in the last several years, since the connection of HPV and HNSCC was made in Europe and world-wide (Attner et al., 2010, Nasman et al., 2015, Nasman et al., 2009, Gillison et al., 2015).

#### 1.2.2.1. Diseases caused by HPV

HPV are mainly known for causing cervical cancers. HPV is also responsible for several other cancers in the anogenital region. These cancers are caused by mucosal HPV types, such as the well-known HPV-16 and -18 (Munoz et al., 2003). Cutaneous HPV types, such as HPV-2, are responsible for causing warts on the skin (predominantly arms, hands, chest and feet), which are known as common or cutaneous warts. These warts can be spread via skin to skin contact or through contact with contaminated objects, through minor breaks in the epidermal barrier (Cubie, 2013).

Mucosal HPVs are separated into two groups: low-risk (LR) and high-risk (HR) HPV. This nomenclature is used to describe the oncogenic potential of the different HPV types. LR-HPV are mainly associated with external genital warts, with the most common being HPV-6 and -11. External genital warts are the most common sexual transmitted disease with a transmission rate of around 60% between partners (Cubie, 2013). High-risk HPV types, such as HPV-16 and -18 have a strong epidemiologic association with cervical cancer (Munoz et al., 2003, Fakhry and Gillison, 2006). The nomenclature of high-risk and low-risk HPV is used for oropharyngeal squamous cell carcinoma (OPSCC) as well, as high risk types are associated with an oncogenic potential, and cause several different types of cancer (Ang et al., 2010, Shaikh et al., 2017). Furthermore, high risk types can cause various anogenital carcinomas or pre-cancers, such as vulval, vagina, anal, penile cancers (Munoz et al., 2003). Associated pre-cancers are found as intraepithelial neoplasia and are named after the site where they are found e.g. cervical intraepithelial neoplasia (CIN). The most common cancer type associated with HPV is cervical cancer. Cancers at the vaginal, vulval or penile region are less common, which might be because of the lack of a cellular transformation zone, where squamous epithelia merge with columnar epithelia (Cubie, 2013).

HPV is responsible for around 99% of cervical cancers worldwide. HPV-16 is responsible for around 50% of invasive cervical cancer, followed by HPV-18 with around 20%. The two most common HPV-types alone are responsible for around 70% of cervical cancer. The remaining HPV cervical cases are caused by other HR HPV types (Walboomers et al., 1999, Munoz et al., 2003). A meta-analysis of HPV prevalence in 2009 investigated HPV prevalence in vulval intraepithelial neoplasia (VIN), vaginal intraepithelial neoplasia (VAIN) and anal intraepithelial neoplasia (AIN) in 93 studies across four continents. The overall HPV prevalence was identified for VIN as 84% and 40.4% for vulvar carcinomas, VAIN as 93.6% and 69.9% vaginal carcinomas and AIN as 92.7% resulting in 84.3% anal carcinomas (De Vuyst et al., 2009).

### 1.2.3. Classification/Taxonomy of papillomaviruses

Papillomaviruses (PV) are included in the family of "Papillomaviridae", a family of small, non-enveloped viruses with double-stranded DNA genomes of 5748 to 8607 bp (Van Doorslaer et al., 2018). Originally PVs were grouped together with the polyomaviruses as "Papovaviridae", as they have similar non-enveloped capsids structure and both have double-stranded DNA. Later discoveries revealed that these viruses differ in genome structure and size as well as possessing different structural and non-structural proteins (de Villiers et al., 2004). The viruses were hence split into two separate families known as "Papillomaviridae" and "Polyomaviridae" (de Villiers, 2013, de Villiers et al., 2004). The new order of the two viruses was officially recognized by the International Committee on the Taxonomy of Viruses (ICTV) (Van Doorslaer et al., 2018).

Papillomaviruses are classified according to the DNA sequence of the L1 open reading frame (ORF). Subfamilies of PV share less than 45% sequence identity. Papillomaviridae are separated into two groups First and Second papillmoaviridae and grouped as genera and species (Van Doorslaer et al., 2018). Within the family of Papillomaviridae, a total number of 16 different genera were identified. Human papillomaviruses are grouped into 5 distinctive genera, known as Alpha-, Beta-, Gamma-, Mu- and Nu-papillomaviruses genera. These genera are further divided into groups and species. The species are then divided into "types" by a diversity of over 10%. Types can be further classified into "sub-types" by a 2-10 % difference at the L1 ORF. Sub-types can be further divided into variants with less than 2 % differences at the L1 ORF region. For example, the most

famous HPV type 16, belongs to the Alpha-papillomaviruses, species 9 (de Villiers, 2013, de Villiers et al., 2004, Chen et al., 2011).

### 1.2.4. HPV genome structure

The Human Papillomaviruses are double stranded DNA viruses of around 8000 base pairs (Doorbar et al., 2015). The HPV genome is divided into a long-control region (LCR)/upstream regulatory region (URR), a late and an early viral region, with a total number of 8 ORFs (Figure 1-2). The LCR region is found between the early (E) and late (L) region genes and is involved in the control of gene expression, replication and transcription of the genome. The L region then encodes for L1 and L2 major and minor capsid proteins (Ramqvist and Dalianis, 2011). One characteristic of L1 is to spontaneously assemble into a 72-pentamer icosahedral structure, either together with L2 or alone. L2 is not required for capsid formation, but has an essential role in encapsidation of the viral genome and plays an important role in the entry pathway of viral infection. The ability of L1 to self-assemble into Virus Like Particles (VLP), is used in production of vaccines against HPV infections (Buck et al., 2008, Harper et al., 2006, Longworth and Laimins, 2004, Villa et al., 2006).

The early region contains 6 ORFs, known as E1, E2, E4, E5, E6 and E7. The ORFs of E5, E6 and E7 are involved in and modulate the transformation process, whereas E1 and E2, known as regulatory proteins, control replication and transcription (Carson and Khan, 2006, de Villiers et al., 2004). The early gene region contains two important promotor sequences. In HPV-16 the early promotor is known as p97 (known as p105 for HPV-18) and is involved in the transcription of E6 and E7 and other viral genes. The p670 promotor (situated in the E7 ORF) which is also involved in differentiation, is the late HPV-16 promoter, which results in the expression of L1 and L2 genes (Longworth and Laimins, 2004, Carson and Khan, 2006).

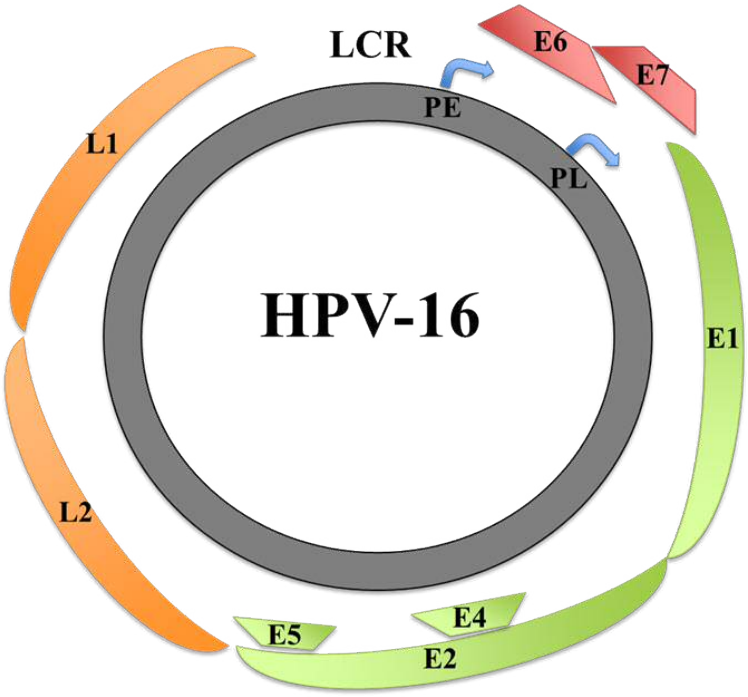



**Figure 1-2: HPV-16 genome structure**

HPV-16 has 6 early genes (viral replication), consisting of E1, E2, E4, E5, E6, E7 and two late genes L1 and L2 (capsid formation). The early gene region contains of two promotors, an early promotor (PE) p97 and a late promotor (PL) p670. The LCR region is responsible for gene regulation.

### 1.2.5. Transmission and rates of infection by HPV

HPV is one of the most common sexually transmitted viruses. The risk of infection increases with number of sexual partners and most sexually active people will be infected with HPV at some point in their lifetime. Most of these HPV infections are transient. Those HPV infections, which are persistent could result in development of neoplasia (Moscicki et al., 2010). The risk of HPV infection in men and women is strongly connected to the number of vaginal intercourse and oral sex partners during a lifetime. A high number of lifetime sex partners (26 or more vaginal-sex partners, 6 or more oral sex partners) are associated with an increased risk of oropharyngeal cancers (D'Souza et al., 2007, D'Souza et al., 2010, Burke et al., 2014). Studies have detected a higher number of HPV infections and OPSCC in men compared to women. This may be related to the higher number of sexual partners reported by males compared to women (Burke et al., 2014). Similar findings have been found for the incidence of oral sex reported among men and women in several countries (Gillison et al., 2015).

Some studies suggest a possibility of non-sexual transmission of HPV, through physical but non-sexual contact or via vertical transmission from mother to child (e.g. vaginal delivery) (Syrjanen, 2010a, Ryndock and Meyers, 2014). Other studies suggest a possibility of infection through contaminated medical equipment. However, these routes are only likely to be responsible for a very small proportion of infections.

### 1.2.6. HPV life cycle – Cell entry to assembly

The main target for HPV infection is the epithelium, specifically the stratified epithelium. The cells attached to the basal membrane in the stratified epithelium are normally the only cells of that epithelium with the potential for proliferation. During proliferation the newly formed daughter cell (after division of the basal cell) detaches from the basal membrane. The daughter cell then starts its differentiation process, exits the cell cycle, proceeds to terminal differentiation and peels off the epithelium. The HPV lifecycle is tightly connected to the differentiation process and is regulated by the host cells (Kajitani et al., 2012). The stages of the lifecycle including entry into the host cell, maintenance stage, proliferation, amplification and final assembly will be described.

#### 1.2.6.1. Cell entry

Viral entry mainly occurs through damaged areas of the epithelium, resulting in infection of the basal cells of the host. However, the mechanism for the viral entry, in terms of receptor of HPV infection has not been fully understood or characterized. The established model of viral entry suggests the requirement for interaction of virus particles with the basal lamina, and the interaction of capsids (late proteins L1 and L2) with heparin sulphate proteoglycans (HSPG) and possibly laminin (laminin-5) (Doorbar et al., 2012). Within the HSPG, syndecans (primarily syndecan-1) are predominant in epithelial cells, the target cells of HPV binding. Syndecan-1 serves as the primary attachment receptor, due to high expression level in the target cells and the involvement in wound healing (Horvath et al., 2010). Binding to the HPSG receptors induces a conformational change in L1, which leads to the exposure of a highly conserved furin convertase recognition site on L2 N-terminus. This exposure leads to a furin mediated cleavage of the L2 N-terminus region. (Day et al., 2008, Horvath et al., 2010, Sapp and Bienkowska-Haba, 2009). Several other HPSGs are implicated to be required for successful entry of HPV into the host cell, such as alpha-6 integrin (Doorbar et al., 2012, Horvath et al., 2010)

Following the binding on the cell surface, the virus gains entry into the host cell via endocytosis (Schelhaas et al., 2012, Allen et al., 2010). After endocytosis, escape of viral DNA from the endosomal compartments, appears to be facilitated by an L2 membrane motif and gamma-secretase. A L2-HPV genome complex is then transferred through the cytoplasm to the nucleus, where transcription of RNA is initiated (Raff et al., 2013), while the L1 protein is retained in the endosome and becomes a subject to lysosomal degradation (Doorbar et al., 2012).

#### 1.2.6.2. Maintenance phase & Viral replication

After entry into the host cells, HPV maintains its genome at a relatively low copy number of approximately 50-100 copies as nuclear plasmids (episomal form) with low expression of early genes (Flores et al., 1999). The number of genomes can vary from lesion to lesion and between different sites. This state is described as a non-productive infectious state (Stubenrauch and Laimins, 1999, Doorbar et al., 2012, Kajitani et al., 2012).

Within the second maintenance phase, the viral genome replicates at a constant copy number in synchrony with the cellular DNA in dividing cells. The continuously dividing

basal cells, function as perfect reservoir for virus during the differentiation process of the epithelium.

The transcription of the expressed E genes is controlled by the p97 promoter, which is located in the LCR region (Kammer et al., 2000). After activation of p97 in the early stage of infection, E1 and E2 are expressed (Wilson et al., 2005). These early gene products are important replication factors of HPV. Both proteins direct the host replication factors to the viral origin of replication, resulting in an increase of viral genome copies per cell (Kadaja et al., 2009). E1 is the primary replication factor that binds to the origin of replication, and exhibits helicase activity, initiating the unwinding or separation of the viral DNA strands/DNA double helix (Lin et al., 2002, Longworth and Laimins, 2004). E2 functions in combination with E1, binding to the recognition site of the LCR next to the E1 binding site (Sedman and Stenlund, 1998). Furthermore, E2 is involved in productive infection, as during cell division (mitosis), E2 ensures efficient segregation of viral genomes into the daughter cells (Van Tine et al., 2004).

Another role of E2 is the positive and negative regulation of the early promoter p97, causing a self-regulation of p97. Low level expression of E2 leads to a positive regulation of the early promotor, whereas a high level expression of E2 causes a negative regulation of p97. Thus differential regulation of E2 can lead to a stable viral copy number in undifferentiated cells and a stable level of the oncogenes of E6 and E7 (Longworth and Laimins, 2004, Steger and Corbach, 1997).

#### 1.2.6.3. Proliferation and amplification

In normal epithelia, basal cells undergo a differentiation process after exiting the cell cycle and migrating into the suprabasal layer. During this process, the newly formed daughter cells leaves the basal membrane after division of the basal cells (Kajitani et al., 2012, Madison, 2003). During HPV infection, cell cycle regulation is disrupted by the expression of E6 and E7, leading to a delay in terminal differentiation (Doorbar, 2005). The E6/E7 proteins are key players in driving cell proliferation in the basal and parabasal cell layers. The proliferation promoting functions of E6 and E7 are linked to the expansion in lesion size (Doorbar et al., 2012).

The main function of E7 is down-regulation of the Retinoblastoma protein (pRb), degrading both the p105 and p107 forms of pRb, which control cell cycle entry in the

basal layer, as well as p130 which is involved in cell cycle re-entry in the upper epithelial layers. Another key role of E7 is to interact with proteins involved in cell proliferation, such as p21 and p27 (Doorbar, 2005). The primary function of E6 is preventing apoptosis by degradation or down-regulation of p53. E6 also maintains telomere integrity during repeated cell division (Doorbar et al., 2012, zur Hausen, 2002, Klingelhutz et al., 1996).

Amplification of the viral genome occurs in the upper epithelium, and is necessary to produce higher number of viral particles in the host cells (Middleton et al., 2003). The expression of E6 and E7 in the upper epithelium, allows infected cells to re-enter S-phase and for the viral genome copy number to increase. During the late stage of viral amplification, the promoter p670 is activated and this leads to up-regulation of E1 and E2 (viral replication), which results in a higher copy number (Doorbar et al., 2012, Flores et al., 1999).

Next to E1 and E2, E4 and E5 indirectly contribute to genome amplification. It is thought that E5 is involved in genome amplification, due to its capacity to stabilize epidermal growth factor receptor, as well as enhancing epidermal growth factor signalling and mitogen-activated protein (MAP) kinase activity and to modulate extracellular-signal-regulated kinase 1/2 (ERK1/2) (Doorbar et al., 2015).

#### 1.2.6.4. Virion assembly

In the final process of HPV infection the viral genome is encapsidated in a protein coat composed of L1 and L2 proteins, then transported to the cell surface. Expression of L1, requires a change in splice sites usage. This leads to an increase of transcription that initiates at the late promoter e.g. p670 in HPV-16, which terminates at the late polyadenylation site. This process is facilitated by high expression of E2 that down-regulates p97. The change in splice sites leads to initiation of genome packaging. Prior to the expression of L1, final encapsidation requires L2 to be recruited by E2. L2 then assembles the virions. E4 is then involved in organisation of virions that disrupts the keratin. This allows the release of virions from the cell surface to allow infection of other cells (Doorbar et al., 2015).

### 1.2.7. Molecular mechanisms of HPV mediated carcinogenesis in HNSCC

Human papillomaviruses are one of the main causes of development of pre-neoplastic and malignant lesions in cervical and head and neck cancer (Narisawa-Saito and Kiyono, 2007, Rampias et al., 2014). The two oncoproteins E6 and E7 play an important role in cellular transformation and formation of cancer, representing significant characteristics of the "hallmarks of cancer" (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). The individual molecular functions of E6 and E7 in carcinogenesis are explained in detail below. The joint effect of the two HPV oncogenes E6 and E7 leads to unregulated cell cycle progression, which results in genomic instability and accumulation of mutations in the host cell. The simultaneous functions of E6/E7 can lead to incorrect chromosome numbers and aberrant mitotic figures (Duensing et al., 2001, Duensing et al., 2000). These induced genomic instabilities facilitate viral integration into the host genome. This occurs mostly at fragile sites in the human genome. Variation can be seen in the location of integration sites, and they are not connected to a specific sequence or site (Thorland et al., 2003). Viral integration may be a significant event in carcinogenesis as it prevents stochastic loss of HPV episomes and deregulates viral gene expression (Pett and Coleman, 2007).

#### 1.2.7.1. Functions of E7 in carcinogenesis

E7 is one of the most important early genes in HR-HPVs. It can target many cellular proteins and dysregulate the cell cycle. Its association with members of the pRb family is well characterized; E7 protein binds and inactivates Rb and associated family members including p103 and p107. Rb is a negative regulator of the cell cycle and prevents S-phase entry by associating with E2F family members (Doorbar, 2005). Through binding and inactivating hypo-phosphorylated Rb, E7 releases E2F, which then has the ability to promote cell cycle progression (Figure 1-3 and Figure 1-4) through activation of many other cell cycle genes like cyclin A and E (Duensing and Munger, 2004). This results in G1/S checkpoint dysregulation. In normal cells phosphorylation of Rb and E2F activity leads to upregulation of p16 which allows the control of kinase activity (cyclin D/CDK4-6) by a negative feedback loop. Since p16 expression is normally repressed, reduction or loss of the hypophosphorylated Rb protein can result in p16 overexpression. The interaction of E7 with Rb, leads to release of active E2F in the absence of CDK4/6 (Boulet et al., 2007, Giarre et al., 2001, Doorbar, 2005, Hanahan and Weinberg, 2000, Hanahan

and Weinberg, 2011). The expression of cyclin E is absolutely necessary for S-phase entry, and is expressed during natural infection as a result of E7 expression and disruption of the E2F/pRb complex. High levels of cyclin dependent kinase inhibitors in differentiated cells can lead to formation of inactive complexes that contain E7, cyclin E/CDK2 and either p21 or p27. These CDKs and the inhibitors are important for cell cycle regulation. During normal cell cycle regulation p21 and p27 are involved in inhibition of G1/S progression. P21 and p27 get inactivated or neutralized by binding of E7, leading to high CDK activity and unregulated G1/S progression (Doorbar, 2005).

#### 1.2.7.2. Functions of E6 in carcinogenesis

Another important function of E6 is the maintenance of telomerase activity, which is known as another key factor of the hallmarks of cancer (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). The maintenance of telomerase activity is highly dependent on activation of the human telomerase reverse transcriptase (hTERT) promoter, which controls levels of the telomerase catalytic subunit. Studies have revealed that E6 mediated activation of the hTERT promoter requires the coordination of an E6/E6-AP complex for engagement of activation. Additionally, activation is dependent on Myc binding sites in the hTERT promoter (Liu et al., 2005, Liu et al., 2008). Secondly hTERT can be activated by degradation of p53 by E6 (Oh et al., 2001). This combined disruption of pRb (E7) and telomerase activity leads to immortalized keratinocytes and cancer formation (Figure 1-3 and Figure 1-4) (Oh et al., 2001).
Beside stabilising telomerase activity, E6 has several other roles in the development of carcinogenesis. Activation of E2F by E7 leads to induction of p14ARF, which stabilizes p53 by inhibiting Mdm2. One of the main functions of E6 is to abrogate the anti-proliferation response, by inducing the degradation of p53 during proliferation (Duensing and Munger, 2004). This degradation occurs after binding of E6-associated protein (E6-AP), which marks p53 for proteosomal degradation. Loss of p53 subsequently facilitates the accumulation of mutations and consequent genomic instability (zur Hausen, 2002) (Figure 1-3 and Figure 1-4).

The capacity of E6 to induce malignancy also comes from its interaction with proteins containing PDZ domains. These proteins are involved in regulation of cell attachment, cell proliferation and cell signalling (Pim et al., 2012, Thomas et al., 2008). The PDZ binding motif is absent from the E6 of low-risk HPV types and thereby represents a

molecular signature for the oncogenic potential in high risk HPV types (Rampias et al., 2014).

Enabling replicative immortality (telomere activity), sustaining proliferative signalling, degradation of pRb blunts the oncogene-induced senescence response (triggered by E7), degradation of pRb serves to evade growth suppressors and resist cell death (p53 degradation by E6) are some important “hallmarks of cancer” which can be found in the above mentioned effects of HPV (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011)

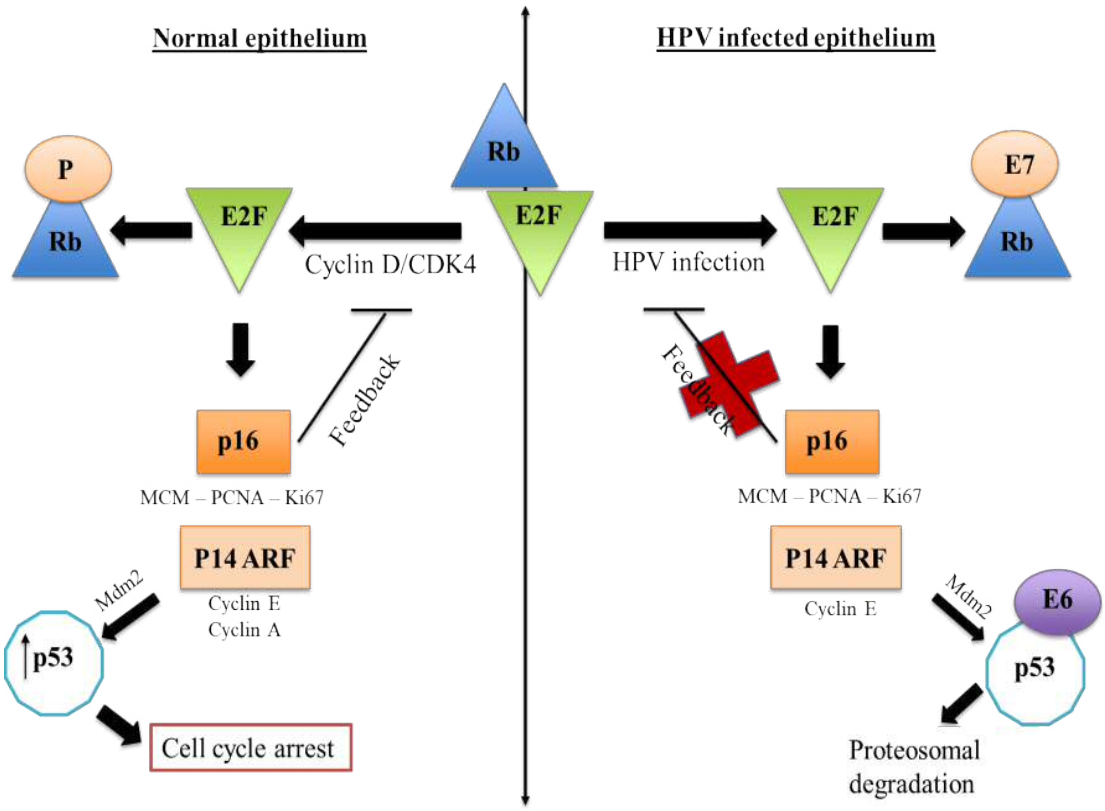



**Figure 1-3: Normal epithelium compared to HPV infected epithelium**

The G1/S cell cycle checkpoint is regulated by the Rb tumour suppressor protein. Compared to HPV infected epithelium E2F is only released in normal epithelium when Rb is phosphorylated by cyclin D/CDK4. A negative feedback loop by E2F is activated, leading to increased expression of p16 which inhibits cyclin D/CDK4. This leads to reduced activity of E2F. Additionally p14-ARF inhibits Mdm2-mediated degradation of p53, which leads to p53 promoted cell cycle arrest. During HPV infection, Rb binds to E7, which causes the release of E2F transcription factor. The negative feedback loops are ineffective as release of E2F is not coordinated by cyclin D/CDK4.

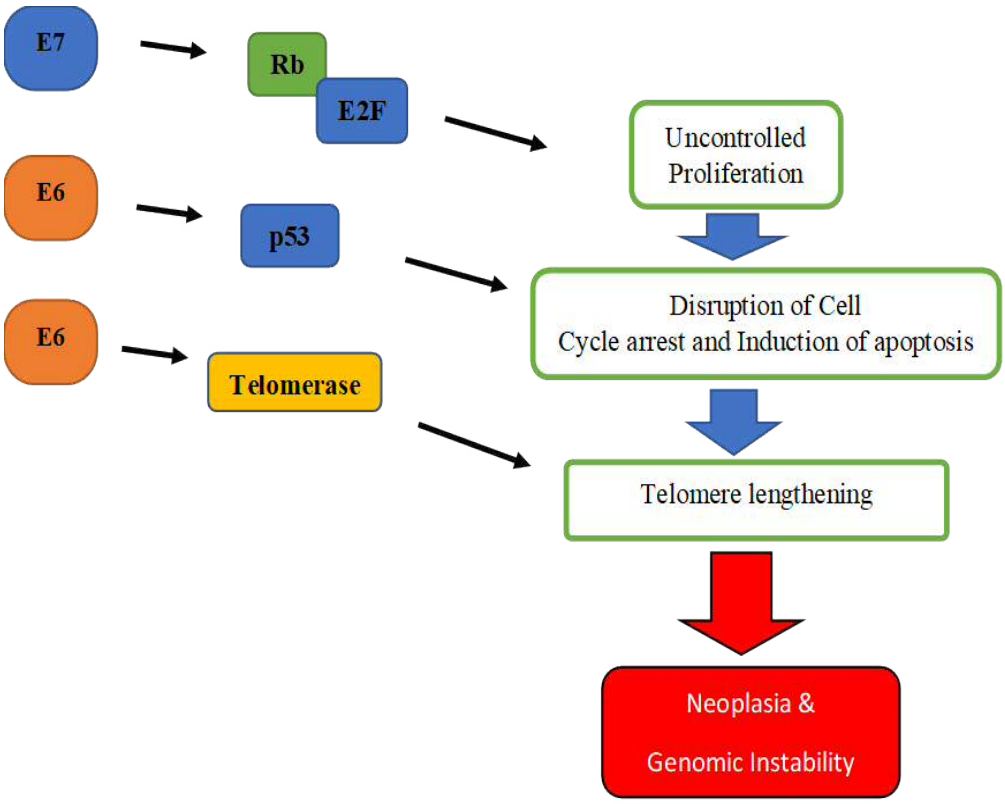



**Figure 1-4: Combined effects of E6 and E7**

E7 downregulates Rb, which results in uncontrolled proliferation. This effect would usually result in p53-mediated cell cycle arrest, but this is prevented by p53 degradation by E6. Another function of E6 is increased telomerase activity. Normally telomeres would shorten due to uncontrolled proliferation and would induce replicative senescence, but this is prevented by E6 mediated increases in telomerase activity.

### 1.2.8. Prevention of HPV infection

Once the infectious cause of cervical cancer was established, prophylactic vaccination to prevent cervical cancers became a possibility. There were two different prophylactic vaccines available before 2014, Gardasil® (Merck, USA) and Cervarix® (GlaxoSmithKline, UK). Both vaccines are based on virus like particles (VLPs), formed by self-assembled L1 protein, together with an adjuvant. Gardasil is a quadrivalent vaccine, protecting against four HPV types (6, 11, 16 and 18) and contains an adjuvant of amorphous aluminium hydroxyphosphate sulfate. Cervarix is a bivalent vaccine which only protects against two HPV types (16 and 18). The adjuvant for the bivalent vaccine is AS04 (aluminium hydroxide and 3-deacylated monophosphoryl lipid A) (Harper et al., 2006, Villa et al., 2006). Both vaccines protect against HPV type 16 and 18, which are responsible for around 70% of cervical cancer. Since 2016 Merck has made a new vaccine available for the public known as Gardasil-9, which covers 9 HPV types (31, 33, 45, 52 and 58) instead of the previous 4 types.

#### 1.2.8.1. Vaccine regulation worldwide

By 2014 according to the World Health Organisation (WHO), HPV vaccination programs were implemented in around 58 (30%) countries with the focus on vaccinating girls (WHO, 2014, WHO, 2015). The number of countries implementing national HPV vaccination programs had increased to 86 countries (40%) by the end of October 2016 (Gallagher et al., 2018). Inclusion of boys in HPV vaccination programs is limited worldwide, which is suggested to be tightly connected to an economic cost factor (Bogaards et al., 2015). Several countries including Australia, Austria, USA and Canada have switched to gender neutral vaccination programmes (Prue, 2014, Elfstrom et al., 2016). The government in England will introduce a vaccination programme for HPV for boys aged 12 and 13 by autumn 2019 (Kmietowicz, 2018).

## 1.3. Natural history of HPV infection in the oropharynx

The connection between HPV and oropharyngeal cancer was established recently and therefore the natural history of HPV infection in the oropharynx is still not well defined. The most common area for OPSCC to develop seems to be the tonsillar crypts (Begum et al., 2005). This means that screening protocols to detect cancerous and pre-cancerous cells, such as for cervical cancer, are most likely to be impractical, due to the difficulty of detection caused by the location of pre-malignant and early stage lesions deep within tonsillar crypts (Franceschi et al., 2015).

HPV infections in the tonsil tissue are unexpected, due to the immunologically exposed environment. However, the specialized epithelium lining the tonsillar crypts might enable HPV infection. Specifically, the epithelial cells of the reticulated epithelium are not a robust barrier as they consist of a mix of infiltrating lymphocytes and epithelial cells. These characteristics of the tonsil tissue may enable virus infection of the epithelial stem cells (Egawa et al., 2015). In general people with persistent HPV infection in the oral cavity, are at higher risk of developing oropharyngeal squamous cell carcinoma. The observed latency period for OPSCC development have been predicted with approximately 10-30 years (Gillison et al., 2015).

OPSCC cases are higher in men, with a 4-fold higher occurrence when compared to women. This observation is found independently of the HPV status of OPSCC. Within the HPV-negative cases this has been tightly connected to the tobacco and alcohol consumption (Evans et al., 2013, Schache et al., 2016). The reason this difference is also present with the HPV-positive tumours is less clear. However, it might be related to the higher prevalence of oral HR-HPV infections within men or the tendency for men to have a higher average number of sexual partners compared to women (Gillison et al., 2012b, D'Souza et al., 2014) .

### 1.3.1. Differences between HPV-positive and HPV-negative tumours

Epidemiologically, patients with HPV-positive OPSCC tend to be younger, fitter and have no or limited history of tobacco and alcohol use, compared to the majority of HPV-negative patients (Deschler et al., 2014). Clinically, HPV-positive tumours respond better to treatment and have better long-term survival (Evans et al., 2013). A recent study aimed to classify OPSCC's based on different histological features classified the histological appearance of HPV-positive OPSCC. The new classification was based on three histologically categories as non-keratinizing cell carcinomas (NKSCC), keratinizing (KSCC) and a hybrid appearance and determined the frequency of HPV in these subsets of tumours. Non-keratinizing cell carcinomas were significantly more likely to be HPV and p16 positive compared to keratinizing OPSCC's ($p<0.001$) with a better disease specific-survival (NKSCC: $p=0.0002$ and KSCC $p=0.0142$).

The standard clinical tumour classification is done according to the tumour-node-metastasis (TNM) staging system (Hermans and Lenz, 1996, Patel and Shah, 2005). It describes the tumour size, involvement of lymph nodes and distant metastasis. Tumours are then classified accordingly, Tx – T4 (Tumour size), Nx – N3 (Nodal stage) and Mx – M1 (Metastasis stage) (Patel and Shah, 2005). The average tumour size at diagnosis tends to be smaller in HPV-positive cases but histologically represents a more advanced stage of neck disease, compared to HPV-negative cases. A similar difference is seen at the nodal stage, with a higher proportion in nodal metastasis at the time of diagnosis (Deschler et al., 2014, Gillison et al., 2012b, Goldenberg et al., 2008).

Differences in mutation pattern have been observed between HPV-positive and HPV-negative HNSCC (Stransky et al., 2011). In a study cohort of 92 HNSCC patients, 74 samples were subjected to whole-exome sequencing of tumour-normal pairs. Half the amount of mutations was found in HPV-positive samples compared to the HPV-negative cohort with a mean of 2.29 mutations/mega base (Mb) to 4.83 mutations/Mb ($p=0.004$). This study confirmed the expected pattern of TP53 mutation between the groups. TP53 mutations are common within HPV-negative HNSCC and OPSCC tumours (Stransky et al., 2011). The Catalogue of Somatic Mutations in Cancer (COSMIC), is a database to explore the impact of somatic mutations on different types of human cancer. However, this tool did not show any mutations for human papillomavirus connected to an associated cancer, only of the associated cancer itself. Cosmic provides the data of common mutations of head and neck cancers, including OPSCC, where TP53 (32%) is the most

common mutated gene, followed by CDKN2A (11%), NOTCH1 (9%) and PIK3CA (7%). All other mutated genes showed a frequency of less than 5% (Data obtained on the 4.2.2019 from https://cancer.sanger.ac.uk).

TP53 mutations in HPV-positive HNSCC are uncommon as a result of the interaction with the oncoprotein E6 during viral integration and replication. TP53 mutations in HNSCC are generally associated with a poor prognosis and lower survival rate (Lindenbergh-van der Plas et al., 2011).

#### 1.3.1.1. Post-transcriptional modifications/regulations in the p53 network

In general, TP53/p53 is of high importance, as it is involved in regulation of the synthesis of mRNA encoding proteins, involved in cellular stress response (cell cycle arrest, apoptosis and senescence). Some of these mRNAs are concurrently regulated at the post-transcriptional level by miRNAs and RNA-binding proteins (RBPs), which can modify the p53 transcriptional program in a cell type and stimulus specific manner (Freeman and Espinosa, 2013).

E3 ubiquitin ligase MDM2 is one of the main factors in post-transcriptional regulation of p53, with the function of targeting p53 for degradation. This interaction with MDM2 controls the level of p53. Several miRNAs are capable of targeting MDM2, which can diminish the amount of MDM2 produced in response to p53 activity, which leads to increased p53 activity in cells. High expression of all these miRNAs can cause a positive feedback loop, which causes p53 to upregulate both its own repressor and multiple negative regulators to enhance its own expression levels and activity (Freeman and Espinosa, 2013, Kubbutat et al., 1997, Momand et al., 1992, Oliner et al., 1993).

In addition to the action of miRNAs, the p53 transcriptional program is also co-regulated, both negatively and positively, by RNA-binding proteins (RBPs) that control the stability and translation of mRNAs induced by p53. p21 (CDKN1A). Furthermore, some RBPs amplify the signal by providing enhanced mRNA transcript (Freeman and Espinosa, 2013, Scoumanne et al., 2011). Another function of p53 to mention is the interaction with 14-3-3σ. This leads to repression of CDC25 transcription, which leads to a p53 induced converging signal, involving transcriptional and post-transcriptional regulation that provokes G2/M arrest (Freeman and Espinosa, 2013, St Clair et al., 2004).

### 1.3.2. Detection of OPSCC and HPV testing

Early diagnosis of OPSCC is a key factor for improved prognosis and treatment. However, it is important to consider the factors associated with late stage diagnosis. In OPSCC, common symptoms such as hoarseness, dysphagia, otalgia and odynophagia can lead to early detection. Oropharyngeal cancer can also be detected in the mouth through routine dental examinations. Lack of routine dental visits can often lead to a late diagnosis (Chen et al., 2007). High-quality imaging is used to determine the stage of the disease, by revealing submucosal tumour spread and detecting subclinical adenopathies (Hermans and Lenz, 1996, Patel and Shah, 2005).

Within OPSCC disease, it is important to define the HPV status of each tumour, as in general, HPV-positive OPSCC patients respond better to treatment and have a better prognosis (Evans et al., 2013). A known biomarker for HPV-positivity is p16 because it is generally upregulated in HPV-positive tumours, due to cell cycle dysregulation by E7 and E6. Immunohistochemistry (IHC) for p16 is routinely performed for OPSCC patients. Besides p16 expression HPV can be identified in tumour tissue through DNA and RNA testing. One of the main goals of HPV detection is to identify if HPV is the driving force behind a certain disease. A standard method to identify HPV within OPSCC and other HPV-driven cancers, is to confirm the presence of the two main oncoproteins E6 and E7, which are the two main factors for carcinogenic transformation within the host (Schache et al., 2011a, Schache et al., 2011b, Wiest et al., 2002).

Common molecular HPV detection methods are: Nucleic acid hybridization assays (e.g. Southern blot or In-situ hybridisation), Signal amplification systems (Hybrid Capture systems) or a target amplification systems known as nucleic acid amplification assays (e.g. PCR, Real-time-PCR, reverse transcriptase-PCR) (Molijn et al., 2005).

For subsequent molecular biological test and diagnostic purposes, tissue samples from extracted biopsies are fixed in formalin and embedded in paraffin (FFPE). This method of storage has been used for decades and can be used for RNA, DNA or protein testing. A disadvantage of FFPE, is the potential negative long-term storage effect on the quality of RNA. RNA isolation from FFPE was shown to have a negative effect on PCR, RT-PCR and cDNA testing. After a year the RNA integrity number (RIN) quality tested for RNA was between 5-6, which would make it unsuitable for genotyping purposes (von Ahlfen et al., 2007). High quality RNA should have a RIN of 8 or above.

As described above, p16 is the most well-known biomarker for HPV-positivity. However, p16 as a surrogate biomarker lacks specificity, as p16 can be upregulated or present in a small percentage (14-17%) of HPV-negative tumours as well (Lewis et al., 2010, Rietbergen et al., 2014). Nevertheless, p16 IHC testing of FFPE slides is still the most common technique to indicate the HPV status of tumours.

One procedure to capture all HPV types and variants present in a biological specimen is DNA sequencing of the viral genome, either after cloning into a plasmid or via polymerase chain reaction (PCR) fragment, for single HPV detection or detection of multiple HPV types (Villa and Denny, 2006).

HPV DNA can be detected by PCR or *in situ* hybridisation (ISH). Different available methods can either detect the presence of a group of HPV types compared to other methods detecting HPV-16 alone. Most of the techniques used are either PCR and Enzyme Linked Immunosorbent Assay (ELISA) based (Robinson et al., 2012). However, tests for p16 or DNA based tests as a standalone test, do not reach the specificity and sensitivity of testing via mRNA detection (RNA qPCR). A combination of these or several other tests would provide a more reliable result (Schache et al., 2011a, Smeets et al., 2007). A study on clinical tests has presented the combinations of p16/HPV DNA by ISH with 88% sensitivity and 90% specificity or p16/HPV DNA qPCR with a 97% sensitivity and 94% specificity as tests with high prognostic values (Schache et al., 2011a).

## 1.4. Treatment of OPSCC

History has shown that patients with OPSCC (HPV-positive and HPV-negative) were often treated when the tumours were large, as smaller tumours did not trigger any symptoms and were rarely detected. Traditional treatment of OPSCC included open surgery which involves large facial and neck incision (Nichols et al., 2013b). Next to surgery, radiotherapy was the standard option for OPSCC treatment in the past. Often a combination of surgery and radiotherapy was used (Parsons et al., 2002, Ramqvist and Dalianis, 2011). Open surgery and radiotherapy (RT) were seen to be equally effective in terms of survival (Nichols et al., 2013b). New approaches in robotic surgery, e.g. trans-oral robotic surgery (TORS), allow the surgeons to overcome visualization and access challenges and avoid incisions in the face and neck. Recent studies have shown that combined treatments of chemo-radiotherapy (CRT), usually with cisplatin, have a better

survival outcome for patients. However, the survival improvements come with increased acute and late toxicities, which can include dysphagia, mucositis, fibrosis, osteoradionecrosis, or neurotoxicity. The combined radio- and chemotherapy also prolongs risk of long-term gastrostomy tube dependency (Nichols et al., 2013b).

### 1.4.1. Ionising radiation

The radiation therapy of cancer typically uses ionising photon beams, X-rays or gamma-ray for local or regional treatment. The cellular damage, measured in Gray (Gy), increases with the amount of absorbed radiation dose. It is defined as the amount of energy absorbed relative to a volume of a tissue (Trikalinos et al., 2009). The radiation used is called Ionising Radiation (IR), as it is capable of forming ions and deposits energy in the cells of the tissue it moves through. The stored energy in the cells is able to kill the cancer cells by damaging DNA, which results in cell death (Baskar et al., 2012).

Caesium-137 (Cs-137) is a common radioactive source used for a wide range of radiotherapy. Cs-137 damages tissue and cells via free radicals. Gamma radiation is used in cancer therapy for its DNA damaging abilities. In general Cs-137 decays indirectly via beta radiation with a 93.5% likelihood to metastable Barium 137 ($^{137m}$Ba) and further decays via gamma radiation into the stable nuclide Barium 137 ($^{137}$Ba). (Lahtz et al., 2012).

### 1.4.2. Radiotherapy for HNSCC/OPSCC in the UK

For head and neck squamous cell carcinoma, including OPSCC, Intensity-Modulated Radiotherapy (IMRT) is the standard therapy for patients undergoing primary or adjuvant radiotherapy (Ho et al., 2009). Due to improvements in dosimetry and imaging techniques, IMRT enables precise delivery of a high radiation dose to the tumour, but a minimal dose to the surrounding tissue. This technique allows more normal tissue to be spared, which reduces the risk of late toxicity (Taylor and Powell, 2004). Internationally, and in some centres in the UK, the definitive treatment dose remains 70Gy in daily fractions of 2Gy over seven weeks. Other centres within the UK use altered fractionation regimens of 65-66Gy in 30 fractions (2 -2.2Gy) over six weeks for HNSCC (Ho et al., 2009).

Combined treatment with cisplatin and radiotherapy is the current standard therapy for stage III and IV patients <70 years of age. As a combined treatment the RT dose is defined as 70Gy in 35 fractions of 2Gy international. However, within the UK a more moderate hypo-fractionated schedule of 65-66Gy in 30 fractions has been incorporated into UK trials and centres (Pignon et al., 2009, Nutting et al., 2011). Current treatment for HPV-positive and HPV-negative OPSCC patients is the same in the UK. The aim of a current clinical trial, PATHOS (Owadally et al., 2015), is to reduce the intensive adjuvant treatment after minimally invasive trans-oral surgery, due to the fact that HPV-positive OPSCC have a more favourable prognosis. HPV-positive patients are included in the PATHOS study, if showing a cancer stage of T1-T3 N0-N2b with a resectable primary tumour via a trans-oral approach. HPV-positivity is tested on biopsy specimens by p16 immunohistochemistry and high risk HPV in-situ hybridisation. These patients are then classified into three pathological risk groups: Patients in the low-risk pathology group receive no adjuvant treatment, as in standard practice. Patients in the intermediate-risk pathology group are randomised to receive either standard dose post-operative radiotherapy (control) or reduced dose radiotherapy. Patients in the high-risk pathology group will either receive either post-operative chemo-radiotherapy (control) or radiotherapy alone. The goal of the study is to standardize the transoral surgery and post-operative intensity-modulated radiotherapy protocols in the UK. Additionally, long-term dysphagia could be decreased by reducing adjuvant therapy without causing a negative clinical outcome. (Owadally et al., 2015).

### 1.4.3. Induction of DNA damage by IR

DNA damage constantly occurs in the human body. These DNA lesions can influence the processes of genome replication and transcription. Incorrect repair or no repair at all, can lead to mutations or causes a wide range of genome aberrations that threatens cell viability. DNA lesions within a cell can occur via endogenous (spontaneous) factors, such as DNA mismatches during DNA replication, hydrolysis, oxidation or alkylation. Secondly, DNA damage can be caused be exogenous (environmental) factors, such as ultraviolet radiation (UV), ionising radiation and various chemical agents (Friedberg et al., 2004, Jackson and Bartek, 2009). To deal with DNA damage, cells have evolved a wide variety of DNA repair mechanisms including: Base excision repair (BER), Mismatch repair (MMR), Nucleotide excision repair (NER), Non-homologous end-joining (NHEJ) and Homologous recombination (HR). These mechanisms are involved

in repair of single strand breaks (SSB) or double strand breaks (DSB) (Jackson and Bartek, 2009).

Ionising radiation, as used in radiotherapy, is the main cause of DSB. The mechanisms of DNA damage caused by IR, will be explained in this section. IR causes different forms of DNA damage, the most toxic are double-strand breaks (DSB) of the DNA (Figure 1-5). This damage can be caused directly or indirectly to DNA molecules. The direct action hits the DNA molecule. For indirect action, radiation hits the water molecules present in cells, which leads to the production of free radicals. These free radicals (unpaired electrons) react with the DNA molecules, causing molecular structural damage. The number of free radicals produced depends on the dose of IR (Desouky et al., 2015).

The DSB induced by IR carry 3' blocking ends (e.g. 3'-phosphate or 3'-phosphoglycolate moieties), and have single-stranded overhangs of variable length. Furthermore, complex induced DSB consist of a high number of oxidised base modifications and abasic (AP) sites directly adjacent of the DSB breaks. In transcriptionally active DNA the damage increases in yield and becomes more complex with IR dose (Lomax et al., 2013).

In response to IR induced DNA damage, several cellular responses are initiated (Figure 1-5). First the MRN (MRE11/RAD50/NBS1) complex binds to DSBs and activates ATM, a key PI3K related kinase in the DNA damage response (DDR). This leads to ATM auto-phosphorylation, inducing activation and phosphorylation of a wide range of substrates in the surrounding chromatin. One of the ATM substrates is a H2A histone variant called γH2AX, which when phosphorylated is one of the first responders in DSB signalling. Phosphorylation is amplified through recruitment of MDC1. Through recruitment of MDC1, several other proteins such as BRCA1 or 53BP1 are engaged as well, triggered by γH2AX. The overall signalling pathway causes downstream phosphorylation of substrates including CHK2 and p53, triggering cell cycle arrest in G1/S and or G2/M, allowing sufficient time for DNA repair (Vignard et al., 2013, Lomax et al., 2013).

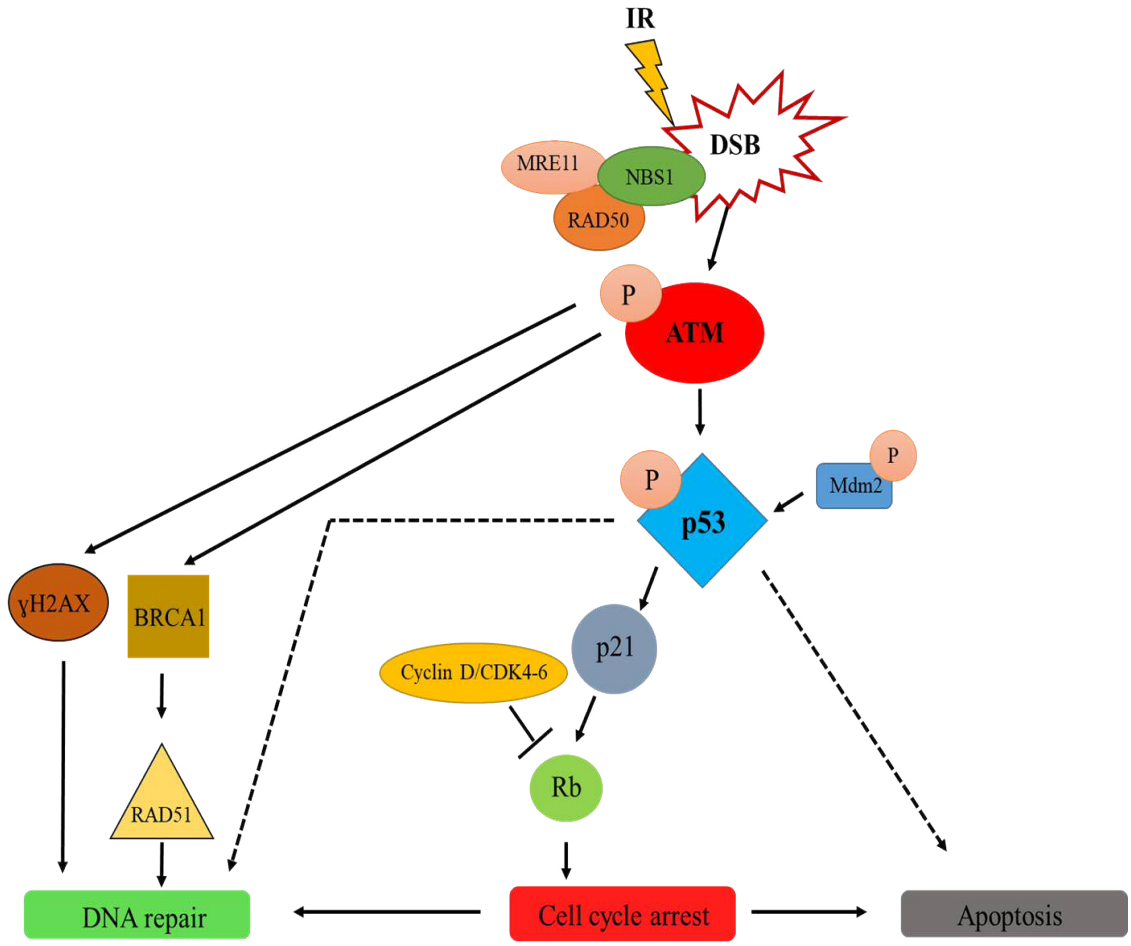



**Figure 1-5: DNA damage response pathway after IR**

One of the most toxic forms of DNA damage by IR are DSBs. The MRN-complex (MRE1-NSB1-RAD50) actives ATM. This causes auto-phosphorylation of ATM, including other surrounding key factors, such as p53, and H2AX. Binding of Mdm2 to p53, induces apoptosis. Through phosphorylation of p53, other substrates are activated in response to DNA damage, such as p21 and Rb, which are involved in cell cycle regulation. This overexpression of Rb, leads to cell cycle arrest. During cell cycle arrest the decision will be made for DNA repair or Apoptosis of the damaged cells. The main DNA repair pathways involved after IR induced DNA damage are HR and NHEJ.

### 1.4.4. DNA repair following IR

DNA double strand breaks are the most common damage caused by IR, which is usually repaired through DSB repair mechanisms, including homologous recombination (HR) and non-homologous end-joining (NHEJ). Other forms of DNA damage are repaired by BER, NER or MMR and are recognized as repair mechanism for single stranded DNA damage (Jackson and Bartek, 2009). The MMR functions mainly on base-base mismatches and is relevant mainly to mis-pairs occurring during DNA replication and recombination (Li, 2008, Jackson and Bartek, 2009). NER recognizes helix-distorting base lesions, and operates via two sub-pathways known as global genomic NER (GG-NER) and transcription coupled NER (TC-NER). TC-NER focuses on the template strand of actively transcribed genes which are connected to actions of RNA molecules down the template strand during transcription. GG-NER addresses the non-template strand of transcribed genes and non-transcribed regions of the genome. After recognition of damage it cleaves the damaged strand upstream and downstream of the damage, yielding a single-strand fragment of 20-30 base nucleotides in length, which is then removed. The gaps are filled by DNA polymerase and associated factor before ligation (Jackson and Bartek, 2009, Weinberg, 2007).

In BER, the damaged or chemically altered base is removed after recognition by a DNA glycosylase enzyme, which cleaves the glycosyl bond linking the altered base and the deoxyribose. The base-free deoxyribosylphosphate is then excised by the enzyme apurinic/apyrimidinic endonuclease (APE). Polymerase is induced to fill the gaps and the repair is completed by ligase proteins (Jackson and Bartek, 2009, Weinberg, 2007).

After induction of DSB by IR, the two main pathways for repair are homologous recombination (HR) and non-homologous end-joining (NHEJ). How it is determined within a cell after DSB, whether HR or NHEJ will be used is still under debate. It is more likely that NHEJ is used outside of S/G2 cell cycle phase, if a homologue is not present near a DSB during cell cycle S/G2 transition. However, if cells are in S-phase and a sister chromatid is close (homology donor), HR will be used (Lieber, 2010). During the NHEJ mechanisms, DSB are recognized by the Ku protein, that binds and activates the protein kinase DNA-PKcs, leading to recruitment and activation of several end-processing enzymes (e.g. X-ray complementing protein 4, XRCC4-like factor, Aprataxin-and-PNK-like factor and DNA ligase IV) to DSBs (Davis and Chen, 2013, Jackson and Bartek, 2009).

HR requires a wide number of proteins in its multistep process, which operates at the S and G2 cell cycle phase. Several sub-pathways exist for HR. However, HR is always induced by ssDNA generation, promoted by a group of proteins including the MRE11-RAD50-NBS1 (MRN) complex. Through RAD51, BRCA1 and BRCA2 the ssDNA invades the undamaged template, which is followed by actions of polymerases, nucleases, helicases and other components. The final step involves DNA ligation and substrate resolution (Jackson and Bartek, 2009).

## 1.5. HPV-positive and HPV-negative OPSCC prognosis

The incidence of OPSCC cases is increasing and HPV is one of the main causes. Several studies have shown that HPV-positive OPSCC patients respond better to treatment compared to HPV-negative OPSCC patients. This better response to treatment is observed with several modalities including surgery followed by RT, CRT or as a stand-alone treatment. In a study in South-Wales, the data indicated a significantly better overall survival for HPV-positive patients with 82.6% survival after 3 years, and 75.4% after 5 years with compared to HPV-negative patients with 32.2% and 25.3%. This corresponds to a reduction of death rate of 78% in HPV-positive cases (HR 0.220, 95% CI; 0.132-0.366, $p<0.001$) (Evans et al., 2013). In a meta-analysis in Canada, a similar range of overall survival was observed with a three year survival rate in HPV-positive patients of 85% versus 49% in HPV-negative patients (Nichols et al., 2013a).

Similar data was obtained from a study in the United States with a 3-year survival rate of 82.4% in HPV-positive patients and 57.1% in HPV-negative patients ($p<0.001$ by the log-rank test). This study further estimated a 9% difference in the overall survival within the HPV-positive OPSCC cases. Death in both HPV-positive and HPV-negative groups increased with each additional pack-year of tobacco smoking (Ang et al., 2010). The adverse influence of smoking on the overall survival of OPSCC patients beside HPV status has been reported in several other studies (Gillison et al., 2012b).

### 1.5.1. Possible reasons for improved prognosis of HPV-positive OPSCC/HNSCC

HPV-positive OPSCC patients respond to treatment better than HPV-negative OPSCC patients. The better responses might be associated with a reduced capacity to repair DNA damage in HPV-positive tumours. During the HPV lifecycle, ATM, ATR, BRCA1, RAD51 or RPA are recruited to viral DNA during genome amplification (Gillespie et al., 2012). Due to the recruitment of DNA repair factors during the viral lifecycle, HPV-positive cells might be deficient in DNA repair.

Deficiency of DNA repair and increased IR sensitivity were suggested by two previous studies using small panels of HNSCC cell lines. These two studies have suggested that HPV-positive cell lines are in general, less able to repair DSB induced by IR. The persistence of DSB after induction of DNA damage correlated with greater radio-sensitivity in HPV-positive cell lines (Kimple et al., 2013, Rieckmann et al., 2013). Both studies agreed that due to the accumulation of DSB after IR, cells fail to progress through the G2/M cell cycle checkpoint, which resulted in a higher number of cells accumulating in G2 after IR (Kimple et al., 2013, Rieckmann et al., 2013). However, the events following G2 arrest are less well understood. One study suggested that accumulation of cells at G2, would give cells sufficient time for DNA repair, before progression through the cell cycle (Rieckmann et al., 2013). However, an extended time course, to investigate a possible decrease of cells in G2, was not performed.

The improved treatment response could be closely connected to the tumour suppressor protein p53. Incomplete loss in HPV-positive tumours might play a role. Even if E6 causes ubiquitination and proteosomal degradation of p53, re-accumulation of p53 may be possible. One of the studies observed that HPV-positive cell lines showed increased levels of p53 and p-p53 following IR, which was associated with apoptosis (Kimple et al., 2013).

Some studies have also suggested the influence of an immune target in better treatment response of HPV-positive OPSCC patients. It has been shown that HPV-positive OPSCC tumours have a higher number of tumour infiltrating lymphocytes (TILs) relative to HPV-negative OPSCC, and furthermore showed an increase in $CD8^+$ cells within HPV-positive tumours (Nasman et al., 2017, Ward et al., 2014).

In summary, several factors could influence the better treatment response of HPV-positive OPSCC patients, and these include deficiency in DNA repair, incomplete degradation and re-accumulation of p53 or an immune response against HPV.

### 1.5.2. Importance of *in-vitro* models

Advances in genomic research have opened new possibilities for translational research and direct evaluation of clinical samples. However, there is still a need for reliable preclinical models to test and improve therapeutic strategies. Currently, human cancer cells lines are the most widely used models to study the biology of cancer and to improve efficacy of treatment. The importance of cell line models is, that according to current knowledge they mimic some of the typical characteristics of tumours, e.g. they have similar genetics and molecular biology to the original tumour. They also have many advantages relating to convenience of use, being a potentially inexhaustible resource that allow investigations to be conducted in controlled conditions and with appropriate repetition. This makes cell lines ideal to test novel or current treatments, as they may retain their sensitivity to relevant target agents, and can be important to identify and characterize target molecules. They also allow study of certain sensitivity and resistance responses (Gillet et al., 2013). Therefore, cell line models are a potentially useful tool to investigate HNSCC and OPSCCs.

The study of OPSCC is limited due to the small number of HPV-positive OPSCC cell lines. A higher number of cell lines is needed, and these should mimic the typical characteristics of HPV-positive OPSCC patients. HPV-positive patients tend to be younger, fitter and are less likely to smoke than HPV-negative patients (Evans et al., 2013, Schache et al., 2016). One study reported that HPV-positive patients had a mean age of 57.4 years compared to HPV-negative patients with a mean age 61.4 years (Schache et al., 2016).

In 2014, only eight published HPV-positive OPSCC cell lines were available, including UMSCC-47, UPCI-SCC-90, UMSCC-104, 93-VU-147T, UD-SCC-2, UT-SCC-45, UPCI-SCC-152 andUPCI-SCC-154. The majority of these cell lines were derived from smokers or former smokers. The vast majority of HNSCC cell lines are HPV-negative cell lines. Some of the HPV-negative HNSCC cell lines were also derived from tumour receiving chemotherapy or radiotherapy, such as UMSCC-4 (derived from tumour receiving post-treatment of 2 chemotherapies) or UMSCC-74a (originated from a

recurrent tumour – after receiving chemotherapy) (Arenz et al., 2014, Bradford et al., 2003, Brennan et al., 1995, Brenner et al., 2010, Descamps et al., 2014, Ferris et al., 2005, Kimple et al., 2013, Lin et al., 2007, Rieckmann et al., 2013, Steenbergen et al., 1995, Tang et al., 2012). The influence of chemotherapy on resistance to treatment was investigated in a study by Di Nicolantonio *et al.* (2005), which showed a slightly increase in resistance in biopsies of various cancer types post-chemotherapy (Di Nicolantonio et al., 2005). Therefore, chemo-resistance could be observed in *in-vitro* experiments.

The number of available OPSCC HPV-positive has since increased to 10 cell lines, because two novel HPV-positive OPSCC cell lines (CU-OP-2 and CU-OP-3) were successfully derived in a previous study in Cardiff (Pirotte, 2017). It is important to develop further new HPV-positive to support research in the field of head and neck cancer, and to study the treatment responses of “typical” head and neck patients (HPV-positive, non-smokers and young)

## 1.6. Aims and hypotheses

According to the literature, there is a shortage of available *in-vitro* models of HPV-positive OPSCC cell lines. This reduces the potential for a complete biological understanding of this disease. It also limits the attempts to understand the molecular mechanisms that might explain the better treatment response of HPV-positive patients. Therefore, a major aim of this thesis was to derive new HPV-positive and HPV-negative OPSCC cell lines and fully characterize them. The new derived cell lines could be then incorporated into a larger panel of cell lines and used to investigate the biology of HPV-positive OPSCC relative to HPV-negative OPSCC.

The investigation of the biology of HPV-positive or HPV-negative cell lines was based on the expected different response to ionising radiation. Therefore, this study focused on identifying differences in survival after IR treatment between the HPV-positive and HPV-negative OPSCC cell lines. This was then followed by investigating p53 accumulation/activation, cell cycle abnormalities and gene expression changes. The current study did not specifically measure DNA damage/repair after each individual experiment. The idea was to identify the individual effect of cell lines in survival, p53 accumulation, cell cycle abnormalities and gene expression changes after IR. Possible changes in gene expression in DNA repair was started to be investigated during basic mRNA sequencing analysis. Future studies on DNA damage and DNA repair will specify the direct involvement in the collected data of the current study.

The study aimed to test the following main hypotheses:

1. Sensitivity will be correlated to the HPV status. Therefore, cell lines derived from HPV-positive OPSCC will be more sensitive to ionising radiation than HPV-negative OPSCC cell lines.
2. HPV-positive cell lines will show accumulation of p53 following ionising radiation, in correlation with accumulation of p21 and p-p53.
3. HPV-positive OPSCC cell lines will show greater variations in cell cycle arrest (e.g. G2 arrest) compared to the HPV-negative OPSCC cell lines
4. Difference in gene expression between HPV-positive and HPV-negative cell lines after IR will be observed.
    a. Differences in gene expression between untreated and treated samples after IR will be observed.

# Chapter 2 : Materials and methods

## 2.1. Cell culture

This section contains information about the cell lines, reagents and culture conditions used in the current study.

### 2.1.1. Established cell lines

#### 2.1.1.1. 3T3 mouse fibroblast

In 1963 the mouse embryo fibroblast cell line J2 3T3 was established by George Todaro and Howard Green at the New York University (Todaro and Green, 1963). This became the standard fibroblast cell line for use in many experimental studies.

Stocks of 3T3 mouse fibroblast (J2 3T3 feeder cells) were received as kind gift from Dr. Sally Roberts (University of Birmingham, UK). They were used in this project to support the growth of Primary Culture Oropharyngeal Cells (PCOC) (known officially as Cardiff University Oropharyngeal (CU-OP) cells) and to supplement growing cells at low cell density during the clonogenic assays (radio-sensitivity assay).

#### 2.1.1.2. HeLa cell line

The HeLa cell line was derived from a cervical cancer in a 31 year-old African-American woman. It is a HPV-18 positive cell line (Scherer et al., 1953). The cell line was received from the American Type Culture Collection, short ATCC (cat. Number: CCL-2).

#### 2.1.1.3. SiHa cell line

SiHa was derived from a cervical squamous cell carcinoma in a 55-year-old Asian woman (Friedl et al., 1970). This cell line is HPV-16 positive and was received from the ATCC (cat. Number HTB-35).

#### 2.1.1.4. CaSki cell line

The CaSki cell line was derived from an epidermoid carcinoma metastasised to the small bowel in a 40 year-old Caucasian woman (Pattillo et al., 1977). CaSki cells contain an integrated HPV-16 genome (about 600 copies per cell). Cells were obtained from the European Collection of Authenticated Cell Cultures, short ECACC (cat. Number: 87020501).

#### 2.1.1.5. UPCI-SCC-90 cell line

The UPCI-SCC-90 cell line is derived from an oral squamous cell carcinoma, from a 44 year-old Caucasian male smoker (Ferris et al., 2005) and carries 100-150 copies of integrated HPV-16 genome. It was received from the Leibnitz-Institute DSMZ Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (cat. Number: ACC 670).

#### 2.1.1.6. UMSCC-47, -4, -6, -19 and -74a

UMSCC-47 (HPV-positive cell lines) and UMSCC-4, UMSCC-6, UMSCC-19 and UMSCC-74a (HPV-negative cell lines) were derived at the University of Michigan (USA) and obtained from Professor Thomas Carey.

UMSCC-47 is a HPV-positive cell line, derived from the oral cavity (tongue squamous cell carcinoma) in a 53-year-old man and carried integrated HPV-16 genome (Lin et al., 2007, Brenner et al., 2010).

UMSCC-4 is a HPV-negative cell line, derived from a tongue squamous cell carcinoma in a 47-year-old woman (Baker, 1985).

UMSCC-6 is a HPV-negative cell line, derived form a tongue squamous cell carcinoma in a 32-year-old male (Baker, 1985).

UMSCC-19 is a HPV-negative cell line, derived from a tongue squamous cell carcinoma in a 66-year-old male (Baker, 1985).

UMSCC-74a is a HPV negative cell line, derived from a tongue squamous cell carcinoma in a 50-year-old male (Takebayashi et al., 2004, Brenner et al., 2010).

### 2.1.2. Human epidermal keratinocytes

Primary human epidermal keratinocytes (HEKn), isolated from neonatal foreskin, were purchased from Thermo Fisher Scientific (Paisley, UK) (cat. Number: C0015C). This keratinocyte cell line was derived from normal (healthy) tissue and is HPV-negative. This cell line was used as a control in various assays including radio-sensitivity experiments, flow cytometry for cell cycle analysis and as a positive control in Western blotting. With the appropriate culture conditions, the cells can sustain around 30 population doublings (PD).

### 2.1.3. CU-OP-2, CU-OP-3, CU-OP-17 and CU-OP-20

CU-OP-2, -3, -17 and -20 were derived at the HPV Group at Cardiff University. CU-OP-17 and -20 were derived during this project. CU-OP-2 and CU-OP-3 were derived during an earlier project by Dr. Evelyne Pirotte. All were of oropharyngeal squamous cell carcinoma origin. CU-OP-2, -3 and -20 are HPV-16 positive cell lines, whereas CU-OP-17 is HPV-negative. All cell lines were derived from male patients. The development and characterisation of these lines is described in detail in section 2.1.6.

### 2.1.4. Cell culture material

#### 2.1.4.1. Culture vessels

Cell lines were grown in tissue culture flasks in a size range from T-25, T-75 to T-175 (Fisher Scientific, Loughborough, UK), or tissue culture plates in a size range from 60x15mm or 100x15 mm (VWR, Lutterworth, UK).

#### 2.1.4.2. Cell culture media

Supplements were added to the compositions of the different cell culture media. For the established cell lines (keratinocytes), Dulbecco's Modified Eagle Medium (DMEM) was used; for novel cell lines (CU-OP cell lines), Glasgow Minimum Essential Medium (GMEM) was used, and for normal human keratinocytes, EpiLife Media with 1% Human Keratinocyte Growth Supplement (HKGS). Supplements added to the media are listed in Table 2-1.

All components were prepared according to supplier's instructions. All components were aliquoted and stored at -20°C or 4°C and thawed at 37°C in a water bath before use.

**Table 2-1: Reagents used for cell culture media**

| Reagents | Supplier |
|---|---|
| Penicillin / Streptomycin (Pen/Strep) | Sigma-Aldrich, Gillingham, UK |
| Hydrocortisone | |
| Epidermal Growth Factor (EGF) | |
| Fetal Bovine Serum (FBS) | |
| Newborn Calf Serum (NCS) | |
| L-Glutamine | |
| Cholera Toxin | |
| Non-essential amino acids (NEAA) | |
| HKGS | Thermo Fisher Scientific |

#### 2.1.4.3. Dulbecco's modified eagle medium for developed keratinocytes and J2 3T3 Feeder cells

The standard Dulbecco's modified eagle medium (DMEM) media (Sigma-Aldrich D6429) was used for following cell lines: UMSCC-47, UPCI-SCC-090, UMSCC-4, UMSCC-6, UMSCC-19 and UMSCC-74a and J2 3T3 feeder cells (different composition listed in Table 2-2 and Table 2-3).

**Table 2-2: DMEM 500ml media for developed cell lines (keratinocytes)**

| Reagents | Concentration |
|---|---|
| NEAA | 0.2% |
| FBS | 10% |
| Penicillin | 100 U/ml |
| Streptomycin | 100 ng/ml |
| L-Glutamine | 2mM |

**Table 2-3: DMEM 500ml media for J2 3T3 Feeder cells**

| Reagents | Concentration |
|---|---|
| NCS | 10% |
| Penicillin | 100 U/ml |
| Streptomycin | 100 ng/ml |

#### 2.1.4.4. Glasgow minimum essential medium

The standard Glasgow minimum essential medium (GMEM) media (without Glutamine) (Sigma-Aldrich G5154) was used for culture of the keratinocytes cell lines of CU-OP-2, CU-OP-3, CU-OP-17 and CU-OP-20. Supplements added to the GMEM media are listed in Table 2-4. All GMEM media were set-up without EGF. EGF was added 24 hours after passaging a CU-OP cell line to support cell growth (Table 2-4).

**Table 2-4: GMEM 500ml media composition**

| Reagents | Concentration |
|---|---|
| Cholera Toxin (CT) | 0.1nM |
| Hydrocortisone (HC) | 1.5nM |
| Penicillin | 100U/ml |
| Streptomycin | 100ng/ml |
| L-Glutamine | 2mM |
| FBS | 10% |
| EGF | 10ng/ml |

#### 2.1.4.4.1. GMEM supplements

##### 2.1.4.4.1.1. Cholera toxin 100nM (1000x)

0.5 mg of Cholera Toxin was dissolved in 0.6mL distilled water. The dissolved solution was made up to 60 mL with GMEM media. 5mL aliquots were stored at 4°C.

##### 2.1.4.4.1.2. Epidermal growth factor (EGF) 1µg/ml (100 x)

0.1 mg of EGF was dissolved in 100 mL GMEM media. Aliquots of 5mL were stored at -20°C.

##### 2.1.4.4.1.3. Hydrocortisone 1mg/mL (2000x)

10 mg of Hydrocortisone was dissolved in 5 mL of 100 % absolute Ethanol (does not dissolve in media or distilled water properly). The dissolved solution was made up to a final volume of 5 mL with distilled water. Aliquots of 250µl were stored at -20°C.

##### 2.1.4.4.1.4. L-Glutamine 200mM (100x)

10 mL media to vial of Glutamine with a final stock concentration of 200 mM and was stored at 4°C.

### 2.1.4.5. Cell culture solutions

A ready-made 1xPhosphate Buffer Saline (PBS) (Gibco, UK) was used to wash cell layers, to dilute solutions and for washing cells after fixation before staining for cell cycle analysis.

To detach cells from the culture flask 1x Trypsin 0.05 % ethylenediaminetetraacetic acid (EDTA) 0.02 % (TE) was used. This was obtained by diluting a 10x TE stock solution (Sigma Aldrich) with 900mL of 1xPBS. TE was aliquoted and stored at -20°C.

### 2.1.5. General cell culture

#### 2.1.5.1. General cell culture conditions

All cell lines were cultured, according to standard cell culture procedure, in a humidified incubator at 5 % carbon dioxide ($CO_2$) and 37°C. Cell culture was performed in a class II safety cabinet using aseptic techniques. Surfaces were cleaned using 70% industrial methylated spirit 74 O.P. (IMS). Cell culture reagents were warmed up for 20-30min in a 37°C water bath prior to use. Liquid waste was inactivated in a sodium dichloroisocyanurate (NaDCC) solution (Presept, Johnson Johnson, USA) (diluted according to manufacturer's instructions). Plastic waste was discarded into yellow waste bags (infectious waste) and stipites and pipette tips collected in Bio-bins designed for infectious waste.

#### 2.1.5.2. Cell counting and viable cell count

Cell numbers were determined by using a Beckman Coulter Vi-cell®. For each count a 1:10 dilution was prepared (50 µl of sample + 450µl of PBS) in a sample cup and placed into the Vi-cell carousel. Sample name and dilution factor was entered into the Vi-cell software and the count started. The live camera option, gave the opportunity to double check, if the appropriate cells (size) were being counted. For each count Vi-cell automatically added Trypan blue for live/dead cell determination, which resulted in a final output of viable cells per millilitre.

#### 2.1.5.3. Freezing and storage

Cell lines were collected and counted using standard cell culture procedure. $2x10^6$ cells were frozen down per cell line in a prepared freezing media, suitable for storage at -80°C or $LN_2$ storage.

Cells were spun down and re-suspended in Freezing media and transferred to cryovials and placed in a Nalgene Mr Frosty freezing device. Tubes were stored at -80°C for 3-7 days and then transferred to liquid nitrogen ($LN_2$) for long-term storage.

##### 2.1.5.3.1. Freezing medium for established cell lines (keratinocytes)

Freezing medium for established immortal cell lines was 80% DMEM media, 10% FBS and 10% DMSO. For the following cell lines UMSCC-47, UPCI-SCC-90, UMSCC-19, UMSCC-4, UMSCC-6 and UMSCC74a, $2x10^6$ cells/ml were frozen. Information of freezing media composition for the established cell lines was provided by the University of Michigan.

##### 2.1.5.3.2. Freezing medium for CU-OP cell lines

Freezing medium for primary cell lines (derived from biopsy tissue at Cardiff University) consisted of 90 % FBS and 10 % DMSO. $2x10^6$ cells/ml were frozen.

#### 2.1.5.4. Thawing of cell lines

Rapid thawing of frozen cell lines was done by briefly placing frozen tubes in the water bath (37°C) until the contents were thawed. These were transferred to a sterile conical tube containing 9 mL of pre-warmed media (to dilute DMSO) then centrifuged for 5min at 180g. The cell pellet was re-suspended in 5mL of fresh media and added to the culture flask containing the appropriate volume of culture media for the flask size.

Established keratinocytes were seeded into a T25 flask containing 5mL DMEM media. HEKn cells were seeded at the defined cell density of $2x10^3$ cells/$cm^2$ into a T175 culture flask. The CU-OP cell lines were seeded into a 10cm tissue culture plate together with $2x10^6$ irradiated J2 3T3 feeder cells in 10 mL EGF-negative GMEM media. The media for the CU-OP cell lines was changed to GMEM EGF-positive after 24 h.

#### 2.1.5.5. Subculture

Cells were split once 80% confluence was reached in the tissue culture flask and re-seeded at a defined ratio (Table 2-5). The media was removed from the culture flask and the cell layer washed with a 1xPBS solution. Cell were then incubated with TE solution until 80-90% of the cells detached from the culture flask (time indicated in Table 2-5). The remaining cells were detached by carefully tapping the culture flask. With an equal amount of media, the TE was inactivated and transferred to a sterile universal container or 15mL/50mL Falcon tube. Cells were centrifuged for 5min at 180g and resuspended in the designated media for established cell lines (DMEM), CU-OP culture (GMEM) or HEKn (EpiLife media).

Established cell lines were split according to a defined ratio (Table 2-5). The appropriate amount of media was added to T25, T80 or T175 culture flasks. Due to their high cell-cell adhesion UPCI-SCC-90 were first incubated for 15-20min with 2mM PBS/EDTA then 5min with TE.

CU-OP subculture was performed in the presence of irradiated J2 3T3 feeder cells to support cell growth. Tissue culture plates were rinsed with PBS, followed with a 2min TE incubation to remove the remaining J2 3T3 feeder cells. Cells were washed again with PBS and incubated with TE (time listed in Table 2-5). CU-OP cells were counted and a cell density of 0.8 to $1x10^6$ plated per 10cm together with $2x10^6$ irradiated J2 3T3 feeder cells in a total of 10ml GMEM EGF-negative media. Media was then replaced every two days with GMEM EGF-positive media and irradiated J2 3T3 feeder cells replaced if cover was sparse (1-2 times per week).

HEKn cells were counted and $2x10^3$ cells/$cm^2$ seeded per T175 flask and 25mL EpiLife media was added. Every two days' media was changed until the flask was 50% confluent then every day until 80% confluent.

**Table 2-5: TE incubation time and split ratio of cell lines.**

| Cell line | TE incubation time (min) | Split ratio |
|---|---|---|
| UMSCC-47 | 10 | 1/3 |
| UPCI-SCC-90 | 5 | ½ |
| CU-OP-2 | 15 | -[A] |
| CU-OP-3 | 15 | |
| CU-OP-20 | 15 | |
| CU-OP-17 | 10 | |
| UMSCC-4 | 10 | 1/3 |
| UMSCC-6 | 10 | ½ |
| UMSCC-74a | 7 | ½ |
| UMSCC-19 | 10 | ½ |
| HEKn | 5 | $2x10^5$ cells/$cm^2$ |
| J2 3T3 feeder cells | 5 | -[A] |

[A] CU-OP cell lines and J2 3T3 feeder cells were re-seeded after defining the cell density by cell counting

### 2.1.6. Primary Culture of Oropharyngeal Cells

#### 2.1.6.1. 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 National Institute for Social Care and Health Research (NISCHR) portfolio. Subsequently any newly developed primary cell lines were changed to the official nomenclature of CU-OP – Cardiff University Oropharyngeal Cancer). Biopsy samples from PCOC-1 to PCOC-12 for the PCOC-study were handled by Dr Evelyne Pirotte, biopsy samples PCOC-13 to PCOC-23 were managed by Stefan Holzhauser.

Patients for this study were recruited and consented by the clinical team (Dr. Evans, Mr. Al-Husseini and Mr Owens), during the head and neck cancer clinics at Velindre Hospital and Heath Hospital in Cardiff. Patients were recruited at different stages of their disease, either at the beginning of the treatment or at the beginning of diagnosis. During surgery an extra biopsy was collected for the PCOC-study during panendoscopy and placed in the provided transport media (for more details of the PCOC study see the designated protocol attached in Appendix 1). The process of handling the biopsy is shown in Figure 2-1.

After receiving the biopsy sample, the transport media was discarded and the biopsy sample placed in the tissue culture plate. The sample was cut in half, one half was used for cell culture and the other half prepared for nucleic acid extraction. The tissue was

washed 3 times with a mix of PBS/Pen-strep and 1x/amphotericin B 5.6mg/l (Sigma Aldrich). The sample was then incubated for 15min in the last wash.

The tissue sample was covered in a thin layer of FBS. The biopsy was placed on a new tissue culture plate and cut into small pieces (≤ 2 mm), using scalpels or curved scissors. Additional FBS was added during this process to prevent the tissues drying out. The small fragments of the biopsy sample were evenly distributed over several 10cm tissue culture plates (number of plates depended on the size of the biopsy) and incubated for 2 hours at 37°C with 5% carbon dioxide in an inverted position. After incubation 10mL of GMEM EGF-negative media containing $2x10^6$ irradiated J2 3T3 cells was added slowly to the plate to avoid detaching the fragments. The plate was then returned to the incubator.

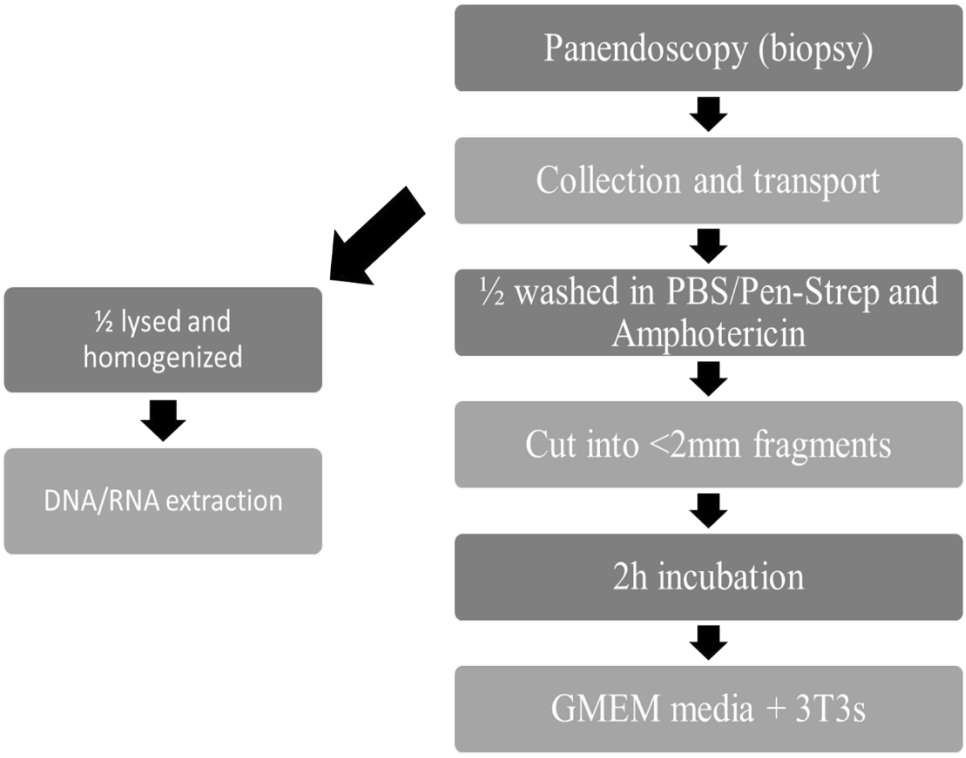



**Figure 2-1: CU-OP (PCOC) biopsy processing**
After surgery, biopsies were collected and placed in transport media until processed. Biopsy was cut into two pieces. One half was lysed and prepared for DNA/RNA extraction. The other half was used for cell culture, washed with PBS/Pen-strep Amphotericin, then cut into small pieces and placed into a FBS coated tissue culture plate. This plate was inverted and incubated for 2h. After 2h incubation GMEM EGF-negative media containing J2 3T3 feeder cells were added.

#### 2.1.6.2. Early cell culture

After two days GMEM EGF-negative media was taken off and switched to GMEM EGF-positive, this was subsequently changed every two days. Irradiated J2 3T3 feeder cells were replaced twice a week, when coverage in the tissue culture plate became sparse. Photographs were taken at the early stage of cell culture using a Zeiss Axiovert 35M inverted bright field microscope and camera to record the development of cell outgrowth (Zeiss, Cambridge, UK). Once some colonies were visible (outgrowth of biopsy) with a diameter of around 3-4mm, cells were passaged the first time. TE incubation times were slowly defined by increasing the incubation time every 5min (Incubation time for early passage of different biopsies were 5-10min). It was important to detach all cells during the first passage. The TE cell suspension was placed in a tube and inactivated with media. Fresh TE was added to the remaining cells in the tissue culture plate. This process was repeated until all cells were detached. The total cell suspension was then centrifuged and the cell pellet re-suspended in fresh GMEM EGF-negative media. The cells were placed into a new tissue culture plate ($2x10^6$ J2 3T3 feeder cells were added) and labelled with passage 1. New GMEM EGF-negative media with $2x10^6$ irradiated 3T3 feeder cells was added to the plates containing biopsy pieces still attached. This process was repeated until biopsy outgrowth stopped.

Early stage primary cultures were passaged once colonies were 50-70% confluent. Cells were frozen down at every passage if enough cells were available, otherwise cells were placed back into a new tissue culture plate and culture continued until a consistent growth was obtained.

### 2.1.7. Culture of J2 3T3 feeder cells (support growth for primary culture)

#### 2.1.7.1. Thawing and seeding of J2 3T3

J2 3T3 feeder cells were thawed and plated as previously described into a T25 flask containing 5ml of DMEM media. After the cells reached 80% confluence they were transferred into a T80 flask (15ml media), then into several T175 flask until 80% confluent.

#### 2.1.7.2. Subculture of 3T3 feeder cells

J2 3T3 feeder cells were passaged until they reached 80% confluence in a T175 flask. Cells were harvested (5-8min TE), counted and a density of 0.5-1x$10^6$ cells re-seeded in T175 flasks containing 25mL DMEM media. The collected J2 3T3 feeder cells were placed into a sterile universal tube and stored at 4°C until irradiation with 60Gy.

#### 2.1.7.3. Irradiation

The J2 3T3 cells were irradiated using a Gammacell® 1000 with a caesium-137 source until a dose of 60Gy was achieved (Best Theratronics, Ottawa, Canada). Irradiation keeps 3T3 cells in a post mitotic stage, which means that they are unable to divide, but remain capable of secretion of extra cellular matrix proteins. Irradiated J2 3T3 feeder cells were used immediately or stored for a maximum of 1 week at 4°C.

### 2.1.8. Mycoplasma detection

#### 2.1.8.1. Reagents preparation and storage

A standardized PCR assay was used to periodically test all cell lines in culture for mycoplasma contamination. This assay was performed by using the Venor®GeM detection kit (Minerva Biolabs, Berlin, Germany), which is specific to the highly conserved 16s rRNA coding region which allows for the detection of a wide range of mycoplasma species.

The reagents were rehydrated using PCR grade water according to manufactures instructions and stored below -18°C. As recommended by the manufacturer, Internal Control DNA and Positive Control DNA were stored in aliquots to avoid repeated freezing and thawing.

#### 2.1.8.2. Sample collection, preparation and storage

Supernatants (500µl) of all the test cultures were collected and transferred to a 1.5mL microfuge tube. 100µl of the cell culture supernatant was transferred to a new microfuge tube and incubated at 95°C for 10 minutes. The remaining supernatant was stored at -20°C. After 10min samples were briefly centrifuged (~5 s at approximately 8000g) to pellet cellular debris and 2µl of supernatant was added to the PCR reaction mix (prepared

in a designated PCR hood under sterile conditions). Ingredients of the PCR reaction mix are shown in Table 2-6.

23µl of this mix was transferred to each tube and 2µl of sample, negative control or positive control were added to each designated tube. DNA polymerase (5 U/µl) was ordered separately as it was not included in the provided kit (Minerva Biolabs, Berlin, Germany). The amplification of the reaction is described in Table 2-7 using a Nexus GSX1 thermal cycler (Eppendorf, Stevenage, UK).

**Table 2-6: Mycoplasma reaction mix**

| Reagents | Volume n=1 (µl) |
|---|---|
| PCR grade water | 15.3 |
| 10x reaction buffer | 2.5 |
| Primer/nucleotide mix | 2.5 |
| Internal control | 2.5 |
| DNA polymerase | 0.2 |
| Total Volume | 23 |

**Table 2-7: Mycoplasma PCR conditions**

| Step | Temperature (°C) | Time | Cycles |
|---|---|---|---|
| Initial denaturation | 94 | 2min | 1 |
| Denaturation | 94 | 30sec | 39 |
| Primer annealing | 55 | 30sec | |
| Extension | 72 | 30sec | |
| Final extension | 72 | 30sec | 1 |

### 2.1.8.3. Gel electrophoresis and gel interpretation

A 1.5% standard agarose gel was made including ethidium bromide (2µl/100mL gel). A detailed description of Agarose gel preparation is shown in section 2.2.7.1.

The electrophoresis was run for 20min at 100V (2cm run distance) and the gel observed under UV light. The success of the reaction was measured according to the size of internal control 191bp, negative control 191bp and positive control 267bp.

## 2.2. Nucleic acid extraction

This section describes DNA and RNA extractions from biopsy samples and cell lines. The working area and equipment was cleaned with RNaseZap (Sigma Aldrich) to eliminate RNase which could degrade the RNA sample, before any RNA extraction process.

### 2.2.1. DNA extraction from cultured cell lines

For all DNA extractions the QIAamp DNA Mini Kit was used (Qiagen, Manchester, UK). This kit contains ready-to-use proteinase K solution (storage 2-8°C), lysis buffer AL, two washing buffers AW1 and AW2, and elution buffer AE. Buffers AW1 and AW2 were prepared according to company's instructions. $2x10^6$ to $5x10^6$ cells were used to perform the DNA extractions. All extractions were carried out at room temperature (15–25°C). Cells were collected as previously described. For DNA extractions the cell pellet was re-suspended in a final volume of 200µl and stored at -20°C.

Blanks were prepared in between each sample. Therefore, 200µl of sterile water were added to each tube to match the volume of the sample. The blanks were treated the same way as the samples. 20µl Qiagen proteinase K were added to each tube and mixed properly. Lysis buffer was added to each tube (200µl Buffer AL) and mixed by pulse-vortexing for 15 seconds. This allowed thorough mixing to ensure efficient lysis. The samples were incubated at 56°C on a preheated heating block for a minimum of 10 minutes. Tubes were briefly centrifuged to remove drops from the inside of the lid. 200µl ethanol (96–100%) were added to each sample, and mixed again by pulse-vortexing for 15 seconds. After mixing, the 1.5 ml micro centrifuge tube were briefly centrifuged to remove drops from the inside of the lid.

For the next step 600µl the mixture was carefully added to the QIAamp Mini spin column, which was provided with the kit (in a 2ml collection tube) without wetting the rim and centrifuged at 8000g for 1min. The QIAamp Mini spin column was placed in a clean 2 mL collection tube (provided). The contained filtrate was discarded.

The QIAamp Mini spin column was carefully opened and 500µl Buffer AW1 were added without wetting the rim. The cap was closed and centrifuged at 8000g for 1min. The QIAamp Mini spin column was placed in a clean 2 mL collection tube, and the filtrate

was discarded. Next 500µl Buffer AW2 were added to the QIAamp Mini spin column without wetting the rim. The cap was closed and centrifuged at 20,000g for 1min.

To help to eliminate the chance of possible Buffer AW2 carry-over the QIAamp Mini spin column was placed in a new 2ml collection tube and centrifuged at 10,000g speed for 3min.

The QIAamp Mini spin column was placed in a clean 1.5ml micro centrifuge tube, and the filtrate was discarded. 200µl Buffer AE was carefully added directly on the filter of the QIAamp Mini spin column. It was incubated at room temperature (15–25°C) for 5min, and then centrifuged at 6000g for 1min. The eluate (final DNA) was stored at -20°C.

### 2.2.2. RNA extraction from cell lines

RNA was extracted from untreated and irradiated (2Gy) cell line samples. Irradiated samples were harvested after 24h and RNA extraction performed as described below.

For all RNA extractions the QIAamp RNA Mini Kit was used (Qiagen). This kit contains ready-to-use RLT buffer, DNase I (550µl RNase-free water added) and RDD buffer, RW1 and RPE buffer (96-100% ethanol added).

Cells were harvested from all cell lines in culture ($2x10^6$ cells) and the defined volume of RLT lysis buffer was added (350µl), according to manufacturer's instructions. 350µl of 70% ethanol was added to the lysate and mixed well by pipetting. 700µl of sample, including any precipitate was transferred to an RNeasy Mini spin column (placed in a 2ml collection tube) and centrifuged for 15 seconds at 8000g and the flow-through discarded.

Steps for DNA digestion:

350µl of RW1 buffer was added to each tube and centrifuged for 15 seconds at 8000g and the flow-through discarded. 10µl DNase I was mixed with 70µl buffer RDD and the 80µl added to each tube and incubated at room temperature for 15min. After 15minutes, 350µl of RW1 was added to each tube and centrifuged at ≥8000g and the flow-through discarded.

700µl RW1 buffer was added to the spin column and centrifuge at 8000g for 15sec and the flow-through discarded. Then 500µl of RPE buffer was added, centrifuge at 8000g for 15sec. This step was repeated with 500µl RPE buffer and centrifuged at 8000g for 2 minutes. The RNeasy spin column was placed into a new 1.5 collection tube and 50µl RNase-free water added directly to the spin column membrane and centrifuged for 1min

at 8000g to elute RNA. This step was repeated by adding the eluted 50µl again to the membrane to increase the yield of RNA.

RNA concentration was measured using spectrophotometry and RNA was stored at -80°C until needed.

### 2.2.3. Genomic DNA and total RNA extraction from oropharyngeal cancer biopsies.

DNA and RNA were extracted simultaneously by a method used combining QIAamp DNA Mini kit and RNeasy fibrous tissue kit (Qiagen).

One half of the collected biopsy was stored for nucleic acid extraction. This half was homogenised using a TissueRuptor (Qiagen) for 10-20sec in 600µl RLT buffer containing 6µl of 14.3 M β-mercaptoethanol (Sigma-Aldrich) and reagent DX (final concentration of 0.5% v/v) (Qiagen) was added to prevent foam formation.

The 600µl sample was divided into two tubes containing 300µl, which were processed in parallel and combined in the final elution steps. The samples were mixed with 708µl of RNase free water and 20µl of Proteinase K (provided with the kit) and incubated for 10min at 56°C, followed by centrifugation for 3min at 8000g and 650µl of supernatant was removed for RNA extraction. The pellet was resuspended in the remaining solution and placed on the heating block at 55°C for later DNA extraction.

### 2.2.4. RNA extraction from tissue sample

RNeasy fibrous tissue kit was used to extract RNA from the biopsy samples. This kit contained the same reagents and columns as the described RNeasy mini kit. However, there were differences in some of the steps and in the volumes of reagents used.

The lysed & homogenized sample (650µl), described in a previous step, was mixed with 1 volume of 70% ethanol and 700µl centrifuged through the RNeasy mini column (provided with the kit) at 8000g for 15sec to bind the RNA to the silica membrane. The centrifugation step was repeated until the entire volume has passed through the mini column. The column was then washed with 350µl of RW1 buffer, centrifuged for 15sec at 8000 g. The flow through was discarded after each step.

DNase I was then added to each sample to eliminate DNA contamination before the final wash and elution of total RNA. The DNase I working solution was prepared by dissolving the DNase powder with 550µl of RNase-free water (provided with the kit), according to manufacturer's instructions. The DNase I solution (10µl) was mixed with 70µl RDD buffer. The mixed solution of 80µl was directly added to the column membrane of the RNeasy mini column and incubated at room temperature for 15 minutes. The next step was to add 350 µl RW1 buffer, followed by two washes of 500 µl with buffer RPE. A longer centrifugation step of 2 minutes at maximum speed was used to dry the membrane before the last wash. The RNeasy mini column was placed on a new collection tube and centrifuged for 1-2 minute to remove any RPE cross over. The RNA was eluted twice with 50µl RNase free water, by centrifugation for 1 minute at 8000g and stored at -80°C.

### 2.2.5. DNA purification from biopsy samples

Following the RNA extraction, the tube containing the homogenized biopsy at 55°C was used for DNA purification. The sample was mixed with 350µl of buffer AL and incubated for 10min at 70°C before mixing with 350 µl of 96-100% ethanol. The sample was then loaded onto the membrane of the DNA spin column with 600µl at a time and centrifuged for 1min at 8000g. The flow through was discarded. The steps were repeated until the remaining sample has passed through the column.

Washing and elution steps were the same as described in the DNA extraction protocol for cell lines using the QIAamp DNA Mini kit in section 2.2.1.

### 2.2.6. Nucleic acid quantification

The purity and concentration of the nucleic acids were measured using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific) according to the manufacturer's instructions. A blank measurement was done with 2µl of the corresponding elution solution. The absorbance at 260nm and 280nm wavelength was measured for each sample (2µl). The concentration was presented in ng/µl. If the ratio between 260/280nm was between 1.8 and 2, it indicated a high purity of nucleic acids.

### 2.2.7. DNA integrity

DNA integrity of the extracted DNA and suitability for further PCR set-ups was measured by PCR amplification of the housekeeping gene β-globin (fragment size: 209bp). Primers were purchased from Sigma-Aldrich (de Roda Husman et al., 1995, Saiki et al., 1988) and are shown in Table 2-8. The overall PCR technique and electrophoresis is described in this section, which is used for all other PCR based techniques (section 2.4 HPV typing HPV-16 E6 and HPV-18).

**Table 2-8: β-globin primer sequence and amplicon size**

| Primers | Forward Primer 5'-3' | Reverse Primer 5'-3' | Amplicon (bp) |
|---|---|---|---|
| PCO3(F)/ PCO5(R) | ACACAACTGTGTTCACTAGC | GAAACCCAAGAGTCTTCTCCT | 209 |

#### 2.2.7.1. General master mix set-up

All PCR reactions were set up in a designated PCR hood. The hood was sterilized for 30 min with UV light before preparation start. The master mix was made of PCR grade water, primers, buffers and Taq polymerase (Sigma Aldrich and QIAGEN). From this master mix 20µl were transferred in the designated PCR tubes. Five microliters of DNA, water for negative control or 5µl of HPV-positive sample for positive control were added to the designated tubes. The concentration and volume of the reagents used is shown in Table 2-9 and used for the remaining PCR reactions (including HPV typing for HPV-16 and HPV-18).

The sealed tubes were placed in a Nexus GSX1 thermocycler (Eppendorf). The reaction (amplification) was performed according to the conditions in Table 2-11. The samples were stored at -20°C for long-term storage or at 4°C for short-term storage.

**Table 2-9: Composition of general PCR-reaction master mix**

| Reagents | Concentration | Volume n=1 (µl) |
|---|---|---|
| Primers (each F + R) | 5 µM | 2.5 |
| Deoxynucleotide | 2 mM | 2.5 |
| 10X reaction buffer | 1x | 2.5 |
| Hotstar Taq polymerase | 5U/µl | 0.1 |
| PCR grade water | - | 9.9 |
| DNA | 1/10 | 5 |
| Total volume | - | 25 |

#### 2.2.7.2. Agarose gel electrophoresis

A 2% w/v agarose gel was prepared using a 1x Tris/Borate/EDTA (TBE) buffer. The TBE buffer was diluted from a ready-to-use 10x solution to a 1x TBE buffer. Agarose and 1µl 10mg/ml ethidium bromide per 100mL gel were added (Sigma-Aldrich). To each sample 5µl of Orange G loading dye (Sigma-Aldrich) was added and properly mixed. The first well and last well before and after samples was loaded with 10µl of PCR Ranger ladder for amplicons up to 1000bp (Norgen Biotek, Thorold, Canada). The well between the ladders were loaded with 15µl of the combined Sample-Orange G mix. The gel was run for 1h at 100 V (3-4 cm run distance) in an electrophoresis tank containing 1xTBE buffer. The bands were visualised under UV-light in a Gel doc (Imager station). Product sizes of 161bp for HPV-16 E6, 209bp for Beta-globin (indicating presence of DNA fragments/suitable for further PCR analysis) and 90bp for HPV-18 E7 were observed.

### 2.2.8. RNA integrity

Total RNA quality and quantity was analysed using an Agilent 2100 Bioanalyser and a RNA Nano 6000 kit, a chip electrophoresis assay, (Agilent Technologies, Santa Clara, USA).

During these assays RNA molecules are separated by size in the chip's micro-fluidic channel and detection is done by using a fluorescence nucleic acid stain. The exact molecular weight (RNA) was determined by comparison with a ladder of known fragment sizes (provided with the RNA Nano 6000 kit). The amount of RNA is equivalent to the amount of fluorescence measured. Results were provided as electropherograms. RNA integrity number (RIN) was automatically calculated by an algorithm developed by the manufacturer. The software automatically assigns an integrity number to a eukaryote total RNA sample (Agilent). RIN values between 8 and 10 were accepted as good quality RNA. A value of 10 indicates un-degraded RNA. Values below 8 indicates degradation of RNA, where 0 represents totally degraded RNA.

The gel dye-mix was brought to room temperature for 30 min and protected from light. The RNA Nano chip was taken out of a sealed bag and placed on the chip priming station. Of the gel-dye mix 9µl were pipetted at the bottom of the indicated wells. The plunger was positioned at 1ml and released after 30 seconds to distribute the dye mix. The chip priming station was opened and 9µl of the gel-dye mix was transferred in each marked

well. Wells labelled with ladder and sample were loaded with 1µl of ladder or 1µl sample (on arrival the ladder was denatured for 2 min at 70°C, the vial immediately cooled down on ice an aliquoted in RNase-free vials with the required volume and stored at -70°, before use thawed and kept on ice). To each well with ladder and samples 5µl of RNA 6000 Nano marker was added. The chip was placed on the mixer and vortexed for 60 s between 2000-2400 rpm. The chip was then inserted in the Agilent 2100 bioanalyzer and the run started within 5min. The chip was then analysed according to manufacturer's instructions.

## 2.3. Short Tandem Repeats – cell line authentication

Short tandem repeats (STR) consist of short repeat units that are 2-6 base pairs (bp) in length. Tandem repeated DNA is widespread throughout the human genome. These repeats can be detected through polymerase chain reaction (PCR), and are effective polymorphic markers (Ruitberg et al., 2001). This assay was used to confirm the identity of the developed primary cell lines with the collected biopsy sample. Furthermore, the identity of all the cell lines in this study was confirmed through the Cellosaurus website (https://web.expasy.org/cellosaurus/) and DSMZ database.

This assay was performed in assistance from Public Health England (PHE) and their cell authentication service in collaboration with NorthGene Ltd., using the Promega Powerplex 16 HS kit (Promega, Southampton, UK). This kit was used to analyse differences at 16 different hypervariable genetic loci (15 STR loci and Amelogenin), which are listed in Table 2-9. Those STR loci are amplified in a single multiplex PCR reaction simultaneously (fluorescence labels for primers are listed in Table 2-10), which results in different product sizes, depended on the number of repeats at different STR loci sites.

DNA samples were sent on FTA cards (GE Healthcare, Amersham, UK) with the recommended 50µl of DNA with a minimum concentration of 10 ng/µl. A minimum yield of 0.5ng of DNA in a 25µl reaction volume was needed for the Promega Powerplex 16 HS system. The components of the PCR amplification mix are listed in Table 2-11. The PCR amplification mix was added in the appropriate wells of the reaction plate and a plunger was used to punch a piece out of the FTA card (size of 1.2mm) into the appropriate well containing the PCR master mix (Table 2-10). Direct amplification of DNA occurred from the FTA card punch in the well with the PCR amplification mix.

The amplification was done using a GeneAmp®PCR System 9700 thermal cycler and cycle conditions shown in Table 2-12.

**Table 2-10: STR loci with chromosomal location in the human genomes**

| STR Locus | Chromosomal Location | Fluorescence label of primer |
|---|---|---|
| D3S1358 | 3p21,31 | Fluorescein (FL) |
| THO1 | 11p15.5 | |
| D21S11 | 21q21.1 | |
| D18S51 | 18q21.33 | |
| PENTA E | 15q26.2 | |
| D5S818 | 5q23.2 | 6-carboxy-4' ,5' –dichloro-2' ,7' –dimethoxy-fluorescein (JOE) |
| D13S317 | 13q31.1 | |
| D7S820 | 7q21.11 | |
| D16S539 | 16q24.1 | |
| CSF1PO | 5q33.1 | |
| PENTA D | 21q22.3 | |
| Amelogenin | X: p22.1-22.3 Y: p11.2 | Carboxy-tetramethylrhodamine (TMR) |
| vWA | 12p13.31 | |
| D8S1179 | 8q24.13 | |
| TPOX | 2p25.3 | |
| FGA | 4q31.3 | |

**Table 2-11: PCR amplification mix for amplification of STR loci.**

| PCR amplification mix - Reagents | Volume per reaction |
|---|---|
| Water, Amplification Grade | 12.5µl |
| PowerPlex® HS 5X Master Mix | 5.0µl |
| PowerPlex® HS 10X Primer Pair Mix[A] | 2.5µl |
| 5X AmpSolution™ Reagent | 5.0µl |

[A] Primer pair mix contains fluorescence labelled primers

**Table 2-12: STR typing thermal cycler conditions**

| Phase | Temperature (°C) | Time (seconds) | Cycles |
|---|---|---|---|
| Initial denaturation | 96 | 120 | 1 |
| Denaturation | 94 | 30 | 10 |
| Primer annealing | 60 | 30 | |
| Extension | 70 | 45 | |
| Denaturation | 90 | 30 | 27 |
| Primer annealing | 60 | 30 | |
| Extension | 70 | 45 | |
| Final extension | 60 | 30 | 1 |

The PCR products were separated using capillary electrophoresis and detected on an API Prism 310 Genetic Analyser (Applied Biosystems, Warrington, UK). The dyes were distinguished using virtual CCD imaging and the light was collected with the ABI Prism 310 data collection software. Electropherograms and a final report with the total counts at each loci sites for each cell line were generated by PHE.

## 2.4. Amplification of HPV-16 E6 gene and HPV-18 E7 gene

**Table 2-13: HPV-16 E6 primer sequence and HPV-18 E7 primer sequence**

| Primers | Forward Primer 5'-3' | Reverse Primer 5'-3' | Amplicon (bp) |
|---|---|---|---|
| HPV-16 E6 | GAACAGCAATACAA CAAACC | GATCTGCAACAAGAC ATACA | 161 |
| HPV-18 E7 | GAACCACAACGTCA CACAATG | CAGAAACAGCTGCTG GAATG | 90 |

The amplification of a 161bp fragment of HPV-16 E6 gene and 90bp fragment of the HPV-18 gene was used to confirm the presence of HPV-16 or HPV-18. Details of the primers are illustrated in Table 2-13.

The general PCR reactions mix is identical for HPV-16 E6, HPV-18 E7 β-globin shown in Table 2-9. Thermal cycler conditions for HPV-16 E6, HPV-18 E7 and are identical and shown in Table 2-14.

**Table 2-14: Thermal cycler program for HPV-16, HPV-18 and β-globin PCRs**

| Assay | Steps | Temperature | Time | Cycles |
|---|---|---|---|---|
| HPV-16 E6 | Initial denaturation | 94 | 15min | 1 |
| | Denaturation | 94 | 30sec | 40 |
| HPV-18 E7 | Primer annealing | 55 | 30sec | |
| | Extension | 72 | 60sec | |
| β-globin | Final Extension | 72 | 4 min | 1 |
| | | | | |

## 2.5. Clonogenic assays for measurement of radio-sensitivity

### 2.5.1. Crystal violet staining

Both the plating density assay (section 2.5.2) and radio-sensitivity assay (section 2.5.3) required staining the cells with a crystal violet solution (Thermo Fisher Scientific), pictures were taken of the plate and the number of colonies were determined using image analysis (section 2.5.4).

Media was removed from the 60x15mm culture dishes using a pipette. Before the media was discarded it was used to remove 3T3 feeder cells from the 60x15 mm culture dishes. 5mL 1xPBS was added to the culture dishes and was slowly pipetted up and down to remove the remaining 3T3 feeder cells and media. The PBS suspension was removed.

In the next step 5mL of 50 % v/v 1xPBS/methanol was added to the plates for 2 min to prepare the cells for the final fixing step with methanol. The mix of 1xPBS/Methanol was removed and absolute methanol added to the plates for 10min and incubated at room temperature. The methanol (Thermo Fisher Scientific) was discarded and the plates left to dry (evaporation of methanol) for 10-20min.

5mL of Crystal violet solution was added per plate. The cells were stained for 10 min at room temperature. The Crystal violet solution was removed, filtered, collected and reused for other experiments. The plates were gently washed with tap water twice and left to dry at RT.

### 2.5.2. Plating density assay

This assay was used to define the cell density to be used in the radio-sensitivity assay**.** Cells were collected according to standard cell culture procedure. The cells were counted as described in section 2.5.4

The collected cells were diluted to a range of cell densities. The range was defined from 20,000 to 1000 cells per 60x15mm cell culture plate. The density test was set-up with and without 3T3 feeder cells to support cell growth. The optimal cell density for each cell line was then defined according to state of confluence within the set-up range of 20,000; 15,000; 10,000; 5,000 and 1,000. and were defined for CU-OP-17 within 5,000 and 10,00

cells and for CU-OP-20 within 5.000 and 1,000 cells. 2500 cells more than the 5.000 cells of the density plate.

Cells were seeded with 5mL media per plate and incubated at 37°C at 5% $CO_2$ for 7 days. After 7 days the media was changed and the plates incubated for another 7 days under the same conditions. After 10-15 days (when colonies with >50 cells were observed) the plates were stained with Crystal violet (section 2.1). Using this assay, defined optimal cell densities were for UMSCC-6 (4000), CU-OP-17 (7000 per plate) and CU-OP-20 (4500).

All other optimal cell densities were defined by Dr. Evelyne Pirotte (Table 2-15).

**Table 2-15: Defined cell densities for clonogenic assays**

| Cell lines | Defined cell densities per plate |
|---|---|
| UMSCC-47 | 2500 |
| UPCI-SCC-90 | 12.500 |
| CU-OP-2 | 4400 |
| CU-OP-3 | 4400 |
| UMSCC-4 | 7500 |
| UMSCC-19 | 7500 |
| UMSCC-74a | 3000 |
| HEKn | 3000 |

### 2.5.3. Radio-sensitivity testing of OPSCC cell lines

All cell lines were plated at low cell density on 60x15 mm culture dishes. 3T3 feeder cells were plated with each cell line to support cell growth. A stock suspension was prepared for each cell lines in three 50ml media Falcon tubes, containing $8x10^6$ J2 3T3 feeder cells ($1x10^6$ cells/mL per plate). 5mL out of this stock suspension was added to each 60x15 mm culture dish and incubated at 37°C with 5 % $CO_2$ for 24 hours. After 24 hours, the cells were firmly attached to the culture plate. The plates were protected by wrapping in Para-film, before transportation to the γ-irradiator, a caesium/137 source (Gammacell-1000 MDS Nordion, Ottawa, CA).

The plates were treated with 0, 0.5, 1, 2, 4 or 6 Gy to assess survival of cells across a range of radiation doses. After irradiation the media was changed and plates incubated at 37°C with 5% $CO_2$ for 7 days. For the CU-OP cell lines media was changed from

GMEM/EGF-negative to GMEM/EGF-positive media. After 7 days the media was changed and plates again incubated at 37°C with 5 % $CO_2$. Depending on the growth characteristics of the cell lines the assay was stopped after 10-15 days as determined during cell density tests. Photographs were taken at different magnifications and cells were stained with crystal violet and colonies counted. All experiments were performed three times in triplicate (3 plates per 3 runs).

### 2.5.4. Colony counting

Pictures of the stained colonies were taken and colonies >50 cells were counted using a Colony counter (ColonyDoc-It Imaging Station, UVP, Cambridge, UK). The plates were labelled on the side with A, B, and C (to confirm exact position at a later point), as three pictures were taken at three different positions, plates being moved at 90° degrees. These pictures were taken with a Canon EOS Rebel T3 camera connected to the counter. White light (base lighting) was used with light coming from below the plates.

The counting area was defined for a 60x15 mm cell culture plate and a template made for each cell line. The template was defined by different colour pixels. The numbers of pixels varied for each cell line. The aim was that almost all colonies were counted automatically after defining the template. A small percentage of colonies had to be added manually or split (as colonies had merged). These were checked using a microscope. The average number of colonies was calculated per run (3 pictures per plate). The plating efficiency and the surviving fraction of each cell line per dose was calculated.

**Plating efficiency**

Plating efficiency was determined using the equation shown below:

$$\frac{\qquad\qquad}{\qquad\qquad}$$

**Surviving fraction**

The surviving fraction was determined using the equation shown below:

$$\frac{\qquad\qquad}{\qquad\qquad}$$

## 2.6. Western blotting

### 2.6.1. Materials and preparation of lysis buffer

All materials used for protein extraction, western blotting and imaging are listed in Table 2-16. Components for lysis buffer are listed in Table 2-17.

**Table 2-16: Materials for Western blot**

| Materials | Manufacturer |
|---|---|
| Mini gel tank | Thermo Fisher Scientific |
| Novex Bolt transfer buffer 20x | |
| Bolt 4-12% Bis-Tris plus gels 15 wells | |
| Novex Bolt Mops SDS running buffer 20x | |
| Novex Bolt LDS sample buffer 4x | |
| Novex Bolt sample reducing agent 10x | |
| Novex Bolt Antioxidant | |
| SeaBlue®Plus-2 Prestained standard | |
| PVDF membrane 0.45µm | GE Healthcare |
| Electrophoresis Power Supply EPS 1001 | Amersham Pharmacia biotech |
| p53 antibody (rabbit) cat:9282 | Cell Signalling, Danvers, USA |
| Phospho-p53 (Ser15) antibody (rabbit) cat:9284 | |
| P21 Waf1/Cip1 (12D1) antibody (rabbit) cat:2947 | |
| B-actin (13E5) antibody (rabbit) cat:4970 | |
| JUNB (C37F9) antibody (rabbit) cat:3753 | |
| NDRG1 (D8G9) XP® anribody (rabbit) cat:9485 | |
| Goat Anti-Rabbit IgG (H+L) AP conjugate cat:1706518 | BioRad |
| CL X-posure Film | Thermo Fisher Scientific |

**Table 2-17: Reagents for phopho-protein lysis buffer**

| Reagents[B] | For 10ml | For 1ml | Manufacturer |
|---|---|---|---|
| 2x Lysis buffer[A] | 5mL | 0.5mL | - |
| 10% NP40 | 1mL | 0.1mL | Fluka Biochemika |
| Protease inhibitor cocktail | 100µl | 10µl | Sigma Aldrich |
| Phosphatase inhibitor cocktail 2 | 100µl | 10µl | |
| Phosphatase Inhibitor cocktail 3 | 100µl | 10µl | |
| PMSF 100mM | 100µl | 10µl | |
| $Na_3VO_4$ 100mM | 100µl | 10µl | |
| $ddH_2O$ | 3.5mL | 0.35 L | - |

**A** 2x Lysis buffer consists: of 100mM Hepes (Sigma Aldrich), 10mM NaF (Sigma Aldrich), 10mM Iodoacetamide (Sigma Aldrich), 150mM NaCl (Fisher Scientific) filled up to a total of 100mL.

**B** The 2x Lysis buffer was prepared separately and stored at 4°C. All other reagents were added immediately before use, because some components in the buffer are not stable for more than 30min in aqueous solution. The protease inhibitor cocktail was added to inhibit enzymes that degrade proteins. Phosphatase inhibitor cocktail 2 inhibits tyrosine protein phosphatases. Phosphatase Inhibitor cocktail 3 inhibits alkaline and serine threonine protein phosphatases.

### 2.6.2. Sample preparation for Western blotting

Cells were plated as described in Table 2-18 and treated with 2 or 6Gy ionising radiation. After a 24h incubation, $2x10^6$-$2.5x10^6$ cells were collected per cell line and treatment dose (2 and 6Gy 24h) and transferred to 1.5 mL tubes for cell lysis. Per $1x10^6$ cells/ml 50 µl of lysis buffer were added. Lysis buffer was added to the cell pellet and cells lysed for 20-30min on ice and vortexed 1-2 times during the 20-30min lysis. The lysate was clarified by centrifugation (10,000g, 20mins, 4'C). The volume of lysis buffer was noted before and after centrifugation to allow determination of exact protein concentration. Samples were stored at -80°C, until determination of protein concentration by performing a Bradford assay.

**Table 2-18: Number of cells seeded for cell cycle analysis, Western blotting and mRNA sequencing set-up**

| Cell lines | Number of cells seeded |
|---|---|
| UMSCC-47 | $4x10^5$ |
| UPCI-SCC-90 | $3.5x10^5$ |
| CU-OP-2 | $1x10^6$ |
| CU-OP-3 | $1.2x10^6$ |
| CU-OP-20 | $1.2x10^6$ |
| CU-OP-17 | $1x10^6$ |
| UMSCC-4 | $4x10^5$ |
| UMSCC-6 | $4.5x10^5$ |
| UMSCC-19 | $4.5x10^5$ |
| UMSCC-74a | $8x10^5$ |
| HEKn | $7x10^5$ |

### 2.6.3. Bradford assay

All samples were slowly thawed on ice. A 1/20 dilution in water of each sample was performed in a 96 well plate. Duplicates of this dilution (10µl) were pipetted in new wells.190µl of Bradford solution (1:1 dilution with double distilled water) was added to each sample. The plate was incubated for 25 min at RT protected from light. The absorbance was measured at a wavelength of 595nm. A standard curve of different protein concentrations of 10µg/ml, 7µg/ml, 4µg/ml, 1µg/ml and 0µg/ml was used to estimate the protein concentration of each sample. The different protein concentration for the standard curve were diluted with distilled water from a stock concentration of 2mg protein/ml (Sigma Aldrich).

### 2.6.4. Protein sample preparation

A final loading concentration of 1000µg/ml was prepared for each sample. Sample volume was calculated, depending on the number of gels operated per experimental run. Reducing agent and LDS loading buffer were added to each sample. All samples were well mixed and briefly centrifuged to remove any liquid from the caps. The samples were then incubated for 5min at 100°C. A brief centrifugation was done to remove all remaining condensation from the caps. All tubes were stored on ice until electrophoresis.

### 2.6.5. Electrophoresis

The gel tank was assembled according to the manufacturer's instructions (Thermo Fisher Scientific) and rinsed with deionized water to remove any residues. The tank was filled with 1xRunning buffer (50mL from 20xMOPS stock solution with 950mL double deionized water). The pre-cast gel cassette (15 wells, Life technologies) was removed from its packaging and rinsed with deionized water. The tape at the bottom of the cassette (back of the gel) was peeled off and the comb gently removed to avoid damage to the wells. The cassette and the wells were rinsed three times with 1x MOPS running buffer. The cassette was placed into the tank and fixed by clamps. The wells were left out of the buffer for loading the samples. Twenty microliters of sample were loaded per well and 5µl of ladder was loaded in the first and last wells. The gel was lowered into the buffer and filled with 1x running buffer until the indicated line of the tank. The electrophoresis was run for 35-40 min at 170 V (constant) and 100mA.

### 2.6.6. Transfer from gel to membrane

Assembly of the Western Blotting Sandwich for protein transfer is illustrated in Figure 2-2. The PVDF membrane and Transfer Buffer were prepared during the electrophoresis run. The PVDF membrane was placed in methanol for 30 seconds and rinsed once with deionized water. The membrane and transfer pads were soaked in 1x transfer buffer (Table 2-19) for several minutes. The filter paper was also soaked in transfer buffer briefly before use. The cathode and anode were rinsed with deionized water and transfer buffer. 5ml of transfer buffer was placed into the cathode chamber. After the electrophoresis the cassette was laid down flat and separated using a gel knife by inserting it into the narrow gap between the two plastic plates. The top plate was gently removed and the wells were cut off with the gel knife. A briefly soaked filter paper was placed on top of the gel and any trapped air bubbles were removed. The plate was turned (gel and filter paper facing downwards) and the gel pushed off the remaining plate and the foot of the gel was cut off. One transfer pad was placed into the cathode and the gel with filter paper (facing downwards) was placed on top of it. The pre-soaked membrane was placed on top of the gel followed by a filter paper and transfer pad. The module was closed with the anode and placed into the electrophoresis tank. The centre of the module was filled with 1x transfer buffer, the outer champers with deionized water just below the electrodes. A 30V constant voltage and 170mA were applied for 1:30 hours.

**Table 2-19: 1xTransfer buffer composition for 1000mL**

| Reagents | Volume (ml) |
|---|---|
| Bolt transfer buffer (20x) | 50 |
| Bolt antioxidant | 1 |
| Methanol | 100 |
| Water | 849 |

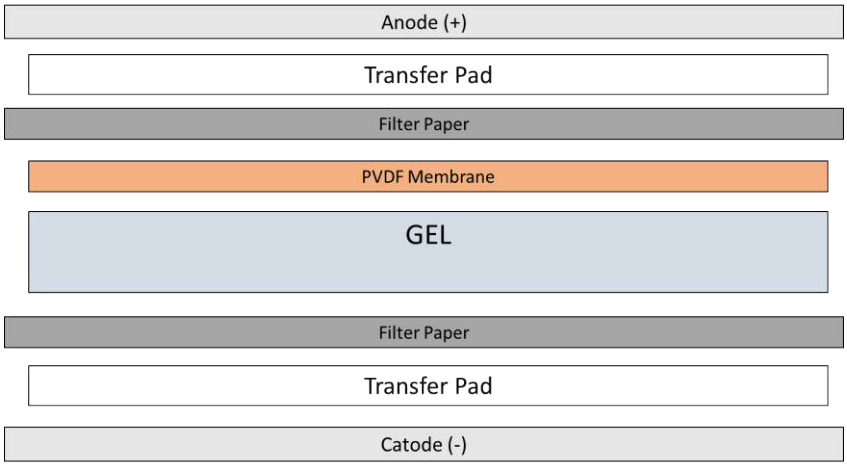



**Figure 2-2: Assembly of the Western blotting transfer sandwich**

### 2.6.7. Blocking and immuno-probing

The membrane was taken out of the blot module and rinsed with PBS-Tween (PBS-T). The membrane was incubated with iBT-buffer (Table 2-20) for a minimum of 1 hour (RT) for blocking. After 1 hour the membrane was removed from the blocking agent and probed with the primary specific antibody (e.g p53 – 1:1000 in iBT) overnight at 4°C.

The membrane was removed from the primary antibody and washed three times (10 min) with PBS-T, followed by probing with the secondary antibody (goat anti-rabbit) for 1 hour. The membrane was removed from the secondary antibody and washed three times (10 min) with PBS-T.

**Table 2-20: Blocking solution**

| Reagent[A] | Concentration | Manufacturer |
|---|---|---|
| Tropix I-Block | 0.2% (1g in 500mL) | Applied Biosystems |
| Tween-20 | 0.1% (0.5mL in 500mL) | Sigma-Aldrich |
| Sodium Azide | 0.4% (2g in 500mL) | Sigma Aldrich |

[A] Solution was prepared in 500mL PBS (Fisher Scientific) and heated up to 80°C to be able to dissolve 1g I-Block powder. As soon as the powder was dissolved and the solution cooled down sodium azide was added to prolong the usage of the buffer for a 2-4 weeks. (I-Block is a highly purified casein-based blocking reagent & provides the highest signal to noise performance in combination with CDP Star substrate for alkaline phosphatase experiments for Western Blotting).

### 2.6.8. Chemiluminescence

The PBS-T was discarded after the last wash and membrane washed for 4min with 1x Alkaline phosphatase (AP) buffer (10x Assay Buffer, Applied Biosystems). This allows enzyme attachment to the goat anti-rabbit–AP conjugate secondary antibody. The membrane was put on tissue to remove the excess buffer and transferred on a plastic sheet. The membrane was incubated for 5min in Tropix CDP-star, a chemiluminescent substrate for alkaline phosphatase stained membranes (Applied Biosystems). The excess detection reagent was removed from the plastic sheet using paper tissues. The plastic sheet was cut to size and placed in a film cassette-using tape to secure. The membrane was exposed to a photographic film (10min initially to define longer exposure time, if necessary), by using a film developer (Xograph Imaging Systems Compact X4). Exposure time points were set, according to the initial results obtained. Images were analysed and quantified using Image Studio Lite (LI-COR Biotechnology, Cambridge, UK).

### 2.6.9. Re-probing of membrane

The membrane was rinsed with PBS-T and washed twice in PBS-T for 5min on a shaker. The membrane was then placed into a plastic bag containing 25mL of 1xMESNA buffer (Table 2-21) for 1 hour at 55°C in a water bath with shaking the plastic bags every 10 min. After stripping off, the membrane was washed for 30min with PBS-T, changing the buffer every 10min. The membrane was then blocked with iBT for 1 hour and re-probed with primary antibody as described in section 2.6.7.

**Table 2-21: Reagents for a 5x MESNA stripping stock solution**

| Reagents | 5x stripping buffer in 100ml[A] | Manufacturer |
|---|---|---|
| Tris-HCI pH 6.8 | 31.25 mL | Sigma Aldrich |
| SDS | 10 g | |
| 2-mercaptoethansulfonate (MESNA) | 4.1 g | |

[A] For stripping off the membrane the solution was diluted to a 1x working solution.

## 2.7. Determination of cell cycle distribution using flow cytometry

### 2.7.1. Sample preparation

Cells were harvested and counted, according to standard cell culture procedure, once 70-80% confluent. The number of cells plated varied with each cell line as previously described (Table 2-18). Cells were seeded in sixteen plates (6cm tissue culture plates) and treated with 2Gy and 6Gy the next day (when ~50-60% confluent). No J2 3T3 feeder cells were used in the set-up of the experiments to avoid misleading cell cycle results. Cells were collected according to the defined time point of 8h, 24h and 48h. 4 plates of untreated cells were collected immediately after treatment (time zero). Reagents used in this assay are described in Table 2-22.

Cells were harvested and plated at a defined cell density in cell culture dishes (60x15 mm). Cells were treated with ionising radiation (IR) when 60-70% confluent and collected at defined time points. 4 plates were set-up for untreated cells, 2 plates per dose (2 and 6Gy) and incubation time point (8h, 24h and 48h). Cells were collected after defined time points and 500,000 cells transferred to labelled Fluorescence-activated cell sorting (FACS) tubes and fixed in 1mL of 70 % ethanol and stored at -20°C for a minimum of 1 hour before propidium iodide staining.

**Table 2-22: Reagents for cell cycle analysis**

| Reagents | Working concentration | Manufacturer |
|---|---|---|
| Propidium Iodide 50µg/ml | 1ug/ml | Sigma Aldrich |
| Ribonucelase A (RNase A) | 10 µg/ml | |
| Ethanol | 70% | Fisher Scientific |

### 2.7.2. Propidium iodide staining

Tubes were centrifuged for 10min at 360g after ethanol fixation. Then cells were washed with PBS and centrifuged for 5min 270g. PBS was removed and 100µl of RNase A was added to all tubes to eliminate any RNA contamination (propidium iodide (PI) binds to RNA). The tubes were incubated at 37°C for 45 minutes. Cells were then re-suspended in 250µl of PI (1ug/ml) and incubated for 15min at 37°C.

### 2.7.3. Initial cell cycle distribution measurement

Cell cycle distribution was then measured using the BD Accuri (BD Biosciences, Wokingham, UK) low pressure flow cytometer. The initial gating was defined on the Accuri C6 software, during data collection with PI staining (PI is excited at 488nm). Each run was stopped after 30,000 events under standard voltage settings on the BD Accuri flow cytometer.

### 2.7.4. Flow cytometry analysis

Data was extracted as FCS files and later analysed using FlowJo analysis (version 10) software with the gating strategy seen in Figure 2-3. The cell cycle pattern was analysed using the Watson pragmatic algorithm (Watson et al., 1987).

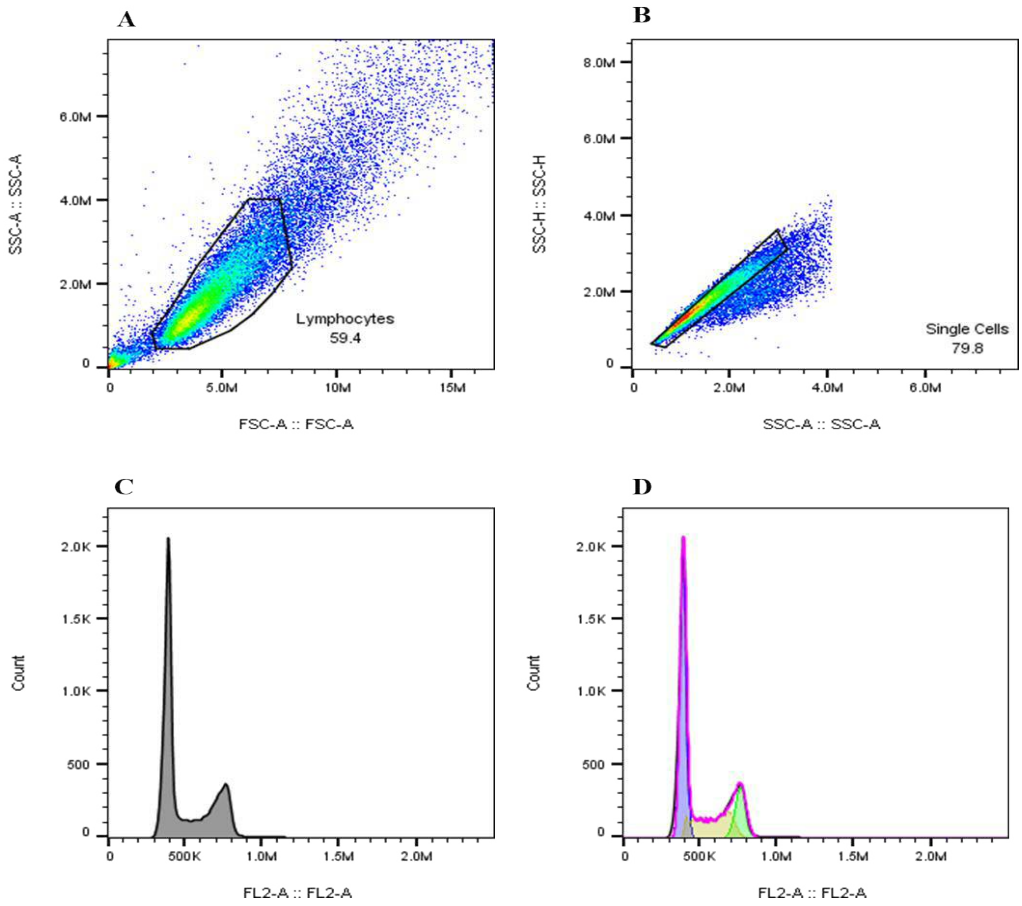



**Figure 2-3: Gating strategy FlowJo (Version 10) for Cell Cycle Analysis**
(A) Axes were defined as SSC-A/FSC-A to separate dead and live cells (Gate defined as Lymphocytes by FlowJo); (B) SSC-H/SSC-A was set to separated doublets and single cells; (C) Cell cycle peaks – axes defined as Y-axe (count) to X-axe (FL2-A) for to determine the number of cells stained with PI; (D) represents the cell cycle phases of G1/S/G2 defined by the Cell Cycle Analysis Tool within FlowJo.

Excel 2016 was used to generate 100% stacked bar charts. Statistical analysis was performed in GraphPad Prism 7 using a two-way ANOVA with a Sidak post-test to the panel of cell lines (treatment dose and time points) with each other to identify any significance.

## 2.8. mRNA sequencing

Experiments for mRNA sequencing were set-up as described in section 2.2.2 RNA extraction from cell lines. Number of cells plated is shown in Table 2-18. RNA was then extracted, quantified and RNA integrity tested as described. Sequencing libraries were prepared using the Illumina® TruSeq® Stranded Total RNA with Ribo-Zero Gold™ kit (Illumina Inc.). The libraries were sequenced using a 75-base paired-end (2x75bp PE) dual index read format on the Illumina® HiSeq2500 in high-output mode according to the manufacturer's instructions.

Library preparation and mRNA sequencing were carried out through commercial service/collaboration with Wales Gene Park. Sequencing data was then analyzed with the support of the bioinformatics service provided by Wales Gene Park (Dr. Kevin Ashelford and Dr. Peter Giles) described in section 2.8.4. Data analysis included quality control of sequencing, analysis of viral genes and integration, differential gene expression, comparison/grouping of non-irradiated and irradiated samples (including different levels of gene expression between non-irradiated and irradiated samples), and Gene Ontology (GO) analysis.

### 2.8.1. Library preparation

Library preparation was performed using the Illumina® TruSeq® Stranded Total RNA with Ribo-Zero Gold™ kit (Illumina Inc.). This kit offers strand information of RNA transcript and capture of both coding, as well as multiple forms of non-coding RNA. Library preparation was done by Wales Gene Park in accordance with manufacturer's instructions.

100-900ng of total RNA with a RIN value of 8 or >8 was depleted of ribosomal RNA. Manufacturer's instructions were followed for the whole library preparation, except for the clean-up after the ribozero depletion step, where Ampure®XP beads (Beckman Coulter) and 80 % Ethanol were used.

This process was then followed by RNA purification, fragmentation and priming for cDNA synthesis (1st strand cDNA synthesis and 2nd strand cDNA synthesis). The process of synthesis of the 1st strand of cDNA, reverse transcribes the cleaved RNA fragments that were primed with random hexamers into first strand cDNA using reverse

transcriptase and random primers. This process removes the RNA template and synthesizes a replacement strand, incorporating dUTP in place of dTTP to generate ds cDNA. AMPure XP beads are used to separate the ds cDNA from the second strand reaction mix to generate a final blunt-ended cDNA. This step is followed by adenylation of 3' ends where a single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction and a 'T' nucleotide provides a complementary overhang for ligating the adapter to the fragment. Ligating adapters to the ends of the cDNA prepares for hybridization onto a flow cell. This step is followed by enrichment of DNA fragments through PCR amplification (12-cycles) which have adapter molecules on both ends. The last step includes library validation before mRNA sequencing.

### 2.8.2. Library validation

The libraries were validated using the Agilent 2100 Bioanalyser and a high-sensitivity kit (Agilent Technologies) to ascertain the insert size, and the Qubit® (Life Technologies) was used to perform the fluorometric quantitation. After validation, the libraries were normalized to 4nM, pooled together and clustered on the cBot™2 (system to amplify hundreds of millions of single-molecule DNA templates simultaneously) following the manufacturer's recommendations.

### 2.8.3. Sequencing

Sequencing was performed using a 75-base paired end (2x75bp PE) dual index read format on the Illumina® HiSeq2500 in high-output mode according to the manufacturer's instructions. It facilitates sequencing of both ends of each fragment generating sequences of high output quality. Primers bound to each end of the template and each single dNTP corresponding to the sequence (labelled with a fluorescence dye) were added, and then cleaved to allow incorporation of the next base. 4 reversible dNTPs are present during each sequencing cycle. After each binding of a nucleotide the clusters were excited by a light source and a characteristic fluorescence signal is emitted and the signal recorded. Each fragment was sequenced both ways and data recorded in pairs. FastQC files were generated as the start of the subsequent bioinformatics analysis.

### 2.8.4. Sequencing data analysis

#### 2.8.4.1. Quality control

For each sample, a standard performance of quality was initiated of the reads using a combination of FastQC and assessing a number of metrics, as number of reads (duplicates), Quality Values (QV) – a per-base estimate of base caller accuracy and insert size. Assessing the metrics was done after mapping 10% of the data using a Burrows – Wheeler Aligner software package. It is a software tool to map low-divergent sequence against a large reference genome. For this QC, BWA-MEM is used as it is recommended for high quality performances as fast and accurate software tool. Having determined that all samples passed quality control checks, the original FastQC files were taken forward for subsequent analysis on the full dataset.

#### 2.8.4.2. Trimming

To remove adapter sequencer and poor quality ends of reads, trimming was performed using Trim Galore, a wrapper tool which uses cutadapt and FastQC to trim and perform QC on the trimmed. Trimming was performed in paired-end mode.

#### 2.8.4.3. Sequence mapping

Reads (trimmed) were mapped against a combined human sequence genome hg19 and HPV16 genome referne sequence NC_001526 using STAR (Alex Dobin, Git Hub). In contrast to many other RNA-sequencing aligners which are extensions to existing short read mappers (and work by aligning short reads to a database of splice junctions and/or or align split-reads before alignment), STAR is designed to align the non-contiguous sequences directly to the reference genome. In order to take a conservative approach to issues around secondary mapping issues within the genome, STAR was run with the MultimapNMax=1 flag, meaning reads mapping to more than 1 location were considered unmapped.

#### 2.8.4.4. Gene expression

Expression counts for both exons and transcripts were calculated, using Subread featureCounts Version 1.5.1 (Liao et al., 2014), a program for assigning sequence reads to genomic features.

Read summarization (counting) was generated for paired end read fragments. It was summarized at exon level and then grouped at transcript level. To provide stringent and robust data, reads overlapping more than one feature were excluded from the count summary (according to authors recommendations (Liao et al., 2014)).

To define the exon and transcript locations (raw reads calculation), the RefSeq gene model was used. This was provided by the UCSC gene model for the human reference hg19 and combined with equivalent information from the NC_001526 HPV16 gene model (NCBI).

#### 2.8.4.5. Principal component analysis

Principal component analysis (PCA) is a statistical procedure, that converts a set of observations of possibly correlated variables into a set of values linearly uncorrelated variables called principal components. PCA was invented by Karl Pearson as an analogue of the principal axis theorem in mechanics (Pearson, 1901).

Top 5000 most variable transcripts were selected according to Median Absolute Deviation (MAD) and PCA applied to the resultant dataset using R environment for statistical computing. Output graphs generated included the percentage (%) of variability capture within each component and graphs coloured by meta-information (HPV source, IR treatment), visualized on plots as component 1 versus component 2 and component 2 versus component 3.

#### 2.8.4.6. ConsensusclusterPlus

To determine the nature of any clustering between samples and the membership and stability evidence of these clusters in an unsupervised analysis, the ConsensusclusterPlus class discovery tool was applied to the dataset. The top 5000 most variable transcripts (identified according to their median absolute deviation) were filtered. This resultant data was used to generate clusters within the dataset (generating consensus cumulative distribution function, delta area and tracking plots), displaying a combination of hierarchical clustering and Pearson's correlation co-efficient as the distance metric.

#### 2.8.4.7. Differential gene expression

Differentially expressed genes were identified using an DEseq2 analysis (statistical analysis of count matrices for systematic changes between conditions), by using the R statistical software/Bioconductor package (Love et al., 2014), on normalised count data. For multiple testing and false discovery issues, the generated p-values were corrected using the FDR method (Benjamini and Hochberg, 1995). It is a method of hypothesising the level of type I errors in null hypothesis testing when conducting multiple comparisons. Designed FDR controlling procedures, control the proportions of „findings "that are false.

Results were presented using heatmaps (a tool to visualise the large and complex matrix of values). To group similar data together while revealing any patterns within the dataset, the generated data was log2 transformed and median centred before hierarchical clustering (using average linkage and Pearson's correlation as the similarity metric). The final values were then converted into 3 colour patterns (green, red and black), where green represented under expressed (compared to median), black represented values around the median expression and red were over expressed values.

#### 2.8.4.8. Gene ontology over-representation analysis

The Gene ontology over-representation analysis (GO ORA) analysis was undertaken using the GOstats bioconductor library. The resultant data was corrected for multiple testing and false discovery using the FDR method.

Genes were classified in three categories, its biological processes (BP), molecular functions (MF) and cellular components (CC).

One of the issues with ORA analysis of very short gene lists was that terms appeared significant because combination of a few observations in a gene set with a limited number of terms (i.e. an issue of rarity). To overcame this limitation common practise was to undertake analysis using a total list of around 200-500 genes in length (even though not all of the genes will have passed any significance threshold) to help this shortcoming of the maths.

#### 2.8.4.9. Human viral fusion transcript plots (Circos plots)

Circos plots were used to vizualize and identify human (hg19) and viral (HPV-16) mRNA fusion transcripts. The software Circos is using a circular layout to display relationships within the genomic data (Krzywinski et al., 2009). Each read mapping to a different chromosome were extracted and only those reads mapping to hg19 and NC_001526 (HPV-16) were filtered, using samtools "bad pairs". These were then converted into circos links format before bundling and plotting the links using Circos against a combined HPV16 and hg19 karyotype.

#### 2.8.4.10. Integrative genomics viewer

The Integrative Genomics Viewer (IGV) enables real-time visualization of aligned sequence reads, different levels of viral gene expression, mutations and mismatches compared to the reference sequence, splicing patterns and gene expression. This software was downloaded from the Broad Institute (Robinson et al., 2011).

#### 2.8.4.11. Geneview

The Genview software was written by Dr. Peter Giles of the Wales Gene Park. It helps to combine expression data alongside annotation data and enabled the rapid visualisation of different patterns of gene expression through histograms or generating heatmaps. Furthermore, it provides Annotation information of every gene, as cellular components, biological processes and molecular functions (incorporated into the software – link to NCBI and other databases).

# Chapter 3 : Derivation and characterization of novel OPSCC cell lines

## 3.1. Introduction

Historically, most OPSCC cases have been associated with tobacco use and alcohol consumption, with patients typically being older with an average age of 59.3 years old (Schache et al., 2016). This picture has widely changed in conjunction with HPV infection, with patients being much younger (mean age for HPV-positive patients 57.4 years compared to 61.4 years for HPV-negative patients) and non-smokers/infrequent smokers with a high number of lifetime sexual partners (Chu et al., 2013, Schache et al., 2016). HPV positive and negative OPSCC are clinically distinct entities, as HPV-positive patients (with a high-risk type of HPV), respond better to radiotherapy compared to HPV-negative patients (Evans et al., 2013). Furthermore, HPV-positive patients showed a better survival rate compared to HPV-negative patients over a period of 5 years (Evans et al., 2013, Lill et al., 2017).

There are no animal models that mimic HPV-associated OPSCC, therefore to study the biology and treatment of this disease, HPV-positive *in-vitro* models are needed. However, the study of head and neck cancer biology is hampered by the limited availability of suitable cell lines. In 2014, there were 8 known published HPV-positive cell lines – consisting of UMSCC-47, UPCI-SCC-90, UMSCC-104, 93-VU-147T, UD-SCC-2, UT-SCC-45,UPCI-SCC-152 and UPCI-SCC-154 (Arenz et al., 2014, Bradford et al., 2003, Brennan et al., 1995, Brenner et al., 2010, Descamps et al., 2014, Ferris et al., 2005, Kimple et al., 2013, Lin et al., 2007, Rieckmann et al., 2013, Steenbergen et al., 1995, Tang et al., 2012). UPCI-SCC-90 and UPCI-SCC-152 were derived from different specimens from the same patient in Pittsburgh (USA). The majority of available head and neck cancer cell lines are HPV-negative. Those cell lines used in publications share all the same characteristics of a HPV-positive cell line, expressing wild-type 53 and p16 (a known surrogate marker for HPV infection). An exception to these published HPV-positive cell lines is 93-VU-147T which was described with genetic alterations and mutated p53 (Steenbergen et al., 1995).

The majority of the known HPV-positive HNSCC cell lines were derived from former smokers. It was established that the smoking characteristics of patients (smoker, non-smoker) have an impact on patient prognosis in terms of survival and recurrence of a tumour. In a study on head and neck cancer it was established that HPV-positive patients

who were also tobacco users were at significantly higher risk of disease recurrence than non-smokers with HPV-positive tumours, with an effect on general survival (Maxwell et al., 2010). All HPV-negative patients in this study were tobacco users and had a significantly reduced disease specific survival compared to HPV-positive patients (Maxwell et al., 2010).

In summary, new cell line models are needed to study the impact of treatment response in new "typical" head and neck patients (HPV-positive, young and non-smoker). Two novel HPV-positive OPSCC cell lines (CU-OP-2 and CU-OP-3) were previously successfully derived in Cardiff by Dr. Evelyne Pirotte. CU-OP-2 was derived from an ex-smoker, whereas CU-OP-3 was derived from a non-smoker. Cell line models, such as CU-OP-2 and CU-OP-3, representing the "typical" characteristics of an HPV-positive patients need to get increased for better comparison to HPV-positive patients and for better comparison to the higher number of HPV-negative head and neck cancer cell lines. Development of new cell line models for HPV-positive head and neck cancer should be continued to a higher number to mimic better the wider variety of treatment response of HPV-positive OPSCC patients.

The aim of this study was to continue deriving novel cell lines from oropharyngeal cancer biopsies, using an explant method (cell outgrowth of <1mm tissue pieces) and to characterize the successfully grown lines as potential models for HPV-positive or HPV-negative OPSCC. The results are described from receiving the biopsy to phenotypical characterisation (including morphology). The cell lines were characterised in terms of HPV status and their identity validated by comparing similarities via STR typing of the original biopsy and the established cell lines.

### 3.1.1. Terminology

The explant method was used to process the biopsies during cell line development. "Explant" describes new cell growth out of the small biopsy pieces. Within this chapter the term primary culture is used to indicate the first passage from the outgrowth (explants). Early passages are classified as short-term culture. A cell line was classified as immortal or established, when it surpassed a certain number of population doublings (PD), compared to characteristics of "normal" epithelium cells e.g. HEKn cells (with an expected proliferative capacity of around 30 PD). Immortal cell lines are likely to have

overcome telomere mediated senescence and programmed cell death, which supports the suggestion that these cell lines are derived from cancer cells present in the biopsy.

The study was named Primary Culture of Oropharyngeal Cells (PCOC), and during collection and growth, biopsies were allocated a "PCOC" number. Once lines were established, they were referred to as "CU-OP#" lines – indicating Cardiff University Oropharyngeal Cancer, with # corresponding to the chronological order of the biopsy collection. The passage number of each culture was recorded and is indicated with the prefix "p", e.g. CU-OP-17 p14 indicates the fourteenth passage of CU-OP-17.

### 3.2. Sample collection & Study samples

Ethical approval for this study was obtained from and NHS Research Ethics Committee and the study was adopted into the National Institute for Social Care and Health Research (NISCHR) portfolio. Approval was also obtained from the Cardiff and Vale University Health Board Research and Development (R&D) Office. Patients were considered for inclusion in the PCOC study, if they were likely to be diagnosed with an oropharyngeal tumour. These inclusion criteria were: presentation at the Head and Neck Oncology clinic at Cardiff and Vale NHS Trust, fine needle aspirate and/or biopsy of a mass in the H&N showing presence of squamous cell carcinoma, and ability to give written informed consent. (full participant entry seen in full PCOC study protocol – Appendix 1). Patients undergoing treatment at the Cardiff and Vale NHS Trust were identified by head and neck surgical consultants or consultant oncologists in the course of their clinical practice. Biopsies were collected during panendoscopy and placed in transport media to prevent or minimize tissue degradation.

Eleven patients were consented between March 2015 and August 2016 and a corresponding biopsy sample collected. All patients were male, except for one patient (PCOC-14) and their mean age was 65.45 years. Compared to a major cohort study of Schachte *et al.,* (2016), the average age was higher for OPSCC patients in the current study. This could be due to the smaller number of patients participating in the current study compared to previous study cohorts. Biopsies were taken from two different sites of the oropharynx: 6 biopsies were collected from the tonsils and 5 biopsies from base of tongue. All biopsies were taken prior to treatment. Characteristics of each biopsy

regarding age of the patient, sex (gender), smoking status, site of the tumour, TNM stage, p16 status, treatment history and outcome are shown in Table 3-1.

**Table 3-1: Details of patients recruited to the PCOC study from March 2015 to August 2016**

| Study number | Sex (M/F) | Age (years) | Smoking status[A] | Site | TNM Stage[B] | p16 IHC | Treatment[C] |
|---|---|---|---|---|---|---|---|
| PCOC13 | M | 78 | Ex-smoker | Tonsil | T4N2cM0 | Y | Radical RT |
| PCOC14 | F | 81 | NK | Tongue base | T4N2bM1 | NK | Best supportive care (palliative) |
| PCOC15 | M | 68 | Y (>20pack/yr) | Tonsil | T3N2bM0 | Y | induction chemo then CRT |
| PCOC16 | M | 64 | Non-smoker | Tonsil | T4aN2bM0 | Y | induction chemo, then CRT, then neck dissection |
| PCOC17 | M | 68 | Y (100 pack/yr ) | Tonsil | T4aN1M0 | Y | induction chemo, then CRT, then neck dissection |
| PCOC18 | M | 68 | Ex-smoker (10pack/yr) | Tonsil | T3N2bM0 | Y | induction chemo, then CRT |
| PCOC19 | M | 70 | Ex (<10pack/yr) | Tongue base | T4aN2cM0 | Y | induction chemo, then CRT |
| PCOC20 | M | 55 | Never | Tonsil | T2N2aM0 | Y | Neck dissection followed by CRT to primary |
| PCOC21 | M | 52 | Ex (<10pack/yr) | Tongue base | T2N2bM0 | Y | Neck dissection followed by CRT to primary |
| PCOC22 | M | 51 | Never | Tongue base | T4aN2cM0 | Y | induction chemo, then CRT |
| PCOC23 | M | 65 | Ex (10pack/yr) | Tongue base | T3N1M0 | Y | CRT |

[A]Pack-years(pack/yr): Number of cigarette packs smoked per day multiplied by the number of years smoked; not known (NK)
[B] (T: size and extent of tumour; N: number of nearby lymph nodes that have cancer; M: cancer metastasized)
[C]Radiotherapy (RT), Chemo-radiotherapy (CRT)

The biopsy pieces varied in size, depending on the size of the tumour within the patient, and were composed of tumour and surrounding tissue. This was seen in growth of different types of cells, during observation within the first 48 hours until the media was replaced. Some of the unidentified non-adherent cells, and red-blood cells (blood clots) were discarded after the first media changes of the culture. Explants were visible in the collected biopsies after 2-7 days of tissue culture. In general, biopsies were considered unsuccessful if they did not produce explants (outgrowths) after 7 days (although media changes were continued until the last biopsy pieces detached). Some tissue fragments did not attach to the plastic tissue culture plate and a few explants were lost due to low adhesion during the first media change after 48 h of explant production. No explants were lost after the second media change. All biopsies received, produced explants (visible new growth) (Table 3-2). Two of these biopsies were successfully grown over multiple passages (PCOC-17 and PCOC-20).

Extended cell growth from explants or after the first 5-6 passages (p) were unsuccessful for several reasons. Explants for PCOC-13 (p3), PCOC-14 (p4), PCOC-15 (p8), PCOC-16 (p9), PCOC-18 (p4), PCOC-19 (p4) and PCOC-22 (p3), were successfully grown for a limited number of passages. However, cell growth slowed down and after the final passage this was followed by senescence (before senescence, cells appeared bigger and heterogeneous). For PCOC-15, PCOC-16 and PCOC-21 cell growth slowed down after around p4-p5, but limited colony growth appeared, which resulted in transfer from a 10cm tissue culture plate into a 6cm tissue culture plate. After p8 for PCOC-15 and p9 for PCOC-16 the last cells appeared senescent.

PCOC-23 showed a yeast contamination in the plates with the tissue fragments. A treatment with amphotericin B appeared to be successful and the explants were passaged the first time. Immediately after the first passage, yeast reappeared. Further treatment with amphotericin B was unsuccessful; it is possible that yeast cells had developed resistance to amphotericin B after the first treatment of the biopsy pieces (Table 3-2).

**Table 3-2: Growth status of PCOC biopsies**

| Biopsy study number | Explants produced | Max. Passage | Final outcome |
|---|---|---|---|
| PCOC-13 | Yes | p3 | slow growth (senescence) |
| PCOC-14 | Yes | p4 | slow growth (senescence) |
| PCOC-15 | Yes | p8 | slow growth (senescence) |
| PCOC-16 | Yes | p9 | slow growth (senescence) |
| ***PCOC-17*** | ***Yes*** | ***p25*** | ***established cell line*** |
| PCOC-18 | Yes | p4 | slow growth (senescence) |
| PCOC-19 | Yes | p4 | slow growth (senescence) |
| ***PCOC-20*** | ***Yes*** | ***p20*** | ***established cell line*** |
| PCOC-21 | Yes | p8 | slow growth (senescence) |
| PCOC-22 | Yes | p3 | slow growth (senescence) |
| PCOC-23 | Yes | p1 | yeast contamination |

### 3.3. Growth of PCOC-17 and PCOC-20

During the early stages of culture of PCOC-17 and PCOC-20, growth characteristics, colony and cell morphology were assessed. These observations provided useful information about the explants and cell culture (e.g. health and cell types present). Cells were checked for any signs of apoptosis or senescence during early culture from tissue fragments.

CU-OP-17 and CU-OP-20 were successfully grown into established cell lines. These two cell lines were grown to around 20 passages (p25 and p20). Cell outgrowth was visible for PCOC-17 (CU-OP-17) after 4 days and for PCOC-20 (CU-OP-20) after 3 days in culture on GMEM-EGF positive media (Figure 3-1). To support cell growth of the primary culture a layer of post-mitotic 3T3 feeder layer was plated. These 3T3 feeder cells were displaced as the colonies of PCOC-17 or PCOC-20 were growing, as a result of division of peripheral cells (Figure 3-2).

The outgrowth for PCOC-17 (CU-OP-17) reached sufficient size to be passaged after 4 days and PCOC-20 (CU-OP-20) after 7 days in culture. During the first passage the appropriate trypsinisation time was established for both cell lines. This was tested with an initial 5 min incubation with TE and inactivation of media, this process was continued until all attached cells (explants) were collected. Cells were then plated the first time as single cells. Both cell lines produced polygonal cells similar to the shape of keratinocytes. PCOC-17 (CU-OP-17) cells were slightly bigger than the PCOC-20 (CU-OP-20) counterpart (Figure 3-2).

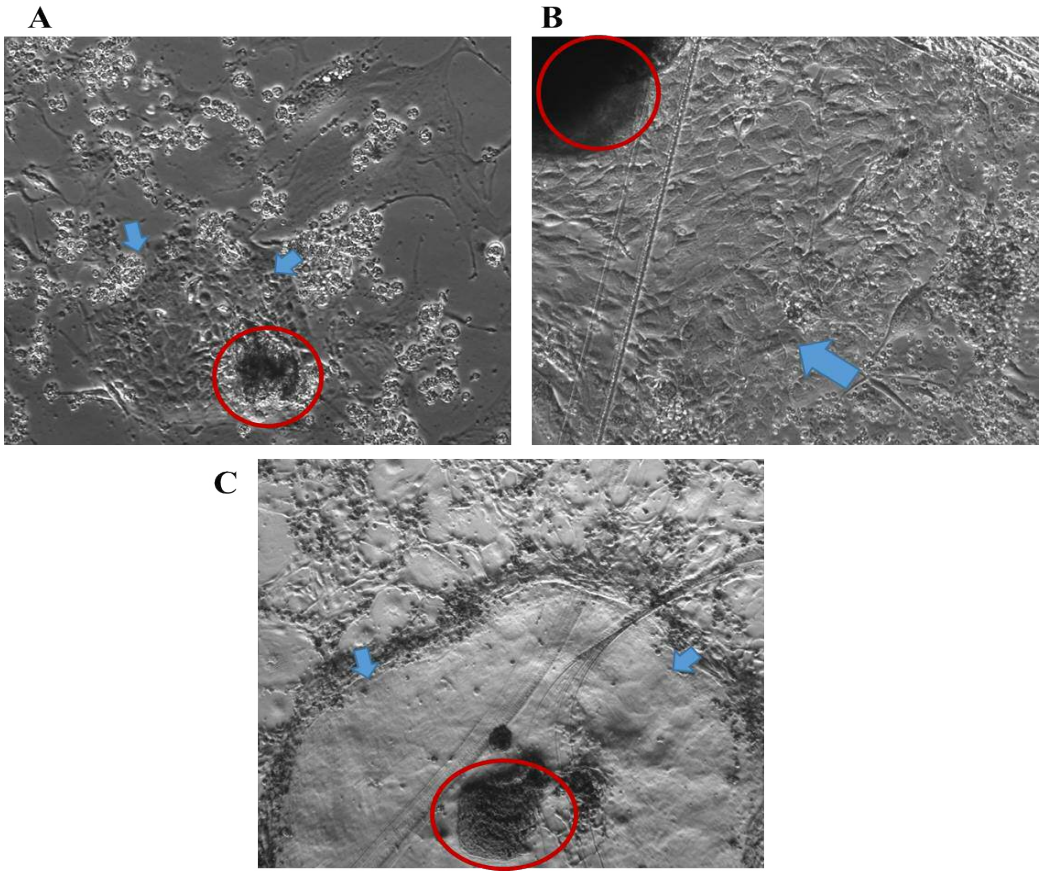



**Figure 3-1: Morphology of PCOC-17 and PCOC-20 explants**

A: PCOC-17 explant growth: Explants, indicated with blue arrows, grew in roughly circular pattern from the biopsy tissue fragment (red circle) – 50X Magnification

B: PCOC-17 outgrowth was of polygonal shape growing tightly – 100X Magnification

C: PCOC-20 explant growth from biopsy tissue fragment in some tight polygonal shape colonies – 50X Magnification

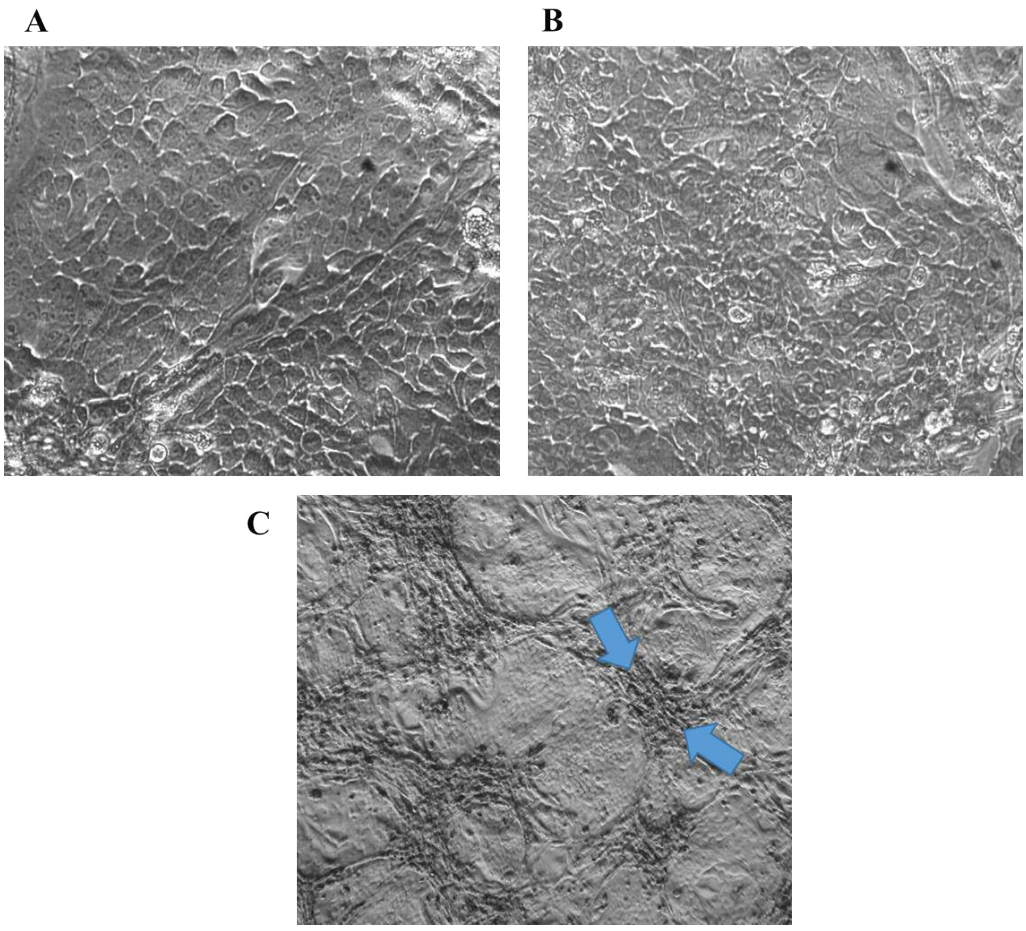



**Figure 3-2: Morphology of CU-OP-17 and CU-OP-20 cells in culture and culture with 3T3 feeder cells**
A: PCOC-17 polygonal cells grew tightly together in a rough circular shape without any gaps between the single cells – 200X Magnification
B: PCOC-20 smaller polygonal cells compared to PCOC-17 (similar circular shape) – 200X Magnification
C: 3T3 feeder cells covering the pates while growing cells out of the biopsy pieces (PCOC) and after support growth for established cell lines (CU-OPs). 3T3 feeder cells are compacted between the grown colonies (e.g. PCOC-17) – indicated with blue arrows.

### 3.4. Differences in colony morphology within cell lines

During tissue culture of PCOC-17 (CU-OP-17) and PCOC-20 (CU-OP-20) two different types of colonies were observed in early passages of these two cell lines. These colonies were classified into two types. Type 1 colonies were characterized by growth as flat monolayers, whereas type 2 colonies comprised more differentiated cells (with colonies growing upwards in multiple cell layers). Type 1 colonies had the tendency to grow faster compared to the type 2 colonies. Type 2 colonies had the appearance of a stratified epithelium, where cells were growing on top of each other. Cells that reached the top of the multilayer detached. Mainly type 1 colonies were observed from early passages in PCOC-17 (CU-OP-17) from p1 onwards with around 80% flat growing colonies. Type 2 colonies were observed for longer in the PCOC-20 (CU-OP-20) cell line until p5 (Figure 3-3). Over time the proportion of flat growing colonies (flat monolayer) increased, and a single pattern of colony growth become dominant; for CU-OP-17 after p5 and for CU-OP-20 for p6 (Figure 3-3).

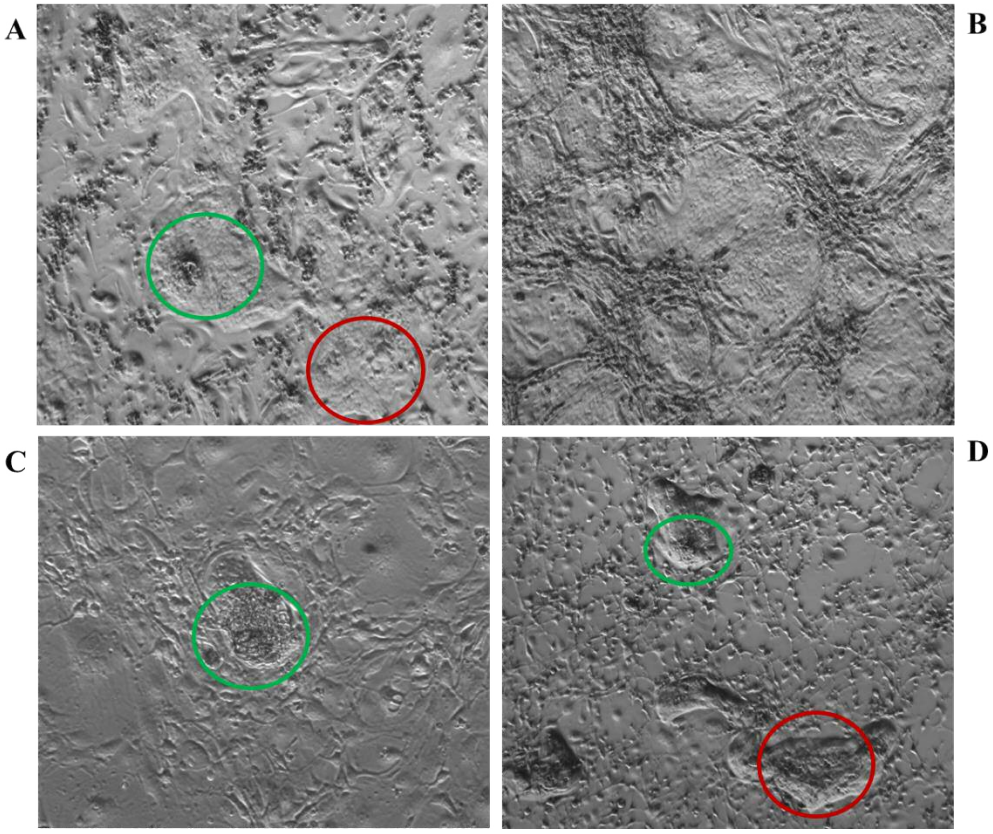



**Figure 3-3: Variation in PCOC-17 (CU-OP-17) and PCOC-20 (CU-OP-20) morphology**

A: PCOC-17 p1: two types of colonies present: 80% of flat colonies, seen with a red circle and 20% differentiated growing (top growing) seen with a green circle. 50X Magnification

B: PCOC17 p5: The amount of multilayer colonies decreased quite quickly and after p5 only flat colonies were present – 50X Magnification

C: PCOC-20 p3: example of multilayer colony after p4 – 50X Magnification

D: PCOC-20 p5: 70% of cells are flat growing – 30% in multilayers – 50X Magnification

### 3.5. Establishing replicative potential of cell lines

It was important to know whether CU-OP-17 (PCOC-17 biopsy) and CU-OP-20 (PCOC-20 biopsy) were immortal cell lines (i.e. with the potential to proliferate indefinitely). Therefore, population doublings (PD) were estimated and compared with the number of PD expected from non-immortal cells. Human Epidermal Keratinocytes (neonatal) (HEKn) from Life Technologies were included in some experiments to allow comparison with non-malignant cells. The manufacturers suggest that these cells can be grown to approximately 30 PD (ThermoFisherScientific, 2009).

Cells were cultured at a range of seeding densities to determine the optimal plating density. The results of these experiments were used to estimate the plating efficiency for CU-OP-17 and CU-OP-20. This indicated a plating efficiency of approximately 10% for both cell lines (i.e. approximately 10% of plated cells divided and produced new colonies). Population doublings of 30 PD for CU-OP-17 and CU-OP-20 were achieved after 6 and 5 passages, respectively. The highest passage achieved for CU-OP17 was p25, which would indicate a PD level of 125. CU-OP-20 was passaged to a maximum of p20, equating to approximately 120 PD.

Stocks of CU-OP-17 and CU-OP-20 were expanded using the earliest available frozen aliquot to obtain a large number of frozen samples with the same passage. At least forty aliquots of CU-OP-17 p11 were frozen. Thirty aliquots of CU-OP-20 p11 were prepared and frozen. As they were expanded from early passages, CU-OP-17 and CU-OP-20 are likely to comprise genetically heterogeneous populations of cells. Long term culture, could result in outgrowth of sub-populations. Experiments were performed using heterogeneous populations as they were judged to be more representative of the original tumour than cultures derived by single cell cloning. Both cell lines showed a similar growth rate as the established cell lines used in this study. This was indicated by the established seeding density for the radio-sensitivity assays, which were in the range of the other known cell lines. The growth rate for the CU-OP cell lines might also be influenced by the support of the 3T3 feeder layer.

## 3.6. Cell line validation

It was important to confirm that the cell lines were derived from the original biopsies. This was achieved using DNA fingerprinting based on PCR amplification Short Tandem Repeat (STR) sequences, which was performed on DNA extracted from the cell lines and from the original biopsies. Mycoplasma testing was undertaken on the established cell lines.

### 3.6.1. STR typing of biopsies & cell lines

STR consist of intronic polymorphisms, which occur frequently in the human genome. They consist of tandemly arranged nucleotide repeat units. The most common approach to identify STR polymorphisms is the amplification of STR loci (microsatellites) by PCR (Tilanus, 2006).

STR typing was performed by Public Health England (PHE) as a paid for service. DNA samples were sent to PHE on Whatman FTA cards with passage number p14 for CU-OP-17 and p12 for CU-OP-20. The data were returned in graphical and tabulated form, with an accompanying report.

To enable future identification of the newly derived CU-OP cell lines (17 and 20), and to confirm that they were derived from the original biopsy, the STR profiles of biopsies and cell lines were compared with each other (Figure 3-4 and Table 3-3; Figure 3-5 and Table 3-3). The STR profiles confirmed that the CU-OP-17 and CU-OP-20 cell lines were derived from the original biopsies. It demonstrated as well that CU-OP-17 and CU-OP-20 differ from each other. The biopsy and cell line of CU-OP-17 share 28 out of 34 alleles, while the CU-OP-20 biopsy and derived cell line share 30 out of 33 alleles. These variations are consistent with loss of heterozygosity (LOH) in cell lines, a variation that is quite common in immortal cell line development (Dracopoli and Fogh, 1983). This genetic instability is common in other head and neck cell lines (Brenner et al., 2010). The LOH took place between the initial outgrowth of explants until the identity of the cell lines was confirmed at a higher passage number through STR typing. The interpretation of the data provided by PHE suggested that this confirmed that both cell lines were generated from the original source material.

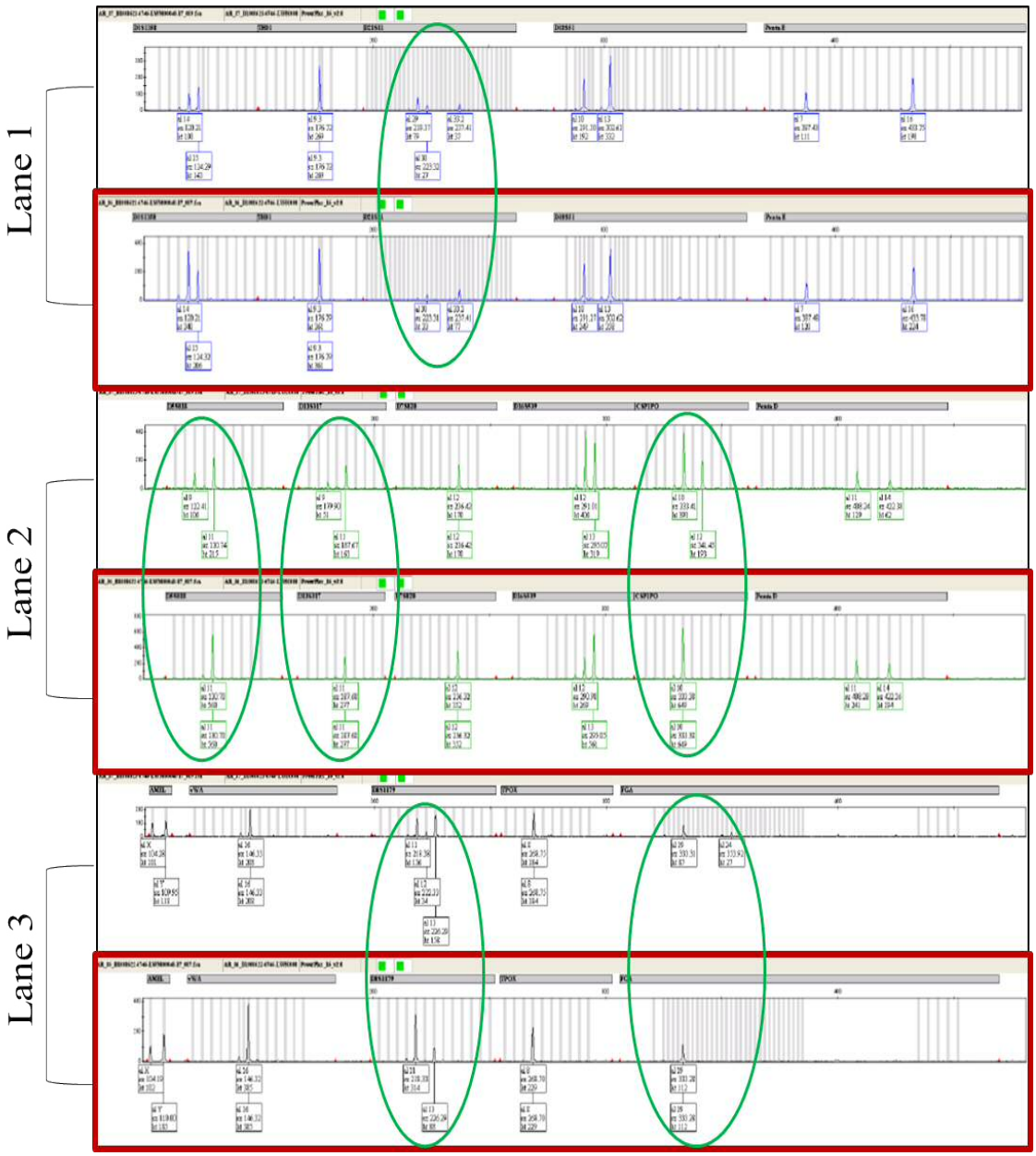



**Figure 3-4: STR profiles of the CU-OP-17 biopsy and derived cell line**
Short tandem repeats were amplified at 16 loci. Alleles are listed in the grey bars above the traces and number of repeats by vertical lines. The biopsy profile is shown in the upper panel of each image pair and outlined in black. The cell line profile is shown on the lower panel of each image pair and outlined in red. The STR profile was separated into three profile lanes The green circles indicate the LOH between the biopsy and the cell lines.

**Table 3-3: STR profiles of the CU-OP-17 biopsy and derived cell line**

| STR locus | Number of repeats at each loci[A] | |
|---|---|---|
| | Biopsy | Cell line |
| D3S1358 | 14, 15 | 14, 15 |
| THO1 | 9.3, 9.3 | 9.3, 9.3 |
| D21S11 | 29, 30, 33.2 | 30, 33.2 |
| D18S51 | 10, 13 | 10, 13 |
| PENTA E | 7, 16 | 7, 16 |
| D5S818 | 9, 11 | 11, 11 |
| D13S317 | 9, 11 | 11, 11 |
| D7S820 | 12, 12 | 12, 12 |
| D16S539 | 12, 13 | 12, 13 |
| CSF1PO | 10, 12 | 10, 10 |
| PENTA D | 11, 14 | 11, 14 |
| AMEL | X, Y | X, Y |
| vWA | 16, 16 | 16, 16 |
| D8S1179 | 11, 12, 13 | 11, 13 |
| TPOX | 8, 8 | 8, 8 |
| FGA | 19, 24 | 19, 19 |

[A] Biopsy and cell line share 28 out of 34 alleles; alleles that were present in the biopsy but were not identified in the cell lines are shown in red.

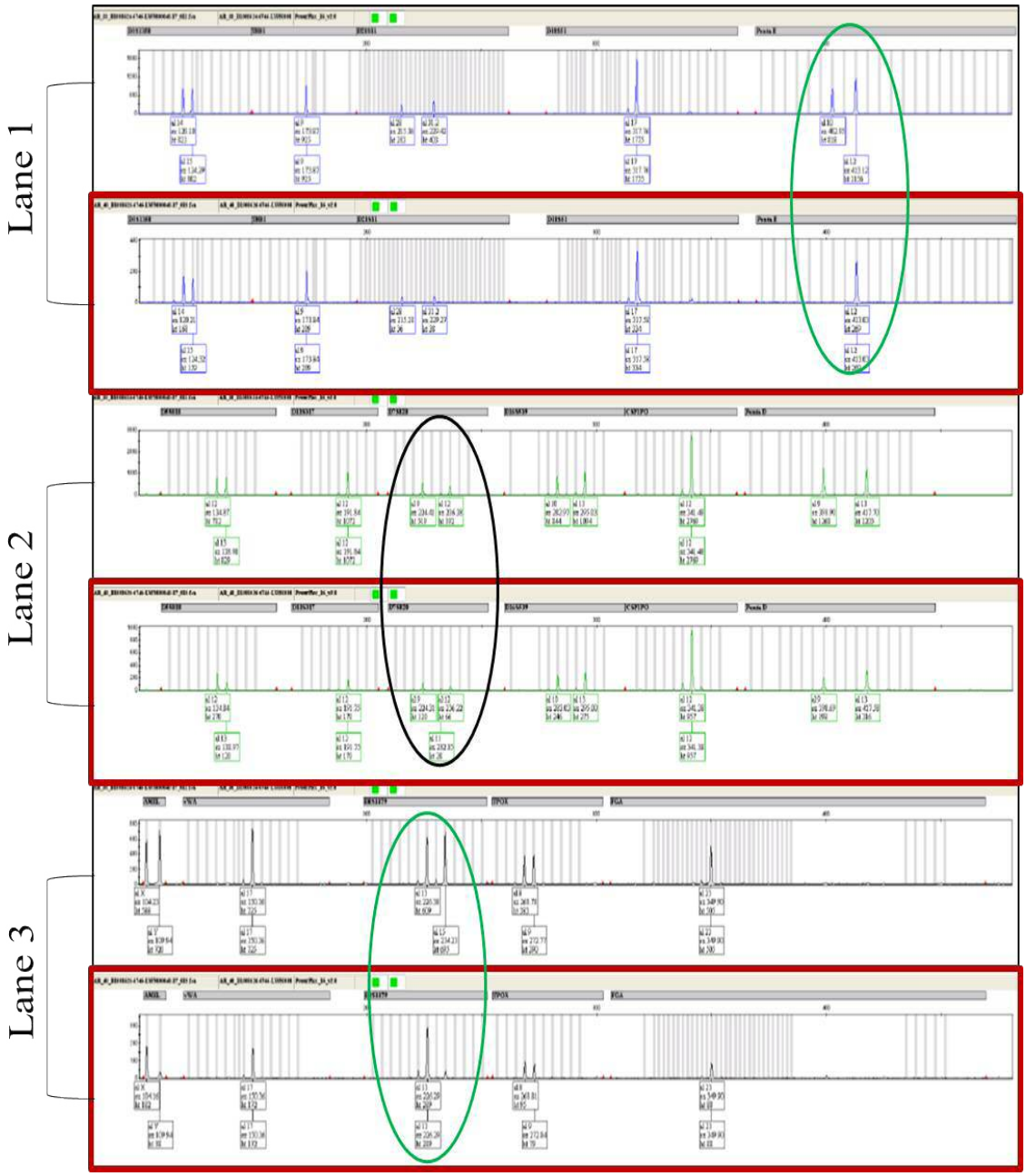



**Figure 3-5: STR profiles of the CU-OP-20 biopsy and derived cell line**
Short tandem repeats were amplified at 16 loci. Alleles are listed in the grey bars above the traces and number of repeats by vertical lines. The biopsy profile is shown in the upper panel of each image pair and outlined in black. The cell line profile is shown on the lower panel of each image pair and outlined in red. The STR profile was separated into three profile lanes The green circles indicate the LOH between the biopsy and the cell lines. The black circle indicates the stutter peak (experimental error), which was not considered as STR peak after receiving the final report.

**Table 3-4: STR profiles of the CU-OP-20 biopsy and derived cell line**

| STR locus | Number of repeats at each loci[A] | |
|---|---|---|
| | Biopsy | Cell line |
| D3S1358 | 14, 15 | 14, 15 |
| THO1 | 9, 9 | 9, 9 |
| D21S11 | 28, 31.2 | 28, 31.2 |
| D18S51 | 17, 17 | 17, 17 |
| PENTA E | 10, 12 | 12, 12 |
| D5S818 | 12, 13 | 12, 13 |
| D13S317 | 12, 12 | 12, 12 |
| D7S820 | 9, 12 | 9, 11, 12 |
| D16S539 | 10, 13 | 10, 13 |
| CSF1PO | 12, 12 | 12, 12 |
| PENTA D | 9, 13 | 9, 13 |
| AMEL | X, Y | X, Y |
| vWA | 17, 17 | 17, 17 |
| D8S1179 | 13, 15 | 13, 13 |
| TPOX | 8, 9 | 8, 9 |
| FGA | 23, 23 | 23, 23 |

[A] Biopsy and cell line share 30 out of 33 allele; alleles that were present in the biopsy but were not identified in the cell lines are shown in red. The extra allele shown in D7S820 is likely to be a stutter peak (Small extra peak immediately before the actual allele. A similar peak is seen in the electropherogram in the biopsy sample, which was not detected).

### 3.6.2. Mycoplasma detection

Mycoplasmas are considered to be the smallest self-replicating organism, with a known size of 0.3 – 0.8 µm in diameter. They are slow growing even under optimal conditions (Drexler and Uphoff, 2002). Mycoplasma are common contaminants in eukaryotic cell cultures, and can cause changes of function, and biochemical and metabolic changes in infected cultures (Drexler and Uphoff, 2002, Miller et al., 2003). CU-OP-17 and CU-OP-20 cultures were tested for mycoplasma contamination by PCR using the Venor GEM kit (Minerva Biosciences). Both lines were confirmed to be free of mycoplasma after amplification and storage.

### 3.7. Assessment of HPV status

HPV status was confirmed via PCR. DNA was extracted from the original biopsy samples and used as template in a PCR targeting a 161 bp region of the HPV-16 E6 gene and a 90 bp of the HPV-18 E7 gene. HPV-16 and HPV-18 are the two most common HPV types for HNSCC worldwide. HPV-16 was chosen as it is the most prevalent, and HPV-18 was chosen because it is the 2$^{nd}$ most prevalent worldwide (Ndiaye et al., 2014). The adequacy of the extracted DNA for PCR based analysis was confirmed through amplification of a 209 bp fragment of the human ß-globin gene. The three PCR reactions were set-up on the same day.

The results of these assays are shown in Figure 3-6. DNA from biopsies PCOC-13 to 23 was tested. Successful amplification of the ß-globin gene showed that all DNA samples were adequate for analysis. All the biopsies tested positive for HPV16 except for PCOC-17 which was negative for both HPV16 and HPV18 DNA.

P16 IHC is now routinely performed on biopsies from patients with suspected OPSCC. Data on p16 IHC for the PCOC patients was obtained from the pathology department at the University Hospital of Wales. p16 is commonly upregulated in following HPV E7 mediated degradation of Rb, and p16 can be used as a surrogate marker for HPV status (although the correlation between HPV status and p16 positivity is not perfect). All biopsies were p16 positive, which is consistent with HPV-16 positive status. However, PCOC-17, tested p16 positive, but was HPV-16 and HPV-18 negative (Figure 3-6).

The HPV status for the established cell lines, of the newly derived cell lines CU-OP-17 (PCOC-17) and CU-OP-20 (PCOC-20) was later confirmed via mRNA sequencing (Chapter 5).

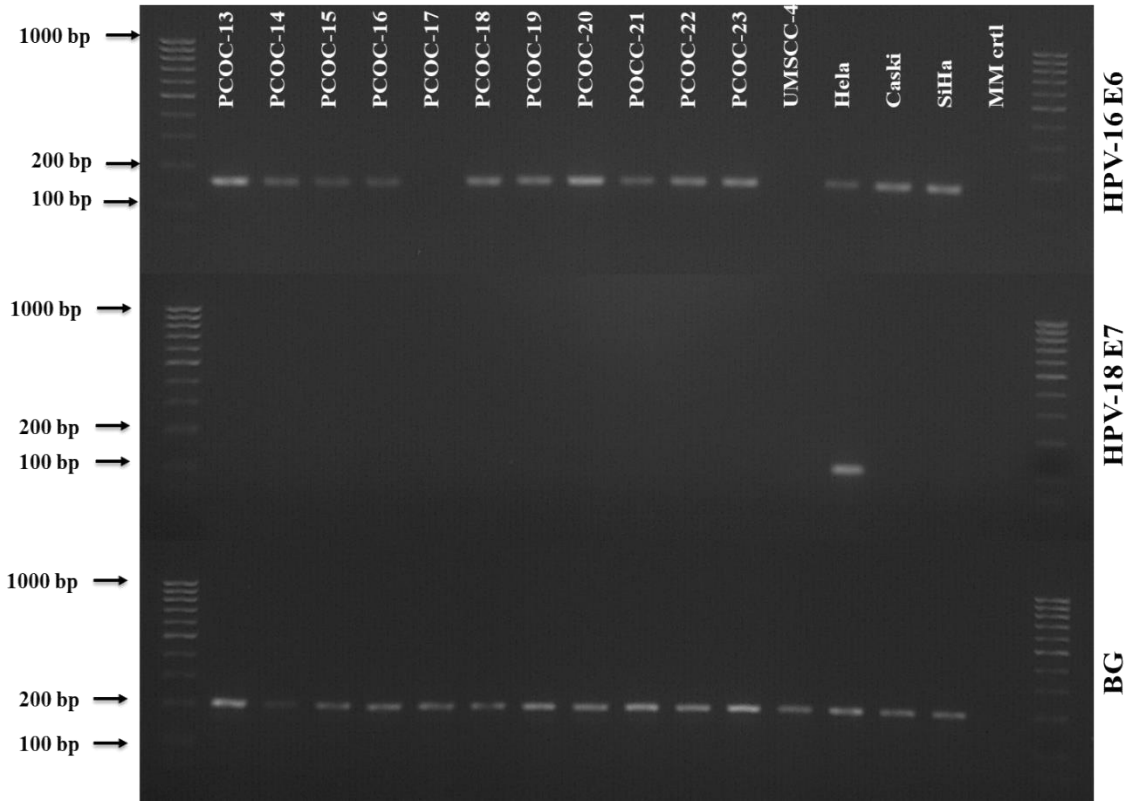



**Figure 3-6: PCR amplification of HPV16 E6, HPV18 E7, and ß-globin fragments in PCOC biopsy DNA**

PCR was performed for each sample for HPV-16 E6 (161 bp), HPV-18 E7 (90 bp) and ß-globin (BG, 209 bp). UMSCC-4 was included as known HPV-negative cell line. A negative control of the corresponding PCR master mix (MM) was added on each gel. HPV-positive controls were Hela (HPV18), Caski and SiHa (both HPV16).

## 3.8. Discussion

One major aims in this study was to develop novel OPSCC cell lines, which was successful with two cell lines (CU-OP-17 and CU-OP-20). The new derived cell lines were helpful to increase the low number of available models. However, there is still a high need to further increase the number of HPV-positive cell lines, as there are still more HPV-negative head and neck cancer cell lines available (Zhao et al., 2011). First of all, generating new HPV-positive head and neck cancer cell lines models should be continued to get closer to the available number of HPV-negative HNSCC cell lines for a better direct comparison of treatment response. Second of all this would give a better reflection and comparison to the general diversity of HPV-positive HNSCC patient in treatment response. Finally, more HPV-positive HNSCC cell lines are needed to mimic the "typical" characteristics of an HPV-positive HNSCC patient (young, fitter, non-smoker). The novel OPSCC cell lines derived in the current and previous study at Cardiff University reflect these characteristics.

For the development of new OPSCC cell lines from biopsies the explant method was used. This method was based on protocols originally developed for culture of HPV positive cervical cancer cells (Stanley, 2002). The success rate for development of new cell lines in the current study was 2 out of 11 received biopsies but this was a labour intensive method. Regular 3T3 feeder change and media change is required to improve successful explant production. The explant method might be improved, by growing tissue pieces using a variety of different media, then choosing the media that provides the best biopsy growth.

The CU-OP-17 (125 PD) and CU-OP-20 cell lines both (120 PD) exceeded the growth span expected of normal human keratinocytes. The cell lines displayed ongoing vigorous growth at 125 and 120 PD respectively, relative to the 30 PD expected lifespan of non-transformed keratinocytes (ThermoFisherScientific, 2009).

The origin of the established cell lines was confirmed via STR typing. Some loss of heterozygosity was observed between the biopsy and cell line profiles, but this is a known phenomenon and the genetic profile of the cell lines can be expected to stabilise over time (Dracopoli and Fogh, 1983). The presence of LOH is usually down to chromosomal loss through non-disjunction or chromosomal rearrangements through double strand breaks (DSB) and Homologous recombination (HR) (Andersen et al., 2008). LOH could be caused by the presence of a tumour suppressor gene in a specific chromosomal region.

However, a gene can only be called a tumour suppressor in many tumour genomes and if the resulting homozygous alleles bear clear and obvious inactivating mutations (Weinberg, 2007). This genetic instability (LOH) was observed in other head and neck cell lines in a study cohort in Michigan (Brenner et al., 2010). Dysregulation of DNA repair factors can cause genomic instability, which is implicated in aging, immune deficiencies or cancer (Broustas and Lieberman, 2014). Genomic instability is a key part of tumour progression (Weinberg, 2007). Another possible cause for the LOH between the cell line and the original biopsy at specific STR loci sites, could be the extended passage of the cell lines. LOH might represent the outgrowth of certain clones present in the heterozygous tumours. *In-vivo* the cancer genome can change during progression, leading to a heterogeneous cell population. However, to maintain an established cell line, routine passaging is essential. In the past serial passaging was not considered as playing a significant role in a "clonal evolution model". In a study with Ishikawa cells a similar observation was made, which showed a heterogeneous copy number in the early passages and distinct profiles in later passages (Kasai et al., 2016). This can be directly connected to the current study, where the LOH could be explained to outgrowth of a novel dominant clone. Therefore, future studies, should consider a model to identify when the LOH has taken place (overgrowth of a dominant clone). Cell lines should be controlled after each passage via STR typing towards the original biopsy. Furthermore, sequencing analysis could identify genomic changes per passage, including differences of single isolated clones of each passage.

It was demonstrated in a previous study that PCOC growth can be successfully supported with post-mitotic 3T3 feeder cells (Pirotte, 2017). This support using 3T3 feeder cells was continued for the PCOC study from PCOC-13 to PCOC-23. Experiments were performed to demonstrate that culture without 3T3 feeders was possible at high cell densities, however cells grew better, especially at low densities with 3T3 support.

Different technologies are available to derive primary tumour cell lines from original tumour samples to be used in research to optimize personalized cancer therapy (Kodack et al., 2017). These technologies include for example, chemical reprogramming, cancer stem cell isolation, sandwich culture, enzymatic degradation and explant (Mitra et al., 2013). More recently, organoid culture has also been developed to support this aim (Drost and Clevers, 2018). In the current study the explant method was used. In this process, the

tissue fragment is cut into small pieces (<1mm) and placed on a FBS coated plate. The explant methods was described in Stanley *et al.,* (2002) and adapted for the start of the PCOC study (Pirotte, 2017, Stanley, 2002). This method was also successfully used for cell line development of vulval cell lines in our laboratory (Bryant et al., 2014a). It was successfully adapted for developing OPSCC cell lines in the same laboratory (Pirotte, 2017). The explant method is also used for culture of keratinocytes isolated from human foreskin (e.g. HEKn) (Orazizadeh et al., 2015). It is a technique that facilitates the retention of native tissue architecture and microenvironment, with better representation of molecular interaction *in-vivo* (Mitra et al., 2013). Furthermore, it has the advantage of reducing the exposure of chemicals such as trypsin to a minimum, as only 1-2 trypsinisation steps are needed. This is desirable as increased chemical exposure could affect cell viability. Compared to the methods using single cells suspension, the explant method provided intact cells within the biopsy piece, even following mechanical disaggregation.

Once cell lines were established, the possibility of single cell cloning was considered, as opposed to retaining cell lines comprising a mixed population of cells. Advantages of the single clone method would have been to investigate differences in HPV integration or expression in various sub clones (Bryant et al., 2014a, Bryant et al., 2014b). However, tumours usually consist of heterogeneous cell populations, and the aim was to mimic this in cell culture with CU-OP-17 and CU-OP-20. This comparison provided the chance to investigate the response of IR on a more heterogeneous cell culture similar to the original tumour.

Long term culture increases the risk of genetic drift (away from the original tumour) but also loss of heterogeneity within the cellular population (Mitra et al., 2013). To reduce the risk of these effects, CU-OP-17 and CU-OP-20 were used at the earliest possible passage numbers (i.e. at the earliest passage after a substantial number of stock aliquots could be generated). In the experiments described in this thesis, the use of cells at p8 to p11, should increase the chance of results being based on a heterologous cell culture population, which better reflects the original tumours. Some loss of heterogeneity did occur though, as evidenced by observing the cell culture microscopically, which showed loss of differentiated multi-layered cell growth (loss of peak shaped colonies – likely to contain normal cells) and establishment of a flat monolayer culture. The loss of colonies showing multilayer cell growth, which possibly contained normal cells, could be due to telomere shortening and senescence. Alternatively, this could reflect diluting out of

differentiated cells that produce fewer progeny over multiple passages. From a cell culture perspective, it would be interesting to clone and characterise the two populations more thoroughly, by HPV typing and/or via deep sequencing.

Senescence, possibly mediated via telomere shortening, may have influenced the success rate of cell line development in the current study. All of the biopsies successfully produced explants (cell outgrowth), and this suggests that the initial biopsy handling and cell culture conditions were appropriate. However, the majority (9/11 – 82%) of the explants did not produce immortal cell lines. Instead they showed reduced growth after p1 to p5 and ultimately stopped proliferating; for example, for PCOC-15 and PCOC-16 at p8 and p9. This could have happened through senescence associated with telomere shortening, alternatively it might reflect stress induced senescence associated with the culture conditions (e.g. due to non-physiological high oxygen tensions). It is interesting to note that many other studies report similarly low success rates in deriving tumour cell lines, though this does appear to vary by tumour site; e.g.(Kodack et al., 2017) report a success rate of 26% based on 568 tumour samples with highest success rates in lung cancers and lowest in breast cancers (29% versus 15%; $p < 0.01$). The success rate of 18% in the current study lies within this range and is consistent with the start of the PCOC study of HPV driven OPSCC tumours (Pirotte, 2017).

CU-OP-17 and CU-OP-20 originated from treatment naïve patients, and so they have not been subject to treatments that are likely to inflict additional mutations or select for subpopulations of possibly resistant cells. The patient from whom CU-OP-20 was derived showed characteristics that are typical of a HPV-positive patient i.e. he was a relatively young (55 years), male and a non-smoker, with an OPSSC of the tonsil. It therefore appears likely to be a valuable new model to investigate the biology of this increasingly common disease.

Conversely CU-OP-17 represents an atypical HPV-negative case, in that it is HPV-negative, but p16 positive (Figure 3-6 and Table 3-1). In a UK and a Dutch study, 14-17% of HPV-negative head and neck cancer patients were identified as p16 positive (Evans et al., 2013, Lewis et al., 2010, Rietbergen et al., 2014). In a later Dutch study, the proportion of p16-positive OPSCC patients (n=388) who tested HPV-negative was 12.4%. This subgroup showed some distinct demographic, clinical and morphologic characteristics. It showed a significantly worse five-year overall survival compared with the HPV-positive OPSCC tumours ($p < 0.001$) (Nauta et al., 2018). It is therefore helpful

to have an additional model for this clinically significant but diagnostically difficult group. P16 is classically regarded as a prominent tumour suppressor gene, and mediator of cellular senescence. It is therefore surprising that p16 expression appears to be upregulated in these tumours, however p16 overexpression has been reported in several other tumours including endometrial, colorectal and basal cell carcinoma (Romagosa et al., 2011), although its function in this context is unclear.

Another interesting point compared to literature was the difference in the average age of OPSCC patients. In a large cohort study (n=1,529) the average age of OPSCC patients was defined as 59.3 years of age compared to this study (n=11) with an average age of 65.45 years of age (Schache et al., 2016). The difference in the average age could be due to the smaller number of patients in the current study. A higher number of patients could better reflect the average age of OPSCC patients, as previously reported (Chu et al., 2013, Schache et al., 2016).

All the PCOC biopsy samples except PCOC-17 tested positive for HPV-16. The high proportion of HPV-positive samples in the current study does not directly reflect the HPV-prevalence in oropharyngeal in the UK. The HPV prevalence in OPSCC in Wales was 55% (2001-2006); in the UK it was 51.8% (2002-2011) (Evans et al., 2013, Schache et al., 2016). However due to the small sample size of the current study, it is not surprising that it does not reflect the general population. It is also possible that patients who were more likely to have HPV-positive tumours may have been selected and consented by the clinical team.

**Strengths and weaknesses**

The methods used in this study had a high success rate in producing explants. This implied that cells had no problems in growing in the culture environment provided over short periods. The low number of established cell lines might be the greatest weakness. The explant increases the chance of the cell lines reflecting at early passages the heterogeneous population of the original tumour, which was considered as a strength in this study. Lower passage numbers were used for all experiments (depending on the number of aliquots). Another strength was the successful confirmation by STR analysis, that the cell lines were derived from the original biopsies. The generation of STR profiles

also facilitates ongoing validation of the lines identities (e.g. when they are provided to collaborators).

## 3.9. Conclusion

Two new OPSCC cell lines were successfully established from fresh biopsies of oropharyngeal cancer. One of the cell lines was identified as HPV-positive OPSCC cell line (CU-OP-20), which reflects the typical HPV-positive patients (relatively young, fit, non-smoker). CU-OP-17 represents one of the few HPV-negative cell lines that is p16 positive. These cell lines are a valuable resource to support investigation of HPV and OPSCC biology.

# Chapter 4 : Investigation of response to ionising radiation in oropharyngeal cancer cell lines

## 4.1. Introduction

The incidence of oropharyngeal squamous cell carcinoma (OPSCC) is increasing and this increase is closely connected to greater incidence of human papillomavirus (HPV)-positive tumours (Evans et al., 2013, Nasman et al., 2015). Most patients with OPSCC will receive radiotherapy as part of their treatment. In general, patients with HPV-positive OPSCC tend to respond better to treatment than patients with HPV-negative OPSCC (Dayyani et al., 2010, Evans et al., 2013).

In the last few years, several studies have investigated the potential mechanisms for the observed superior patient survival associated with HPV-positive OPSCC, by testing responses to radiotherapy using *in-vitro* models. These *in-vitro* models consisted of different panels of HPV-positive and HPV-negative head and neck cancer (HNSCC) cell lines. The main findings were that as a group, HPV-positive cell lines showed greater sensitivity to radiation compared to HPV-negative cell lines, and this was associated with more substantial G2 cell cycle arrest in HPV-positive cell lines after treatment (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013). It has been suggested that the increased sensitivity to radiation was connected to a deficiency in double strand break (DSB) repair in HPV-positive cell lines, which leads to accumulation of unrepaired breaks and apoptosis (Arenz et al., 2014, Rieckmann et al., 2013).

The primary aim of the work described in this chapter was to use clonogenic assays, which assess proliferative capacity after irradiation, to investigate the radio-sensitivity of a panel of OPSCC cell lines, including both HPV-positive and HPV-cell lines. Increased sensitivity to IR might also be partly attributable to residual p53 activity in HPV-positive cases. For example, it is possible that the tumour cell lines used in *in-vitro* studies might have low levels of functional p53 present due to incomplete degradation by the HPV-E6 protein, as opposed to entirely inactive p53 protein following p53 gene mutation (Kimple et al., 2013). Therefore, levels of p53 were also assessed in the current study.

In terms of the link between cell cycle arrest and p53 status, some discrepancies have been reported in the literature. Two studies reported similar findings, where they observed a direct correlation between cell cycle arrest and p53 status (Arenz et al., 2014, Rieckmann et al., 2013), whereas a third study did not observe this connection (Kimple

et al., 2013). This may be explained by different panels of cell lines being in the various studies, although there was some overlap in the lines used. Therefore, effects on cell cycle distribution (G2 arrest) were investigated in this study. Overall, this study aimed to clarify the relationship between HPV, p53 status and radio-sensitivity, *in vitro* using a cell line panel including established and newly generated OPSCC cell lines. However, the intensity of DNA damage caused by IR was not measured, as residual p53 activation was associated with greater sensitivity after IR. After investigation of p53 accumulation/activation, abnormalities cell cycle arrest was investigated (e.g. G2 arrest). The role of DNA repair and DNA damage should be included in future studies to confirm the collected data of survival, p53 activation/accumulation and cell cycle arrest.

***It was hypothesised that***:

1. Sensitivity to ionising radiation would correlate with HPV status; specifically, HPV-positive OPSCC cell lines would show a greater sensitivity to radiation compared to HPV-negative OPSCC cell lines.
2. Increased radiation sensitivity would be associated with accumulation of p53, due to incomplete degradation by HPV E6, in HPV-positive OPSCC cell lines.
    a. Accumulation of p53 will be directly correlated to accumulation of p-p53 and p21.
3. HPV-positive OPSCC cell lines would show greater G2 cell cycle arrest following IR compared to HPV-negative OPSCC cell lines.
    a. Greater G2 arrest would be expected to be greater in HPV-positive OPSCC cell lines due to the involvement of HPV E6 (p53) and HPV E7 (pRb) on cell cycle regulations.

## 4.2. Defining seeding density of OPSSC cells for clonogenic assays

Before clonogenic assays could be used to determine radio-sensitivity, it was necessary to determine the optimal number of cells to use in these assays (i.e. the seeding density). Too low a seeding density would result in too few colonies to achieve statistically significant results, while too high a seeding density would lead to errors in counting due to control plates becoming confluent, and also from colonies not emerging from single cells. The seeding density was defined for the majority of the panel of cell lines during an earlier study (Pirotte, 2017). For the three new cell lines (UMSCC-6, CU-OP-17, CU-OP-20) the seeding density was optimized, as it represents a key parameter for clonogenic assays. The differences observed between the seeding densities per cell line in the current study indicate a difference in doubling time. This could be a possible confounder, which might result in different length of the assay or state of confluence per cell line used. As described above, too confluent plates would lead to counting errors. The effect or influence of differences in doubling time on the proliferation assay was tried to be reduced by defining a specific seeding density. This resulted in a similar length of the assay per cell line (10-15 days),

As previously reported, none of the OPSSC cell lines could initiate and sustain growth from low plating densities (Rieckmann et al., 2013). In a previous study (Pirotte, 2017), it was concluded that post-mitotic 3T3 cells were needed to address the issue of limited cell growth at low seeding density. These “feeder” cells can promote growth by providing cell-cell contact, ECM and growth factors (Rheinwald and Green, 1975b, Rheinwald and Green, 1975a). Use of feeders supports plating of OPSCC cells at low density and so allows clonogenic assays to be assessed after 10-15 days of culture (Pirotte, 2017). Therefore, in this study, all clonogenic assays were set-up with the support of 3T3 feeder cells. Seeding density tests for UMSCC-6, CU-OP-17 and CU-OP-20 were set-up with a range of cell densities, on a layer of post-mitotic 3T3 feeder cells (Figure 4-1). Frequency of media changes, duration of tests and evaluation of the cell lines are described in detail in the Material and Methods chapter. Based on the results of the density tests, optimal seeding densities were defined as 4000 cells for UMSCC-6, 4500 cells for CU-OP-20 and 7000 cells for CU-OP-17 per 10cm culture plate. A 10-15 days’ assay duration then resulted in countable colonies for all cell lines (mimicking the control plate conditions).

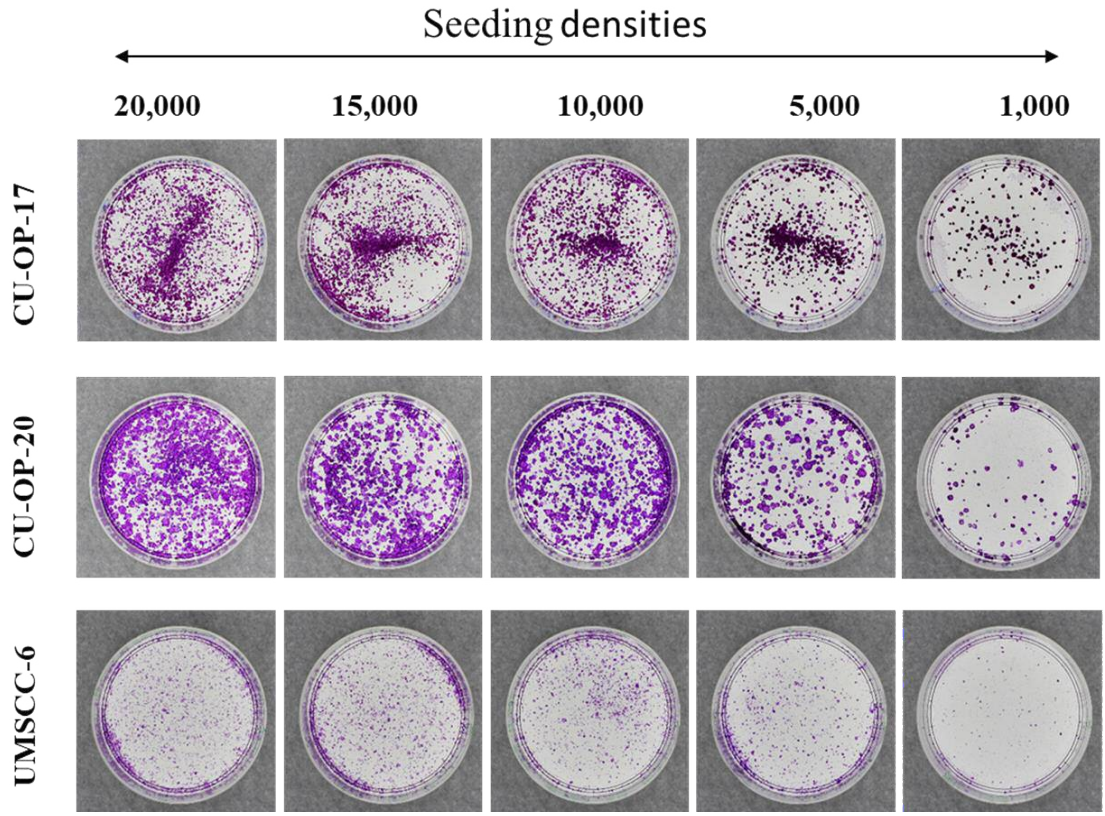



**Figure 4-1: Effect of seeding density with the presence of 3T3 feeder cells on colony formation with UMSCC-6, CU-OP-17 and CU-OP-20**

Cells were seeded at densities between 20,000 and 1,000 cells with the support of 3T3 feeder cells. 3T3 feeder cells were washed off with PBS before crystal violet staining. Plates were incubated at 37°C for 10-15 days, then stained with crystal violet and photographed (images above). Colonies with >50 cells were counted.

## 4.3. Colony formation capacity after ionising radiation

Cell lines were defined as sensitive or resistant to ionising radiation on the basis of the surviving fraction obtained after clonogenic assays. Clonogenic assays were specifically designed for testing response to ionising radiation in cell lines, and are the standard assay used worldwide. They test the ability of cells to proliferate following radiation exposure, i.e. to retain reproductive ability to form large colonies or a clone (hence the term "clonogenic") (Munshi et al., 2005). Clonogenic assays can also be used to test drug sensitivity on cell lines or drugs in combination with radiation (Franken et al., 2006, Munshi et al., 2005, Rafehi et al., 2011).

In a study by West *et al.,* (1993), cell lines derived from early stage cervical cancers were classified as radio-sensitive if they displayed a surviving fraction (SF) of less than 40% (SF < 0.40) and radio-resistant if they showed a SF of more than 40% (SF > 0.40), after 2Gy treatments. This classification appeared to be clinically relevant as survival of cancer patients was shown to correlate with the response of their cultured cells to this dose (West et al., 1993). A similar approach was used in the current study: cell lines were classified as radiation sensitive if the SF post-treatment with 2Gy IR was above or below 40%, In other studies on OPSCC cell lines, lines were simply defined as "less" or "more radio-sensitive" (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013).

To assess response to ionising radiation (IR), cells were treated with IR and the surviving fractions were compared. Examples of the experimental plates are shown in Figure 4-2. The remaining cell lines are shown in Appendix 2. A wide range of responses was observed across the cell lines panel, with greater variation within the HPV-positive OPSCC cell lines (Figure 4-3). The surviving fraction of all cell lines combined (Figure 4-3) showed an overlap between HPV-positive and HPV-negative OPSCC cell lines. The normal human keratinocytes cell lines showed an intermediate response to radiation after 2Gy with a SF of 32.7% (classified as sensitive).

This greater variation in sensitivity was especially observed in HPV-positive OPSCC compared to HPV-negative OPSCC cell lines, following treatment with 2Gy IR, (Figure 4-3). The surviving fraction indicated radio-resistant and radio-sensitive cell lines in the HPV-positive group (Figure 4-3A). At 2Gy, the surviving fractions within the HPV-positive OPSCC cell lines were UMSCC-47 (SF=14.1%), UPCI-SCC-90 (SF=28.3%) and CU-OP-20 (SF=20.4%), therefore classified as radio-sensitive, whereas CU-OP-2

(SF=44.7%) and CU-OP-3 (SF=60.4%) were classified as radio-resistant. CU-OP-2 and CU-OP-3 were only sensitive to IR after 4Gy to 6Gy treatment (Table 4-1).

Significant differences in SF were measured between the HPV-positive OPSCC cell lines at 2Gy treatment between the most resistant and sensitive cell lines, UMSCC-47 to CU-OP-2 ($p=0.0001$, two-way ANOVA), UMSCC-47 to CU-OP-3 ($p<0.0001$, two-way ANOVA), UPCI-SCC-90 to CU-OP-3 ($p<0.0001$, two-way ANOVA), CU-OP-3 to CU-OP-20 ($p<0.0001$, two-way ANOVA) and CU-OP-2 to CU-OP-20 (as shown in Figure 4-3.

**Table 4-1: Surviving fraction of HPV-positive OPSCC cell lines after treatment doses from 0.5Gy to 6Gy**

| Treatment dose (Gy) | Surviving fraction | | | | |
|---|---|---|---|---|---|
| | UMSCC-47 | UPCI-SCC-90 | CU-OP-20 | CU-OP-2 | CU-OP-3 |
| 0.5 | 78.4 | 72.5 | 66.8 | 93.6 | 90.5 |
| 1 | 44.4 | 48.5 | 46.9 | 64.1 | 83.3 |
| 2 | 14.1 | 28.3 | 20.4 | 44.7 | 60.4 |
| 4 | 1.3 | 7.5 | 3.8 | 12.2 | 19.1 |
| 6 | 0.3 | 1.7 | 0.8 | 7.2 | 12.1 |

HPV-negative OPSCC cell lines shared a similar pattern after IR treatment with the majority being radio-resistant after 2Gy treatment (Table 4-2). The only HPV-negative cell line with a SF below the 40% threshold defined at the beginning of the assay, was UMSCC-19 with a SF of 32.4%, which was therefore classified as radio-sensitive. All other HPV-negative OPSCC cell lines were classified as radio-resistant with an SF > 40% (Table 4-2; Figure 4-3).

Within the HPV-negative OPSCC treated with 2Gy IR, the greatest significant difference was observed between the most sensitive cell line UMSCC-19 and the most resistant cell lines CU-OP-17 ($p<0.0001$, two-way ANOVA). Treatment at 0.5 to 1Gy did not show an obvious response to IR, and a limited number of colonies were detected after 4Gy and 6Gy in the majority of cell lines. Therefore, the focus was placed on IR response after 2Gy treatment within the panel of cell lines (Figure 4-4). The mean SF for HPV-positive

OPSCC cell lines was 33.6% and for HPV-negative OPSCC cell lines was 51.4%; this difference was not statistically significant (t-test, p = 0.1379).

**Table 4-2: Surviving fraction of HPV-negative OPSCC cell lines and HEKn after treatment doses of 0.5-6Gy**

| Treatment dose (Gy) | Surviving fraction | | | | | |
|---|---|---|---|---|---|---|
| | **UMSCC-19** | **UMSCC-6** | **UMSCC-74a** | **UMSCC-4** | **CU-OP-17** | **HEKn** |
| 0.5 | 81.7 | 87.4 | 85.3 | 92.3 | 94.3 | 79.8 |
| 1 | 69.1 | 62.9 | 70.1 | 85.1 | 84.2 | 70.6 |
| 2 | 32.4 | 44.9 | 47.7 | 60.1 | 71.8 | 32.7 |
| 4 | 15.5 | 11.3 | 8.9 | 15.6 | 28.8 | 17.6 |
| 6 | 3.8 | 2.4 | 0.9 | 2.3 | 10.3 | 1.6 |

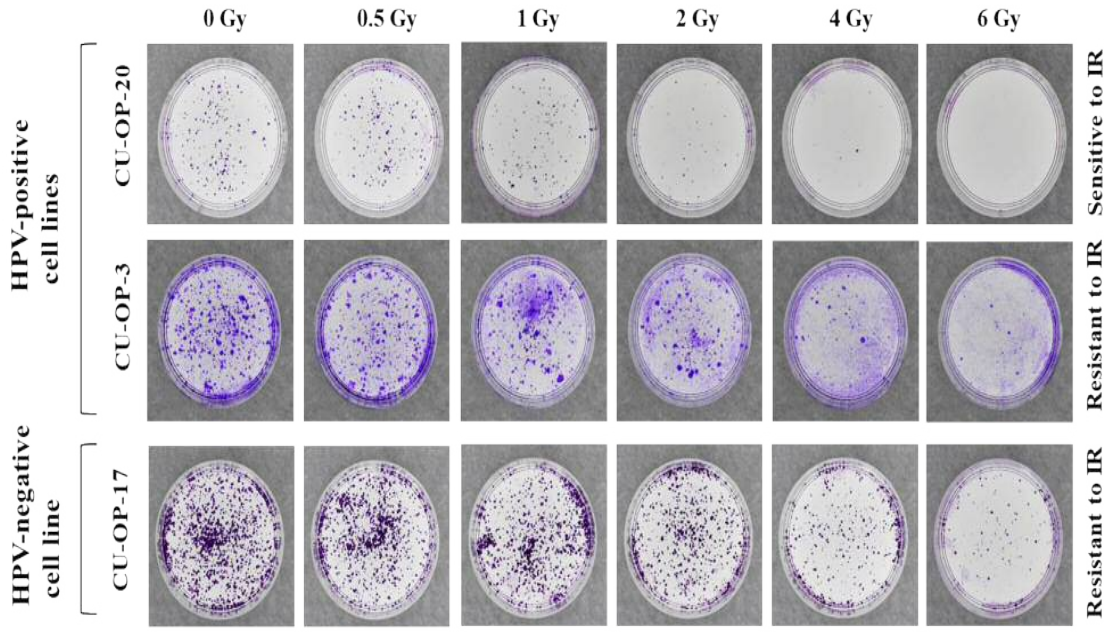



**Figure 4-2: Survival and proliferation following IR in CU-OP-20, CU-OP-3 (HPV-positive) and CU-OP-17 (HPV-negative) cell lines**
CU-OP-20 is representative of a sensitive cell line after 2Gy treatment; CU-OP-3 is a resistant cell line after 2Gy treatment. CU-OP-17 represents an HPV-negative cell line with a high resistance after 2Gy treatment. Cells were plated at low density for 24 hours before treatment with varying doses of IR. The media was changed and the plates were incubated for 10-15 days. Plates were then washed, stained with crystal violet, and photographed (image above). Colonies >50 cells were counted.

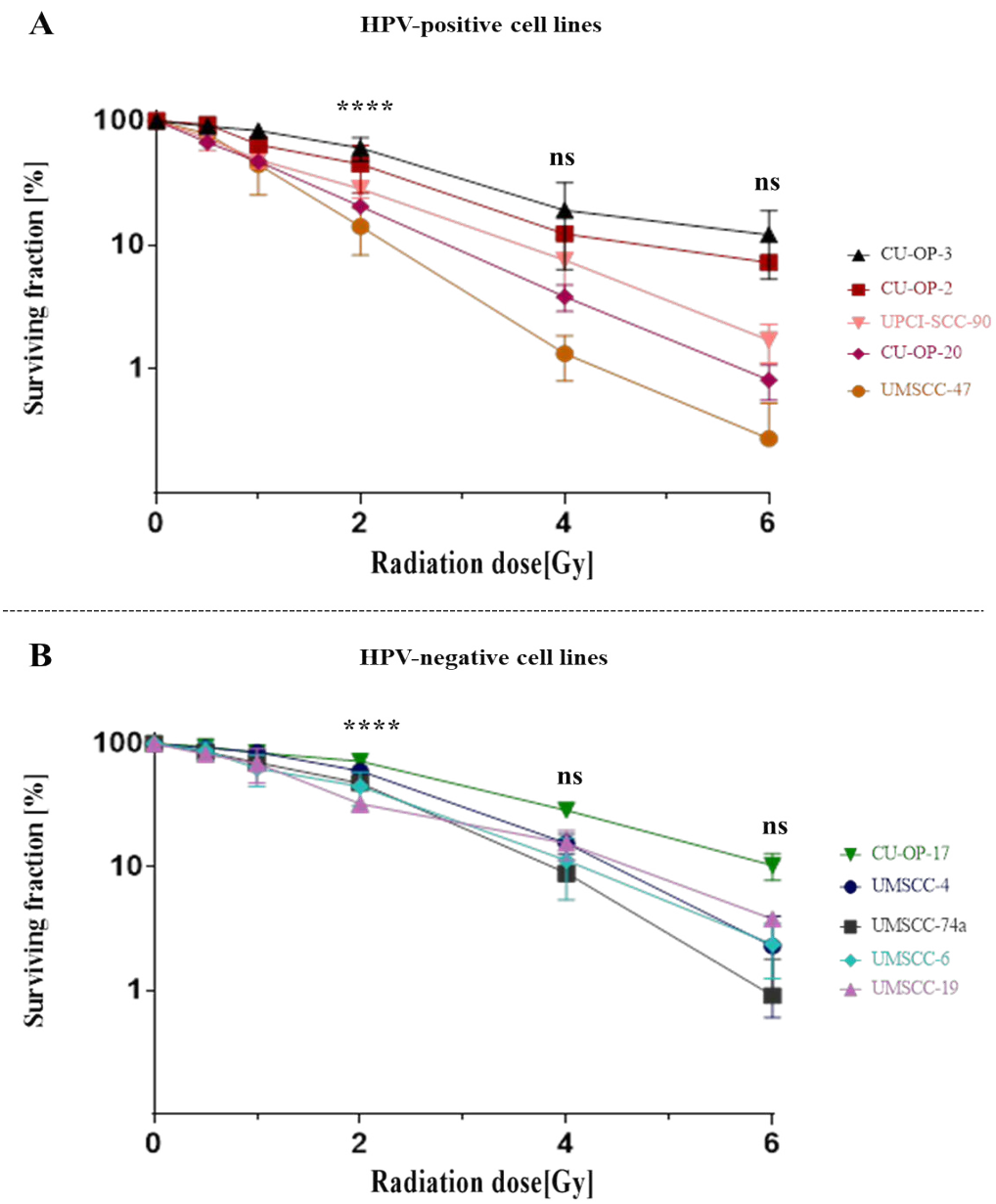



**Figure 4-3: Radiation response of HPV positive and negative OPSCC cell lines**
Surviving fraction of cells after varying doses of IR was measured in a clonogenic assay.
The mean results presented are representative of at least 3 experiments, including triplicates within each experiment. A two-way ANOVA with a Sidak post-test was used to test for statistical significance. Error bars indicated standard deviation (SD). ns stands for not significant.
(A) Surviving fraction of all HPV-positive OPSCC cell lines: Greater variation in SF was observed among the HPV-positive OPSCC cell lines in response to IR, relative to the HPV-negative lines. The colony formation capacities of UMSCC-47, UPCI-SCC-90 and CU-OP-20 were reduced after 2Gy treatment. An example of a comparison between UMSCC-47 and CU-OP-3 is illustrated (****p-value: <0.0001 at 2Gy). ns stands for not significant
(B) Surviving fraction of all HPV-negative OPSCC cell lines: HPV-negative cell lines shared a similar pattern of response to IR. UMSCC-19 was the most sensitive cell line within HPV-negative cell lines after 2Gy. As an example the result of the comparison between UMSCC-19 and CU-OP-17 is illustrated (****p-value <0.0001 at 2Gy). ns stands for not significant

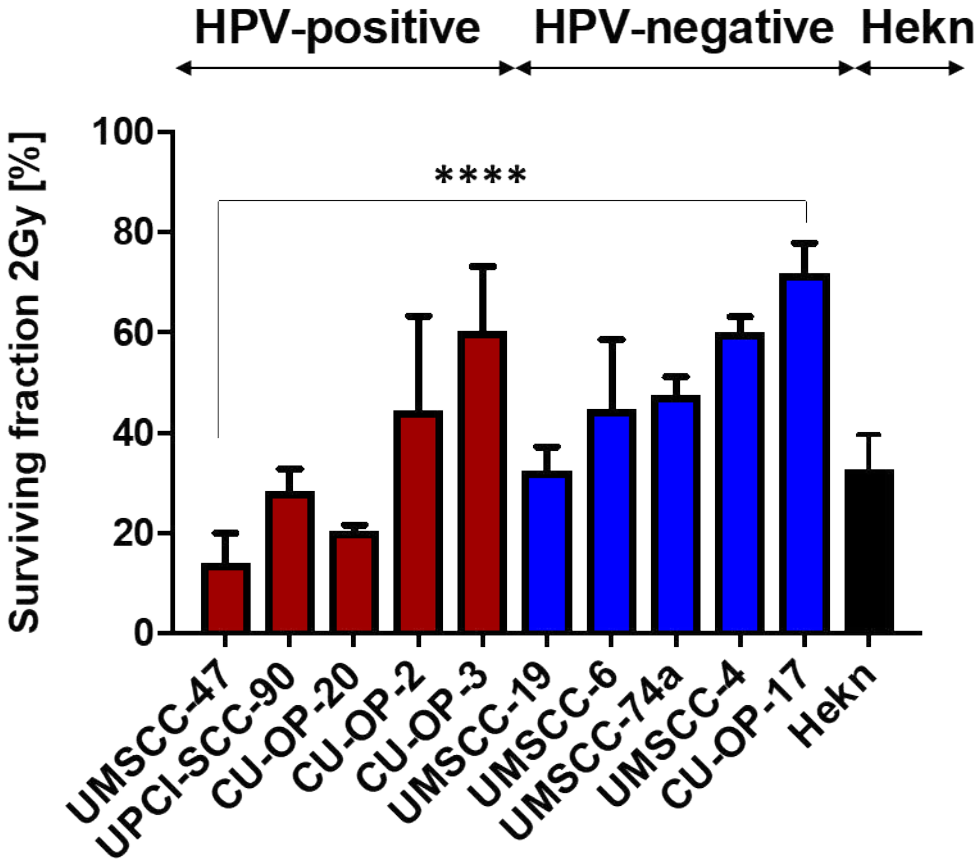



**Figure 4-4: Summary of response to 2Gy irradiation in OPSCC cell lines**
Depending on their response to treatment with 2Gy IR, the cell lines were classified as sensitive or resistant. Within the HPV-positive cell lines 3 cell lines were defined as sensitive to IR (UMSCC-47, UPCI-SCC-90 and CU-OP-20). The HPV-negative cell lines shared a similar pattern after 2Gy, 4 cell lines were defined as resistant, and only UMSCC-19 was defined as radio sensitive cell line. Significant differences were observed between the most sensitive HPV-positive cell line to the most resistant with a p-value range of <0.0001 to 0.0216. Statistical significance was assessed using two-way ANOVA with a Sidak post-test. Error bars indicate standard deviation. An example comparison between the most sensitive HPV-positive OPSCC cell line (UMSCC-47) and the most resistant HPV-negative cell line (CU-OP-17) is illustrated (**** p-value: <0.0001).

## 4.4. Molecular effect of ionising radiation

Cells respond to IR, an important physiological stress factor, by activating multiple signalling pathways, including induction of DNA repair or apoptosis (Braunstein et al., 2009). DNA damage, as induced by IR, usually triggers an increase in p53 protein, which leads to three major possible outcomes: cell cycle arrest, DNA repair and/or apoptosis (Blagosklonny, 2002, Haupt et al., 2003). The induced DNA damage leads to phosphorylation of p53 at Ser15 and Ser20 which reduces interaction with the negative regulator Mdm2. This promotes p53 accumulation and activation in response to DNA damage (Shieh et al., 1997). Phosphorylation at Ser15 of p53 has a key role in up-regulation and activation of p53 (Tibbetts et al., 1999). This phosphorylation was reported to happen in conjunction with CDKs at S and G2/M cell cycle phase, which promotes p53 binding to Waf1/p21 binding sites. Waf1/p21/Cip1, which is induced by p53 is found in several kinase complexes, including those in the S and G2 phase (Wang and Prives, 1995). Therefore, levels of p53, its activated form phospho-p53 (Ser15), and CDKN1A (p21) (a cyclin dependent kinase inhibitor whose transcription is regulated by p53) were assessed in the panel of OPSCC cell lines. This analysis was undertaken to determine if response to IR was mediated through p53 dependent or independent pathways.

The time point of 24h was chosen to investigate a long term effect of p53 after IR towards a correlation of greater sensitivity to IR. This was correlated with literature, where stronger bands were seen after 24h after IR (Kimple et al., 2013). Activation of p53 was detected via p-p53. To investigate the 24h effect on p53 activation the same time point was chosen for p-p53.

### 4.4.1. p53, phosho-p53 and p21 expression in OPSCC cell lines after IR treatment

Expression of p53 was detected in all 5 HPV-positive OPSCC cell lines (Figure 4-5 to 4-8), but with substantial variation. By contrast, expression of phospho-p53 (p-p53) was not reliably detected by western blot in any of the HPV-positive cell lines.

In UMSCC-47, equally strong expression of p53 was found before and after irradiation. However, in this cell line the strongest p21 expression was observed in the non-irradiated (untreated) sample. The p21 signal expression became weaker after treatment of 2Gy and 6Gy (approximately 50% and 25% of the original signal respectively). No expression of phospho-p53 was detected in UMSCC-47. CU-OP-2 showed a similar expression patterns to UMSCC-47 for p53, p21 and phospho-p53.

Weaker expression of p53 was seen in UPCI-SCC-90, CU-OP-3 and CU-OP-20. Each of these 3 cell lines demonstrated stronger expression of p21 compared to levels of p53. No activation of p53 was detected, as indicated by the absence of phospho-p53. Expression of p53, phospho-p53 and p21 was detected in all HEKn 10Gy samples (positive control).

Among the HPV-negative cell lines variation in the levels of p53 was observed as well (Figures 4-7 and 4-8). For the cell lines UMSCC-4, -6 and -19, no visible expression of p53 and phospho-p53 was seen whereas UMSCC-74a and CU-OP-17 showed some accumulation of p53 and activation of p53 (p-p53) after 2Gy and 6Gy treatment. Weak expression of phospho-p53 was seen in both untreated samples. p21 expression was observed in all HPV-negative cell lines, despite the absence of p53.

In HEKn cells, accumulation of p53 was evident, as well as p-p53 expression after 2Gy and 6Gy treatment. P21 expression appeared to be stronger than p53 in every HEKn sample tested.

In summary, p53 was detected in all HPV-positive cell lines (5 out of 5 cell lines), whereas only two HPV-negative cell lines demonstrated detectable expression of p53 (2 out of 5 cell lines). Phospho-p53 expression was difficult to detect in the majority of cell lines tested, however there was accumulation of phospho-p53 in UMSCC-74 and CU-OP-17 after 2Gy and 6Gy treatment. p53-independent transcription of p21 was seen in 3 out of 5 HPV-negative cell lines. In the majority of cell lines tested, p21 expression became weaker after IR treatment, which might indicate a negative relationship between IR treatment and p21 expression in the majority of the HPV-negative cell lines (Stuart

and Wang, 2009). The majority of HPV-positive cell lines showed a positive association between IR treatment and levels of p21.

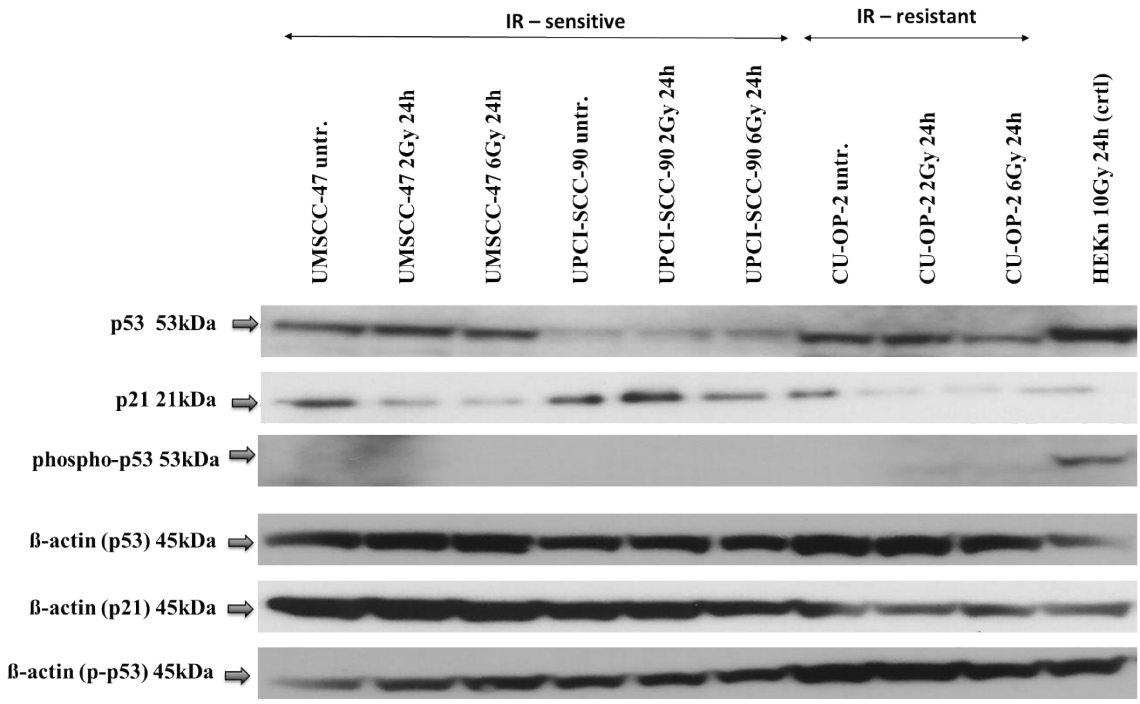



**Figure 4-5: Expression of p53, p21 and phospho-p53 for UMSCC-47, UPCI-SCC-90 and CU-OP-2 (HPV-positive) cell lines**

Western blot images. Cell line samples were arranged in order of treatment dose i.e. untreated to 2Gy and 6Gy IR treatment (collected after 24h incubation after IR treatment). HEKn treated with 10Gy and harvested after 24h was added as positive control. p53, p21 and phospho-p53 were run on three individual gels/membranes (3 membranes used). Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) was tested for all three individual runs. p53 (45min exposure) bands are seen in all three cell lines, with minor differences in expression level. The second membrane was probed for p21 (25-40 min exposure). The third membrane was probed with phospho-p53 (45min exposure).

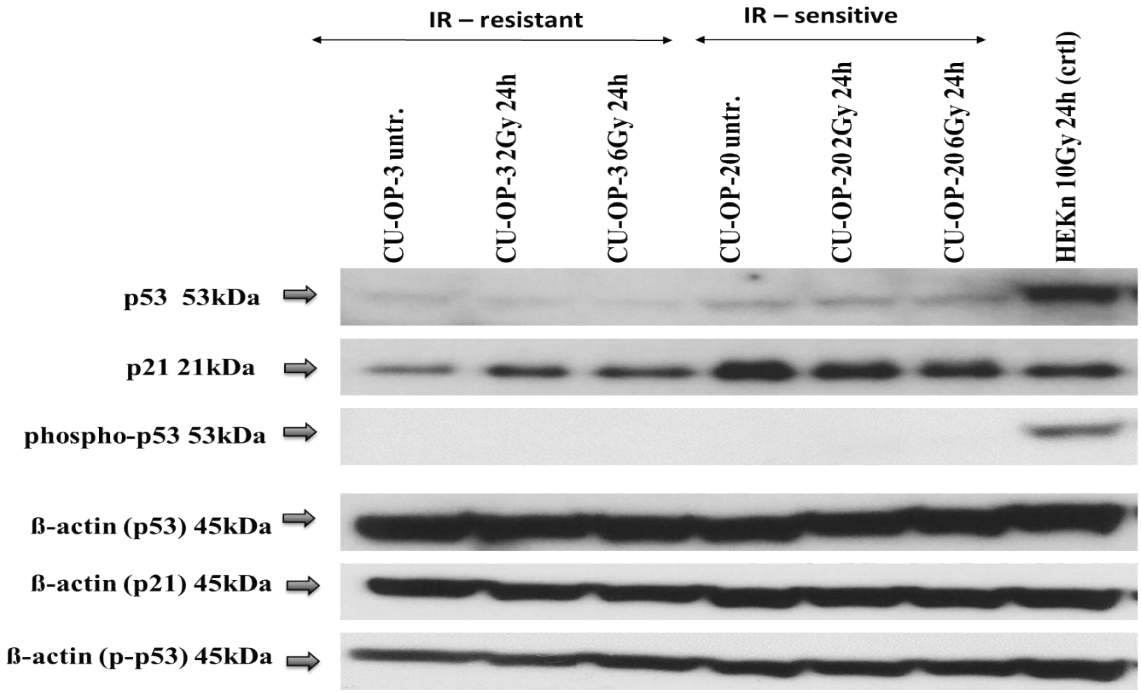



**Figure 4-6: Expression of p53, p21 and phospho-p53 for CU-OP-3 and CU-OP-20 (HPV-positive) cell lines**

Western blot images. Cell line samples were arranged in order of treatment dose i.e. untreated to 2Gy and 6Gy IR treatment (collected after 24h incubation after IR treatment). HEKn treated with 10Gy and harvested after 24h was added as positive control. p53, p21 and phospho-p53 were run on three individual gels/membranes (3 membranes used). Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) was tested for all three individual runs. p53 (45min exposure) bands are seen in all three cell lines, with minor differences in expression level. The second membrane was probed for p21 (25-40 min exposure). The third membrane was probed with phospho-p53 (45min exposure).

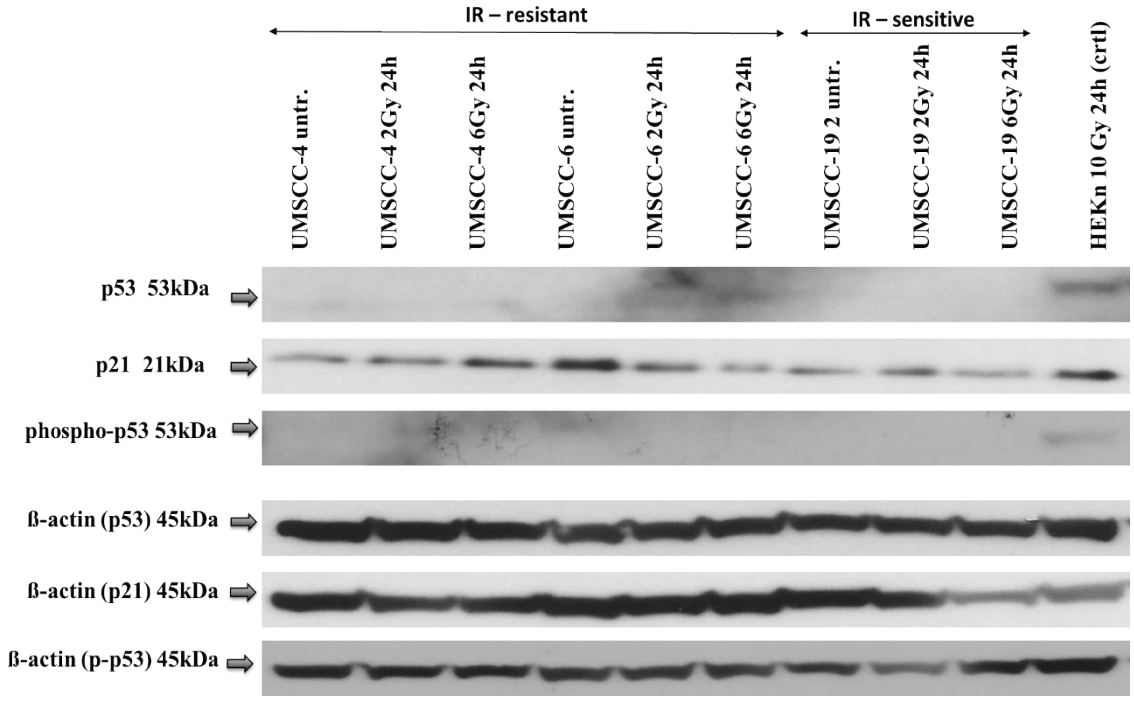



**Figure 4-7: Expression of p53, p21 and phospho-p53 for UMSCC-4, UMSCC-6 and UMSCC-19 (HPV-negative) cell lines**

Western blot images. Cell line samples were arranged in order of treatment dose i.e. untreated to 2Gy and 6Gy IR treatment (collected after 24h incubation after IR treatment). HEKn treated with 10Gy and harvested after 24h was added as positive control. p53, p21 and phospho-p53 were run on three individual gels/membranes (3 membranes used). Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) was tested for all three individual runs. p53 (45min exposure) bands are seen in all three cell lines, with minor differences in expression level. The second membrane was probed for p21 (25-40 min exposure). The third membrane was probed with phospho-p53 (45min exposure).

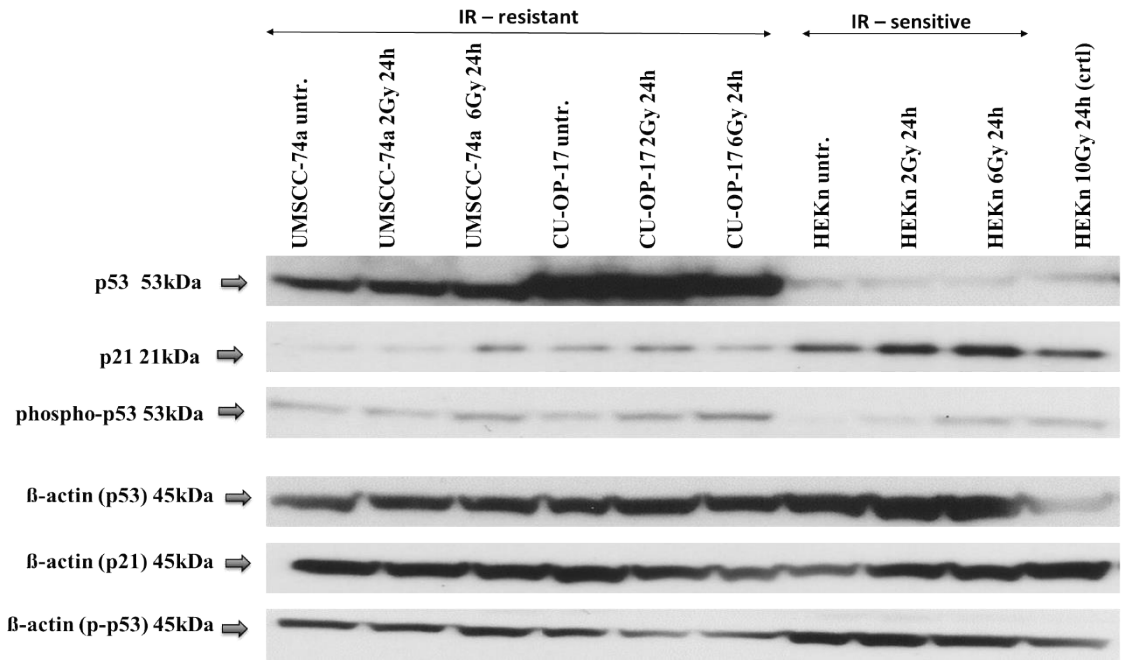



**Figure 4-8: Expression of p53, p21 and phospho-p53 for UMSCC-74a, CU-OP-17 (HPV-negative) and HEKn (normal human keratinocytes) cell lines**
Western blot images. Cell line samples were arranged in order of treatment dose i.e. untreated to 2Gy and 6Gy IR treatment (collected after 24h incubation after IR treatment). HEKn treated with 10Gy and harvested after 24h was added as positive control. p53, p21 and phospho-p53 were run on three individual gels/membranes (3 membranes used). Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) was tested for all three individual runs. TP53 (45min exposure) bands are seen in all three cell lines, with minor differences in expression level. The second membrane was probed for p21 (25-40 min exposure). The third membrane was probed with phosho-p53 (45min exposure).

## 4.5. Cell cycle response following IR

The original hypothesis, that HPV-positive cell lines would show an increased G2 cell cycle arrest following IR compared to HPV-negative OPSCC cell lines. It was based on the expected clear difference between HPV-positive and HPV-negative OPSCC cell lines after the radio-sensitivity assays (Chapter 4). The hypothesis was slightly revised according to the data obtained within the radio-sensitivity assays, where a greater variation within the HPV-positive cell lines and some overlap with the HPV-negative cell lines in response to IR was seen. Furthermore, no clear difference was observed between the HPV-positive and HPV-negative OPSCC cell lines in terms of levels of phospho-p53 (sensitive and resistant cell lines), or p21. Accumulation of p53 was seen in all HPV-positive cell lines with some variation, whereas only 2 out of 5 HPV-negative cell lines showed accumulation of p53. Differences in doubling time of each cell line, as indicated in different seeding densities, might play a role in differences in cell cycle arrest.

**Hypothesis:**

- Sensitivity to IR within the HPV-positive cell lines is associated with greater G2 arrest after IR, due to absence of functional p53 (p-p53).
    - This predicts that the IR sensitive cell lines UMSCC-47, UPCI-SCC-90 and CU-OP-20 will differ from the IR resistant cell lines CU-OP-2 and CU-OP-3.
- HPV negative OPSCC cell lines with absence of p53 and p-p53 will have a greater accumulation of cells in S-phase or G2 phase after IR.

### 4.5.1. IR leads to cell cycle arrest in OPSCC cell lines

The effect of ionising radiation induced DNA damage on progression through the cell cycle was investigated using flow cytometry (detailed description of set-up in Chapter 2). The mean distribution of cells between the phases of the cell cycle was calculated based on three experimental runs. The experiment was designed to compare the data obtained for each treatment dose and time point (8, 24 and 48h) to the untreated sample of "time zero", to identify changes in cell cycle distribution after IR treatment. Therefore, no additional untreated samples were included after 8, 24 and 48h. The collection time of the

untreated sample is equal to the time of treatment of the reaming plates for the cell cycle analysis.

The data for HPV-positive OPSCC cell lines are presented in Figure 4-9, for HPV-negative OPSCC cell lines in Figure 4-10 and for HEKn cell line in Figure 4-11. One representative flow cytometry experiment for each cell line is shown in Figure 4-12 for HPV-positive OPSCC cell lines, in Figure 4-13 for HPV-negative OPSCC cell lines and in Figure 4-14 for the HEKn cell line.

A similar distribution between the phases of the cell cycle was observed in the untreated samples across the panel of OPSCC cell lines with an average percentage in G1 around 50-60% (Figure 4-9). Interestingly, HEKn showed a slightly higher proportion of cells in G1 than the other cell lines (both HPV-positive and HPV-negative cell lines), which suggests a lower proportion of actively cycling cells (Figure 4-11). Other cell lines, such as UMSCC-19, UMSCC-4, CU-OP-17and CU-OP-3 showed fewer cells in G1 phase compared to the untreated sample, indicating a higher proportion of proliferating cells within these cell lines. (Figure 4-9 and Figure 4-10). However, a significant difference in G1 arrest was seen between the untreated control sample and 2Gy 8h for UMSCC-47 (two-way ANOVA, $p=0.0031$), UMSCC-74a (two-way ANOVA, $p=0.0077$) and UMSCC-4 (two-way ANOVA, $p<0.0001$). UMSCC-74a was the only cell line, which showed a significant increase of cells in G1 compared to the untreated control sample after 2Gy and 24h (two-way ANOVA, $p=0.0037$). CU-OP-17 showed a significant decrease in G1 compared between the untreated control sample and 2Gy 24h (two-way ANOVA, $p=0.0402$). After 6Gy treatment UMSCC-47 (two-way ANOVA, $p=0.0056$), UPCI-SCC-90 (two-way ANOVA, $p<0.0001$), UMSCC-4 (two-way ANOVA, $p=0.0307$) and UMSCC-19 (two-way ANOVA, $p=0.0312$), showed a significant difference after 24h compared to the untreated control sample.

Eight hours after treatment with IR, the proportion of cells in S-phase increased in all cell lines compared to the corresponding untreated sample with significant differences after 2Gy for UMSCC-47 (two-way ANOVA, $p=0.0002$) and UMSCC-4 (two-way ANOVA, $p<0.0001$). Significant differences were observed after 6Gy for UMSCC-47 (two-way ANOVA, $p=0.0004$), UPCI-SCC-90 (two-way ANOVA, $p=0.0123$), CU-OP-2 (two-way ANOVA, $p=0.0144$) and UMSCC-4 (two-way ANOVA, $p=0.0061$), but subsequently decreased over 24-48 hours (Figure 4-9 to 4-10) with no significant differences after 2Gy and 6Gy treatment in the entire cell line panel. This would be consistent with an initial arrest, with a subsequent return to cell cycle following DNA repair, or with apoptosis of

arrested cells. The highest proportion of cells in S-phase were detected in HPV-negative OPSCC cell lines e.g. UMSCC-4, only had a relatively limited number of cells in G1 and G2 after 8 hours (2 and 6Gy) (Figure 4-10). This decrease of cells in G1 was significant for UMSCC-4 between the untreated sample (time zero) and for 2Gy (two-way ANOVA, $p<0.0001$) and 6Gy treatment (two-way ANOVA, $p=0.0077$).

An increase of cells in G2 was observed in all HPV-positive OPSCC cell lines between the untreated and the 2Gy 24 hours, as well as untreated and 6Gy 24 hours. A significant increase in cells in G2 after 2Gy was apparent only for UMSCC-47 (two-way ANOVA, $p=0.0135$) and CU-OP-3 (two-way ANOVA, $p=0.0075$). UPCI-SCC-90, CU-OP-20 and CU-OP-2 did not show a significant difference in the proportion of cells in G2 between the 2Gy treatment and the untreated sample.

All HPV-positive samples showed a significant increase in cells in G2 after 24 hours, when comparing 6Gy treated cells and the untreated sample. This was indicated for UMSCC-47 (two-way ANOVA, $p<0.0001$), CU-OP-20 (two-way ANOVA, $p=0.0004$) UPCI-SCC-90 (two-way ANOVA, $p<0.0001$), CU-OP-2 (two-way ANOVA $p=0.0019$) and CU-OP-3 (two-way ANOVA, $p=0.0647$) (Figure 4-9).

Within the HPV-negative cell lines, only UMSCC-4 demonstrated a significant difference between the untreated and 2Gy 24 hours' sample (two-way ANOVA, $p= 0.0222$). After 6Gy treatment CU-OP-17 was the only HPV-negative cell line that showed a significant difference compared to the untreated sample after 24 hours (two-way ANOVA, $p=0.0052$) (Figure 4-10).

The normal human keratinocytes (HEKn) showed no significant changes in the proportion of cells in G1, S and G2 between the untreated samples and the 2Gy and 6Gy samples (8, 24 and 48h) (Figure 4-11). A decrease of cells in G1 with a corresponding increase in S and G2 was observed after 6Gy treatment for all time points. The cell cycle distribution varied for the 2Gy treated samples, with a decrease of cells in G1 after 8 hours, a minor increase after 24 hours and a decrease after 48 hours, with an opposite effect of G2.

Overall, there was no significant difference in cell cycle distributions for any of the HPV-negative cell lines including the normal human keratinocytes after 48 hours. In a comparison between the samples after 24 and 48 hours, UMSCC-74a was the only cell line where a clear visual decrease of G2 was observed, compared to the 24 hours. However, this decrease did not show a significant difference. UMSCC-6 did not show an

effect of IR on cell cycle distribution throughout the entire panel of treatment doses and time points.

All HPV-positive OPSCC cell lines showed a decrease of G2 after 48 hours compared to the 24 hours' samples. Only minor decreases were observed after 2Gy treatment, which was confirmed after statistical analysis with no identified significant differences between 24 hours and 48 hours. A significant difference was measured after 6Gy treatment between 24 hours and 48 hours for UMSCC-47, UPCI-SCC-90, CU-OP-2. CU-OP-3 did not show a significant difference after 6Gy treatment between 24 hours and 48 hours. The only significant increase of cells in G1 was shown in UMSCC-47 and CU-OP-2.

In summary, all HPV-positive cell lines showed a G2 cell cycle arrest after 24 hours (2Gy and 6Gy) and a decrease of cells in G2 after 48 hours. Three out of 5 HPV-negative cell lines showed a similar pattern to the HPV-positive cell lines. However, these differences were only statistically significant in 2 HPV-positive OPSCC and 1 HPV-negative OPSCC cell line after 2Gy 24h and statistically significant in all HPV-positive cell lines after 6Gy 24h.

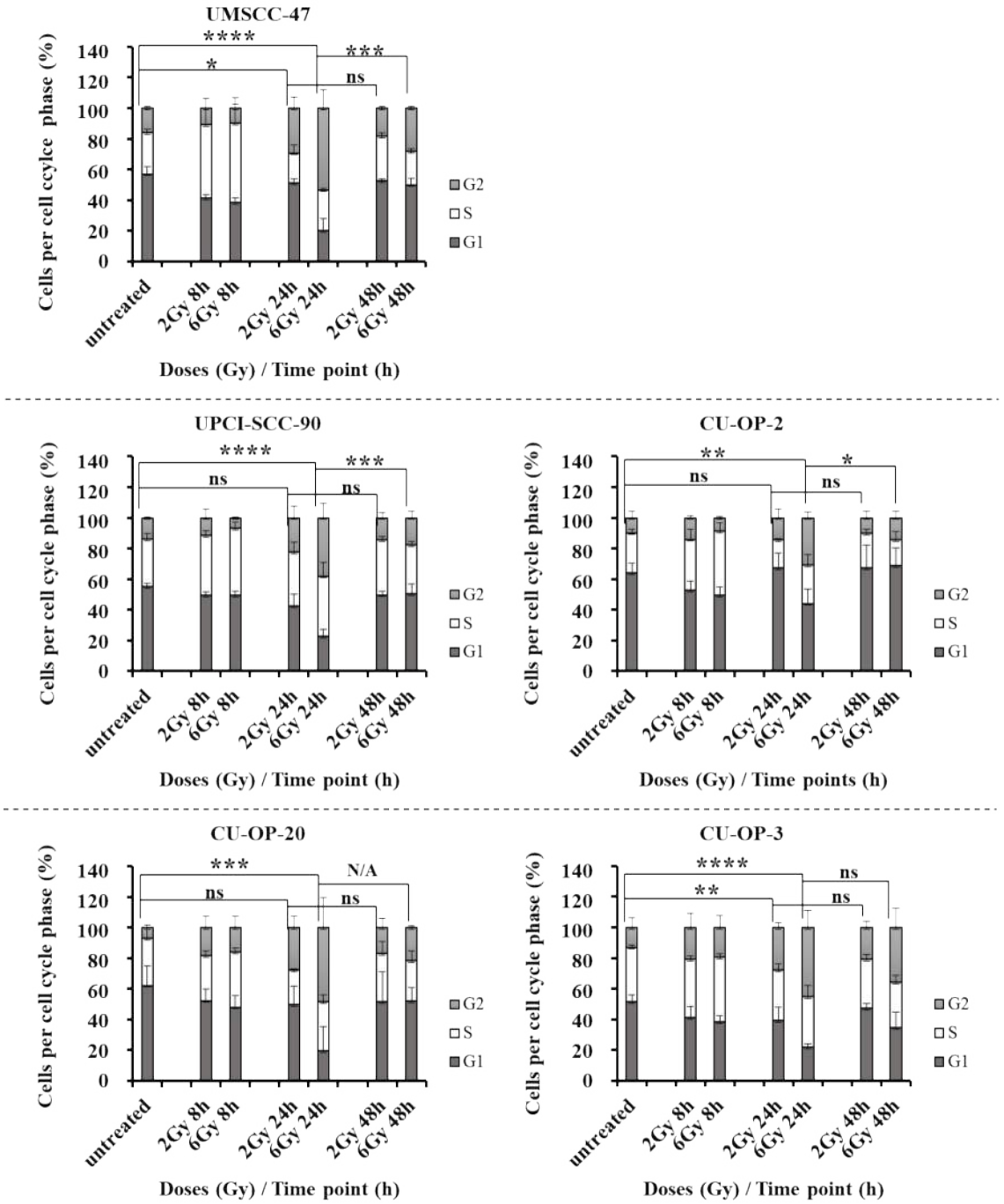



**Figure 4-9: Cell cycle distribution in HPV-positive OPSCC cell lines in response to IR**
All cell lines showed a similar pattern of cell cycle arrest after IR. The untreated time point was collected at the beginning of the assay and defines time zero of the cell cycle phases. Cells were treated with 2Gy and 6Gy and collected after 3 different time points (8, 24 and 48 hours); they were then fixed with 70% Ethanol for a minimum of 1 hour. Cell pellets were stained with PI and the cell cycle distribution assessed using a BD Accuri Flow Cytometer. Cell cycle was analysed using FlowJo software utilising the Watson pragmatic algorithm to define the proportion of cells per cell cycle phase. All experiments were performed in triplicated; the error bars represent the standard deviation between the three experimental runs. Statistical significance was assessed using a two-way ANOVA with a Sidak post-test per cell line to compare the proportion of cells in G2 phase between untreated and 24-hour time points, between 24 hours and 48 hours. The stars representing the strength of significance (e.g: 0.00332 (*), 0.0021 (*), 0.0002 (***), <0.0001 (****). Ns stands for not significant and N/A stands for not available. For CU-OP-20 one data set of 6Gy 48 hours is missing and a two-way ANOVA could not be assessed.

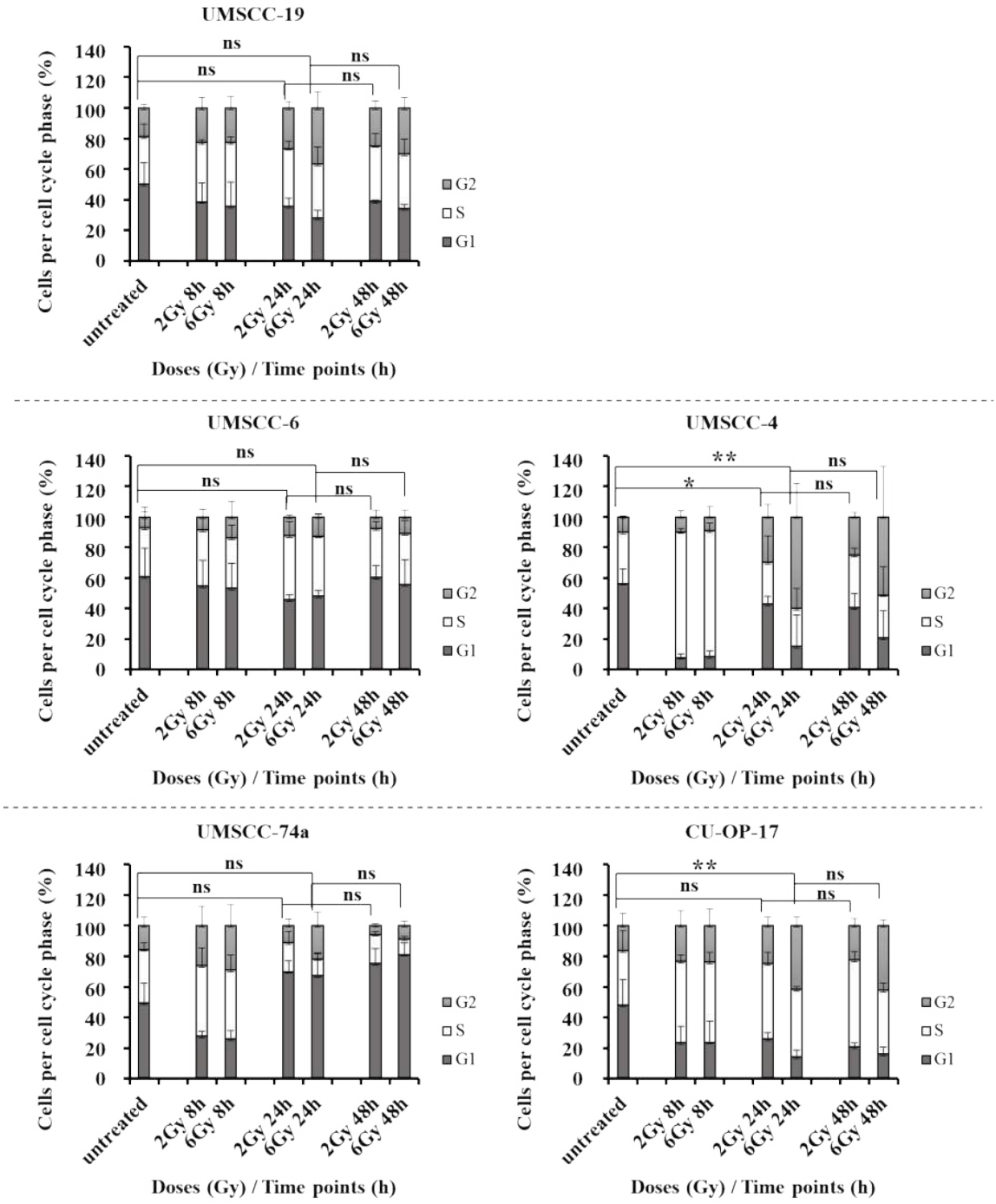



**Figure 4-10: Cell cycle distribution in HPV-negative OPSCC cell lines following IR**
Three out of 5 cell lines showed a similar pattern as the HPV-positive OPSCC cell lines after IR (CU-OP-17, UMSCC-19 and UMSCC-4). Less G2 arrest was observed in UMSCC-74 and UMSCC-6. The untreated time point was collected at the beginning of the assay and defines time zero of the cell cycle phases
Cells were treated with 2Gy and 6Gy and collected after 3 different time points (8, 24 and 48 hours); they were then fixed with 70% Ethanol for a minimum of 1 hour. Cell pellets were stained with PI and the cell cycle distribution assessed using a BD Accuri Flow Cytometer. Cell cycle was analysed using FlowJo software utilising the Watson pragmatic algorithm to define the proportion of cells per cell cycle phase. All experiments were performed in triplicate; the error bars represent the standard deviation between the three experimental runs. Statistical significance was assessed using a two-way ANOVA with a Sidak post-test per cell line to compare the proportion of cells in G2 phase between untreated and 24-hour time points, between 24 hours and 48 hours. The stars representing the strength of significance (e.g: 0.00332 (*), 0.0021 (*), 0.0002 (***), <0.0001 (****). Ns stands for not significant.

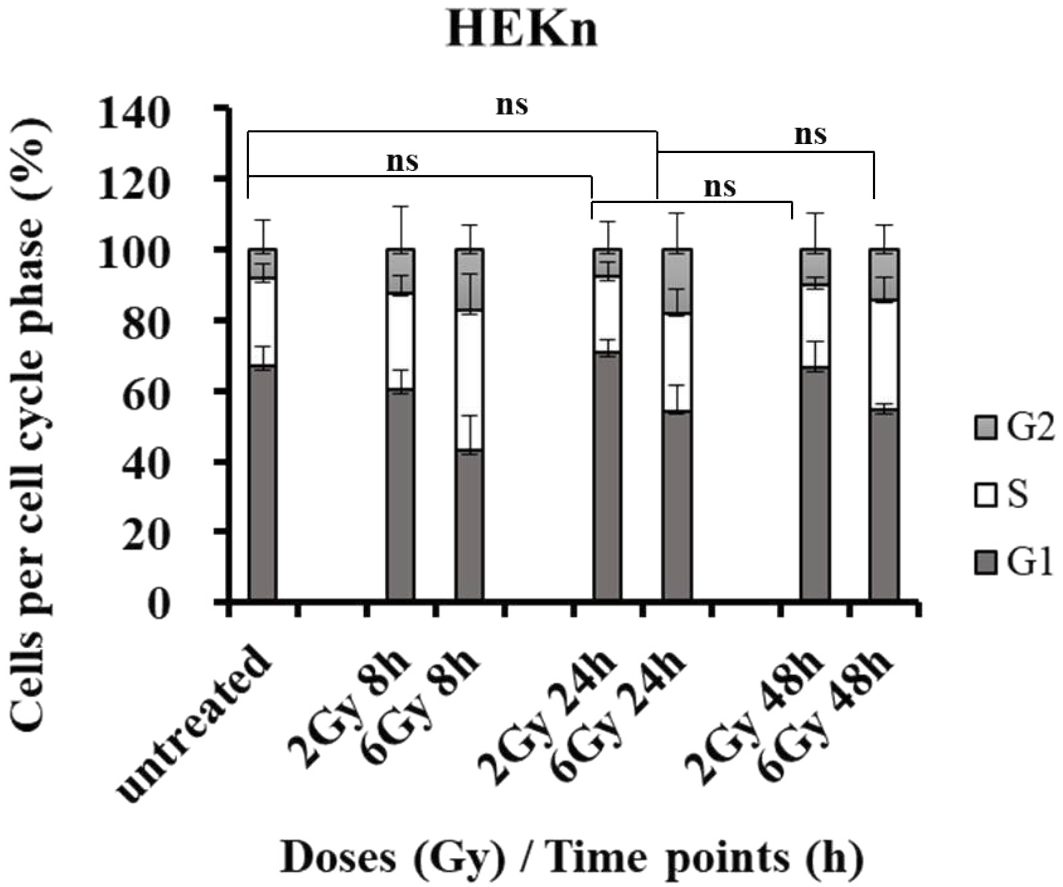



**Figure 4-11: Cell cycle distribution in HEKn cells following IR**
The untreated time point was collected at the beginning of the assay and defines time zero of the cell cycle phases.
Cells were treated with 2Gy and 6Gy and collected after 3 different time points (8, 24 and 48 hours); they were then fixed with 70% Ethanol for a minimum of 1 hour. Cell pellets were stained with PI and the cell cycle distribution assessed using a BD Accuri Flow Cytometer. Cell cycle was analysed using FlowJo software utilising the Watson pragmatic algorithm to define the proportion of cells per cell cycle phase. All experiments were performed in triplicated; the error bars represent the standard deviation between the three experimental runs. Statistical significance was assessed using a two-way ANOVA with a Sidak post-test per cell line to compare the proportion of cells in G2 phase between untreated and 24-hour time points, between 24 hours and 48 hours. The stars representing the strength of significance (e.g: 0.00332 (*), 0.0021 (*), 0.0002 (***), <0.0001 (****). Ns stands for not significant.

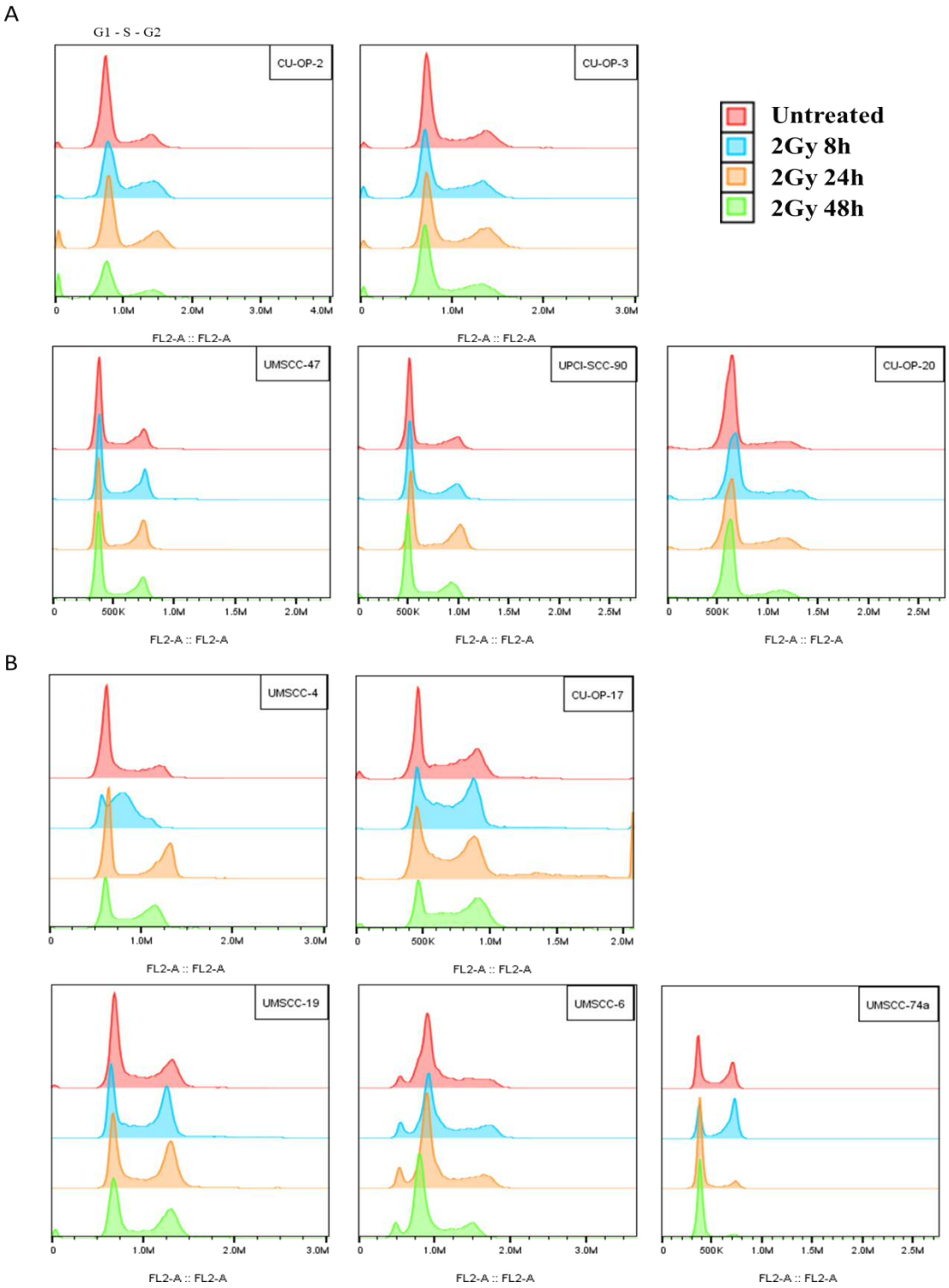



**Figure 4-12: Cell cycle analysis for HPV-positive and HPV-negative OPSCC cell lines after 2Gy**
Overlay expression of one representative run of the cell cycle phases after 2Gy treatment and 8, 24 and 48-hour time points. (A) HPV-positive cell lines after 2Gy treatment; (B) HPV-negative cell lines after 2Gy treatment.
Cells were treated with 2Gy and 6Gy and collected after 3 time points (8, 24 and 48 hours) and fixed with 70% ethanol for a minimum of 1 hour. Cell pellets were stained with PI and cell cycle distribution measured on a Flow Cytometer (BD Accuri). Cell cycle was analysed in FlowJo using the Watson pragmatic algorithm.

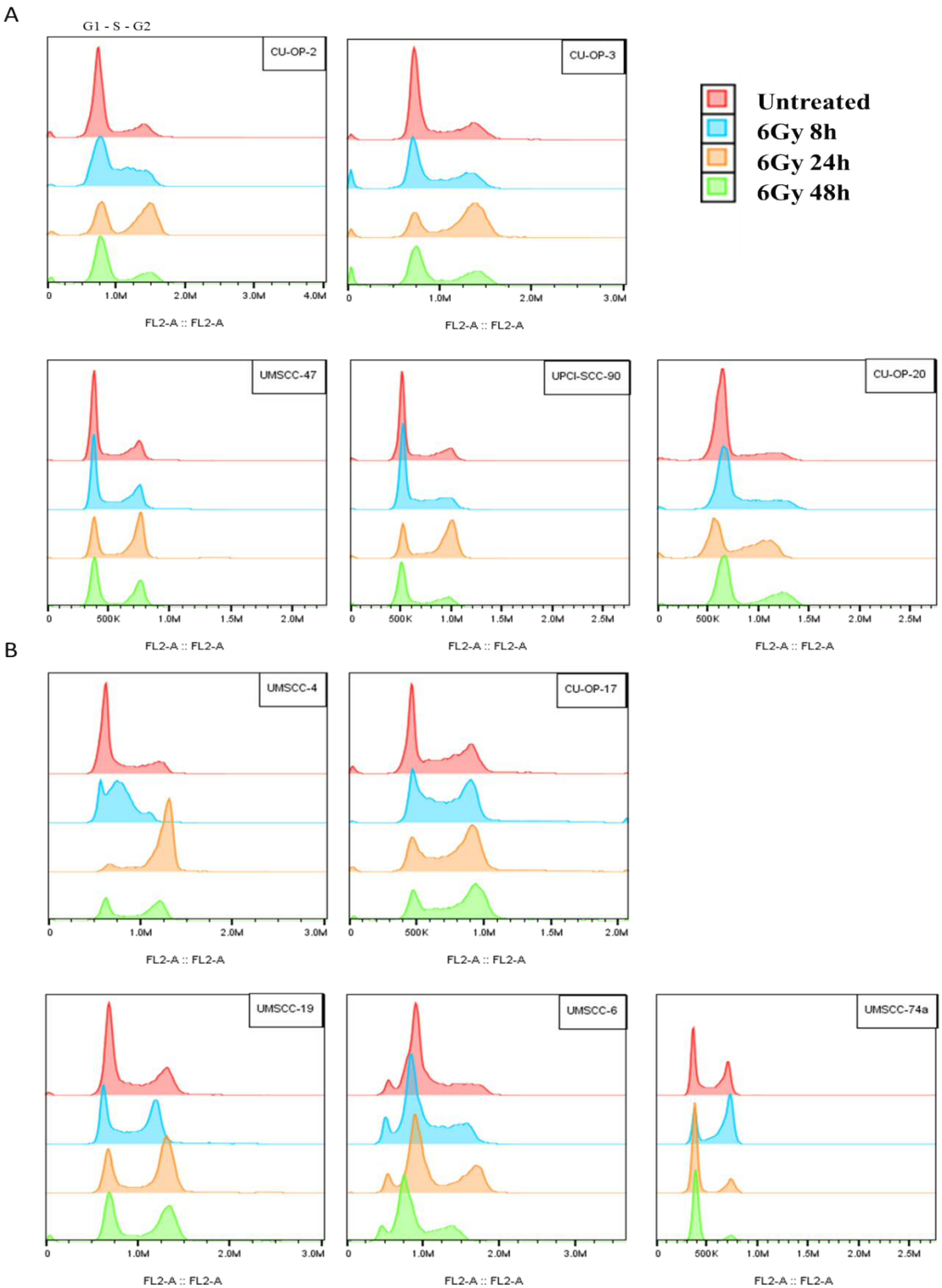



**Figure 4-13: Cell cycle analysis for HPV-positive and HPV-negative OPSCC cell lines after 6Gy**
Overlay expression of one representative run of the cell cycle phases after 6Gy treatment and 8, 24 and 48-hour time points. (A) HPV-positive cell lines after 6Gy treatment; (B) HPV-negative cell lines after 6Gy treatment.
Cells were treated with 2Gy and 6Gy and collected after 3 time points (8, 24 and 48 hours) and fixed with 70% ethanol for a minimum of 1 hour. Cell pellets were stained with PI and cell cycle distribution measured on a Flow Cytometer (BD Accuri). Cell cycle was analysed in FlowJo using the Watson pragmatic algorithm.

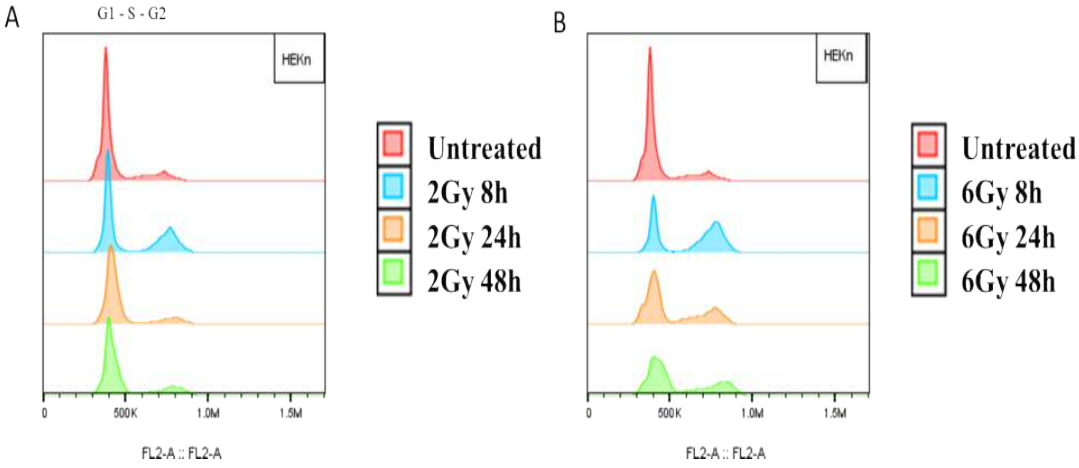



**Figure 4-14: Cell cycle analysis for Hekn cell line after 2Gy and 6Gy**
Overlay expression of one representative run of the cell cycle phases of HEKn after (A) 2Gy and (B) 6Gy. Cells were treated with 2Gy and 6Gy and collected after 3 time points (8, 24 and 48 hours) and fixed with 70% ethanol for a minimum of 1 hour. Cell pellets were stained with PI and cell cycle distribution measured on a Flow Cytometer (BD Accuri). Cell cycle was analysed in FlowJo using the Watson pragmatic algorithm.

## 4.6. Summary of main findings

With reference to the original hypothesis and the revised hypotheses after the protein expression experiments, the investigations in this chapter established that:

- At 2 and 4Gy doses, HPV-positive OPSCC cell lines showed a greater variation in sensitivity to IR than HPV-negative OPSCC cell lines
    - UMSCC-47, UPCI-SCC-90 and CU-OP-20 were the most sensitive HPV-positive cell lines. CU-OP-2 and CU-OP-3 were the most resistant HPV-positive cell lines
- HPV-status correlated only partially with radiation sensitivity
    - 3 HPV-positive cell lines were sensitive to IR (2Gy)
    - Only one HPV-negative cell line was sensitive to IR (2Gy)
    - HEKn were also sensitive to IR (2Gy)
- HPV-positive cell lines did not show a greater p53, p21 or p-p53 accumulation after IR relative to HPV-negative cell lines.
- 2/5 HPV-negative cell lines and HEKn showed accumulation of p-p53 after IR
- p53 independent transcription of p21 was detected in 3 HPV-negative OPSCC cell lines
- p21 demonstrated a greater variation in expression with up- or down-regulation after IR dependent on the cell line tested.
- Degree of G1, S and G2 cell cycle arrest post irradiation did not correlate with HPV status – example listed for G2 arrest:
    - 2/5 HPV-positive cell lines showed a significant G2 arrest after 2Gy 24 hours
    - 5/5 HPV-positive cell lines showed a significant G2 arrest after 6Gy 24 hours
    - 1/5 HPV-negative cell lines showed a significant G2 arrest after 2Gy 24 hours
    - 2/5 HPV-negative cell lines showed a significant G2 arrest after 6Gy 24 hours
- After 48h a decrease of cells in G2 phase was detected in 3/5 HPV-positive cell lines after 6Gy and no significant decrease of cells in G2 was detected within the HPV-negative cell lines after 48h

The main findings of the clonogenic, western blot and cell cycle analyses in the OPSCC cell line panel are summarized in Table 4-3.

**Table 4-3: Summary of OPSCC responses to IR**

| | Cell lines | Clonogenic assay [A] | | Protein Expression [B] | | | | | | | | | Cell cycle arrest [C] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | p53 | | | p-p53 | | | p21 | | | G2-phase | | | |
| | | 2Gy | 6Gy | untreated | 2Gy | 6Gy | untreated | 2Gy | 6Gy | untreated | 2Gy | 6Gy | 2Gy | 6Gy | 2Gy | 6Gy |
| | | 10-15 days | | 24 hours | | | | | | | | | 24 hours | | 24 – 48 hours | |
| HPV-positive | UMSCC-47 | + | + | ++ | ++ | ++ | - | - | - | ++ | + | + | + | + | ns | - |
| | UPCI-SCC-90 | + | + | + | + | + | - | - | - | ++ | ++ | + | ns | + | ns | - |
| | CU-OP-20 | + | + | + | + | + | - | - | - | ++ | ++ | ++ | ns | + | ns | N/A |
| | CU-OP-2 | - | + | ++ | ++ | ++ | - | - | - | ++ | + | + | ns | + | ns | - |
| | CU-OP-3 | - | + | + | + | + | - | - | - | + | ++ | ++ | + | + | ns | ns |
| HPV-negative | UMSCC-19 | + | + | - | - | - | - | - | - | + | + | + | ns | ns | ns | ns |
| | UMSCC-6 | - | + | - | - | - | - | - | - | ++ | + | + | ns | ns | ns | ns |
| | UMSCC-74a | - | + | ++ | ++ | ++ | + | + | + | + | + | + | ns | ns | ns | ns |
| | UMSCC-4 | - | + | - | - | - | - | - | - | + | + | ++ | + | + | ns | ns |
| | CU-OP-17 | - | + | ++ | ++ | ++ | + | + | + | + | + | + | ns | + | ns | ns |
| | HEKn | + | + | + | + | + | + | + | + | ++ | ++ | ++ | ns | ns | ns | ns |

[A] Clonogenic survival assay result: Green indicates the sensitivity of the cell lines (SF<40%) and red indicates resistant cell lines (SF>40%) after 2Gy and 6Gy

[B] Western blotting result: Green indicates indicates that the cell lines showed p53, p-p53 or p21 protein (+, ++ or – represent the strength of the band)

[C] Focus in the summary table was put on signifcant difference in G2 (based on previous findings in litearture). Green represents a signifcant accumulation of more cells in G2 phase after 2Gy and 6Gy compared to the untreated sample after 24h. Red represent a signifcant decrease of cell in G2, by comparing 24h and 48h after 2Gy and 6Gy treatment (+ indicates an increase; – indicates a decrease; ns stands for not signifcant; N/A stands for not available)

## 4.7. Discussion

### 4.7.1. Clonogenic assays

The 10 OPSCC cell lines used in this study showed a wide range in sensitivity to single doses (0.5-6Gy) of IR. Three HPV-positive cell lines (UMSCC-47, UPCI-SCC-90 and CU-OP-20) showed a high sensitivity at the IR dose to 2Gy IR (this dose was used to classify sensitivity) and 1Gy, whereas CU-OP-2 and CU-OP-3 showed greater resistance at the same dose. However, CU-OP-2 was at the borderline of being a sensitive or resistant cell line. Minimal effects on survival were observed on all cell lines with doses of 0.5Gy. Therefore, 2Gy IR treatment was chosen as the main dose to analyse sensitivity of cell lines to IR, as at lower doses cells did not show a significant response and at higher doses most cells did not proliferate (resulting in a very limited number of colonies). A similar pattern of response was observed in all HPV-negative cell lines, except for UMSCC-19 which was the only HPV-negative cell line to have a SF below the set threshold of 40%. One interesting finding was that the SF of HEKn after 2Gy was below 40%. It was expected for the normal human keratinocytes cell line to show less sensitivity to intermediate IR doses (2Gy), as they are presumed (as a normal cell line) to be competent in DSB repair. In a previous study, human keratinocytes showed a SF of 49+/-8%, at 2Gy (Panteleeva et al., 2003). The difference between the two studies could be due to the different culture conditions used. The experiment in the study by Panteleeva *et al.,* (2003) was performed in 25 $cm^2$ flasks, incubated for several days and irradiated before reaching 70% confluence and the culture medium did not cover the cells when using the X-ray beam (only the plastic walls were covered) (Panteleeva et al., 2003). Different culture conditions and radiation sources might have resulted in a higher survival fraction in the Panteleeva study. The SF at 2Gy for the HEKn cells was above the three most sensitive HPV-positive cell lines and below the most resistant HPV-positive cell lines. Therefore, the IR response of HEKn was intermediate compared to the HPV-positive cell lines, but were more sensitive at the 2Gy dose than the majority of HPV-negative cell lines.

The IR dose range for this study was chosen on the basis of the available literature (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013). To test a potential intermediate effect of IR on the clonogenic assays, lower doses of 0.5Gy and 1Gy were included as well. The effect of IR on HEKn, used as a non-transformed epithelial cell control, was novel and a strength of the current study. As a normal cell line, HEKn cell line expresses

functioning p53 (wild-type) and other DNA repair related mechanisms (including cell cycle control) are presumably operating. Furthermore, it provided a comparison between normal keratinocytes and cancer cell lines (transformed/mutated keratinocytes), a feature that was lacking in previous studies of OPSCC (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013). The sensitive response to IR seen in the HEKn cells, could be an indication of a functional response to DNA damage (Willers et al., 2002), which could explain the minor increase in G2 cell cycle phase following a functioning DNA repair and a general decrease of G2 after 48 hours (return to normal cell distribution).

According to the data obtained in this study, there was no clear difference detected between HPV-positive and HPV-negative OPSCC cell lines' response to IR between a dose range of 4-6Gy. After 6Gy treatment, there was a limited capability for cell lines to form colonies, because the majority of the cells had died. This indicated that higher treatment doses are not ideal to investigate response to IR and an intermediate dose such as 2Gy is more appropriate. The cell lines classified as resistant to IR included both HPV-positive and HPV-negative lines. An overlap in sensitivity to radiation among HPV-positive and HPV-negative head and neck cancer cell lines (HNSCC) has been reported previously (Rieckmann et al., 2013). However, other studies have documented clear differences in sensitivity to IR between HPV-positive and –negative HNSCC cell lines (Arenz et al., 2014, Kimple et al., 2013). These differences might reflect the fact that different cell lines were used in the studies by Arenz and Kimple *et al.,* (Arenz et al., 2014, Kimple et al., 2013).

A similar level of sensitivity to 2Gy IR treatment was measured for UMSCC-47, UPCI-SCC-90 and UMSCC-6 in three other studies (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013). UMSCC-47 and UPCI-SCC-90, were reported as being the most sensitive cell lines and UMSCC-6, as reported in this study, was a less sensitive cell line. For clarification, the term resistant was not used in these three studies, the terms of "most and least sensitive" cell lines were used (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013).

Some studies have investigated a combination of radiotherapy with other DNA damaging agents; for example, cisplatin with IR. Busch *et al*., (2016) investigated the effects of cisplatin and radiation and reported that the combined treatments resulted in only a minor enhancement in sensitivity compared to the expected large enhancement. Although they suggest that results may be dependent on the experimental set-up, which can result in variations of SF (Busch et al., 2016). Another study tested whether a CHK1 inhibitor

(roscovitine) in combination with radiation could enhance the sensitivity of HNSCC cell lines to IR in clonogenic assays. HPV-negative HNSCC cell lines showed enhanced sensitivity whereas the combined treatment had no effect on HPV-positive HNSCC cell lines (Ziemann et al., 2017). For breast cancer cell lines, an enhanced effect of combining olaparib (BRCA2 mutation effector) and radiotherapy has been seen (Jang et al., 2015). This indicated that some combination of DNA damaging agents or cell cycle inhibitors can enhance the effect of radiation in some transformed cell lines, but this has not been consistently shown for specific groups of HPV-positive or HPV-negative HNSCC cell lines.

In general, the findings in this study are consistent with previous radio-sensitivity assays using the same cell lines, but differ slightly in the surviving fractions calculated. However, unlike the majority of previous studies, no clear differences in the response to IR were seen between HPV-positive and HPV-negative cell lines. This could be the result of differences in methodology (e.g. length of experiment, seeding density) or the inclusion of feeder cells in the current studies. These feeder cells gave OPSCC cells a better chance of survival. Another variable that needs to be considered are difference of doubling time of each cell line, indicated by different seeding densities defined, that could influence survival of cells after IR. However, the influence of doubling time was intended to be reduced a defined seeding density per cell lines before the start of the radio-sensitivity assay. In this study the length of the experiments was set for the panel of cell lines between 10-15 days, whereas clonogenic assays were set-up in one study for 14-18 days using a different radiation technique; cells were irradiated with 6 MeV photons using a linear accelerator (Elekta Supernova, Elekta, Stockholm, Sweden) (Arenz et al., 2014).

Two of the HPV-negative HNSCC cell lines used in the current study were derived from patients receiving chemotherapy, UMSCC-4 (derived from tumour receiving post-treatment of 2 chemotherapies) and UMSCC-74a (originated from a recurrent tumour – after receiving chemotherapy). In a study on adaption of cancer cells to chemotherapy, a slight increase of resistance in biopsies post-chemotherapy was observed (Di Nicolantonio et al., 2005). This might be seen in UMSCC-4 and UMSCC-74 in the current study. However, this would need to get confirmed by comparing it to an untreated tumour sample of these two cell lines, if available.

Compared to cervical cancer cell lines, the OPSCC cell lines used in the current study are, in general, more sensitive to IR at 2Gy. In a previous study (Banath et al., 2004), the

HPV negative cervical line C33A was the only cell line with a similar sensitivity to the HPV positive OPSCC cell lines (UPCI-SCC-90 and CU-OP-20) used in the current study.

**Strengths and weaknesses of the clonogenic assays**

One of the strengths of the current study, i.e. a robust clonogenic assay, was achieved through careful definition of culture conditions and seeding density for each cell line. The most appropriate seeding density for 7 cancer cell lines and one HEKn cell line was defined previously (Pirotte, 2017), whereas the seeding densities for UMSCC-6, CU-OP-17 and CU-OP-20 were defined in the current study. Furthermore, the optimal duration of experiment to measure IR response was defined for each individual cell line in the current study. Cell lines were treated with IR, 24 hours after seeding to minimize the effect of IR on cell adhesion. Image analysis software was used to count the colonies after crystal violet staining, to minimise potential errors generated by manual counting. A potential weakness is the error rate of the colony counter, as counts could vary, if pictures were taken from different angles on the same plate. This weakness was however minimized, by performing three experimental runs and three replicates per treatment dose. Photographs were taken for each plate from three different angles and the mean value calculated.

Another strength of this study was the panel of cell lines used, that included 2 novel cell lines, and focussed specifically on OPSCC cell lines, with a total of 10 OPSCC cell lines and one normal cell line used. In contrast, other studies used a wide range of HNSCC cell lines derived from different anatomical sites with a cell line panel size of 8 cell lines in two studies (Arenz et al., 2014, Kimple et al., 2013) and 10 cell lines in another study (Rieckmann et al., 2013). A normal cell line control was not included in any of the three published studies (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013).

### 4.7.2. Assessment of protein levels of p53, phospho-p53 and p21

There was no apparent correlation between levels of p53, p-p53 or p21 expression and HPV-status. All HPV-positive cell lines expressed p53 and p21. However, only 2 out of 5 HPV-negative cell lines showed detectable levels of p53. Interestingly, p21 was detected in all HPV-negative cell lines. Expression of phospho-p53 was only detected in three cell lines CU-OP-17, UMSCC-74a and HEKn, which can contribute to p53 stabilization in those cell lines by preventing Mdm2-mediated degradation (Ashcroft et al., 1999, Shieh et al., 1997). One of the main sites of phosphorylation of p53 after DNA

damage is Ser15, which was the main reason to include it in the current study (Ashcroft et al., 1999, Shieh et al., 1997).

Higher basal levels of p53 were expected and increased sensitivity to radiation was predicted due to the p53 mutation status in some of the HPV-negative cell lines (UMSCC-4 with a point mutation and UMSCC-19 with a 10bp deletion). The absence of HPV in the remaining cell lines (CU-OP-17, UMSCC-6 and UMSCC-74a) combined with the expression of WT p53 might also lead to similar results. It might also be predicted that there will be a higher expression of p53 after irradiation based on the properties of these cell lines, as no p53 degradation by HPV E6 is possible in these HPV-negative cell lines.

For the HPV-positive cell lines, p53 might be predicted to be undetectable because of the degradation of p53 mediated by HPV E6, and these cell lines would have been expected to be resistant to radiation due to loss of p53 mediated cell cycle arrest. However, p53 was detected in all the HPV-positive lines, which is consistent with incomplete degradation of p53 and this might have caused increased sensitivity in the majority of the HPV-positive cell lines (Bradford et al., 2003, Ferris et al., 2005, Mandic et al., 2005, Somers et al., 1992). Interestingly, levels of p53 varied between the HPV-positive cell lines, but levels of p53 detected in each line did show any relationship with radiation dose. Presence of detectable p53 in the untreated samples for all HPV-positive lines would be consistent with incomplete E6-mediated degradation of p53, but may also suggest that the cells were inherently stressed as a result of the culture conditions (e.g. by low oxygen levels). Similarly, p53 levels did not always correlate with IR sensitivity as assessed in the clonogenic assays. For example, UMSCC-47 and UPCI-SCC-90 were classified as sensitive cell lines, but UMSCC-47 showed high levels of p53 protein while UPCI-SCC-90 showed lower levels.

Phosphorylation of p53 at Ser15 helps modulate the stability of p53 (Ashcroft et al., 1999). Levels of p-p53 did not correlate with radio-sensitivity as assessed in the clonogenic assays. Neither was there a correlation between p53 levels and p-p53, which might have been expected, as phosphorylation at ser15 reduces Mdm-mediated degradation of p53. This might be due to a lack of transcription of Mdm, associated with non-functional p53.

In the current study, UMSCC-47 showed strong expression of p53 but no p-p53 in all conditions. This cell line was also used in two previous studies that investigated p53 and p-p53 expression after IR treatment (Arenz et al., 2014, Kimple et al., 2013). These

studies found different expression levels of p53 after IR compared to the current study. One study reported no accumulation of p53 after IR (4Gy) although a weak p53 band was detected after 24 hours, and no phosphorylation of p53 was detected (Kimple et al., 2013). Another study detected weak bands of p53 and p-p53 1 and 4 hours following 6Gy IR (Arenz et al., 2014). This study (Arenz et al., 2014) is in agreement with the p53 results in this chapter. However, in the current study incubation time was longer than 4 hours, cells were collected after 24 hours, which might have an effect on stronger accumulation of p53. However, different results for phosphorylation of p53 were obtained when comparing the study of Arenz *et al.,* (2014) with the current study (i.e. they were able to detect phosphorylation on Ser15). The reasons for this difference are unclear, as the p-p53 (Ser15) antibody was the same between studies. This did not appear to be associated with a general failure of the antibody/conditions as the same antibody detected p-p53 in UMSCC-74a, CU-OP-17 and HEKn (normal cell line control), with stronger bands detected after 2Gy to 10Gy treatment. However, for the majority of the cell lines p-p53 might be degraded after 24h, which might cause different results compared to previous studies. Strong bands of p53, p21 and weak bands of p-p53 could indicate stable and functioning p53 in UMSCC-74a, CU-OP-17 and HEKn (Ashcroft et al., 1999).

The identities of the cell lines used in the western blotting were confirmed by short tandem repeat (STR) typing, as it was in the studies of Kimple and Arenz (Arenz et al., 2014, Kimple et al., 2013). Findings for UMSCC-6 (Arenz et al., 2014) and UPCI-SCC-90 (Kimple et al., 2013) were similar in this study for p53 and p-p53. However, the greater contradictory findings for UMSCC-47 might be a result of differences in methodology. For example, culture conditions, absence of positive controls in the previous studies, differences in lysis conditions (inhibitors used) and differences in exposure of films. Furthermore, cells could have been stressed by Trypsin-EDTA (TE) before cell lysis, which might explain the stronger p53 bands in the untreated UMSCC-47 sample in the current study. To compensate for the different findings, lysate samples of cell lines used in both studies would need to be tested in a new separate experimental run.

p21 is generally regarded as a p53 regulated gene, the normal function of which is to cause cell cycle arrest in response to DNA damage. The role of p21 is directly connected to p53 transcription. P21 (cyclin-dependent kinase inhibitor p21WAF/CIP1) is a cell-cycle checkpoint effector and inducer of senescence and is regulated by p53 (Weinberg,

2007). In normal cells, genotoxic stress activates the ATM-p53 pathway that upregulates the expression of p21, leading to cell cycle arrest (Galanos et al., 2016).

p21 was detected in all cell lines in both untreated and treated samples even if there were no visible p53/p-p53 bands detected. In three HPV-negative OPSCC cell lines (UMSCC-4, UMSCC-6 and UMSCC-19) where p53 was absent, p21 protein expression was detected, which might indicate p53-independent transcription. Independent transcription of p21 and no direct transcription through p53 was suggested in a recent study on a wide range of human tumours (Galanos et al., 2016). The strong expression of p21 in the majority of the HPV-positive OPSCC cell lines could be connected to better prognosis and IR sensitivity in UMSCC-47, UPCI-SCC-90 and CU-OP-20. This was observed in a study on tonsillar squamous cell carcinoma to investigate the prognostic value of key cell cycle proteins in the pRb and p53 pathway. Overexpression of p21 was determined as a favourable prognostic factor for tonsillar squamous cell carcinoma (Hafkamp et al., 2009).

Ionising radiation can be responsible for the degradation or decrease of p21 in the majority of HPV-negative cell lines, as suggested through studies on HEK293 cells (Stuart and Wang, 2009). This decrease in p21 expression might be caused by IR and could be an explanation for a decrease in band intensity of p21 in some HPV-negative OPSCC cell lines, for example UMSCC-19 and UMSCC-6. The mechanism behind the decrease of levels of p21 after IR in some cell lines in this panel is unclear at the moment.

A study by Galanos *et al.,* (2016) of a wide set of human tumours (HSNCC, lung cancer, urothelial carcinomas and colon precancerous lesions) suggested that p53 independent expression of p21 can indicate more aggressive tumour cells, increased genomic instability and chemo-resistance. (Galanos et al., 2016). Sustained accumulation of p21, which inhibits CRL4-CDT2 ubiquitin ligase (a substrate recognition factor that prevents degradation of cell cycle regulated proteins) leads to replication stress and genomic instability and more aggressive tumour cells (Galanos et al., 2016, Abbas and Dutta, 2011, Abbas et al., 2008). This might be the case for these three HPV-negative cell lines, as they are less sensitive to IR than most of the HPV-positive OPSCC cell lines. The two most resistant HPV-positive OPSCC cell lines (CU-OP-2 and CU-OP-3) show both p53 and p21 bands. CU-OP-3 (the most resistant cell line) showed weak accumulation of p53 and stronger p21 expression, which might explain its radio-resistance and might indicate

that additional mechanisms besides p21 independent transcription lead to more radio-resistance tumours.

In a study in T-cell lymphoma parental lines, expression of p53 and p21 was directly connected to the wild-type or mutation (mut) allele of p53. Cells expressing p21 after IR showed bands of p53 (wild-type). On the other hand cell lines with mutations in p53 had weak or no expression of p21 (el-Deiry et al., 1994). Comparing this study of different cell types with OPSCC cell lines in the current study is not directly possible, but provides a framework for possible interpretations. In the current study, no connection between p53 and p21 expression to its wild-type or mut-type of p53 can be directly drawn. For example, UMSCC-6 (wt-p53) and UMSCC-4 (mut) did not show bands of p53, but both express p21. However, the majority of the HPV-positive OPSCC cell lines (wild-type p53) showed p53 and p21 bands together. A greater variation in strength of bands for p53 and p21 is observed within the HPV-negative OPSCC lines, which express either wild-type and mutated alleles of p53. Furthermore, these cell lines showed levels of p21 even if expressing a p53 mutation.

In summary, base-line p21 and p53 levels did not appear to be directly correlated with cell lines being more sensitive or resistant to IR. Hence the data did not support the original hypothesis that increased radiation sensitivity would be associated with accumulation of p53 in HPV-positive OPSCC cell lines. Other molecular factors in the cell cycle and DNA repair process, might play a more significant role in terms of sensitivity to IR and DNA repair response, which was not investigated further on protein level through Western blot assays. DNA damage and DNA repair after IR on protein levels should be determined in a future study.

**Strengths and weaknesses of the Western blotting experiments**

One of the strengths of this assay was the range of conditions assessed. At relatively low IR doses of 0.5Gy and 1Gy, minimal effects on proliferation were seen, while at higher doses between 4Gy and 6Gy limited colony formation was observed. The greatest variation in response to IR was observed at 2Gy, therefore this dose was used in the protein-based assays. The time point of 24 h post IR treatment was chosen to assess the short-term effect of IR, however a weakness of this assay is that these data cannot be compared directly to the clonogenic assays which reflect a response over 10-15 days.

The strengths of this assay include use of an appropriate control, i.e. a normal human keratinocyte cell line with “normal” patterns of expression of p53 and p21 (Munro et al.,

1999). Although the presence of p53, p-p53 and p21 in the untreated HEKn may suggest that these cells were stressed by the culture conditions. A potential weakness might be the absence of a positive control for induction of apoptosis and DNA damage specific to each cell line. The positive control of $_{10Gy}$ irradiated HEKn (normal human keratinocytes) cell line for p53, p-p53 and p21 expression, partially compensated for this weakness. Another potential weakness in this study is the leak of measurement of the strength of DNA damage induced before the start of this assay. For example, yH2AX could have been used as such a marker. This would have helped to choose additional markers, beside p53, p-p53 and p21, to identify a responsible factor for the greater radio-sensitivity of HPV-positive OPSCC cell lines.

### 4.7.3. Analysis of cell cycle distribution before and after IR

The original hypothesis at the start of the study, was that following exposure to IR, a greater G2 arrest would be seen in HPV-positive OPSCC cell lines compared to HPV-negative OPSCC cell lines. This hypothesis was subsequently revised to take account of the wide range of sensitivity to IR seen in the HPV-positive cell lines, and was reframed as: greater G2 arrest would be seen in the OPSCC cell lines with the greatest sensitivity to IR. G1 and S-phase arrests were observed as well to identify an initial response to IR. This was done to identify a clear difference in terms of IR response on cell cycle distribution between G1, S or G2 arrest. A G2 arrest was previously reported in eukaryotic cells in response to IR (Bernhard et al., 1995, Branzei and Foiani, 2008).

Cell cycle distribution was assessed at three different time points (8, 24 and 48 hours). Out of these three time points, 24 hour appeared the best possible indicator for IR induced cell cycle arrest. The reason for this, was that the highest G2 arrest was observed after 24 hours, whereas no obvious G2 arrest was observed after 8h or 48h. G1 arrest after IR can be regulated by p53, which causes an induction of p21, which leads to a delay in S-phase and a further delay in G2. G2 arrest was previously demonstrated in eukaryotic cells after IR (Bernhard et al., 1995, Branzei and Foiani, 2008). The effect of IR (2Gy and 6Gy) on cell cycle distribution in terms of G1 and G2 arrest were most apparent after 24 hours. The panel of HPV-positive OPSCC cell lines showed an increase of cells in G2 after 24 hours in 2Gy treated samples and a more substantial arrest in 6Gy treated samples, with a decrease in G1 phase after 24 hours. A decreased proportion of the cells were in G2 after 48 hours. The decrease of G1 was not significant for the HPV-positive OPSC cell lines after 2Gy and 24h, but was significant for 2/5 cell lines after 6Gy 24h. However, only 2/5 HPV-negative OPSCC cell lines significant difference to the untreated sample after 2Gy 24h. Out of these two cell lines UMSCC-74a showed a significant increase after 2/6Gy 24h, which could indicate a greater response to IR induced DNA damage. This induced DNA damage would be needed to get confirmed via an immunofluorescence assay for γH2AX at a later stage. A prolonged G2 phase after 24 hours on the other side could indicate a delay in entering mitosis, to give the cells enough time to repair any DNA damage, which was suggest in a previous study by Rieckmann *et al.,* (2013). This decrease observed after 48-hours provided a possible indicator of induction of DNA repair mechanisms and resumption of normal cell cycling. The shorter time point of 8 hours did not show any significant increase in G2 after IR, even at higher treatment doses of 2Gy or 6Gy. Induction of DNA repair between 24h to 48h was not

measured in the current study. The correlation of decrease of cells within G2 and DNA repair needs to get confirmed in a future study.

Interestingly, UMSCC-4 showed a significant increase in S-phase after 8 hours, followed by an increase of cells in G2 after 24 hours. This is difficult to interpret but could be explained by a dysfunctional G1/S checkpoint, coupled with retention of cells in S-phase (although there is no published precedent for this). High p21 expression seen in the protein assay (Western blotting) could be responsible for the S-phase delay through the inhibition of cyclins(Bernhard et al., 1995, Branzei and Foiani, 2008).. By 24 hours however, cells appeared to accumulate in G2, which suggests the presence of a functioning G2/M checkpoint.

Compared to the untreated sample (time zero), cells increase in S-phase after 8h (2 and 6Gy) and decreased after 24 and 48h in the majority of cell lines. This initial increase after 8h, especially seen and described for UMSCC-4, could be a response to IR with a possible DNA damage response (not confirmed during this study), which keeps the cells in S-phase and hold replication. Cells could either retreat into G1 phase, continue into G2 until the G2/M cell cycle checkpoint where cells finally arrest in response to IR. Another study associated a prolonged S-phase in response to stress or DNA damage by cells being trapped in a "S-phase stasis" due to the loss of CDKs and associated proteins, making cells unable to move forward or retreat into G1 (Borel et al., 2002). The observed decrease after 8h following 24h and 48h samples, suggest cells retreating into G1 (checkpoint), prolong replication or induced death due to incomplete replication.

Greater G2 arrest in HPV-positive cell lines is consistent with previous findings for UMSCC-47 and UPCI-SCC90, where an increase of G2 after 4Gy treatment was reported (Kimple et al., 2013). However, in the current study no obvious difference between HPV-positive and HPV-negative cell lines was detected; rather, 1/5 HPV-negative OPSCC cell lines showed a significant increase in G2 arrest after 2Gy and 24h, which was similar to the HPV-positive cell lines with 2/5 cell lines with a significant increase after 2Gy 24h. At a lower dose of 2Gy this was inconsistent with their findings in G2 arrest between HPV-positive and HPV-negative cell lines. However, this could be due to the lower treatment dose of 2Gy. At a dose of 6Gy, the HPV-positive cell lines showed a clear

difference in G2 arrest, as observed after 4Gy treatment in the Kimple study (Kimple et al., 2013).

Normal human keratinocytes (HEKn) showed a consistent pattern in terms of cell cycle arrest throughout the experiment after 2Gy (8 – 48 hours) and 6Gy (8 – 48 hours) with no significant difference compared to the untreated sample. This leads to the conclusion that the cell cycle processes (e.g. cell cycle checkpoints and DNA repair mechanisms) are operating and prevents cells from further continuing in the cycle if damaged. Functioning cell cycle processes, involving cell cycle checkpoint and DNA repair mechanisms would need to get confirmed in future studies.

The original hypothesis predicted that HPV-positive cell lines would show an increased G2 cell cycle arrest following IR, compared to HPV-negative OPSCC cell lines. All HPV-positive OPSCC cell lines showed an increase of cells in G2 after 24 hours after 2Gy IR. This increase was statistically significant in only 2/5 cell lines (UMSCC-47 and CU-OP-3) after 2Gy 24 hours. However, all HPV-positive OPSCC cell lines showed a significant G2 arrest after 6Gy 24 hours. A greater G2 arrest might indicate a functioning G2/M checkpoint, and this delay in DNA repair function, could be connected to the greater sensitivity of cells towards IR seen in UMCC-47 and CU-OP-20 observed in the clonogenic assays. However, a similar G2 arrest after 2Gy 24 hours was observed in CU-OP-3, which was defined as a resistant cell line. This suggests that sensitive and resistant cell line cannot be distinguished by the degree of G2 arrest.

Additionally, differences in doubling time have to be considered as a possibility in influencing the outcome in cell cycle arrest of this study. This differences in doubling time are indicated by the different seeding densities at the experimental set-up, Therefore, cells could have been at a different stage of the cell cycle.

In summary, 2/5 HPV-positive OPSCC cell lines showed a significant G2 arrest after 2Gy and 24 hours, whereas HPV-negative cell lines only showed a G2 arrest in 1/5 cell lines, compared to the untreated samples at time zero. However, a clearer difference of significant G2 arrest between HPV-positive and HPV-negative OPSCC cell lines was observed after 6Gy 24 hours. A similar pattern was observed in a cervical cancer study with cells arresting in G2 for the majority of the cell lines tested after 2Gy/5Gy treatment after 24 hours (Banath et al., 2004). In the current study, the difference between the mean

values at 2Gy 24 hours of all HPV-positive versus HPV-negative OPSCC cell lines was not significant (two-way ANOVA, $p>0.05$).

**Strength and weaknesses of the cell cycle analysis**

The strengths of this experiment include the use of a panel of 11 cell lines, compared to previous studies with fewer cell lines and no normal cell lines; all experiments were performed in triplicate, and the mean value and SD were calculated. The same number of cells for each cell line were analysed to minimise variation. The main weakness is the lack of untreated control samples for each time point, which in retrospect would have provided a useful comparison. Untreated control samples for each time point, would have provided the information of changes in cell cycle distribution in an untreated sample. Effects of confluence of cells in the individual untreated samples per time point could have been directly correlated to the treated counterpart of each time point. However, inclusion of time points at 8, 24 and 48 hours for each treatment dose, provides a useful comparison for increase or decrease of cells in G1, S and G2 after IR.

Overall, a significant G2 arrest was observed in 2/5 HPV-positive OPSCC cell lines after 2Gy and 5/5 HPV-positive OPSCC cell lines after 6Gy. Less G2 arrest was seen in the HPV-negative OPSCC cell lines with 1/5 HPV-negative OPSCC cell lines after 2Gy and 2/5 HPV-negative OPSCC cell lines after 6Gy 24 hours. The decrease in G2 after 48 hours in some HPV-positive cell lines (UMSCC-47, UPCI-SCC-90 and CU-OP-2) is suggested to be associated with DNA repair. However, a weakness of this study is that the involvement of DNA repair factors was not specifically analysed during the cell cycle analysis. Therefore, functioning DNA repair after 48 hours is considered as a theory which needs to be confirmed in future studies. This effect is observed mainly in UMSCC-47, UPCI-SCC-90, CU-OP-20 and CU-OP-2. The study could have been improved by testing additional cell lines to strengthen the conclusions regarding the differences between HPV-positive and HPV-negative OPSCC cell lines.

## 4.8. Conclusion

There are several conclusions that can be drawn from the experiments performed. One of the main conclusions is that there is a greater variation in sensitivity to IR among the HPV-positive cell lines than among HPV-negative lines. However, a consistent difference between HPV-positive and HPV-negative OPSCC was not observed, as the resistant cell lines of the HPV-positive group overlap with the HPV-negative OPSCC cell lines. The data obtained were partly consistent with the original hypothesis, as more HPV-positive OPSCC cell lines are more sensitive to IR compared to HPV-negative OPSCC cell lines. A greater number of *in-vitro* models in both groups (HPV-positive and HPV-negative), might provide a clearer picture. This does not mean that HPV does not play a role, but it is not the dominant factor determining survival in these *in-vitro* models. A novel finding in this study was the observed variation in sensitivity to IR within the HPV-positive cell lines. This has never been previously reported, and suggests that in future, larger panels of cell lines need to studied to make more robust conclusions about the link between HPV and sensitivity/resistance to therapy. The heterogeneity of cell lines might reflect the clinical situation where 80% of the HPV-positive patients respond better to treatment after 3 years. Thus even within the HPV-positive patients a variation in treatment response is observed (Evans et al., 2013). At the time of this study it was not known whether the cell lines were derived from patients in the favourable response group (80%), or the unfavourable (20%) response group.

There were also no clear differences between HPV-positive and negative OPSCC cell lines after IR, in terms of cell cycle or levels of p53, p21 and p-p53. Furthermore, HPV-positive OPSCC cell lines showed p53 expression, hence p53 is not completely degraded by HPV E6. The observed diverse response between the two groups of HPV-positive versus HPV-negative cell lines (and within the groups themselves) appears to be a reflection of biological heterogeneity of OPSCC cell lines. This is similar to the diversity observed in terms of patient survival. This suggests that conclusions drawn from studies with a limited number of cell line models should be regarded with caution. Therefore, it is important to increase the number of cell lines models in further studies or to develop methods to research primary patient material directly and follow-up these observations by tracking the patient response to therapy.

# Chapter 5 : Investigation of gene expression in OPSCC cell lines using mRNA sequencing

## 5.1. Introduction

mRNA sequencing or transcriptome sequencing (RNA-sequencing) is a high-throughput method that provides a quantitative and sensitive insight into the whole transcriptome of cells at a specific time point or treatment dose (Kukurba and Montgomery, 2015, Mortazavi et al., 2008). Furthermore, it has evolved into a technique for tumour characterization, and plays a key role in enhancing our fundamental understanding of the biology of certain cancers and their connection with virus integration (HPV integration). Cell lines with and without integrated HPV are characterized by different patterns of gene expression (Parfenov et al., 2014). One main approach of mRNA sequencing in the current study is to identify differences of gene expression, splice variant or mutations.

In general, mRNA transcripts are controlled by post-transcriptional regulation, which is mediated by RNA-binding proteins (RBPs) that regulate specific subsets of mRNAs. Together with small interfering RNAs, microRNAs and protein effector complexes, degradation or translation of transcripts is controlled (Mata et al., 2005). Later transcripts assemble with translation factors and ribosomes for protein synthesis. Proteins can then be further controlled through modifications of amino acids, cleavage by site-specific proteases or degradation of proteasome. However, the protein level does not have to directly correlate with the mRNA expression level (Halbeisen et al., 2008). As mentioned above mRNA sequencing can be a helpful tool in identifying changes in gene expression, but it does not show protein level change in activity.

In the current study, mRNA sequencing was used to investigate differences in gene expression between HPV-positive and HPV-negative OPSCC cell lines. The effect of IR on gene expression was also investigated.

These was done based on the following aims:

- To assess HPV integration and gene expression in HPV-positive cell lines, and to assess gene expression and p53 mutations in the whole panel of cell lines.
- To investigate if there is a response to IR treatment in changes in gene expression between HPV-positive and HPV-negative OPSCC cell lines (e.g. DNA repair).

**The hypotheses**:

1. By consideration of the HPV status, differences in gene expression will be observed between HPV-positive and HPV-negative OPSCC cell lines after IR.
2. Variation in radio-sensitivity among the HPV-positive OPSCC cell lines will be identified by different up-/down-regulations of genes after treatment.
    a. These genes could be associated with DNA repair pathways.

### 5.1.1. Experimental design

Ten cell lines, 5 HPV-positive and 5 HPV-negative OPSCC cell lines, were selected to use as a study panel. This study panel was extended by the corresponding 2Gy treated sample for each cell line. Samples were collected after 24 hours, to identify a long-term effect of changes in gene expression between untreated and 2Gy treated samples. This provided the opportunity to correlate it to the obtained radio-sensitivity assays (10-15 days), by choosing a relative longer time point for the mRNA sequencing study. Later the data obtained from the radio-sensitivity data was used to classify some analysis in this chapter according to the sensitivity status to IR.

The experiment was performed with a single sample of each cell line and treatment dose. Due to the direct comparison of untreated with treated samples of each cell line, one replicate was considered to be sufficient.

The aim of this study was to select OPSCC cell lines of a wide diversity. The HPV-positive cell lines consisted of CU-OP-2, CU-OP-3 (derived from heavy smokers – data not shown) (Pirotte, 2017). CU-OP-20 was derived from a patient who has never smoked (described in Chapter 3). UMSCC-47 and UPCI-SCC-90 were also included. All HPV-positive cell lines harboured wild-type p53 and were identified with the greatest variation to IR (Radio-sensitivity assay Chapter 4). CU-OP-20 was later identified during mRNA sequencing to harbour wild-type p53. The p53 status for the HPV-negative cell lines varied from UMSCC-19 and UMSCC-4 (derived from tumour receiving post-treatment of 2 chemotherapies) described previously as showing mutations for p53 and UMSCC-6 and UMSCC-74a (originated from a recurrent tumour – after receiving chemotherapy) harbouring wild-type 53 (Bradford et al., 2003, Brennan et al., 1995, Brenner et al., 2010, Ferris et al., 2005, Lin et al., 2007, Mandic et al., 2005, Somers et al., 1992). The novel cell line CU-OP-17 (described Chapter 3) was later identified during mRNA sequencing to harbour wild-type p53 in the coding region, but represents a likely pathogenic SNP at the UTR 5' region of p53.

The experiment was used to test the general differences in gene expression between HPV-positive and negative cell lines. Furthermore, this experiment was designed to test changes in gene expression after IR treatment (e.g. DNA repair pathways) in the entire panel of OPSCC cell lines. Differences of gene expression were further investigated regarding the sensitivity to IR within the HPV-positive cell lines.

## 5.2. Data and quality control

Preliminary bioinformatics quality control analyses were performed by Dr. Peter Giles (Wales Gene Park). Sample quality control was performed as described in section 2.8.4 (Material and Methods). Reads were mapped against the HPV-16 reference sequence (NC_001526) and the human reference sequence hg19 (genome build 19) (Table 5-1 for untreated and 2Gy IR treated samples).

**Table 5-1: Mapped reads for HPV-16 and human (hg19) sequences from untreated and treated samples**

| | | Cell line | Mapped reads (hg19) | Mapped reads (HPV-16) |
|---|---|---|---|---|
| HPV-positive | untreated | UMSCC-47 | 55,704,368 | 1336 |
| | | UPCI-SCC-90 | 49,934,662 | 11252 |
| | | CU-OP-20 | 43,679,896 | 11040 |
| | | CU-OP-2 | 25,625,574 | 2156 |
| | | CU-OP-3 | 25,594,880 | 9409 |
| | 2Gy IR treated | UMSCC-47 | 37,853,816 | 1690 |
| | | UPCI-SCC-90 | 34,994,300 | 15703 |
| | | CU-OP-20 | 33,492,546 | 8110 |
| | | CU-OP-2 | 48,538,120 | 2484 |
| | | CU-OP-3 | 45,589,324 | 8506 |
| HPV-negative | untreated | UMSCC-19 | 39,432,500 | 1 |
| | | UMSCC-6 | 45,644,480 | 0 |
| | | UMSCC-74a | 52,895,862 | 0 |
| | | UMSCC-4 | 31,573,578 | 0 |
| | | CU-OP-17 | 51,291,862 | 0 |
| | 2Gy IR treated | UMSCC-19 | 47,035,588 | 0 |
| | | UMSCC-6 | 58,038,150 | 0 |
| | | UMSCC-74a | 60,398,218 | 0 |
| | | UMSCC-4 | 49,670,746 | 0 |
| | | CU-OP-17 | 50,831,804 | 0 |

## 5.3. Characterization of novel and established cell lines

### 5.3.1. HPV gene expression in HPV-positive OPSCC cell lines

For all 5 HPV-positive OPSCC cell lines, RNA transcripts mapped to the HPV-16 sequence (untreated and treated sample) (Figure 5-1). The total number of reads can vary, depending on the cell line (hg19_hpv16 combined) (Figure 5-1 & Table 5-1). HPV-16 reads provided an additional confirmation of the HPV-status of the novel cell line of CU-OP-20. HPV status was tested as well for the two novel cell lines CU-OP-17 and CU-OP-20 by PCR testing. HPV status of the other HPV-positive cell lines was confirmed in the HPV group during a previous study. The data provided a confirmation of HPV status, type and integration sites of the known cell lines of UMSCC-47 and UPCI-SCC-90. (Pirotte, 2017).

Figure 5-2 summarises the Reads Per Kilobase of transcripts per Million mapped reads (RPKM) of HPV gene expression within the panel of HPV-positive OPSCC cell lines. The expression pattern of individual HPV genes varied between the cell lines. The oncogenes E6 and E7 showed the strongest expression in three HPV-positive OPSCC cell lines (UMSCC-47, UPCI-SCC-90 and CU-OP-20) in the untreated and 2Gy treated samples. Furthermore, CU-OP-20 showed high expressions of E2 and E5 genes. No or low expression of the late genes of L1/L2 were present in all HPV-positive OPSCC cell lines indicating a non-productive HPV infection. However, a low level of expression of L1/L2 genes was seen in UPCI-SCC-90 (untreated & treated 2Gy IR) with a range of 11 (untreated) and 22 (treated) reads.

The highest levels of E6 and E7 gene expression were observed in UPCI-SCC-90, CU-OP-3 and CU-OP-20. The highest expression of E4 was observed in CU-OP-20 and comparably higher expression of E4 was observed as well in CU-OP-2 and CU-OP-3. Except for UMSCC-47, all other HPV-positive OPSCC cell lines showed expression of E4, E5 and E2. A disruption of E2, E4 and E5 gene expression suggests a fully integrated HPV virus in UMSCC-47.

IR has an effect on the expression of HPV open-reading frames (ORF) in certain cell lines. A decrease of a fold-change between untreated and 2Gy treated sample was observed in CU-OP-20 for E2 (-0.9858), E4 (-0.9647) and E5 (-1.6646). A fold-change increase of HPV ORF's after IR, was detected in UPCI-SCC-90 for E5 (1.02368), L2 (1.12758) and L1 (1.05117). CU-OP-3 showed a fold-change increase in L1 (1.24533).

All other cell lines and HPV ORF's showed a minimal fold-change difference after IR. Fold changes in E4, E5, L2 and L1 could indicate an episomal HPV expression (no total HPV integration).

### 5.3.2. Splicing of HPV-16 transcripts

HPV transcription is regulated from two main promoters, the early p97 and the late p670 promoter. Activation of these promoters leads to synthesis of polycistronic mRNAs. These mRNAs are then further regulated by differential splicing. High differential or alternative splicing is found in the HPV-16 genome and other high risk HPV types e.g. HPV-18 and HPV-31. Different oncogenic potential has been ascribed to the resulting transcripts (e.g. splicing at E6) (Rosenberger et al., 2010).

Splicing patterns were assessed in the current panel of cell lines (including the novel line CU-OP-20) and are presented as Sashimi plots (Katz et al., 2015) in Figure 5-3. Splice forms were identified using information from the Papillomavirus Episteme website (https://pave.niaid.nih.gov/#explore/transcript_maps/hpv16 - on the 14/07/2018 (Van Doorslaer et al., 2017). Different splicing results across the panel of HPV-positive cell lines (transcripts (isoforms) and position) are shown in Table 5-1 (untreated samples) and Table 5-2 (2Gy treated samples).

All HPV-positive cell lines showed the most common transcripts of E6 in the untreated samples (E6*I and E6*II). CU-OP-2 only showed a transcript of one of the E6 isoforms, E6*I. The most common splicing site in the panel of HPV-positive cell lines is E1^E4, which suggest a transcription via episomal HPV. Expression of E2 genes, suggests no full integrated HPV into the host genome (Figure 5-2) (Dona et al., 2013). The E1^E4 site is shown in all cell lines except for UMSCC-47. CU-OP-20 showed the greatest variation of splicing sites in terms of transcripts within the panel of cell lines. All cell lines derived at Cardiff University also showed transcripts at E2 or E5. The lowest number of splicing sites was detected in UMSCC-47 (main sites are E6*I and E6*II).

A change of alternative splicing sites was documented (mapped reads hg19_HPV16 combined (bam files)) in most cell lines after 2Gy IR treatment, which might be a result of IR. An increase of splicing sites is documented in all OPSCC cell lines after IR (counts with 10 reads included): UMSCC-47 from 2 splicing sites to 3 (one identical), UPCI-SCC-90 from 3 splicing sites to 4 (3 identical), CU-OP-20 from 7 splicing sites to 10 (6

identical), CU-OP-2 from 4 splicing sites to 5 (3 identical) and CU-OP-3 from 4 splicing sites to 7 (3 identical). Those identical splicing sites vary in number of counts. The counts decrease in number in all identical splicing sites, except for CU-OP-2 and CU-OP-3, were an increase of counts was detected for E1^E4 (880 (start nt) – 3356 (end nt). CU-OP cell lines showed additional splicing sites for E5 or E2 transcripts. UMSCC-47 showed a novel splicing site in the E6 region and E2 (10 counts) after 2Gy IR treatment.

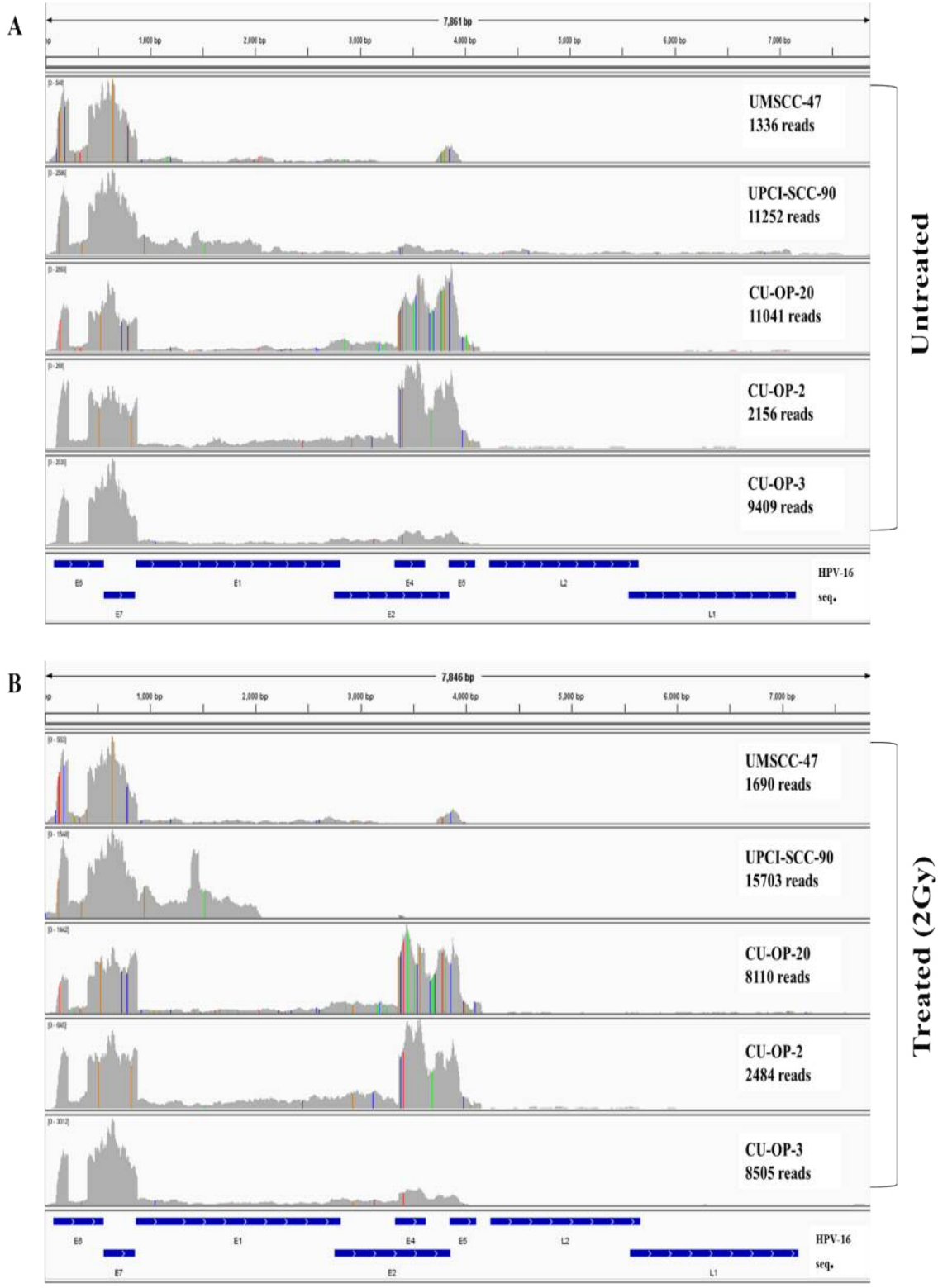



**Figure 5-1: Mapping of reads against HPV-16**

(A) Raw reads against HPV-16 in untreated samples (B) Raw reads against HPV-16 in treated samples (2Gy). The HPV-16 genome with Open reading frames (ORF's) (E6, E7, E1, E2, E4, E5 and the L1, L2) shown below the mapped reads. For each cell line, the number of reads mapped to the HPV-16 (NC_001526) are listed above. The coloured lines within the sequence indicate mismatches with the reference sequence. This graph was obtained by using the IGV software (IGV version 2.3.65, Broad Institute) (Robinson et al., 2011).

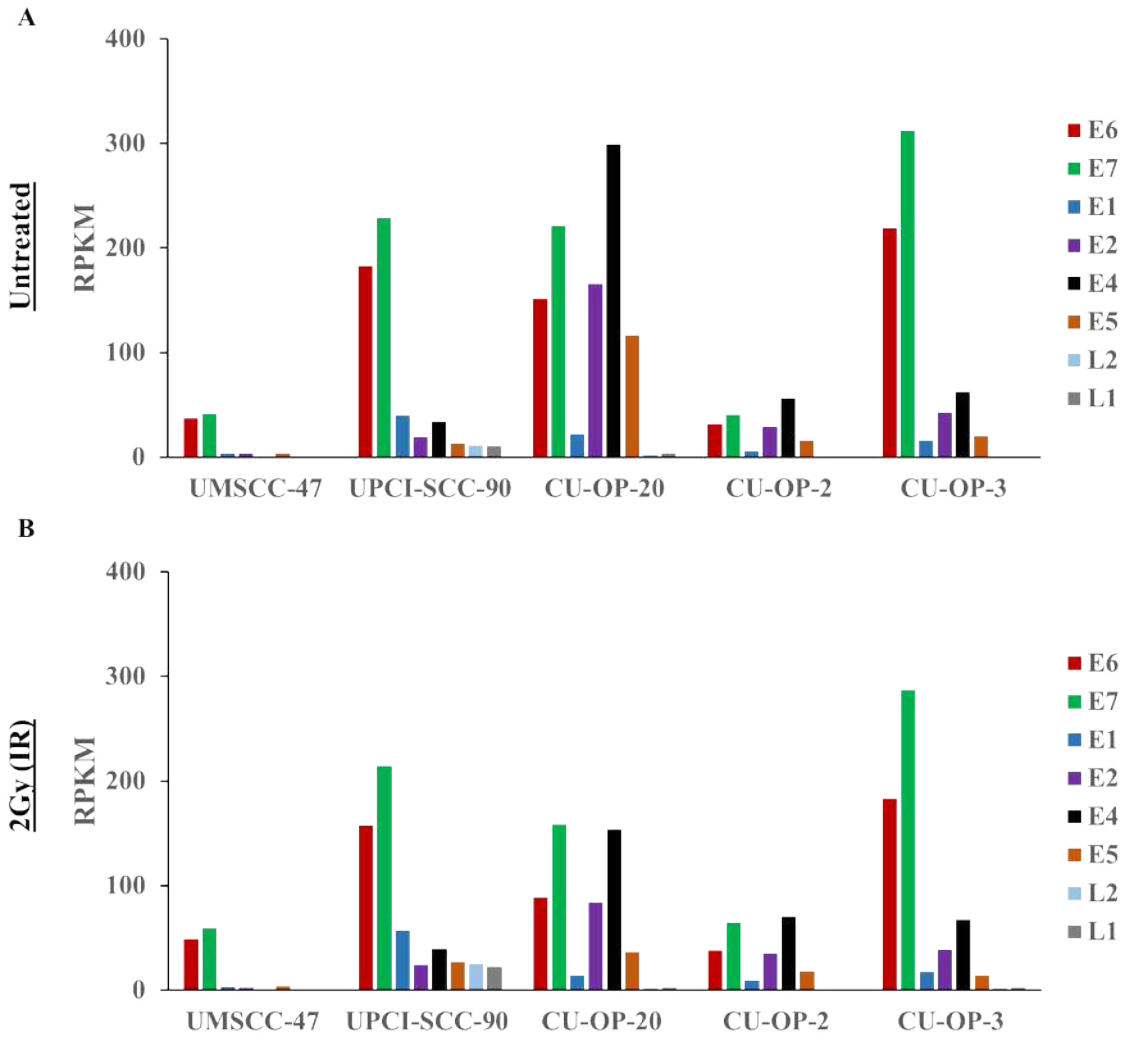



**Figure 5-2: HPV gene expression in UMSCC-47, UPCI-SCC-90, CU-OP-20, CU-OP-2 and CU-OP-3** Reads per Kilobase of transcript per million Mapped reads (RPKM) are plotted for each HPV Open Reading Frame (ORF) (per cell line). (A) HPV gene expression in untreated samples (B) HPV gene expression in 2Gy treated samples. Greatest fold-change seen in E7, E4 and E5 expression in CU-OP-20.

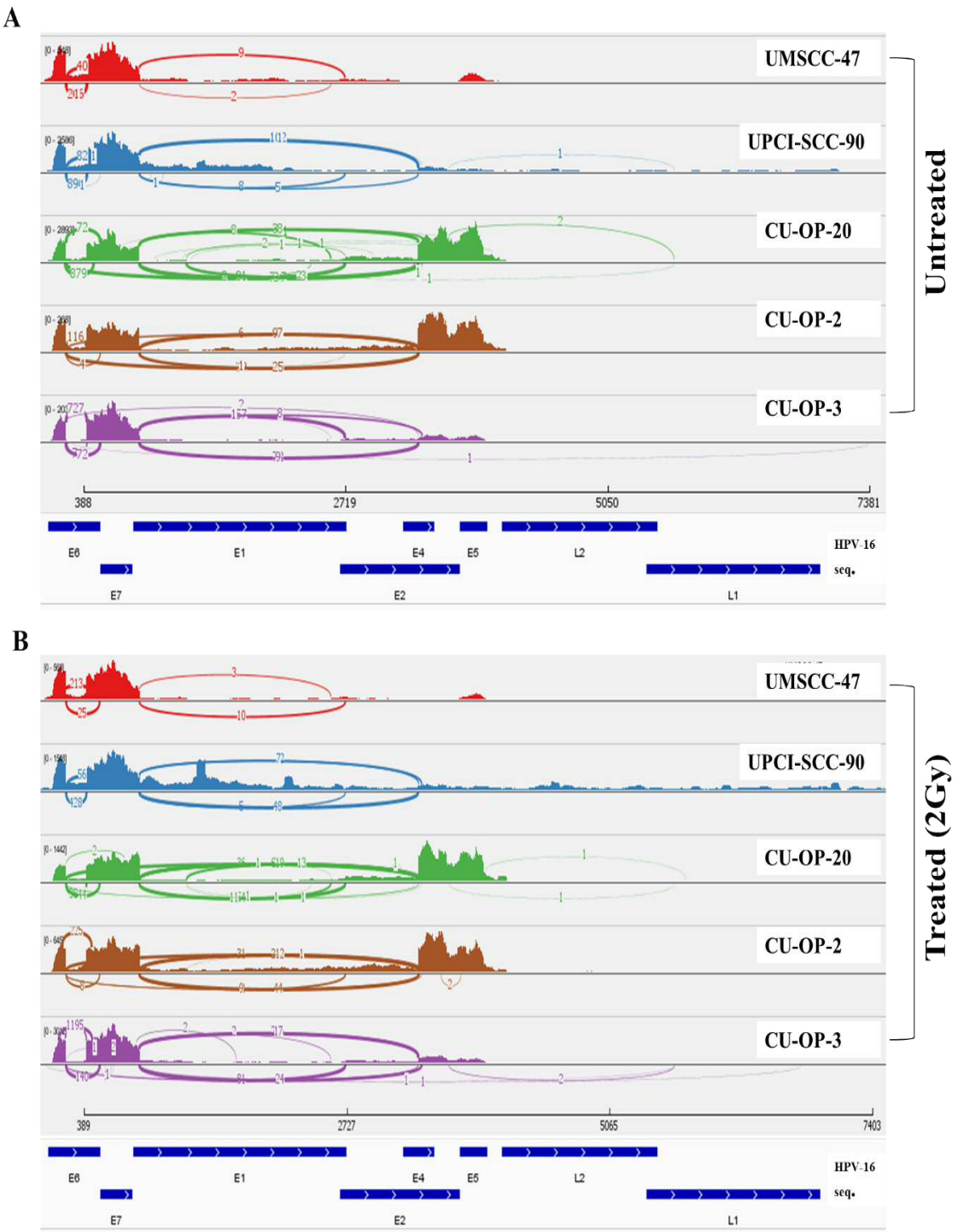



**Figure 5-3: Sashimi plots (Splice Junctions) in HPV-positive OPSCC cell lines**

Figures represent reads with ends mapping at different sites of the HPV genome, visualised using IGV. HPV transcription is regulated by two promoters. Activation of these promoters leads to synthesis of polycistronic mRNAs. These mRNAs are then further regulated by differential splicing. The loops in the graphs from one HPV site to another HPV site of the sequence (e.g. E6*I). It is representing different splicing sites of the HPV genome (A) shows the Sashimi plot of untreated samples (B) shows the Sashimi plot of treated samples (2Gy).

**Table 5-2: Quantification of HPV-16 spliced transcripts in untreated samples**

| | Cell line | Start (nt)[A] | End (nt) | Count[B] | Transcript[C] |
|---|---|---|---|---|---|
| Untreated | UMSCC-47 | 226 | 408 | 206 | E6*I |
| | | 226 | 525 | 40 | E6*II |
| | UPCI-SCC-90 | 226 | 408 | 890 | E6*I |
| | | 226 | 525 | 82 | E6*II |
| | | 880 | 3356 | 102 | E1^E4 |
| | CU-OP-20 | 216 | 472 | 879 | E6*I |
| | | 226 | 3356 | 64 | E5 |
| | | 226 | 525 | 72 | E6*II |
| | | 880 | 2707 | 81 | E2 |
| | | 1300 | 3356 | 23 | E5 |
| | | 880 | 3356 | 767 | E1^E4 |
| | | 880 | 3359 | 32 | E1^E4 |
| | CU-OP-2 | 226 | 408 | 116 | E6*I |
| | | 226 | 3356 | 39 | E5 |
| | | 880 | 3359 | 99 | E1^E4 |
| | | 880 | 3356 | 25 | E1^E4 |
| | CU-OP-3 | 226 | 408 | 72 | E6*I |
| | | 226 | 525 | 727 | E6*II |
| | | 880 | 2707 | 57 | E2 |
| | | 880 | 3356 | 72 | E1^E4 |

[A] Nucleotide (nt)
[B] Splice sites with a minimum of 10 reads were considered to avoid confounding due to sequencing artefacts
[C] Splice forms are named with reference to the Papillomavirus Episteme database (Van Doorslaer et al., 2017).

**Table 5-3: Quantification of HPV-16 spliced transcripts in 2Gy treated samples**

| | Cell line | Start (nt)[A] | End (nt) | Count[B] | Transcript[C] |
|---|---|---|---|---|---|
| Treated 2Gy (IR) | UMSCC-47 | 226 | 525 | 25 | E6*II |
| | | 175 | 474 | 213 | novel |
| | | 880 | 2707 | 10 | E2 |
| | UPCI-SCC-90 | 226 | 408 | 428 | E6*I |
| | | 226 | 525 | 56 | E6*II |
| | | 880 | 3356 | 48 | E1^E4 |
| | | 880 | 3359 | 12 | E1^E4 |
| | CU-OP-20 | 226 | 408 | 385 | E6*I |
| | | 226 | 525 | 44 | E6*II |
| | | 226 | 3356 | 36 | E5 |
| | | 880 | 2579 | 11 | E2 |
| | | 880 | 2707 | 54 | E2 |
| | | 1300 | 3356 | 13 | E5 |
| | | 226 | 3356 | 36 | E5 |
| | | 880 | 3389 | 19 | E1^E4 |
| | | 880 | 3356 | 616 | E1^E4 |
| | | 880 | 3359 | 34 | E1^E4 |
| | CU-OP-2 | 193 | 450 | 225 | E6*I |
| | | 226 | 3356 | 31 | E1^E4 |
| | | 880 | 3359 | 44 | E1^E4 |
| | | 880 | 3356 | 263 | E1^E4 |
| | | 880 | 3389 | 12 | E5 |
| | CU-OP-3 | 226 | 525 | 140 | E6*II |
| | | 161 | 462 | 1195 | E6 |
| | | 221 | 408 | 10 | E6*I |
| | | 880 | 2707 | 31 | E2 |
| | | 880 | 3389 | 24 | E5 |
| | | 880 | 3356 | 206 | E1^E4 |
| | | 880 | 3359 | 17 | E1^E4 |

[A] Nucleotide (nt)
[B] Splice sites with a minimum of 10 reads were considered to avoid confounding due to sequencing artefacts
[C] Splice forms are named with reference to the Papillomavirus Episteme database (Van Doorslaer et al., 2017).

### 5.3.3. HPV-variants

The genome sequence of HPV types is highly conserved, however minor differences within the HPV-16 sequence can be found. These minor differences are termed HPV-variants, and different variants are connected to geographical and ethnic distribution. The variants are classified into four main groups of European, Asian, American and African, with subtypes for each group (Burk et al., 2013, Fontecha et al., 2015). Variations at specific loci in the highly conserved E6 and LCR region are used to identify the HPV-16 specific variants listed in Table 5-3. The E6 nucleotide sequence of the new developed CU-OP-20 cell line was compared to mismatched sites of known HPV-16 variants; this analysis is presented in Table 5-3. CU-OP-20, showed sequence variation consistent with Asian-American origin. UPCI-SCC-90, CU-OP-2 and CU-OP-3 showed the main European HPV-16 variant, whereas UMSCC-47 showed an African 2A variant. UCPI-SCC-90 and UMSCC-47 were derived in the United States of America (USA).

**Table 5-4: Different HPV variants & comparison to HPV-16 sequence of the panel of HPV-positive cell lines**

| Variant or cell line | Nucleotide position (HPV-16 E6 sequence)[A] | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **109** | **131** | **132** | **143** | **145** | **178** | **286** | **289** | **295** | **335** | **350** | **403** | **532** |
| **HPV-16 reference** | T | A | G | C | G | T | T | A | T | C | T | A | A |
| **EUR (A1/A2)[B]** | /C | /G | - | - | - | /A | - | - | - | /T | /G | - | - |
| **EAS[B]** | - | - | - | - | - | **G/C** | - | - | - | - | - | - | - |
| **AFR1a[B]** | - | - | **C** | G | T | - | A | G | - | T | - | - | - |
| **AFR1b[B]** | - | G | - | G | T | - | A | G | **G** | T | G | - | - |
| **AFR2a[B]** | C | - | **T** | G | T | - | A | G | - | T | - | **G** | - |
| **AFR2b[B]** | - | - | - | G | T | - | A | G | - | T | - | - | - |
| **NA[B]** | - | - | - | - | T | - | A | G | - | T | G | - | - |
| **AA1[B]** | - | - | - | - | T | - | A | G | - | T | G | - | G |
| **AA2[B]** | - | - | - | - | T | - | A | G | - | T | G | - | /G |
| **UMSCC-47** | C | A | T | G | T | T | A | G | T | T | T | G | A |
| **UPCI-SCC-90** | T | G | G | C | G | T | T | A | T | C | G | A | A |
| **CU-OP-20** | T | A | G | C | T | T | A | G | T | T | G | A | G |
| **CU-OP-2** | T | A | G | C | G | T | T | A | T | C | T | A | A |
| **CU-OP-3** | T | A | G | C | G | T | T | A | T | C | G | A | A |

[A] A dash indicates no change to the HPV-16 reference sequence and a mismatch of nucleotide position is indicated in **bold red**. Nucleotides separated by a slash indicate a different nucleotide position for a given sub-lineage (e.g. A1/A2). The top half of the table shows the known HPV-16 variants described by Cornet et al. (Cornet et al., 2012).

[B] Abbreviations: EUR – European; EAS – European-Asian; AFR – African; NA – North-American: AA – Asian-American.

### 5.3.4. HPV integration sites

Integrated virus is present in the majority of HPV-positive tumours and integration of the HPV DNA into the host genome confers a proliferative advantage *in-vitro* (Jeon et al., 1995, Parfenov et al., 2014). HPV-integration has a major impact on the human genome as alteration in DNA copy number or mRNA transcript abundance or splicing have been previously reported (Olthof et al., 2015, Parfenov et al., 2014). Viral integration can potentially change the expression of genes in the human genome and lead to carcinogenesis (Olthof et al., 2015). This effect might also be influenced by IR treatment. In combination, these effects can cause genetic instability and disease progression (Jeon et al., 1995, Parfenov et al., 2014). Integration sites for UMSCC-47, UPCI-SCC-90, CU-OP-2 and CU-OP-3 have been previously reported and described in untreated samples (Olthof et al., 2015, Pirotte, 2017).

### 5.3.5. Fusion transcripts

The presence of integrated virus, and the site of integration into the human genome, can be detected by analysis of fusion transcripts that include both human and viral sequences (i.e. reads with one end mapping to human sequence, and the other mapping to viral sequence). Fusion transcripts were visualized by Circos plots, and insertion sites were then further analysed by using IGV (Figure 5-4 and Table 5-5). Circos plots were used to indicate HPV integration (NC_001526) into the human genome hg19 before and after 2Gy IR treatment in all HPV-positive OPSCC cell lines. HPV-integration sites were of interest, as there might be an association between integration and adjacent host genomic structural variation in cancer cell lines and primary tumours (Akagi et al., 2014). Gene information for the main integration site was obtained from the GeneCards Human Gene Database (http://www.genecards.org/- on the 04/07/2017) and previous publications on mRNA sequencing data for UMSCC-47 and UPCI-SCC-90 (Akagi et al., 2014).

One main integration site was detected in UMSCC-47, which was located at chromosome 3, with fusion transcripts at E6, E7 and a few in E5 (less than 5 reads – not listed in table). The main integration site within chromosome 3 was located at Exon 11 and some additional reads at Exon 9 at the human gene TP63. TP63 is a key regulator of epithelial differentiation (Bergholz and Xiao, 2012) and it encodes a member of the p53 family of transcription factors. Alternative splicing of TP63 and the use of alternative promoters

can lead to different isoforms of the gene with variations in functional properties. Additionally, it may be required for initiation of p53/TP53 dependent apoptosis in response to genotoxic insults. TP63 also activates transcription of p21. Diseases connected to TP63 include Adult syndrome (ectodermal dysplasia e.g. excessive freckling) and Limb-Mammary Syndrome (ectodermal dysplasia e.g. severe hand and/or foot anomalies).

The greatest number of fusion transcripts among the remaining cell lines, was detected in the UPCI-SCC-90 cell line as indicated in Figure 5-4. The main integration site was detected at chromosome 9 of the gene C9orf156 (Chromosome 9 Open Reading Frame 156) also known as TRMO gene, with reads mapping at E6 and E7 for Exon 3 and Exon 4 and E1 for Exon 1 and Intron 2. Another integration site was detected around 200 bp before FOXE1 from E7. This integration site has been previously reported for UPCI-SCC-90 (Bergholz and Xiao, 2012, Pirotte, 2017). C9orf156 is a poorly characterized gene, but is known as a protein coding gene and associated with diseases as isolated cleft lip and cleft palate. FOXE1 belongs to the forkhead family and functions as a thyroid transcriptions factor connected to thyroid morphogenesis.

The main integration site for CU-OP-2 mapped to chromosome 10 with reads originated at E2. Reads mapped with YME1L1 gene at Introns 6 and 7. YME1L1 is the human orthologue to Yme1p (yeast mitochondrial AAA metalloprotease). This protein is localized in the mitochondria and the gene might play a role in mitochondrial metabolism and could be involved in mitochondrial pathologies. Diseases associated with the gene include Optic Altrophy (loss of some or all nerve fibres) and Asphyxia Neonatorum (affects babies, if not enough oxygen during the birth process).

The main integration site for CU-OP-3 was located at chromosome 20 at the gene CEBPB (CCAT/Enhancer Binding Protein Beta). The majority of fusion transcripts mapped to E7 and a lower number mapped to E1 at Exon1. Interestingly, the CEBPB gene is only formed of one Exon with several fusion transcripts appear to be linked to E7 splice donor region. It is an important transcription factor regulating genes involved in immune and inflammatory responses. Furthermore, it is involved in adipogenesis, as well as in the gluconeogenic pathway, liver regeneration and hematopoiesis. Diseases associated with mutation of CEBPB included Acute Promyelocytic Leukemia.

Main integration sites for UMSCC-47, UPCI-SCC-90, CU-OP-2 and CU-OP-3 did not change after ionising radiation with 2Gy. The number of minor HPV-integration sites

increased after IR (Figure 5-4 and Figure 5-5). Three HPV-OPSCC cell lines showed a greater increase of chromosomal integration sites UMSCC-47, CU-OP-20 and CU-OP-3. UMSCC-47 showed three chromosomal integration sites in the untreated samples, which increased to 16 different chromosomal integration sites after IR (difference of 13). However, the number of total integration reads decreased by 1% after IR. CU-OP-20 showed 7 different chromosomal integration sites in the untreated sample, which increased to 20 sites after IR (difference of 13 with an increase of reads by 922,2%). CU-OP-3 showed 6 different chromosomal integration sites in the untreated sample, which increased to 21 after IR (difference of 15 with an increase of reads by 76,19%). UPCI-SCC-90 (difference of 3 chromosomal sites and a decrease of reads by 50%) and CU-OP-2 (difference of 9 with an increase of 167,64%) showed an increase of chromosomal integration sites, with a minor difference to the untreated sample.

The novel OPSCC cell line CU-OP-20 did not show any main integration site. Minor integration sites (less than 5 mapped reads) were detected in the untreated and treated samples, for example: from E7 into Exon 3 of ZC3H12C in chromosome 11 or from E2/E4 into Exon 8 of EEF1A1 in chromosome 6.

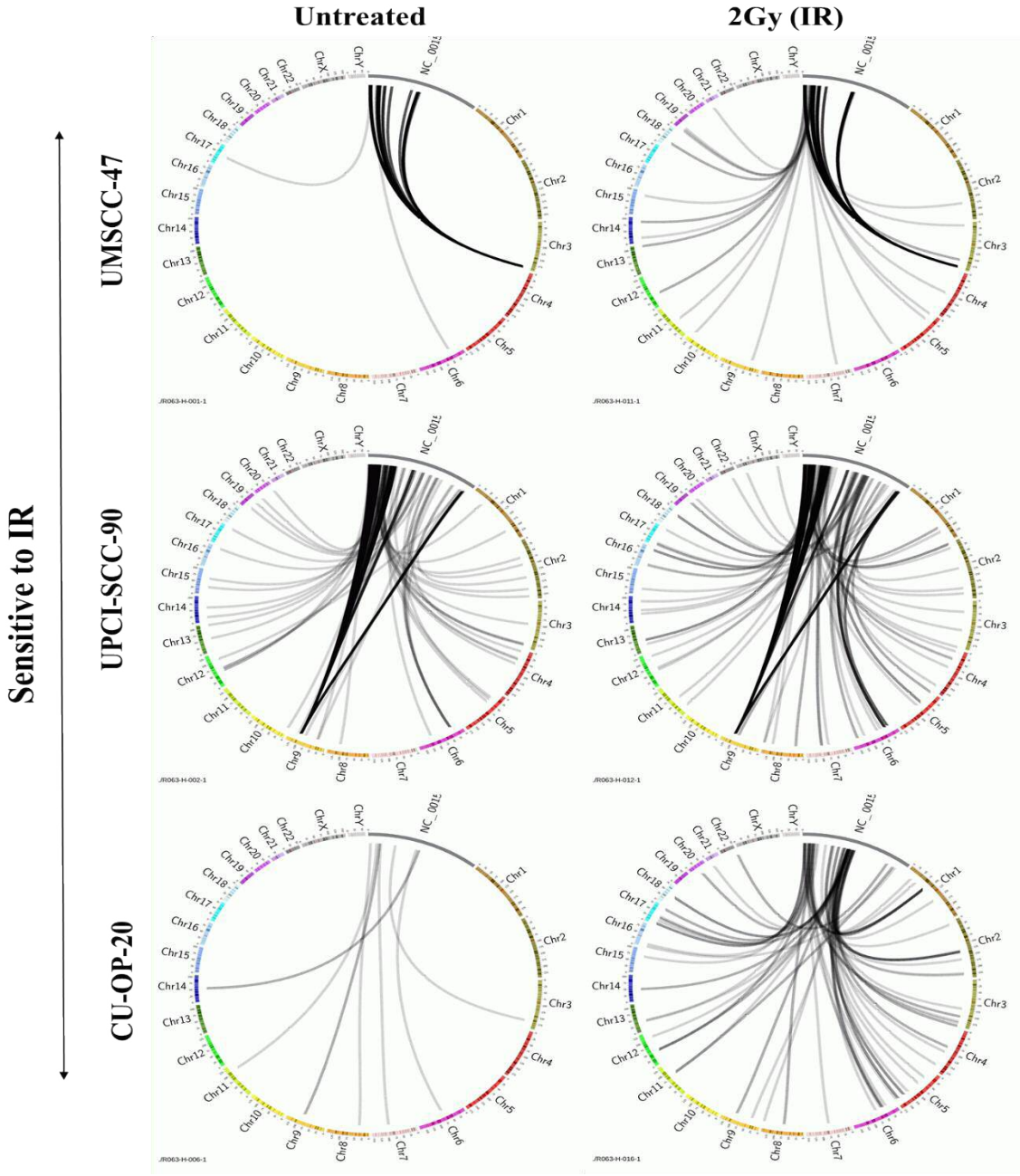



**Figure 5-4: Circos plots showing HPV:human fusion transcripts in untreated and treated (2Gy) samples (sensitive to IR)**

The circle is comprised of HPV16 sequence in grey and human chromosomes as coloured bars. Fusion transcripts are represented as grey lines joining the loci to which they map in human and HPV sequence. Thicker/darker lines indicate multiple reads mapping to the same loci.

The number of fusion transcripts appeared to increase following IR treatment of 2Gy. Multiple integration sites were detected in CU-OP-20 in untreated and treated samples. An increase of minor integration sites (less than 5 mapped reads at different chromosome sites) was detected after IR or 2Gy in all OPSCC cell lines.

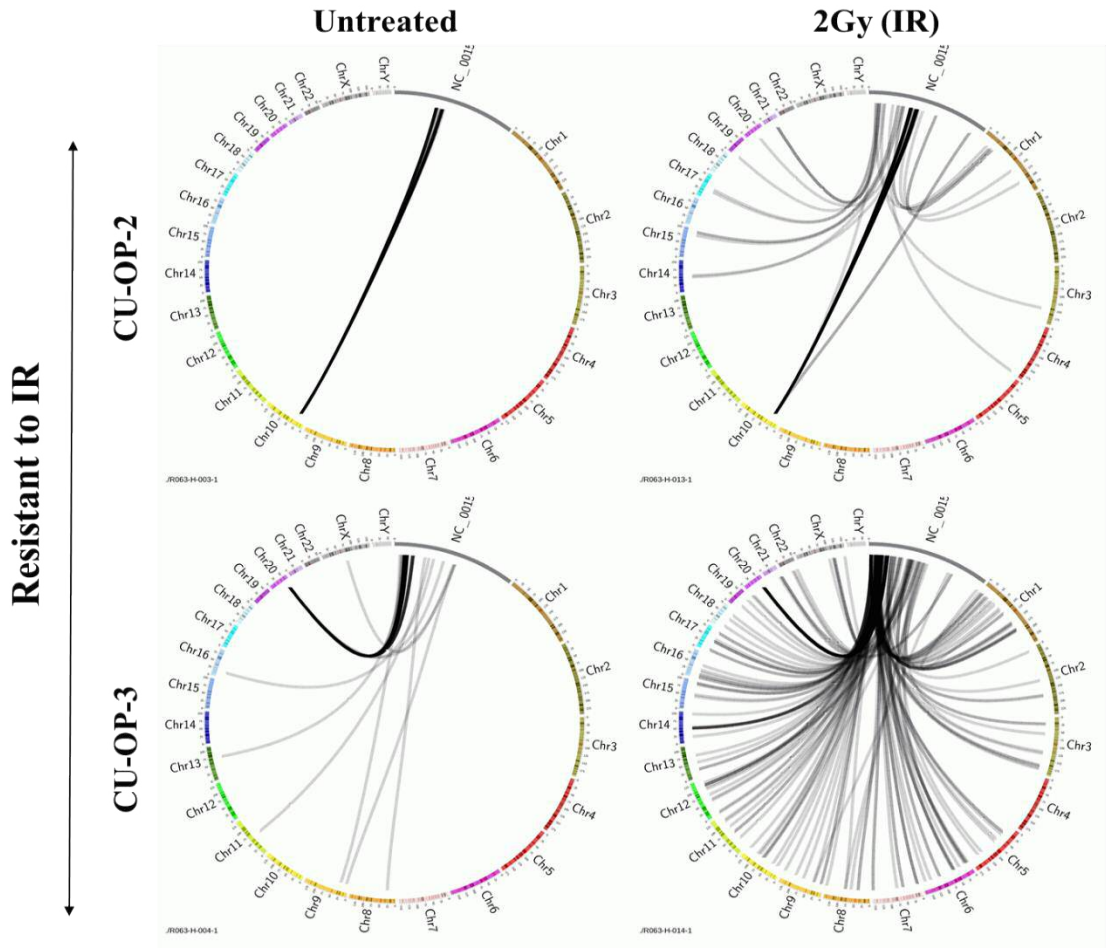



**Figure 5-5: Circos plots showing HPV: human fusion transcripts in untreated and treated (2Gy) samples (resistant to IR)**

The circle is comprised of HPV16 sequence in grey and human chromosomes as coloured bars. Fusion transcripts are represented as grey lines joining the loci to which they map in human and HPV sequence. Thicker/darker lines indicate multiple reads mapping to the same loci.

The number of fusion transcripts appeared to increase following IR treatment of 2Gy. An increase of minor integration sites (less than 5 mapped reads at different chromosome sites) was detected after IR or 2Gy in all OPSCC cell lines.

**Table 5-5: HPV-16 integration sites for HPV-positive OPSCC cell lines in untreated and 2Gy (IR) treated samples.**

| | Cell line | HPV gene | Human chromosome | Human gene[A] |
|---|---|---|---|---|
| Untreated | UMSCC-47 | E7 | 3 | TP63: Exon 9 ; Exon 11 |
| | UPCI-SCC-90 | E7 | 9 | C9orf156 Exon 3 |
| | | E7 | 9 | Non-coding +200 bp before FoxE1 |
| | | E7 | 9 | C9orf156 Exon 4 |
| | | E6 | 9 | C9orf156 Exon 3 |
| | | E6 | 9 | C9orf156 Exon 4 |
| | | E1 | 9 | C9orf156 Intron 2 |
| | CU-OP-2 | E2 | 10 | YME1L1 (intron 6 and 7) |
| | CU-OP-3 | E7 | 20 | CEPBP Exon 1 |
| 2Gy (IR) treated | UMSCC-47 | E7 | 3 | TP63: Exon 9 ; Exon 11 |
| | UPCI-SCC-90 | E7 | 9 | C9orf156 Exon3 |
| | | E6 | 9 | C9orf156 Exon 4 |
| | | E6 | 9 | C9orf156 Exon 3 |
| | | E1 | 9 | C9orf156 Intron 2 |
| | | E7 | 9 | C9orf156 Intron 3 |
| | | E7 | 9 | C9orf156 Exon 4 |
| | | E7 | 9 | Non-coding +200 bp before FoxE1 |
| | CU-OP-2 | E2 | 10 | YME1L1 (intron 6 and 7) |
| | CU-OP-3 | E7 | 20 | CEPBP Exon 1 |

[A] Sites with at least 5 reads mapped are listed – CU-OP-20 did not show a main integration site (i.e. => 5 mapped reads) and is therefore not listed in this table.

### 5.3.6. P53 status (SNPs and mutation) of OPSCC cell lines

It has been suggested that p53 gene mutation is common in HPV-negative and rare in HPV-positive OPSCC cell lines (Guerrero-Preston et al., 2014). Published literature suggests that p53 is wild-type in the HPV-positive cell lines UMSCC-47 and UPCI-SCC-90 as well as in two HPV-negative cell lines UMSCC-6 and UMSCC-74, while two HPV-negative cell lines UMSCC-4 and UMSCC-19 were described as harbouring mutations of p53 (Bradford et al., 2003, Ferris et al., 2005, Mandic et al., 2005, Somers et al., 1992). The p53 status for CU-OP-2 and CU-OP-3 were confirmed as wild-type p53 (Pirotte, 2017). P53 status for the novel cell lines of CU-OP-17 and CU-OP-20 was identified by using IGV. To determine if the cell lines used in this study conform to these previous results, and therefore are appropriate models of OPSCC, it was necessary to determine the p53 mutation status.

mRNA sequence data was examined to identify mismatches with the reference sequence, these were then individually investigated to see whether they were likely to be SNPs or mutations. To be considered genuine, at least five reads were required to show a difference from the reference sequence; this was based on the same process as for HPV-integration (to avoid a confounding effect by counting isolated sequencing errors). Confirmed mismatches were searched against the NCBI dbSNP database to determine the allele frequency and clinical significance.

Single nucleotide polymorphism (SNP) or mismatches were identified for all cell lines except for UMSCC-6 and UMSCC-19. However, a deletion in UMSCC-19 for p53 was detected, which consisted of a gap of 22 nucleotides in exon 8. This mutation has been previously reported (Somers et al., 1992).

In both untreated and 2Gy treated UMSCC-4, a pathogenic mutation in the DNA binding domain (heterozygous G>A mutation) at chromosome 17:7578212 with the SNP reference number of rs397516436 was detected. The detectable read of the untreated sample only consisted of one read. The identification of this mutation was consistent with previous publications (Pirotte, 2017, Mandic et al., 2005)). This was considered as an acceptable exception to the threshold of 5 reads, as the mutation was confirmed with more reads in the treated sample. For the rest of the cell line panel the minimum number of 5 reads was followed.

In the novel cell line CU-OP-17 (HPV-negative), a SNP at chromosome location 17:7578478 with the SNP reference number of rs1057520000 was detected and was described as occurring in the UTR variant 5 prime domain, and as "likely pathogenic (missense mutation)". More information about this SNP was not available on the NCBI dbSNP database. Minor Allele Frequency (MAF) was not available, which indicates that it is uncommon (rare). On this basis, CU-OP-17 can be considered as wild-type for p53 within the coding region.

One SNP, rs1042522 (Minor Allele Frequency (MAF)/Minor Allele count: G=0.4571/2289 (1000 Genomes)), was detected in the whole panel of HPV-positive OPSCC cell lines except for UMSCC-6 (no detectable SNP above 5 reads). The MAF/Minor Allele count showed that this SNP is quite common. A study on lung cancer identified the rs1042522 genotype as being more likely resistant to chemotherapy (Han et al., 2008). The role of p53 polymorphism at codon 72 (rs1042522) in head and neck cancer is unclear, as expression of this SNP did not lead to greater resistance in all cell lines to IR. Furthermore, p53 polymorphism at codon 72 (rs1042522) in cervical cancer was associated with increased risk (Rosenthal et al., 1998, Storey et al., 1998). This was later disproved by other larger studies with no association between rs1042522 and cancer risk (Koshiol et al., 2009).

In summary the novel cell lines of CU-OP-17 and CU-OP-20 the p53 status was investigated through the mRNA sequencing data. CU-OP-17 showed likely pathogenic missense mutation at a UTR region but no mutations (SNP) within the coding region. CU-OP-20 appeared to carry no mutations in p53, and no pathogenic SNP's were detected. The p53 status of the remaining cell lines, presented in previous studies, was confirmed in the current study.

In terms of SNPs and mutation there appeared to be no differences between untreated and 2Gy treated samples. No detection of additional pathogenic mutations was made.

## 5.4. Analysis of differential gene expression

The mRNA sequencing data was used to compare gene expression between three different sub-groups of cell lines. Comparisons were made between:

- HPV-positive versus HPV-negative OPSCC cell lines
- HPV-positive OPSCC: cell lines sensitive to IR versus cell lines resistant to IR (untreated comparison & 2Gy treated comparison)
- OPSCC cell lines: untreated cell lines versus treated cell lines with IR (2Gy)

Differentially expressed genes were identified using a DEseq2 analysis on normalized data (Love et al., 2014). The resultant p-values were corrected for multiple testing and false discovery using the FDR method (FDR $p<0.05$) (Benjamini and Hochberg, 1995). Data was then presented as heatmaps based on a hierarchical clustering (using average linkage and Pearson's correlation as the similarity metric). DEseq2 analysis and generation of heatmaps are described in more detail in chapter 2 section 2.8.4.

### 5.4.1. Differences in gene expression of HPV-positive versus HPV-negative OPSCC cell lines

Differential gene expression between the 5 HPV-positive OPSCC and 5 HPV-negative cell lines was investigated. The prediction was that there would be differential gene expression between HPV-positive and HPV-negative OPSCC cell lines, due to differences in radio-sensitivity (Chapter 4) and differences in clinical response of HPV-positive and HPV- negative patients (Evans et al., 2013). This comparison was done, before focusing on sensitivity comparison and effect of IR (paired comparison). Within this section HPV-positive and HPV-negative OPSCC cell lines were compared as untreated samples and 2Gy treated samples with each other (Figure 5-6 and Figure 5-7).

Data was generated as described in the Introduction of section 5.4.

The generated heatmaps in both untreated and treated (2Gy IR) comparison indicated that HPV-positive and HPV-negative OPSCC cell lines formed discrete groups with differential gene expression according to the HPV-status. The top 20 significant genes (FDR $p<0.05$) did not indicate any connection to DNA repair, DNA damage, cell death or apoptosis. (Table 5-6 and Table 5-7).

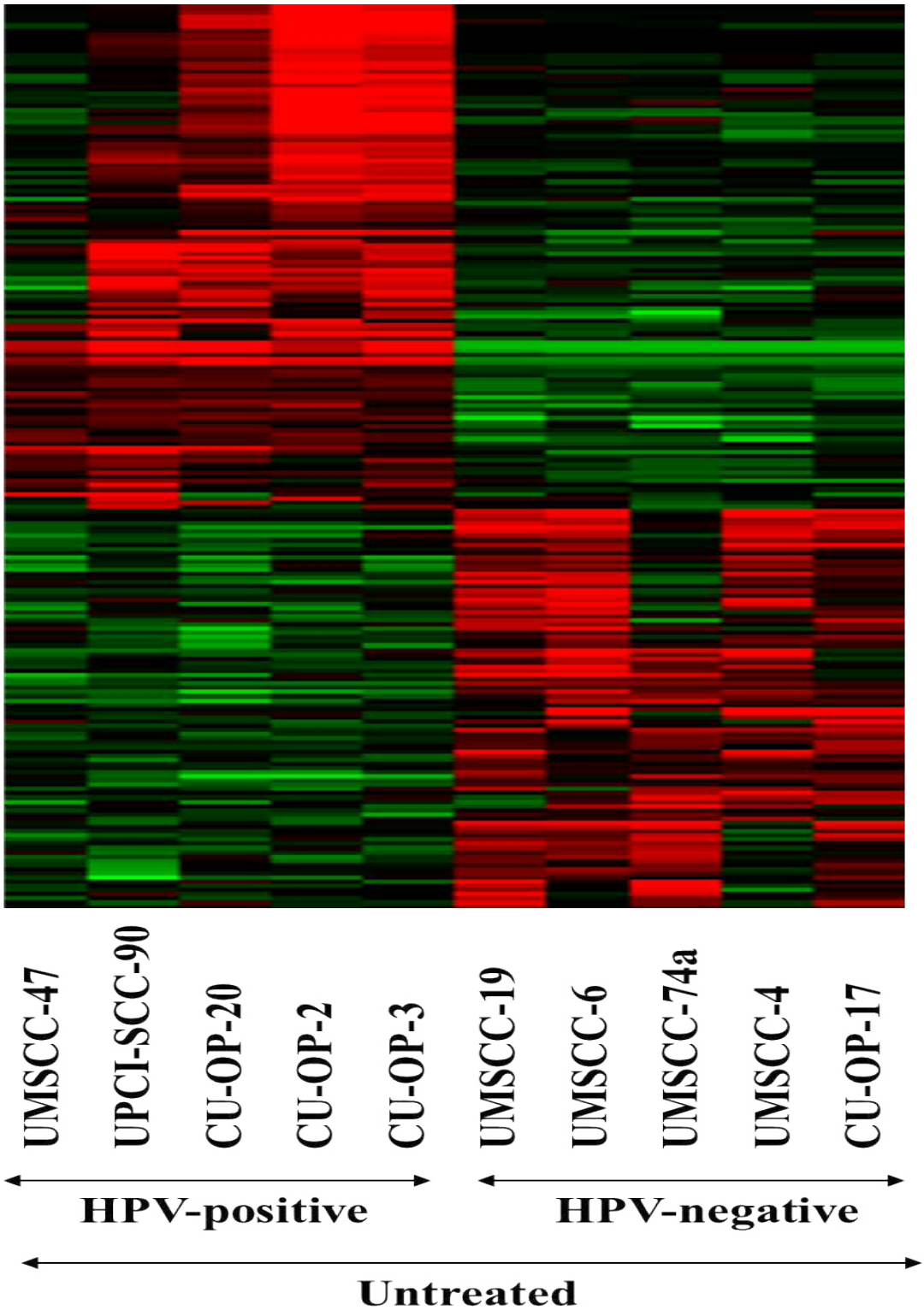



**Figure 5-6: Heatmap representing differential gene expression between untreated HPV-positive and HPV-negative OPSCC cell lines.**
Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (i.e. red = higher expression). After correction for multiple testing, 209 genes showed a significant difference between the two groups (FDR $p<0.05$**).** The top 20 most significant genes are listed in Table. 5-6. Untreated samples were collected at the same time as the treated counterpart.

**Table 5-6: Differential gene expression between untreated HPV-positive and HPV-negative OPSCC cell lines (untreated)**

| Gene[A] | Description | ID | fdr p | log2FC[B] |
|---|---|---|---|---|
| GPR158 | G Protein-coupled Receptor 158 | NM_020752_chr10 | 1.28E-12 | -6.17 |
| SYCP2 | Synaptonemal Complex Protein 2 | NM_014258_chr20 | 2.86E-12 | 4.06 |
| MMP13 | Matrix Metallopeptidase 13 | NM_002427_chr11 | 1.38E-09 | -6.37 |
| LOC729683 | Uncharacterized LOC729683 | NR_046273_chr17 | 1.38E-09 | 4.08 |
| KLHL35 | Kelch Like Family Member 35 | NM_001039548_chr11 | 4.75E-09 | 3.68 |
| ITGA4 | Integrin Subunit Alpha 4 | NM_000885_chr2 | 6.62E-09 | -6.19 |
| ZNF788 | Zinc Finger Family Member 788 | NR_027049_chr19 | 1.01E-08 | -5.46 |
| TCAM1P | Testicular Cell Adhesion Molecule 1, Pseudogene | NR_002947_chr17 | 7.78E-08 | 4.88 |
| LINC00662 | Long Intergenic Non-protein Coding RNA 662 | NR_027301_chr19 | 9.87E-08 | -5.55 |
| ZNF541 | Zinc finger protein 541 | NM_001101419_chr19 | 2.98E-07 | 5.58 |
| LOC375196 | Uncharacterized LOC375196 | NR_028386_chr2 | 7.21E-07 | 5.31 |
| MEI1 | Meiotic Double-stranded Break Formation Protein 1 | NM_152513_chr22 | 1.14E-06 | 4.96 |
| RASIP1 | Ras Interacting Protein 1 | NM_017805_chr19 | 1.25E-06 | 4.35 |
| NEFH | Neurofilament Heavy | NM_021076_chr22 | 1.86E-06 | 5.72 |
| NMNAT2 | Nicotinamide Nucleotide Adenylyltransferase 2 | NM_015039_chr1 | 4.73E-06 | 5.03 |
| ANXA2P1 | Annexin A2 Pseudogene 1 | NR_001562_chr4 | 4.73E-06 | 5.04 |
| SHCBP1L | SHC Binding And Spindle Associated 1 Like | NM_030933_chr1 | 5.74E-06 | 5.50 |
| XCL1 | X-C Motif Chemokine Ligand 1 | NM_002995_chr1 | 7.27E-06 | -5.12 |
| LCP1 | Lymphocyte Cytosolic Protein 1 | NM_002298_chr13 | 1.45E-05 | -5.06 |
| LDOC1 | Leucine Zipper Down-regulated In Cancer 1 | NM_012317_chrX | 1.52E-05 | -4.88 |

[A] The 20 most significant genes are listed in this table - additional GO clustering was performed to group genes according to their function.

[B] log2FC: Fold change between the comparison of two groups of this significant presented gene (increase or decrease of expression)

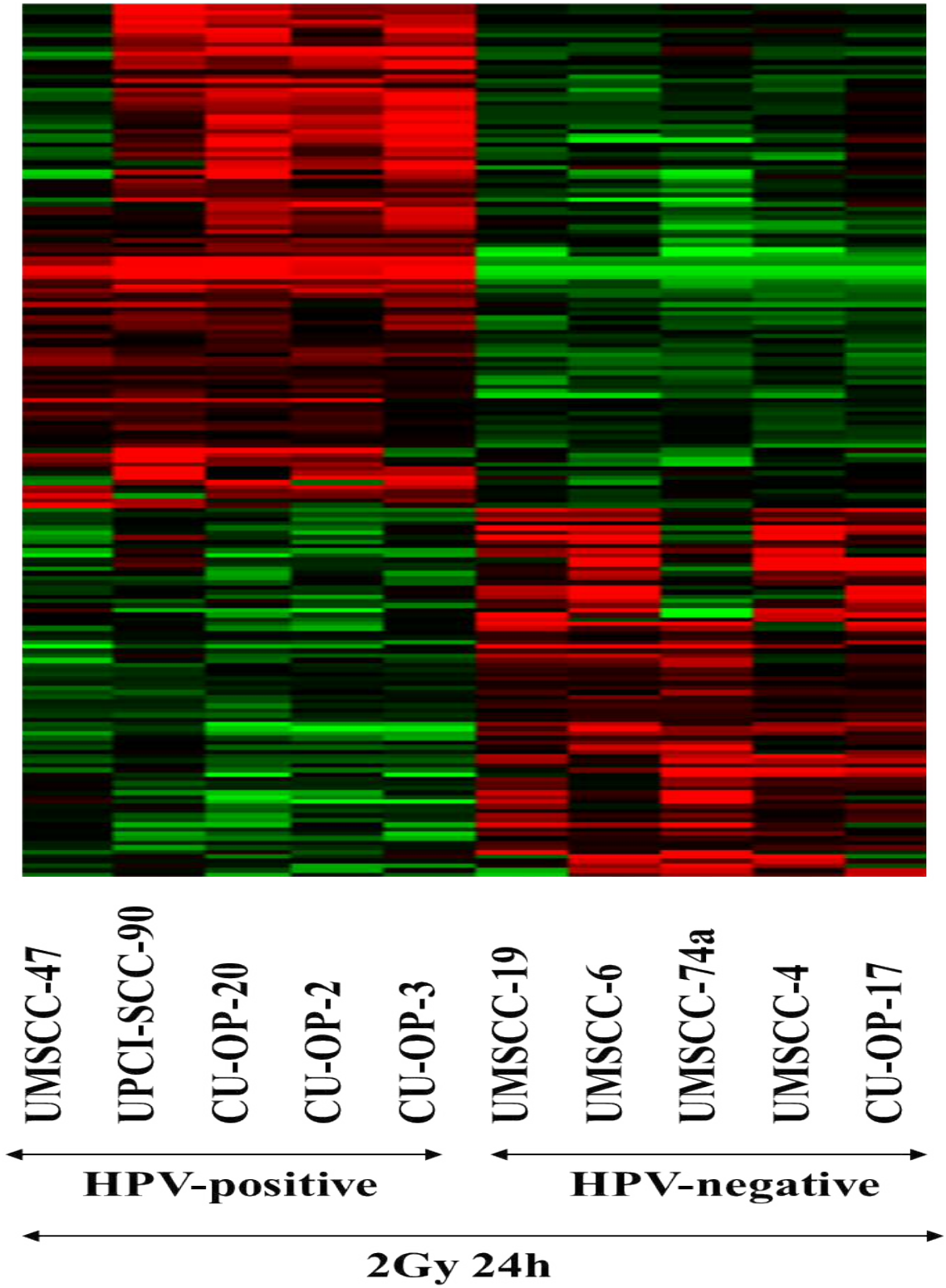



**Figure 5-7: Heatmap representing differential gene expression between 2Gy IR treated HPV-positive and HPV-negative OPSCC cell lines.**
Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (i.e. red = higher expression). After correction for multiple testing, 209 genes showed a significant difference between the two groups (FDR $p<0.05$**).** The top 20 most significant genes are listed in Table. 5-7. Untreated samples were collected at the same time as the treated counterpart.

**Table 5-7: Differential gene expression between 2Gy IR treated HPV-positive and HPV-negative OPSCC cell lines**

| Gene[A] | Description | ID | fdr p | log2FC[B] |
|---|---|---|---|---|
| SYCP2 | Synaptonemal Complex Protein 2 | NM_014258_chr20 | 2.7E-19 | 4.40 |
| LOC729683 | Uncharacterized LOC729683 | NR_046273_chr17 | 2.0E-10 | 3.10 |
| ZNF541 | Zinc finger protein 541 | NM_001101419_chr19 | 4.8E-10 | 5.32 |
| KLHL35 | Kelch Like Family Member 35 | NM_001039548_chr11 | 4.8E-10 | 2.86 |
| LCP1 | Lymphocyte Cytosolic Protein 1 | NM_002298_chr13 | 2.4E-09 | -6.29 |
| LOC375196 | Uncharacterized LOC375196 | NR_028386_chr2 | 5.7E-09 | 5.43 |
| ITGA4 | Integrin Subunit Alpha 4 | NM_000885_chr2 | 9.9E-09 | -5.78 |
| MMP13 | Matrix Metallopeptidase 13 | NM_002427_chr11 | 2.7E-08 | -4.72 |
| GPR158 | G Protein-coupled Receptor 158 | NM_020752_chr10 | 6.1E-07 | -5.03 |
| ZNF788 | Zinc Finger Family Member 788 | NR_027049_chr19 | 8.0E-07 | -4.59 |
| XCL1 | X-C Motif Chemokine Ligand 1 | NM_002995_chr1 | 3.0E-06 | -5.09 |
| TCAM1P | Testicular Cell Adhesion Molecule 1, Pseudogene | NR_002947_chr17 | 4.7E-06 | 4.34 |
| RAB3IL1 | RAB3A Interacting Protein Like 1 | NM_013401_chr11 | 6.3E-06 | -2.20 |
| OSGIN2 | Oxidative Stress Induced Growth Inhibitor Family Member 2 | NM_001126111_chr8 | 5.9E-05 | -2.00 |
| JAKMIP3 | Janus Kinase And Microtubule Interacting Protein 3 | NM_001105521_chr10 | 7.1E-05 | -3.70 |
| MEI1 | Meiotic Double-stranded Break Formation Protein 1 | NM_152513_chr22 | 1.1E-04 | 4.29 |
| CCDC155 | Coiled-coil Domain Containing 155 | NM_144688_chr19 | 1.4E-04 | 4.50 |
| DPYSL4 | Dihydropyrimidinase Like 4 | NM_006426_chr10 | 1.7E-04 | -4.70 |
| MMP1 | Matrix Metallopeptidase 1 | NM_002421_chr11 | 1.7E-04 | -4.60 |
| NDUFA4L2 | Mitochondrial Complex Associated Like 2 | NM_020142_chr12 | 1.7E-04 | 3.89 |

[A] The 20 most significant genes are listed in this table - additional GO clustering was performed to group genes according to their function.
[B] log2FC: Fold change between the comparison of two groups of this significant presented gene (increase or decrease of expression)

#### 5.4.1.1. Identification of gene ontology groups between HPV-positive OPSCC cell lines versus HPV-negative OPSCC cell lines

To allow analysis of large numbers of differentially regulated genes, gene ontology analysis was performed. Gene ontology highlighted several differences between HPV-positive and HPV-negative OPSCC cell lines, in the comparisons of untreated or 2Gy IR treated samples (Table 5-8 and Table 5-9).

The gene ontology overexpression analysis (GO ORA) was undertaken to gain insight into the biological consequences of differential expression of large numbers of genes. During this analysis genes were grouped into three categories: biological processes (BP), molecular functions (MF) and cellular component (CC). The data were corrected for multiple testing and false discovery using the FDR method (Falcon and Gentleman, 2007, Young et al., 2010). A detailed description of the methodology is given in the Material and Methods chapter section 2.8.4.8.

The number of significant overrepresented GO groups (using the top 500 transcripts) for the untreated comparisons of HPV-positive versus HPV-negative OPSCC cell lines and the 2Gy IR treated counterpart (GO ORA) are presented in Table 5-8 and Table 5-9. The number of significant overrepresented GO groups are presented with the adjusted p-value (FDR $p<0.05$) as following for the untreated comparison: 377 (BP), 65 (MF) and 40 (CC). The number of significant overrepresented GO groups are presented with the adjusted p-value (FDR $p<0.05$) as following for 2Gy treated comparison between HPV-positive and HPV-negative OPSCC cell lines: 132 (BP), 44 (MF) and 20 (CC). The top 10 significant groups for BP, MF and CC were presented in Table 5-8 and Table 5-9 as example GO groups.

The BP ontologies that appeared to be differentially regulated between the untreated HPV-positive and HPV-negative OPSCC cell lines included several groups relating to regulation of cell communication, signalling and regulation of cellular protein metabolic processes. The MF ontologies showed differentially expressed processes, including enzyme, collagen and kinesin binding overrepresented GO groups. The CC ontologies showed differential expressed regulation, including within the Ku70:Ku80 complex, a dimeric protein complex binding to DNA double-strand breaks ends (Non homologous end joining) (Table 5-8).

The BP ontologies within the 2Gy treated comparison between HPV-positive and HPV-negative OPSCC cell lines, appeared to have differentially regulated processes, including

glutamine family amino acid metabolism or mRNA 3'-splice site recognition. The MF ontologies showed differentially regulated processes, including activities in phosphoglycerate dehydrogenase and glycine-gated chloride ion channel activity. The CC ontologies showed the top 10 significant overrepresented GO groups. Compared to the untreated comparison and differentially regulated Ku70:Ku80 complex, was not significant in the entire CC ontology comparison within the 2Gy IR treated samples (Table 5-9).

**Table 5-8: Significant gene ontology categories between untreated HPV-positive OPSCC and HPV-negative OPSCC cell lines**

| Category[A] | GO ID | Term | fdr p |
|---|---|---|---|
| BP | GO:0060255 | regulation of macromolecule metabolic process | 0.042 |
| BP | GO:0009719 | response to endogenous stimulus | 0.042 |
| BP | GO:0010646 | regulation of cell communication | 0.042 |
| BP | GO:0023051 | regulation of signaling | 0.042 |
| BP | GO:0009057 | macromolecule catabolic process | 0.042 |
| BP | GO:0030182 | neuron differentiation | 0.042 |
| BP | GO:0032268 | regulation of cellular protein metabolic process | 0.042 |
| BP | GO:0051094 | positive regulation of developmental process | 0.042 |
| BP | GO:0042325 | regulation of phosphorylation | 0.042 |
| BP | GO:0009887 | organ morphogenesis | 0.042 |
| MF | GO:0019899 | enzyme binding | 0.018 |
| MF | GO:0004222 | metalloendopeptidase activity | 0.018 |
| MF | GO:0005518 | collagen binding | 0.018 |
| MF | GO:0019894 | kinesin binding | 0.018 |
| MF | GO:0050750 | low-density lipoprotein particle receptor binding | 0.018 |
| MF | GO:0005161 | platelet-derived growth factor receptor binding | 0.018 |
| MF | GO:0004070 | aspartate carbamoyltransferase activity | 0.018 |
| MF | GO:0004087 | carbamoyl-phosphate synthase (ammonia) activity | 0.018 |
| MF | GO:0004088 | carbamoyl-phosphate synthase (glutamine-hydrolyzing) activity | 0.018 |
| MF | GO:0004151 | dihydroorotase activity | 0.018 |
| CC | GO:0031090 | organelle membrane | 0.034 |
| CC | GO:0005773 | vacuole | 0.034 |
| CC | GO:0031233 | intrinsic component of external side of plasma membrane | 0.034 |
| CC | GO:0043564 | Ku70:Ku80 complex | 0.034 |
| CC | GO:1990597 | AIP1-IRE1 complex | 0.034 |
| CC | GO:0016935 | glycine-gated chloride channel complex | 0.034 |
| CC | GO:0070743 | interleukin-23 complex | 0.034 |
| CC | GO:1990630 | IRE1-RACK1-PP2A complex | 0.034 |
| CC | GO:1990332 | Ire1 complex | 0.034 |
| CC | GO:1990425 | ryanodine receptor complex | 0.034 |

[A] Gene ontology categories = Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) – Top 10 significant groups for BP, MF and CC are listed as representative figures in this table.

**Table 5-9: Significant gene ontology categories between 2Gy IR treated HPV-positive OPSCC and HPV-negative OPSCC cell lines**

| Category[A] | GO ID | Term | fdr p |
|---|---|---|---|
| BP | GO:0009064 | glutamine family amino acid metabolic process | 0.037 |
| BP | GO:0036498 | IRE1-mediated unfolded protein response | 0.037 |
| BP | GO:0032981 | mitochondrial respiratory chain complex I assembly | 0.037 |
| BP | GO:0010171 | body morphogenesis | 0.037 |
| BP | GO:0060325 | face morphogenesis | 0.037 |
| BP | GO:0098810 | neurotransmitter reuptake | 0.037 |
| BP | GO:0000389 | mRNA 3'-splice site recognition | 0.037 |
| BP | GO:0003149 | membranous septum morphogenesis | 0.037 |
| BP | GO:0003307 | regulation of Wnt signaling pathway involved in heart development | 0.037 |
| BP | GO:0018401 | peptidyl-proline hydroxylation to 4-hydroxy-L-proline | 0.037 |
| MF | GO:0051879 | Hsp90 protein binding | 0.023 |
| MF | GO:0004617 | phosphoglycerate dehydrogenase activity | 0.023 |
| MF | GO:0015105 | arsenite transmembrane transporter activity | 0.023 |
| MF | GO:0017096 | acetylserotonin O-methyltransferase activity | 0.023 |
| MF | GO:0045517 | interleukin-20 receptor binding | 0.023 |
| MF | GO:0022852 | glycine-gated chloride ion channel activity | 0.029 |
| MF | GO:0045518 | interleukin-22 receptor binding | 0.029 |
| MF | GO:0048257 | 3'-flap endonuclease activity | 0.029 |
| MF | GO:0036094 | small molecule binding | 0.031 |
| MF | GO:0031406 | carboxylic acid binding | 0.031 |
| CC | GO:0005739 | mitochondrion | 0.019 |
| CC | GO:0016935 | glycine-gated chloride channel complex | 0.019 |
| CC | GO:1990630 | IRE1-RACK1-PP2A complex | 0.019 |
| CC | GO:1990332 | Ire1 complex | 0.019 |
| CC | GO:1990425 | ryanodine receptor complex | 0.019 |
| CC | GO:0016222 | procollagen-proline 4-dioxygenase complex | 0.021 |
| CC | GO:1990597 | AIP1-IRE1 complex | 0.021 |
| CC | GO:0005658 | alpha DNA polymerase:primase complex | 0.021 |
| CC | GO:1990604 | IRE1-TRAF2-ASK1 complex | 0.021 |
| CC | GO:0031975 | envelope | 0.026 |

[A] Gene ontology categories = Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) – Top 10 significant groups for BP, MF and CC are listed as representative figures in this table.

### 5.4.2. Differences in gene expression of HPV-positive OPCC cell lines sensitive/resistant to IR

According to the obtained data from the radio-sensitivity assay, basic comparison of gene expression was performed on the defined sensitive and resistant cell line (untreated and treated comparison). Genes presented in the heatmaps were normalised and significant according to the FDR value (FDR $p<0.05$).

These comparisons were done to identify possible key factors, which might explain the greater variation within the HPV-positive OPSCC cell lines (HPV-negative OPSCC cell lines have shown a similar IR response). A heatmap of sensitive cell lines (UMSCC-47, UPCI-SCC-90 and CU-OP-20) versus resistant cell lines (CU-OP-2 and CU-OP-3) showed clear differences between the two groups (Figure 5-8 and Figure 5-9). The two resistant cell lines showed high expression of the majority of the differential expressed genes compared to a median or lower expression of these genes in the three sensitive cell lines. The most significantly differentially expressed transcripts map to genes involved in signal transduction, protein phosphorylation & de-phosphorylation, regulation of transcription, cell cycle arrest, or cell adhesion (Table 5-10).

A similar analysis was then performed using the mRNA sequencing data from treated cells (Figure 5-9). This comparison showed only 19 significant differences in gene expression (genes listed in Table 5-11) (FDR $p<0.05$). A number of genes were detected in both groups: TXLNG2P, ZNF253, USP32P1, PPP1R3C, EDIL3 and FAM50B.

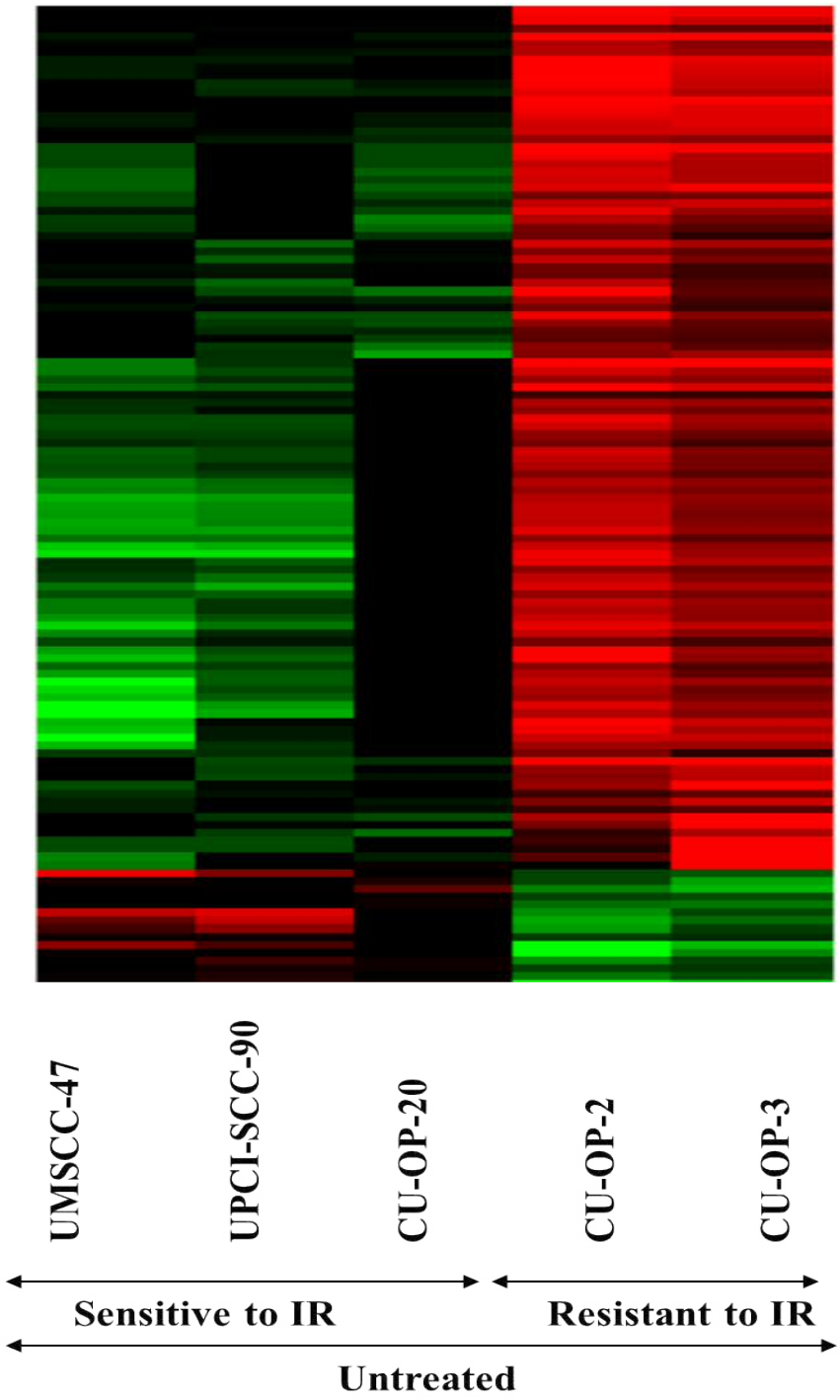



**Figure 5-8: Heatmap representing differential gene expression between the most sensitive and resistant HPV-positive cell lines (untreated cells)**
Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (i.e. red = higher expression). After correction for multiple testing, 123 genes showed a significant difference between the two groups (FDR $p<0.05$**).** The top 20 most significant genes are listed in Table. 5-10. Untreated samples were collected at the same time as the treated counterpart.

**Table 5-10: Differential gene expression between sensitive and resistant cell lines (untreated)**

| Genes[A] | Description | ID | fdr p | log2FC[B] |
|---|---|---|---|---|
| PRKG2 | Protein Kinase, CGMP-Dependent, Type II | NM_006259_chr4 | 2.6E-12 | -3.80 |
| TXLNG2P | Taxilin Gamma Pseudogene, Y-linked | NR_045128_chrY | 2.6E-12 | -6.57 |
| PPP1R3C | Protein Phosphatase 1 Regulatory Subunit 3C | NM_005398_chr10 | 3.1E-10 | -5.93 |
| ZNF253 | Zinc Finger Protein 253 | NM_021047_chr19 | 1.5E-08 | -5.72 |
| RCN3 | Reticulocalbin 3 | NM_020650_chr19 | 2.4E-07 | -3.43 |
| FAM50B | Family with Sequence Similarity 50 Member B | NM_012135_chr6 | 1.8E-06 | 4.37 |
| USP32P1 | Ubiquitin Specific Peptidase 32 Pseudogene 1 | NR_003190_chr17 | 1.3E-05 | -3.54 |
| BRINP3 | BMP/Retinoic Acid Inducible Neural Specific 3 | NM_199051_chr1 | 2.5E-05 | -4.90 |
| DENND2A | DENN Domain Containing 2A | NM_015689_chr7 | 2.7E-05 | -4.07 |
| CCND2 | Cyclin D2 | NM_001759_chr12 | 8.2E-05 | -4.68 |
| TXLNG2P | Taxilin Gamma Pseudogene, Y-linked | NR_045129_chrY | 8.2E-05 | -4.71 |
| COL5A2 | Collagen Type V Alpha 2 Chain | NM_000393_chr2 | 1.1E-04 | -4.09 |
| EDIL3 | EGF Like Repeats And Discoidin Domains 3 | NM_005711_chr5 | 2.6E-04 | 4.28 |
| CRACR2A | Calcium Release Activated Channel Regulator 2A | NM_001144958_chr12 | 3.0E-04 | -4.10 |
| HIST2H2BF | Histone Cluster 2 H2B Family Member F | NM_001024599_chr1 | 3.7E-04 | -4.50 |
| H3F3C | H3 Histone Family Member 3C | NM_001013699_chr12 | 4.6E-04 | -4.12 |
| KALRN | Kalirin RhoGEF Kinase | NM_001024660_chr3 | 4.8E-04 | -4.46 |
| EIF5AL1 | Eukaryotic Translation Initiation Factor 5A-Like 1 | NM_001099692_chr10 | 5.0E-04 | -3.67 |
| STAT5A | Signal Transdurcer and Activator Of Transcription 5A | NM_003152_chr17 | 7.4E-04 | -2.29 |
| PCOLCE | Procollagen C-Endopeptidase Enhancer | NM_002593_chr7 | 9.7E-04 | -3.25 |

[A] The 20 most significant genes are listed in this table - additional GO clustering was performed to group genes according to their function.
[B] log2FC: Fold change between the comparison of two groups of this significant presented gene (increase or decrease of expression)

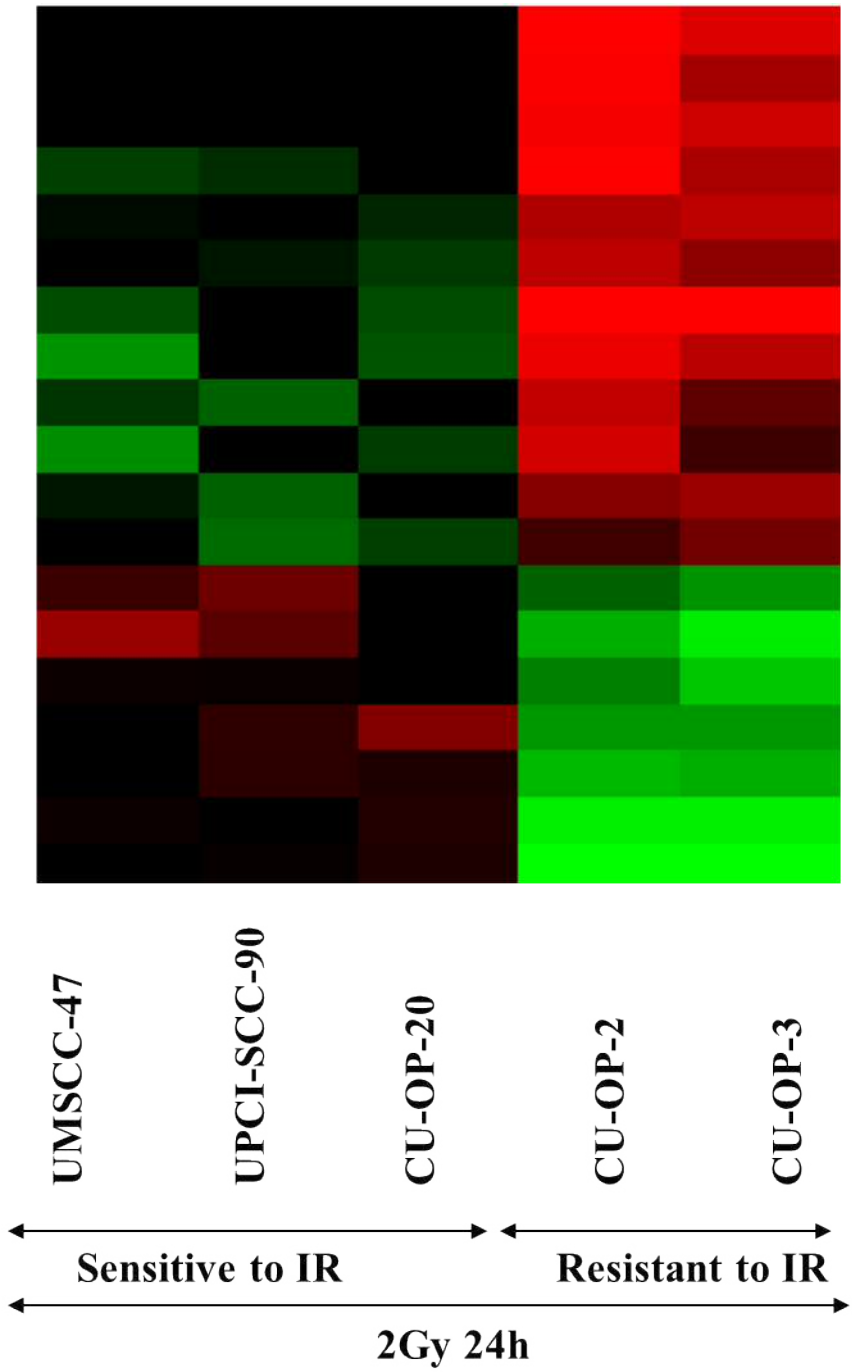



**Figure 5-9: Heatmap representing differential gene expression within the 2Gy IR treated most sensitive and resistant HPV-positive cell lines (treated sample 2Gy 24h)**
Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (higher expression. After correction for multiple testing between sensitive and resistant groups within HPV-positive OPSCC, 19 transcripts (genes) showed a significant difference between the two groups (FDR $p<0.05$**).** The top 20 genes are listed as representative genes in Table. 5-11. Untreated samples were collected at the same time as the treated counterpart.

**Table 5-11: Differential gene expression between sensitive and resistant cell lines (2Gy treated)**

| Genes[A] | Description | ID | fdr p | log2FC[B] |
|---|---|---|---|---|
| FAM50B | Family with Sequence Similarity 50 Member B | NM_012135_chr6 | 2.1E-10 | 5.31 |
| TXLNG2P | Taxilin Gamma Pseudogene, Y-linked | NR_045128_chrY | 2.6E-10 | -6.12 |
| USP32P1 | Ubiquitin Specific Peptidase 32 Pseudogene 1 | NR_003190_chr17 | 2.1E-08 | -4.24 |
| GPR114 | Adhesion G Protein-Coupled Receptor G5 | NM_153837_chr16 | 2.9E-08 | 5.21 |
| PPP1R3C | Protein Phosphatase 1 Regulatory Subunit 3C | NM_005398_chr10 | 1.4E-07 | -5.12 |
| ZNF253 | Zinc Finger Protein 253 | NM_021047_chr19 | 6.1E-06 | -4.77 |
| EDIL3 | EGF Like Repeats And Discoidin Domains 3 | NM_005711_chr5 | 1.6E-05 | 4.78 |
| TXLNG2P | Taxilin Gamma Pseudogene, Y-linked | NR_045129_chrY | 1.2E-04 | -4.80 |
| TSPAN12 | Tetraspanin 12 | NM_012338_chr7 | 4.4E-03 | 3.50 |
| NCRNA00185 | Testis Specific Transcript, Y-linked 14 (Non-protein coding) | NR_001544_chrY | 6.2E-03 | -4.15 |
| HOXD9 | Homeobox D9 | NM_014213_chr2 | 1.2E-02 | -3.38 |
| PCDH7 | Protocadherin 7 | NM_002589_chr4 | 1.2E-02 | 3.01 |
| INHBE | Inhibin Beta E Subunit | NM_031479_chr12 | 1.4E-02 | -3.82 |
| ISL1 | ISL LIM Homeobox 1 | NM_002202_chr5 | 2.4E-02 | 3.55 |
| PLAC1 | Placenta Specific 1 | NM_021796_chrX | 2.6E-02 | -3.53 |
| SUSD2 | Sushi Domain Containing 2 | NM_019601_chr22 | 3.0E-02 | -2.83 |
| SULT4A1 | Sulfotransferase Family 4A Member 1 | NM_014351_chr22 | 3.6E-02 | -3.41 |
| EYA2 | EYA Transcriptional Coactivator And Phosphatase 2 | NM_005244_chr20 | 4.1E-02 | 3.60 |
| TTTY14 | Testis Specific Transcript, Y-linked 14 (Non-protein coding) | NR_001543_chrY | 4.2E-02 | -3.73 |
| RHOBTB2 | Rho Related BTB Domain Containing 2 | NM_001160036_chr8 | 5.0E-02 | -3.45 |

[A] The 20 most significant genes are listed in this table - additional GO clustering was performed to group genes according to their function.

[B] log2FC: Fold change between the comparison of two groups of this significant presented gene (increase or decrease of expression)

#### 5.4.2.1. Identification of gene ontology groups within the HPV-positive OPSCC cell lines sensitive/resistant to IR

To allow analysis of large numbers of differentially regulated genes, gene ontology analysis was performed. Gene ontology highlighted several differences between the sensitive and resistant lines, but none of these related to DNA repair or ionising radiation response (Table 5-12 and 5-13).

The number of significant overrepresented GO groups (using the top 500 transcripts) for the untreated comparisons of sensitive versus resistant to IR and treated counterpart (GO ORA) are presented in Table 5-8 and Table 5-9. The numbers of significant overrepresented GO groups are presented with the adjusted p-value (FDR $p<0.05$) as following for the untreated comparison: 193 BP, 29 MP and 42 CC. The numbers of significant GO groups are presented with the adjusted p-value (FDR $p<0.05$) as following for the 2Gy comparison: 122 BP, 34 MF and 18 CC. The top 10 significant groups for BP, MF and CC were presented in Table 5-12 and Table 5-13 as example GO groups. The p-value represented how likely it is that this over representation has occurred by chance.

The BP ontologies that appeared to be differentially regulated between sensitive and resistant HPV-positive OPSCC cell lines (untreated samples) included several groups relating to cell signalling and communication. The MF ontologies that showed differentially expressed processes between sensitive and resistant cell lines (untreated samples), including differences in transcription factor activity, RNA polymerase II transcription factor recruiting, glycogen synthase activity or laminin and fibronectin binding. Differences in the CC ontologies included factors as macromolecular complex or different expression of lateral plasma membrane. These seem consistent with a difference between the two groups in terms of cell communication.

The BP ontologies within the treated (2Gy) comparison, appeared to show differentially regulation between sensitive and resistant HPV-positive OPSCC cell lines, and included several groups relating to cellular localization or lipid metabolic process. The only similarity shown between the untreated and 2Gy treated comparison in both 10 significant listed GO groups identified different regulation in the BP ontology of homeostasis. Different GO ontologies groups for MF (e.g. arrestin family protein binding) and CC (e.g. external side of plasma membrane) were identified compared to the untreated counterparts.

**Table 5-12: Significant gene ontology categories between untreated sensitive and resistant HPV-positive OPSCC cell lines**

| Category[A] | GO ID | Term | fdr p |
|---|---|---|---|
| BP | GO:0010646 | regulation of cell communication | 0.044 |
| BP | GO:0023051 | regulation of signalling | 0.044 |
| BP | GO:0042592 | homeostatic process | 0.044 |
| BP | GO:2000026 | regulation of multicellular organismal development | 0.044 |
| BP | GO:0048583 | regulation of response to stimulus | 0.044 |
| BP | GO:0055082 | cellular chemical homeostasis | 0.044 |
| BP | GO:0061458 | reproductive system development | 0.044 |
| BP | GO:0030111 | regulation of Wnt signalling pathway | 0.044 |
| BP | GO:0009746 | response to hexose | 0.044 |
| BP | GO:0043588 | skin development | 0.044 |
| MF | GO:0001968 | fibronectin binding | 0.011 |
| MF | GO:0043236 | laminin binding | 0.022 |
| MF | GO:0046966 | thyroid hormone receptor binding | 0.022 |
| MF | GO:0001135 | transcription factor activity, RNA polymerase II transcription factor recruiting | 0.022 |
| MF | GO:0004373 | glycogen (starch) synthase activity | 0.022 |
| MF | GO:0045519 | interleukin-23 receptor binding | 0.022 |
| MF | GO:0061547 | glycogen synthase activity, transferring glucose-1-phosphate | 0.022 |
| MF | GO:0005518 | collagen binding | 0.025 |
| MF | GO:0004521 | endoribonuclease activity | 0.025 |
| MF | GO:0034618 | arginine binding | 0.025 |
| CC | GO:0016328 | lateral plasma membrane | 0.011 |
| CC | GO:1902911 | protein kinase complex | 0.020 |
| CC | GO:0043265 | ectoplasm | 0.020 |
| CC | GO:1990332 | Ire1 complex | 0.020 |
| CC | GO:1990425 | ryanodine receptor complex | 0.020 |
| CC | GO:0034667 | integrin alpha3-beta1 complex | 0.025 |
| CC | GO:0070743 | interleukin-23 complex | 0.025 |
| CC | GO:1990630 | IRE1-RACK1-PP2A complex | 0.025 |
| CC | GO:0032991 | macromolecular complex | 0.028 |
| CC | GO:0032299 | ribonuclease H2 complex | 0.028 |

[A] Gene ontology categories = Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) – Top 10 significant groups for BP, MF and CC are listed as representative figures in this table.

**Table 5-13: Significant gene ontology process (BP, MF, CC) between 2Gy treated sensitive and resistant HPV-positive OPSCC cell lines**

| Category | GO ID | Term | fdr p |
|---|---|---|---|
| BP | GO:1902578 | single-organism localization | 0.047 |
| BP | GO:0051641 | cellular localization | 0.047 |
| BP | GO:0070887 | cellular response to chemical stimulus | 0.047 |
| BP | GO:0071702 | organic substance transport | 0.047 |
| BP | GO:0006629 | lipid metabolic process | 0.047 |
| BP | GO:0032870 | cellular response to hormone stimulus | 0.047 |
| BP | GO:0002520 | immune system development | 0.047 |
| BP | GO:0009628 | response to abiotic stimulus | 0.047 |
| BP | GO:0042592 | homeostatic process | 0.047 |
| BP | GO:0010243 | response to organonitrogen compound | 0.047 |
| MF | GO:0008599 | protein phosphatase type 1 regulator activity | 0.016 |
| MF | GO:0015485 | cholesterol binding | 0.026 |
| MF | GO:0004352 | glutamate dehydrogenase (NAD+) activity | 0.026 |
| MF | GO:0004353 | glutamate dehydrogenase [NAD(P)+] activity | 0.026 |
| MF | GO:1990825 | sequence-specific mRNA binding | 0.026 |
| MF | GO:1990763 | arrestin family protein binding | 0.026 |
| MF | GO:0004656 | procollagen-proline 4-dioxygenase activity | 0.029 |
| MF | GO:0008381 | mechanically-gated ion channel activity | 0.029 |
| MF | GO:0015137 | citrate transmembrane transporter activity | 0.029 |
| MF | GO:0030628 | pre-mRNA 3'-splice site binding | 0.029 |
| CC | GO:0071438 | invadopodium membrane | 0.040 |
| CC | GO:0032541 | cortical endoplasmic reticulum | 0.040 |
| CC | GO:0033263 | CORVET complex | 0.040 |
| CC | GO:0016222 | procollagen-proline 4-dioxygenase complex | 0.040 |
| CC | GO:0034667 | integrin alpha3-beta1 complex | 0.040 |
| CC | GO:1990589 | ATF4-CREB1 transcription factor complex | 0.040 |
| CC | GO:0043231 | intracellular membrane-bounded organelle | 0.045 |
| CC | GO:0031300 | intrinsic component of organelle membrane | 0.045 |
| CC | GO:0009897 | external side of plasma membrane | 0.045 |
| CC | GO:0005666 | DNA-directed RNA polymerase III complex | 0.045 |

[A] Gene ontology categories = Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) – Top 10 significant groups for BP, MF and CC are listed as representative figures in this table.

### 5.4.3. Differential gene expression of all OPSCC cell lines after IR

To understand the effect of IR on the whole panel of cell lines HPV-positive and HPV-negative cell lines were grouped together and compared between untreated and 2Gy treated samples (Figure 5-10). After comparing all cell lines, the comparison was later focused on the difference in gene expression within the HPV-positive OPSCC cell lines (greater IR variation after the Radio-sensitivity assay).

Differentially expressed genes were identified using DEseq2 analysis on normalized data (Love et al., 2014). The resultant p-values were corrected for multiple testing and false discovery issues using the FDR method (FDR $p<0.05$) (Benjamini and Hochberg, 1995). Data was then presented as heatmaps based on a hierarchical clustering (using average linkage and Pearson's correlation as the similarity metric). Within this comparison a paired test approach was performed to compare untreated and treated samples. A paired comparison can be defined as location test to compare two sets of comparisons (untreated and treated). This repeated "location test" or "repeated measurements" are more robust to compare untreated and treated samples, because it can be paired with observations in the other sample. This analysis was performed to identify differences in gene expression between untreated and treated samples to identify possible down or up regulation of genes after IR.

The heatmap indicates a significant increase in 519 genes after IR by comparing all OPSCC cell lines (Figure 5-10). Table 5-14 showed the top 20 significant genes out of the whole panel of significant genes (as representative figure). The differentially expressed genes contributing to the heatmap (Figure 5-10) did not include genes involved in DNA repair, DNA damage or stress response, which could have explained greater sensitivity or resistance within the panel of cell lines. This paired comparison indicated no clear difference between the HPV-positive and HPV negative cell lines in terms of low or higher expression of genes. Therefore, the investigation was focused on the paired comparison of HPV-positive OPSCC cell lines (untreated versus 2Gy treated samples) after section 5.5.1.1 (GO of all paired OPSCC cell lines), for possible significant genes which could explain variation in IR response (Radio-sensitivity assay) within this subset.

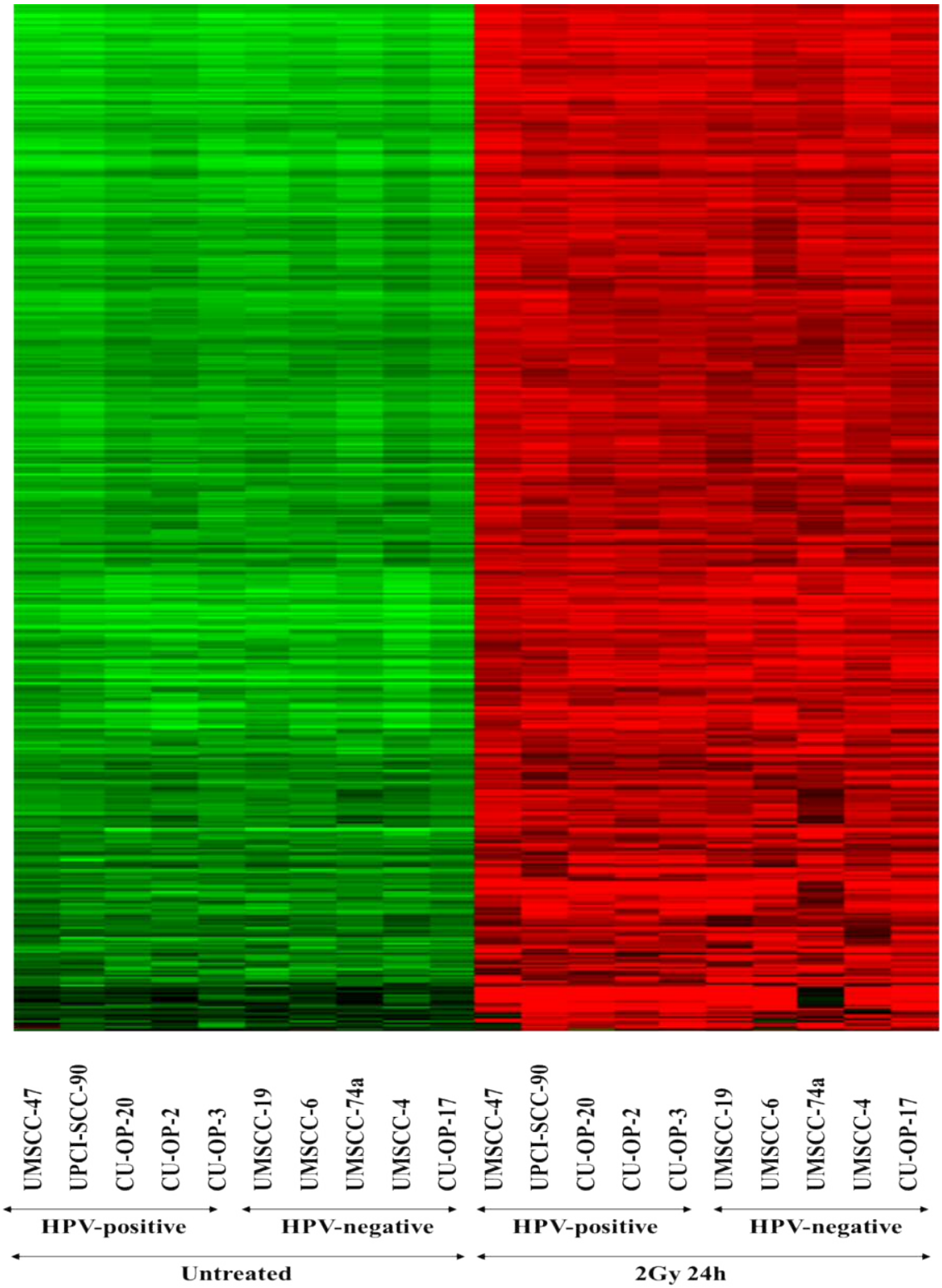



**Figure 5-10: Heatmap representing differential gene expression between untreated and treated cell lines (paired comparison)**

Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (i.e. red = higher expression). After correction for multiple testing, 519 transcripts (genes) showed a significant difference between the two groups (FDR $p<0.05$**).** The 20 genes showing the most significant difference in transcript levels are listed in Table 5-14. Untreated samples were collected at the same time as the treated counterpart.

**Table 5-14: Differential gene expression of paired test of untreated versus treated cell lines**

| Genes[A] | Description | ID | fdr p | log2FC[B] |
|---|---|---|---|---|
| EEF1E1 | Eukaryotic Translation Elongation Factor 1 Epsilon 1 | NM_004280_chr6 | 1.1E-05 | 0.66 |
| ITPRIP | Inositol 1,4,5-Trisphosphate Receptor Interacting Protein | NM_033397_chr10 | 1.7E-05 | -0.59 |
| NDUFAF4 | NADH:Ubiquinone Oxidoreductase Complex Assembly Factor 4 | NM_014165_chr6 | 1.7E-05 | 0.69 |
| SH3TC1 | SH3 Domain And Tetratricopeptide Repeats 1 | NM_018986_chr4 | 5.0E-05 | -0.72 |
| GNB4 | G Protein Subunit Beta 4 | NM_021629_chr3 | 6.4E-05 | 0.58 |
| PNPT1 | Polyribonucleotide Nucleotidyltransferase 1 | NM_033109_chr2 | 7.1E-05 | 0.59 |
| POLR3G | RNA Polymerase III Subunit G | NM_006467_chr5 | 1.7E-05 | 0.83 |
| MTR | 5-Methyltetrahydrofolate-Homocysteine Methyltransferase | NM_000254_chr1 | 3.6E-04 | 0.38 |
| MIR3687-1 | MicroRNA 3687-1 | NR_037458_chr21 | 3.6E-04 | -0.98 |
| RPL22L1 | Ribosomal Protein L22 Like 1 | NM_001099645_chr3 | 3.6E-04 | 0.68 |
| GNPNAT1 | Glucosamine-Phosphate N-Acetyltransferase 1 | NM_198066_chr14 | 4.0E-04 | 0.51 |
| AGPAT5 | 1-Acylglycerol-3-Phosphate O-Acyltransferase 5 | NM_018361_chr8 | 4.0E-04 | 0.53 |
| CASD1 | CAS1 Domain Containing 1 | NM_022900_chr7 | 8.4E-04 | 0.52 |
| RPF2 | Ribosome Production Factor 2 Homolog | NM_032194_chr6 | 9.1E-04 | 0.50 |
| NDRG1[C] | N-Myc Downstream Regulated 1 | NM_001135242_chr8 | NA | -1.04 |
| EPHB3 | EPH Receptor B3 | NM_004443_chr3 | 9.1E-04 | -0.62 |
| TWISTNB | TWIST Neighbour | NM_001002926_chr7 | 9.5E-04 | 0.51 |
| P3H2 | Prolyl 3-Hydroxylase 2 | NM_018192_chr3 | 9.5E-04 | -0.73 |
| GSN | Gelsolin | NM_000177_chr9 | 1.4E-03 | -0.95 |
| OTUD6B | OTU Domain Containing 6B | NM_016023_chr8 | 2.4E-03 | 0.57 |

[A] The 20 most significant genes are listed in this table - additional GO clustering was performed to group genes according to their function.
[B] log2FC: Fold change between the comparison of two groups of this significant presented gene (increase or decrease of expression)
[C]NDRG1 listed in the table – chosen as gene candidate for testing mRNA sequencing data.

#### 5.4.3.1.Identification of gene ontology groups of all OPSCC cell lines after IR

The genes that were differentially expressed between treated and untreated samples were further classified according to gene ontology groups; genes involved in different BP, MF or CC are listed in Table 5-15 and 5-16.

The number of significant overrepresented GO groups for the all OPSCC cell lines in comparison of treated (2Gy) and untreated samples (GO ORA) are presented in Table 5-15 The numbers of significant GO groups with the adjusted p-value (FDR $p<0.05$) were as follows: 92 BP, 26 MP and 12 CC. The top 10 GO groups for BP, MF and CC are listed in Table 5-15

The BP ontologies that appeared to be differentially regulated between untreated and treated OPSCC cell lines included a BP connected to cell cycle regulation. This BP specified the BP of "positive regulation of G0 to G1 transition" of the cell cycle. This positive regulation of the G0 to G1 transition causes the cell to return to G1 and resume growth and division. This significant GO might indicate, that cells even, if damaged are being processed through the cell cycle after initial 2Gy treatment and induction of DNA damage.

MF ontologies appeared to be differentially regulated between untreated and treated OPSCC cell lines, included several groups connected to different phosphatase activities. CC ontologies showed differentially regulated processes in a total number of 12 GO groups (Table 5-15).

**Table 5-15: Significant gene ontology categories between untreated/treated OPSCC cell lines**

| Category[A] | GO ID | Term | fdr p |
|---|---|---|---|
| BP | GO:0035690 | cellular response to drug | 0.032 |
| BP | GO:0035094 | response to nicotine | 0.032 |
| BP | GO:0006102 | isocitrate metabolic process | 0.032 |
| BP | GO:0010891 | negative regulation of sequestering of triglyceride | 0.032 |
| BP | GO:0036018 | cellular response to erythropoietin | 0.032 |
| BP | GO:0044268 | multicellular organismal protein metabolic process | 0.032 |
| BP | GO:0070318 | positive regulation of G0 to G1 transition | 0.032 |
| BP | GO:2000535 | regulation of entry of bacterium into host cell | 0.032 |
| BP | GO:0021913 | regulation of transcription from RNA polymerase II promoter involved in ventral spinal cord interneuron specification | 0.032 |
| BP | GO:0034287 | detection of monosaccharide stimulus | 0.032 |
| MF | GO:0004449 | isocitrate dehydrogenase (NAD+) activity | 0.016 |
| MF | GO:0008240 | tripeptidyl-peptidase activity | 0.016 |
| MF | GO:0008969 | phosphohistidine phosphatase activity | 0.016 |
| MF | GO:0031711 | bradykinin receptor binding | 0.016 |
| MF | GO:0003920 | GMP reductase activity | 0.016 |
| MF | GO:0008241 | peptidyl-dipeptidase activity | 0.016 |
| MF | GO:0034417 | bisphosphoglycerate 3-phosphatase activity | 0.016 |
| MF | GO:0052826 | inositol hexakisphosphate 2-phosphatase activity | 0.016 |
| MF | GO:0097020 | COPII adaptor activity | 0.016 |
| MF | GO:0017176 | phosphatidylinositol N-acetylglucosaminyltransferase activity | 0.022 |
| CC | GO:0043194 | axon initial segment | 0.040 |
| CC | GO:0000124 | SAGA complex | 0.040 |
| CC | GO:0000506 | glycosylphosphatidylinositol-N-acetylglucosaminyltransferase (GPI-GnT) complex | 0.040 |
| CC | GO:0035253 | ciliary rootlet | 0.040 |
| CC | GO:0032584 | growth cone membrane | 0.040 |
| CC | GO:1902560 | GMP reductase complex | 0.040 |
| CC | GO:0000145 | exocyst | 0.045 |
| CC | GO:0005892 | acetylcholine-gated channel complex | 0.045 |
| CC | GO:0002102 | podosome | 0.045 |
| CC | GO:0034451 | centriolar satellite | 0.045 |

[A] Gene ontology categories = Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) – Top 10 significant groups for BP, MF and CC are listed as representative figures in this table.

### 5.4.4. Gene expression differences between untreated and 2Gy treated samples in HPV-positive cell lines

After comparing the entire panel of OPSCC cell lines between untreated and treated samples. By grouping all OPSCC cell lines together to investigate the effect of IR, no obvious up- or down-regulation of gene expression was identified between the cell lines according to their HPV-status. Whereas comparison of HPV-positive and negative cell lines showed a differential gene expression, if not grouped in untreated and treated groups.

Therefore, the analysis focused again on the HPV-positive cell lines where the greatest difference had been seen in response to IR and compared untreated and treated samples within this group.

Differently expressed genes were again identified using DEseq2 analysis on normalized data (Love et al., 2014). The p-value was corrected the same way as described in the previous sub-chapters. During this comparison within the HPV-positive OPSCC cell lines only two significant genes were identified: JUNB and KRTN4 (Figure 5-11 and Table 5-16). All other genes within this comparison were not significant (FDR $p>0.05$).

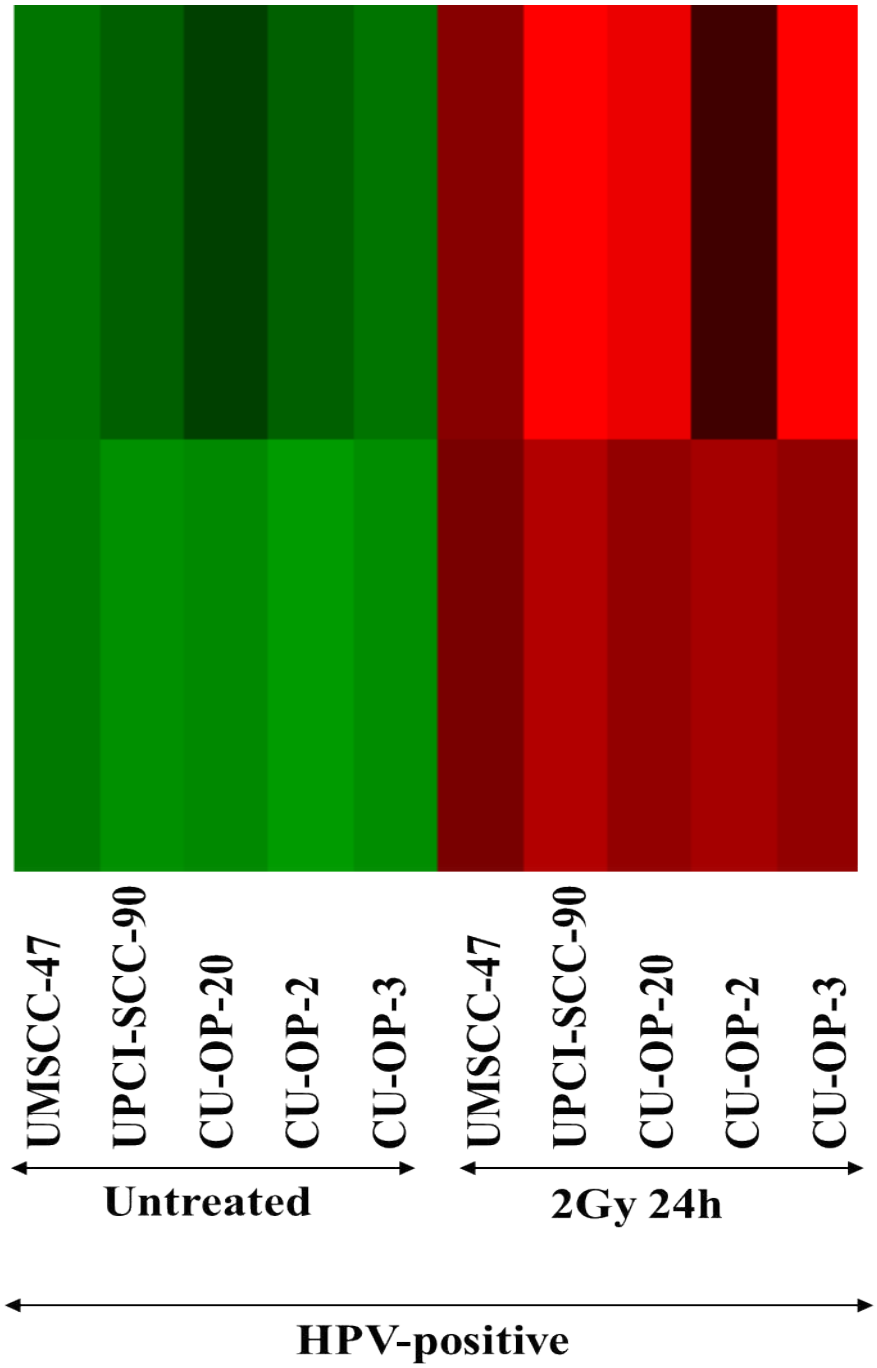



**Figure 5-11: Heatmap represent differential gene expression of all HPV-positive OPSCC cell lines**
Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (higher expression. After correction for multiple testing between sensitive and resistant groups within HPV-positive OPSCC, 2 genes showed a significant difference between the two groups (FDR $p<0.05$**).** These 2 genes are listed in Table. 5-16. All other genes within this comparison were not significant FDR $p>0.05$. Untreated samples were collected at the same time as the treated counterpart.

**Table 5-16: Differential gene expression of paired test of HPV-positive OPSCC cell lines**

| Genes[A] | Description | ID | fdr p | log2FC |
|---|---|---|---|---|
| KRT4 | Keratin 4 | NM_002272_chr12 | 0.049 | -1.02 |
| JUNB | JunB Proto-Oncogene, AP-1 Transcription Factor Subunit | NM_002229_chr19 | 0.049 | -0.83 |

[A] only 2 genes showed a significant difference within the HPV-positive OPSCC cell lines (FDR p<0.05)

#### 5.4.4.1. Identification of gene ontology groups within the HPV-positive OPSCC cell lines after IR

The number of significant overrepresented GO groups (GO ORA) for the HPV-positive OPSCC cell lines, comparing untreated and treated (2Gy) samples are presented in Table 5-17. The numbers of significant GO groups are presented with the adjusted p-value (FDR p<0.05) as following for the untreated comparison: 121 BP, 48 MP and 33 CC. The top 10 significant GO groups of BP, MF and CC were listed as representative groups (Table 5-17).

The BP ontologies that appeared to be differentially regulated between untreated and treated samples within the HPV-positive OPSCC cell lines, included several groups relating to cell death and regulation of apoptosis (GO connection between JUNB and p53 – both involved in cell death and apoptosis).

The MF ontologies showed a differentially regulated process, including fibronectin binding (connection tissue matrices) and activity of receptor signalling of protein tyrosine kinase inhibitor, which stops and prevents or reduces the activity of a receptor signalling protein tyrosine kinase.

The CC ontologies showed significant differentially regulated processes in membrane enclosed lumen, intracellular organelle lumen and nucleolar ribonuclease P complex.

**Table 5-17: Significant gene ontology categories between untreated/treated HPV-positive OPSCC cell lines**

| Category[A] | GO ID | Term | fdr p |
|---|---|---|---|
| BP | GO:0010941 | regulation of cell death | 0.039 |
| BP | GO:0042981 | regulation of apoptotic process | 0.039 |
| BP | GO:1901615 | organic hydroxy compound metabolic process | 0.039 |
| BP | GO:0051047 | positive regulation of secretion | 0.039 |
| BP | GO:0042692 | muscle cell differentiation | 0.039 |
| BP | GO:0010976 | positive regulation of neuron projection development | 0.039 |
| BP | GO:0008202 | steroid metabolic process | 0.039 |
| BP | GO:0051216 | cartilage development | 0.039 |
| BP | GO:0055002 | striated muscle cell development | 0.039 |
| BP | GO:0009062 | fatty acid catabolic process | 0.039 |
| MF | GO:0001968 | fibronectin binding | 0.016 |
| MF | GO:0031432 | titin binding | 0.016 |
| MF | GO:0008111 | alpha-methylacyl-CoA racemase activity | 0.016 |
| MF | GO:0030294 | receptor signaling protein tyrosine kinase inhibitor activity | 0.016 |
| MF | GO:0033791 | 3alpha,7alpha,12alpha-trihydroxy-5beta-cholestanoyl-CoA 24-hydroxylase activity | 0.016 |
| MF | GO:0034417 | bisphosphoglycerate 3-phosphatase activity | 0.016 |
| MF | GO:0052826 | inositol hexakisphosphate 2-phosphatase activity | 0.016 |
| MF | GO:0097020 | COPII adaptor activity | 0.016 |
| MF | GO:0001105 | RNA polymerase II transcription coactivator activity | 0.021 |
| MF | GO:0016402 | pristanoyl-CoA oxidase activity | 0.021 |
| CC | GO:0031974 | membrane-enclosed lumen | 0.040 |
| CC | GO:0070013 | intracellular organelle lumen | 0.040 |
| CC | GO:0030016 | myofibril | 0.040 |
| CC | GO:0044438 | microbody part | 0.040 |
| CC | GO:0005782 | peroxisomal matrix | 0.040 |
| CC | GO:0005655 | nucleolar ribonuclease P complex | 0.040 |
| CC | GO:0030677 | ribonuclease P complex | 0.040 |
| CC | GO:0030891 | VCB complex | 0.040 |
| CC | GO:0046581 | intercellular canaliculus | 0.040 |
| CC | GO:0070937 | CRD-mediated mRNA stability complex | 0.040 |

[A] Gene ontology categories = Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) – Top 10 significant groups for BP, MF and CC are listed as representative figures in this table.

## 5.5. Identification of gene candidates for confirming mRNA sequencing findings on molecular level

Gene candidates for were identified by comparing the most significant genes after the paired comparisons of all OPSCC cell lines and within the HPV-positive cell lines. Those two comparisons were used to identify gene products for analysis at the protein level to test the effect of IR. The focus was put on the HPV-positive cell lines as these cell lines showed the greatest variation in response after the Radio-sensitivity assay. The cell lines were then grouped for Western blot experiments according to the previously identified groups of cell lines that were sensitive or resistant to IR.

**Reasons for choosing gene candidates:**

- Identifying differences in genes involved in cell death, apoptosis, DNA repair or DNA damage at molecular level within HPV-positive cell lines
- Confirming fold-change differences of gene expression within the HPV-positive cell lines at molecular level after IR

**Process of identification of gene candidates:**

To start the process of identifying gene candidates, a broad identification of genes was done on the entire OPSCC cell lines panel (including the HPV-negative cell lines). These genes were listed in order of significance and limited to 20 representative genes listed in the individual sections above.

The next step included, listing the significant genes of HPV-positive cell lines (untreated vs treated), together with their BP processes.

The top significant genes of the HPV-positive cell line comparison and all OPSCC cell line comparison were then analysed according to their biological process (Gene Ontology), to identify genes involved in cell death, apoptosis, DNA repair or DNA damage to narrow down the number of genes candidates: Out of the HPV-positive cell line comparison and investigating the BP of all cell lines, JUNB and NDRG1 were chosen as candidates. Both genes are connected to cell death, apoptosis or DNA damage response.

The last step included a check of the number of reads of the gene candidates on Genview, to confirm an increase or decrease (fold-change) of reads after IR. This was seen as an exploratory mechanism (using RPKM reads).

At the end of this process, two genes, JUNB and NDRG1, were selected for confirmation of expression status (Figure 5-12).

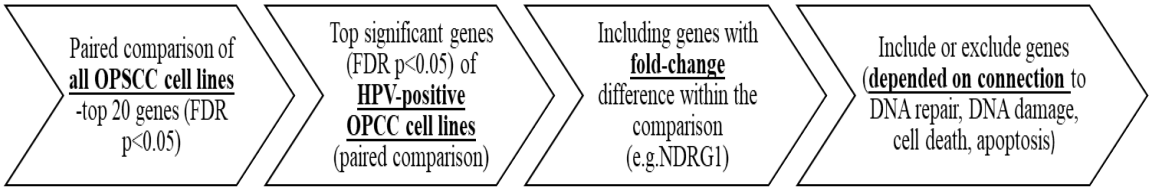



**Figure 5-12: Process for identification of gene candidates for confirming status on molecular level**

### 5.5.1. JUNB

Four out of five HPV-positive cell line showed an increase in fold-change for JUNB after IR (Figure 5-13). The Biological Process ontology showed that JUNB is involved in regulation of cell death (apoptosis), response to radiation and regulation of cell cycle.

### 5.5.2. NDRG1

NDRG-1 showed the greatest fold-change (2.6x) between untreated and treated (IR) for CU-OP-2 within the HPV-positive cell lines. The Biological Process ontology showed that NDRG1 is connected to DNA damage response (Signal transduction by p53 class mediator) and is involved in regulation of apoptotic process (Figure 5-13).

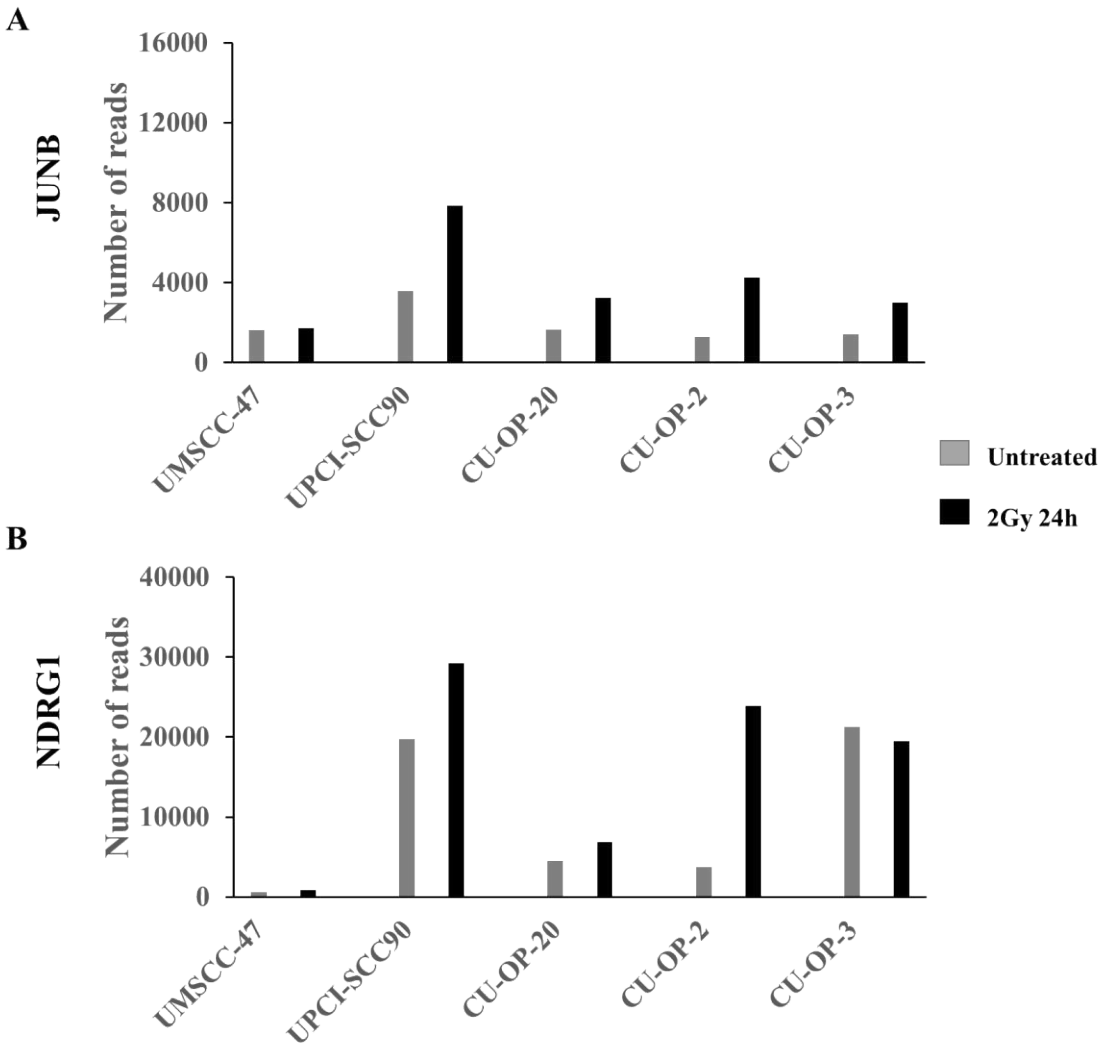



**Figure 5-13: Gene expression of JUNB and NDRG1 in untreated and treated (2Gy) HPV-positive cell lines**

Histograms were generated using Genview software, and show number of reads, expressed as RPKM. The histograms show (A) increased expression of JunB in four of the five cell lines following IR; and (B) increased expression of NDRG1 in three of five cell lines following IR, with the greatest induction in CU-OP-2.

### 5.5.3. Expression of JUNB and NDRG1

Cells respond to IR, an important physiological stress factor, by activating multiple signalling pathways (Braunstein et al., 2009) including induction of DNA repair or apoptosis. Considerable variation in response to IR was seen within the panel of HPV-positive cell lines (Radio-sensitivity assay). During the mRNA sequencing analysis, this variation was investigated further via sensitivity comparison and pair wise comparison within the HPV-positive cell lines. Different expression of numerous genes was detected following exposure to IR. However, only two genes fulfilled some or all of the previous mentioned criteria for further investigation. The protein levels of these two candidates, JUNB and NDRG1, were investigated (Figure 5-14 to Figure 5-17) and compared to the corresponding mRNA sequence data. Cell lysate was collected from untreated sample and treated sample after 24 hours (2Gy and additionally 6Gy). The same treatment doses and time point were chosen to be consistent with previous Western blots. Additionally, 24 hours was the same time point chosen for the mRNA sequencing assay. This time to be more consistent with the order of the cell lines used in mRNA sequencing, cell lines were ordered as following: UMSCC-47, UPCI-SCC-90 and CU-OP-20 (one gel) and CU-OP-2 and CU-OP-3 (second gel). Cell lines were not ordered according to sensitivity as in previous Western blotting assays (example seen in Chapter 4, Figure 4-5).

JUNB expression decreased after IR treatment for most of the cell lines (3 out of 5), whereas NDRG1 showed an increase in band intensity after 2Gy and 6Gy treatment (3 out of 5) (Figure 5-14 to Figure 5-17). Both resistant cell lines CU-OP-2 and CU-OP-3 showed an increase of NDRG1 after IR. The mRNA sequencing data did not directly correlate with the protein expression observed after western blot.

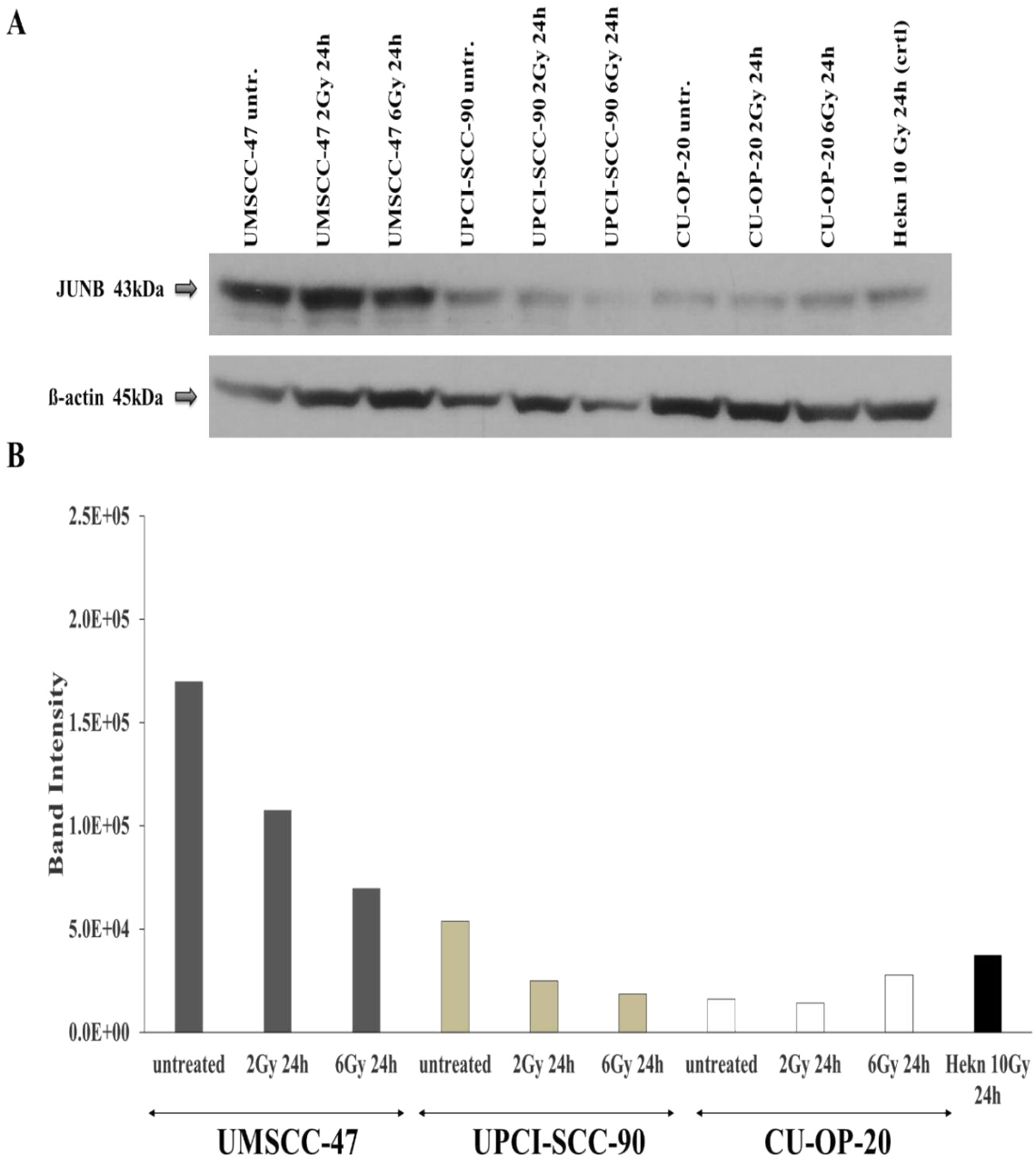



**Figure 5-14: Western blot results for JUNB (sensitive cell lines to IR)**
(A) Protein bands of JUNB for UMSCC-47, UPCI-SCC-90 and CU-OP-20. Cell lines were ordered from untreated to 2Gy and 6Gy IR treated samples (collected after 24h incubation after IR treatment). HEKn 10Gy 24h was added as positive control. Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) JUNB (30-35min exposure) bands are seen in all three cell lines.
(B) Quantification of western blot intensities of JUNB: The band intensity of the western blotting bands as quantified in relation to ß-actin level using Image Studio Lite.

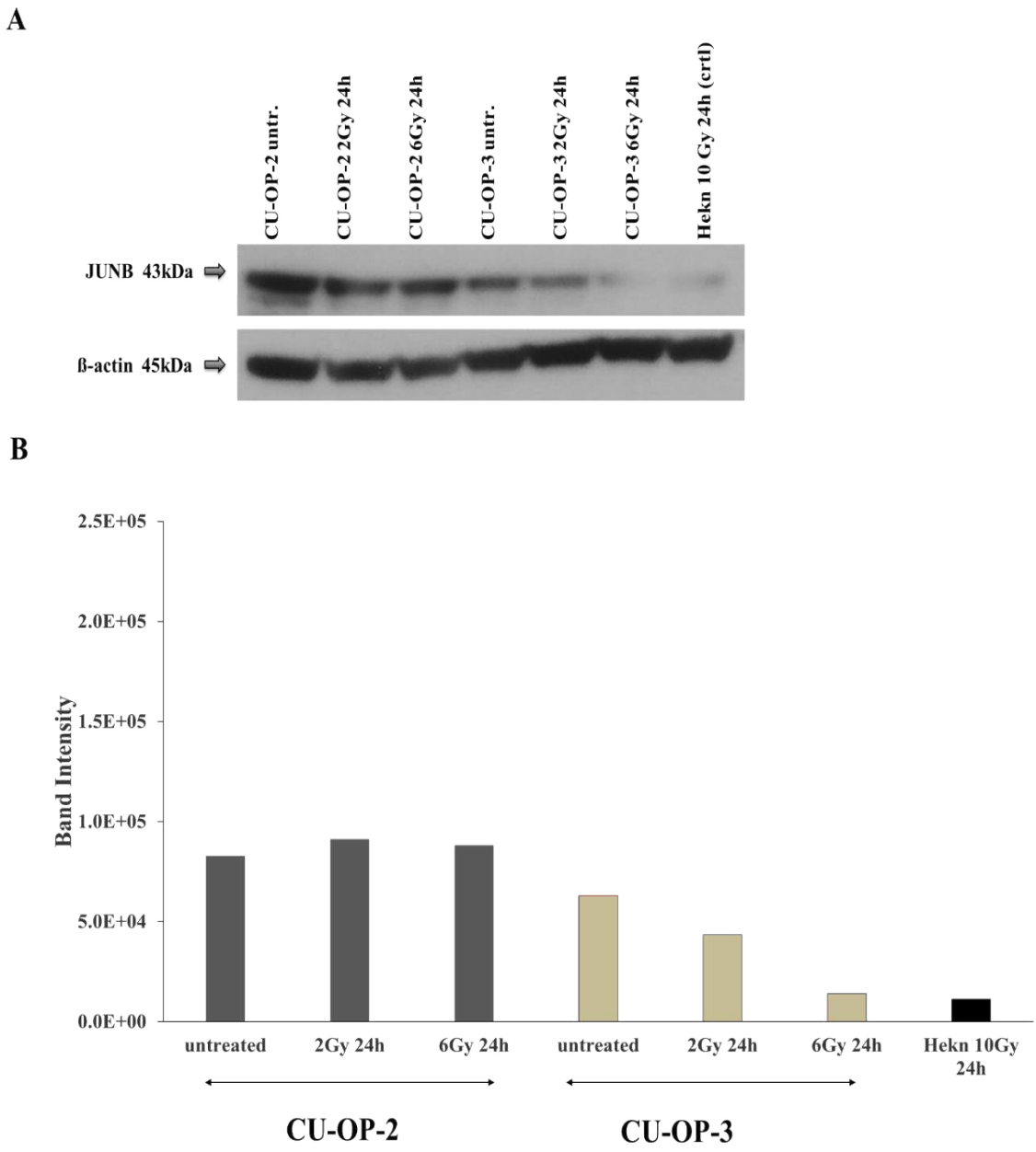



**Figure 5-15: Western blot results for JUNB (resistant cell lines to IR)**
(A) Protein bands of JUNB for CU-OP-2 and CU-OP-3. Cell lines were ordered from untreated to 2Gy and 6Gy IR treated samples (collected after 24h incubation after IR treatment). HEKn 10Gy 24h was added as positive control. Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) JUNB (30-35min exposure) bands are seen in all three cell lines.
(B) Quantification of western blot intensities of JUNB: The band intensity of the western blotting bands as quantified in relation to ß-actin level using Image Studio Lite.

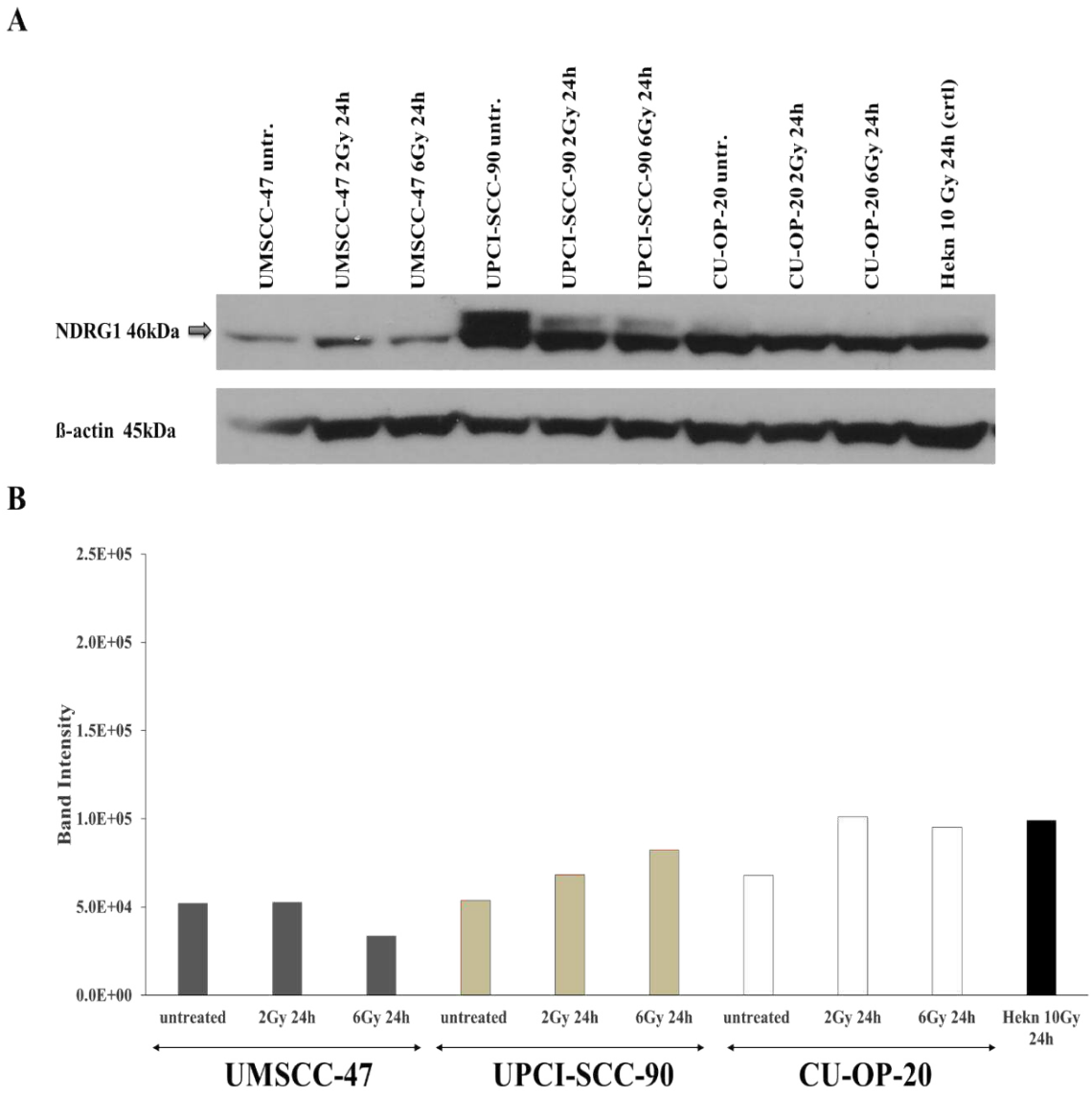



**Figure 5-16: Western blot results for NDRG1 (resistant cell lines to IR)**
(A) Protein bands of NDRG1 for UMSCC-47, UPCI-SCC-90 and CU-OP-20. Cell lines were ordered from untreated to 2Gy and 6Gy IR treated samples (collected after 24h incubation after IR treatment). HEKn 10Gy 24h was added as positive control. Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) NDRG1 (1-15min exposure) bands are seen in all three cell lines.
(B) Quantification of western blot intensities of NDRG1: The band intensity of the western blotting bands as quantified in relation to ß-actin level using Image Studio Lite.

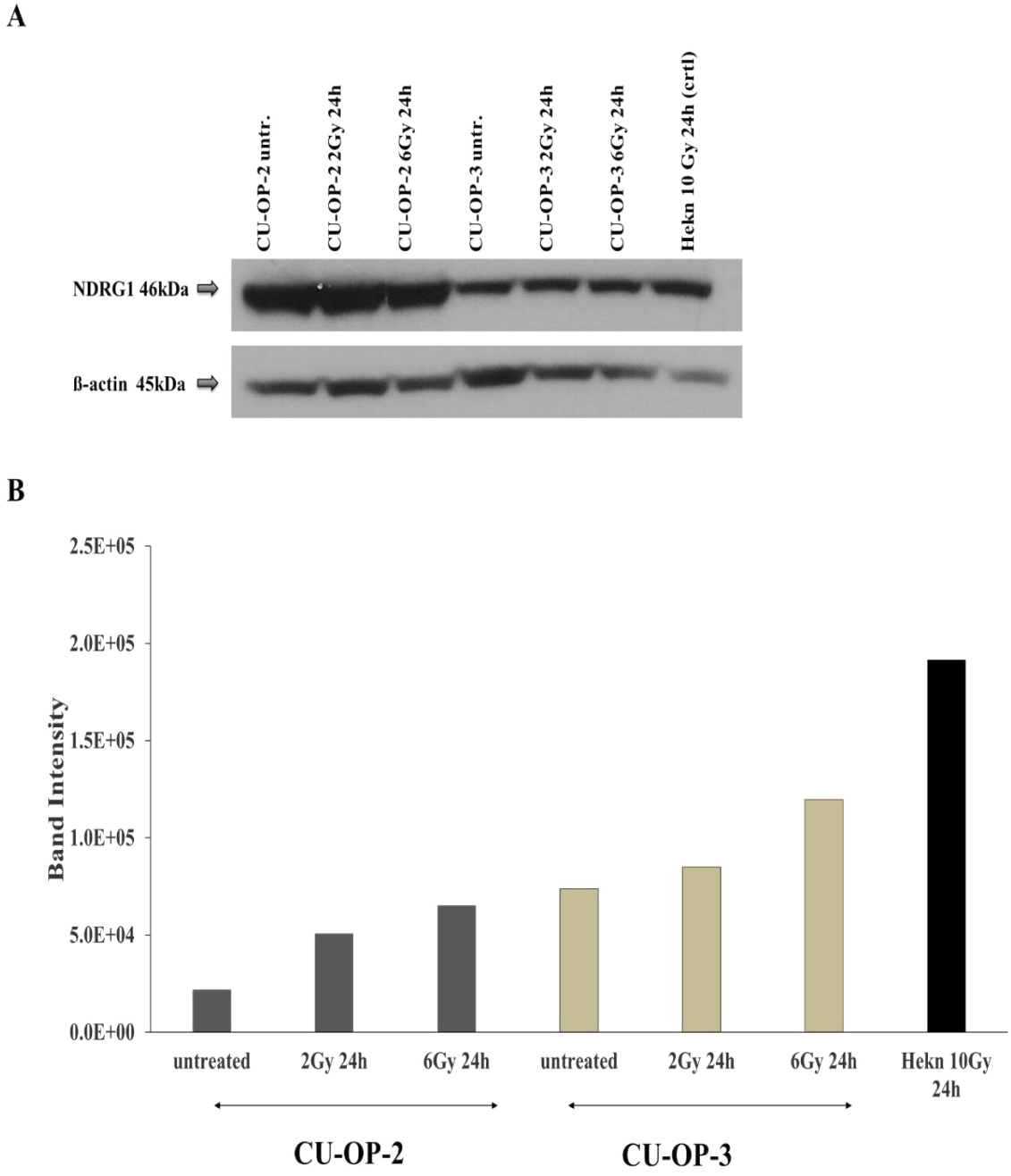



**Figure 5-17: Western blot results for NDRG1 (resistant cell lines to IR)**
(A) Protein bands of NDRG1 for CU-OP-2 and CU-OP-3. Cell lines were ordered from untreated to 2Gy and 6Gy IR treated samples (collected after 24h incubation after IR treatment). HEKn 10Gy 24h was added as positive control. Membrane was stripped and re-probed with ß-actin antibody. ß-actin (1-2min exposure) NDRG1 (1-15min exposure) bands are seen in all three cell lines.
(B) Quantification of western blot intensities of NDRG1: The band intensity of the western blotting bands as quantified in relation to ß-actin level using Image Studio Lite.

## 5.6. Genetic pathways of interest

It was hypothesized that differences in DNA repair mechanism might explain the differences in IR sensitivity within the HPV-positive cell lines. The expression of genes in main DNA repair/damage pathways were investigated using GenView, a GO directed analysis.

These comparisons were performed using "Subread featureCounts" Version 1.5.1 (Liao et al., 2014) a software package for assigning sequence reads to genomic features. Read summarization (counting) was generated for paired end read fragments. It was summarized at exon level and then grouped at transcript level. To provide stringent and robust data, reads overlapping more than one feature were excluded from the count summary, because any single fragment must originate from only one of the target genes but the identity of the true target gene cannot be confidently determined (as per the authors' recommendations (Liao et al., 2014)). Heatmaps were then obtained using the Genview package. The heatmap outputs for Double-Strand Break Repair via Homologous Recombination (DSBR-HR), Base Excision Repair (BER) and Double-Strand Break Repair via Non-Homologous End Joining (DSBR-NHEJ) ontologies are shown in Figure 5-18, Figure 5-19 and Figure 5-20.

The heatmaps for genes involved in DSBR-HR, BER and DSBR-NHEJ showed no obvious differences between HPV-positive versus HPV-negative cell lines, either with untreated or irradiated cells (Figure 5-18, Figure 5-19 and Figure 5-20).

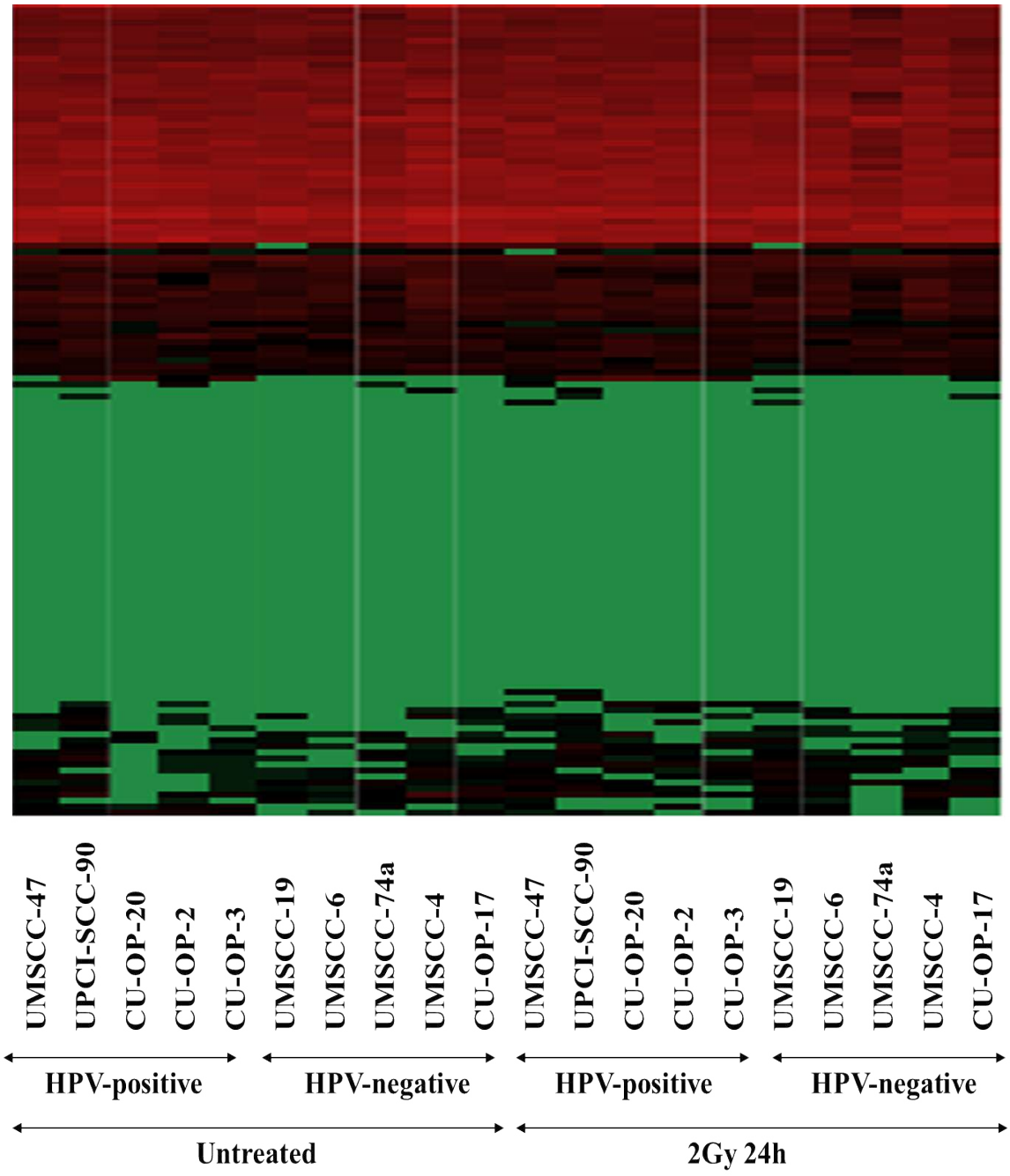



**Figure 5-18: Differential expression of genes involved in the DSBR-HR pathway between HPV-positive and HPV-negative OPSCC cell lines**
Cell lines were listed from left to right, as untreated (HPV-positive and HPV-negative) and treated 2Gy 24h (HPV-positive and HPV-negative). Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (higher expression). The heatmap presented is based on the analysis of 135 genes. Untreated samples were collected at the same time as the treated counterpart.

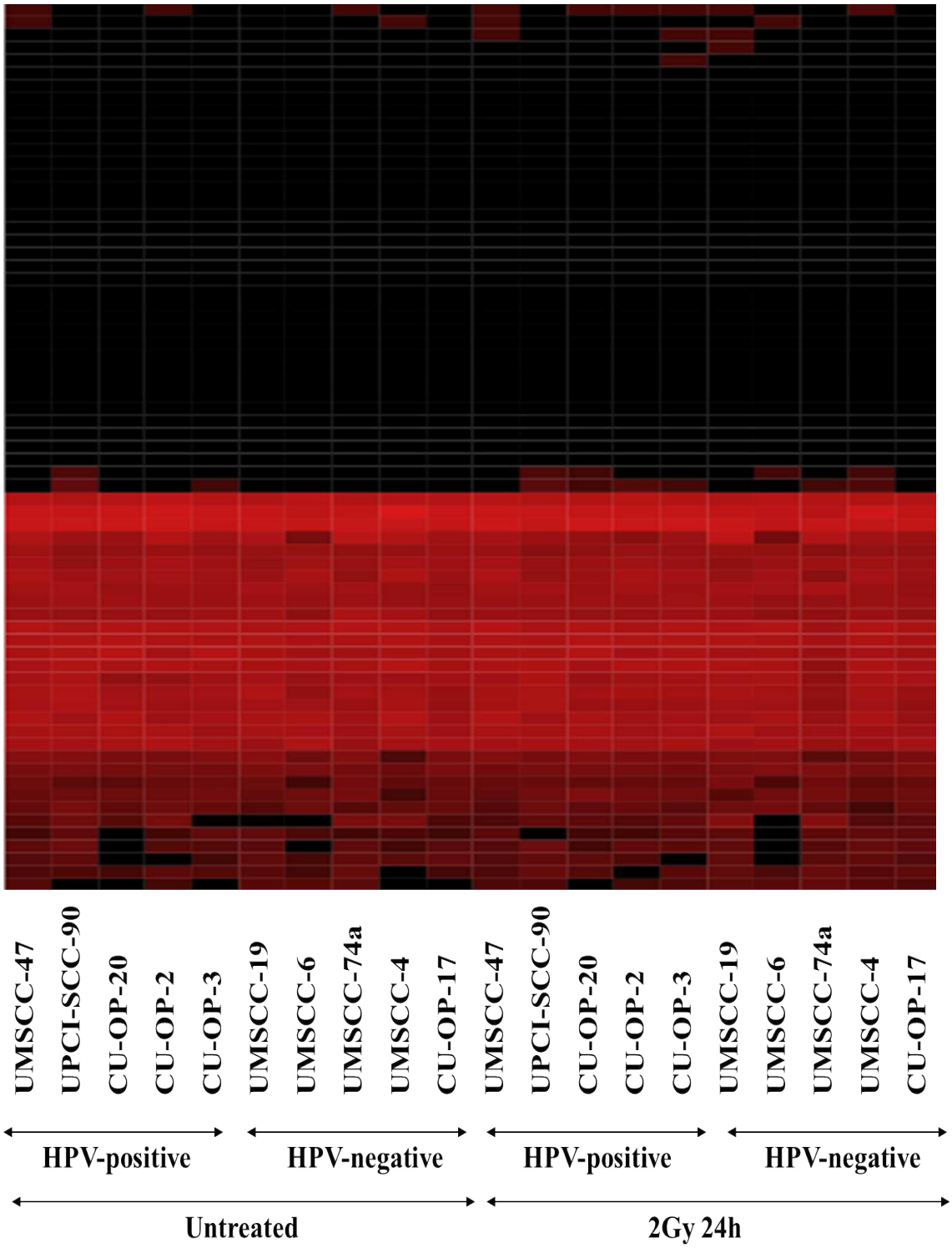



**Figure 5-19: Differential expression of genes involved in the BER pathway between HPV-positive and HPV-negative OPSCC cell lines**
Cell lines were listed from left to right, as untreated (HPV-positive and HPV-negative) and treated 2Gy 24h (HPV-positive and HPV-negative). Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (higher expression). Based on the analysis of 69 genes. Untreated samples were collected at the same time as the treated counterpart.

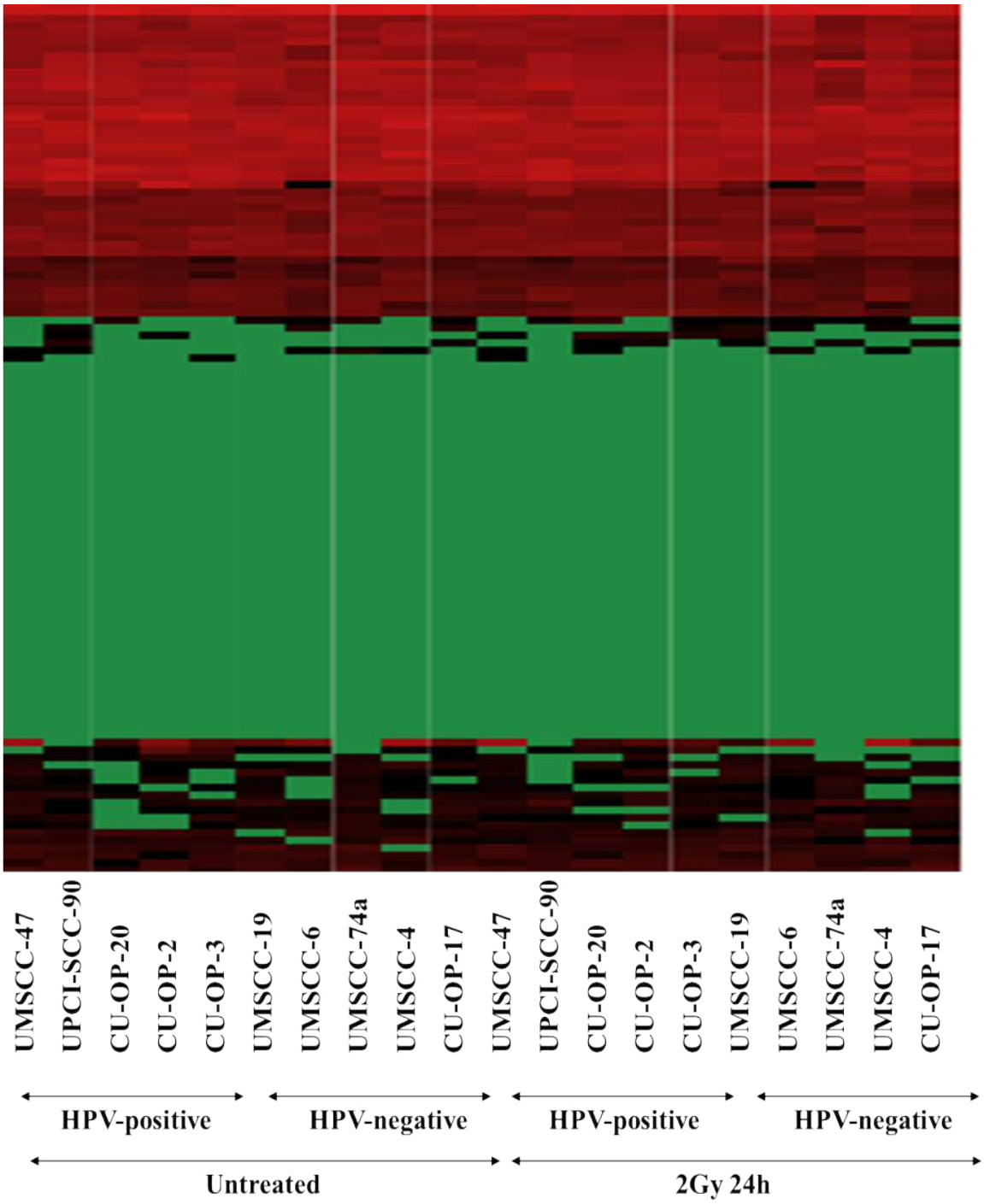



**Figure 5-20: Differential gene expression DSBR-NHEJ pathway between HPV-positive and HPV-negative OPSCC cell lines**
Cell lines were listed from left to right, as untreated (HPV-positive and HPV-negative) and treated 2Gy 24h (HPV-positive and HPV-negative). Gene expression was normalised across the tested samples and colour coded (green, black, red). Black represents a median value, green is below the median value and red is above the median value (higher expression). Based on the analysis of 116 genes. Untreated samples were collected at the same time as the treated counterpart.

## 5.7. Relative expression of genes of interest in terms of IR sensitivity

Based on the hypotheses predicting roles for p53, p21 and p16 in determining resistance to IR, an exploratory gene expression analysis was carried out using the mRNA sequencing data. These genes are presented (showing fold-changes after IR) in histograms using RPKM (multiple reads) in Genview (Figure 5-21).

### 5.7.1. TP53 transcript levels

TP53 plays a key-role in DNA repair and apoptosis. The Western blotting analysis performed earlier (Chapter 4) showed no obvious differences in p53 expression between the HPV-positive cell lines and HPV-negative cell lines after IR treatment (2Gy and 6Gy). Transcript levels of TP53 were investigated to shed further light on regulation of p53 activity in OPSCC cell lines.

After 2Gy treatment there was no obvious overall increase or decrease in p53 transcript level between the untreated versus treated samples. However, at the individual cell line level, CU-OP-2 showed an increase in fold-change after 2Gy treatment. UMSCC-47, CU-OP-3 and CU-OP-20 also showed a minor increase in number of reads of p53 after 2Gy treatment whereas UPCI-SCC-90 showed a decrease of p53 transcription after 2Gy treatment (low fold-change) (Figure 5-21).

Low transcription of p53 was shown in UMSCC-4, UMSCC-19 and UMSCC-6, where UMSCC-4 and UMSCC-19 had a mutated allele of p53 compared to the HPV-positive cell lines (Figure 5-21).

### 5.7.2. CDKN1A (p21) transcript levels

Between the paired comparison of untreated samples versus 2Gy treated samples no obvious differences in fold-changes were identified. However, all OPSCC cell lines showed a transcription of p21 in untreated and treated samples. A greater fold-change was seen in UMSCC-19 (FC: 1.6) and UMSCC-74a (FC: 1.9) after 2Gy IR treatment. Minor increases were detected on transcription level in all other cell lines (minor fold-changes) (Figure 5-21).

### 5.7.3. CDKN2A (p16) transcript levels

P16 a surrogate marker for HPV-16 infection (Schache et al., 2011a), was detected in all HPV-positive OPSCC cell lines. IR treatment had no obvious impact on fold-changes on p16 after IR treatment. Number of reads for p16 were measured as well in two HPV-negative cell lines (UMSCC-19 and CU-OP-17) with a low number of transcription (no obvious fold-change). Detection of p16 in HPV-negative cancers is possible and therefore p16 expression is not a complete guarantee of HPV infection (Evans et al., 2013, Pirotte, 2017) (Figure 5-21).

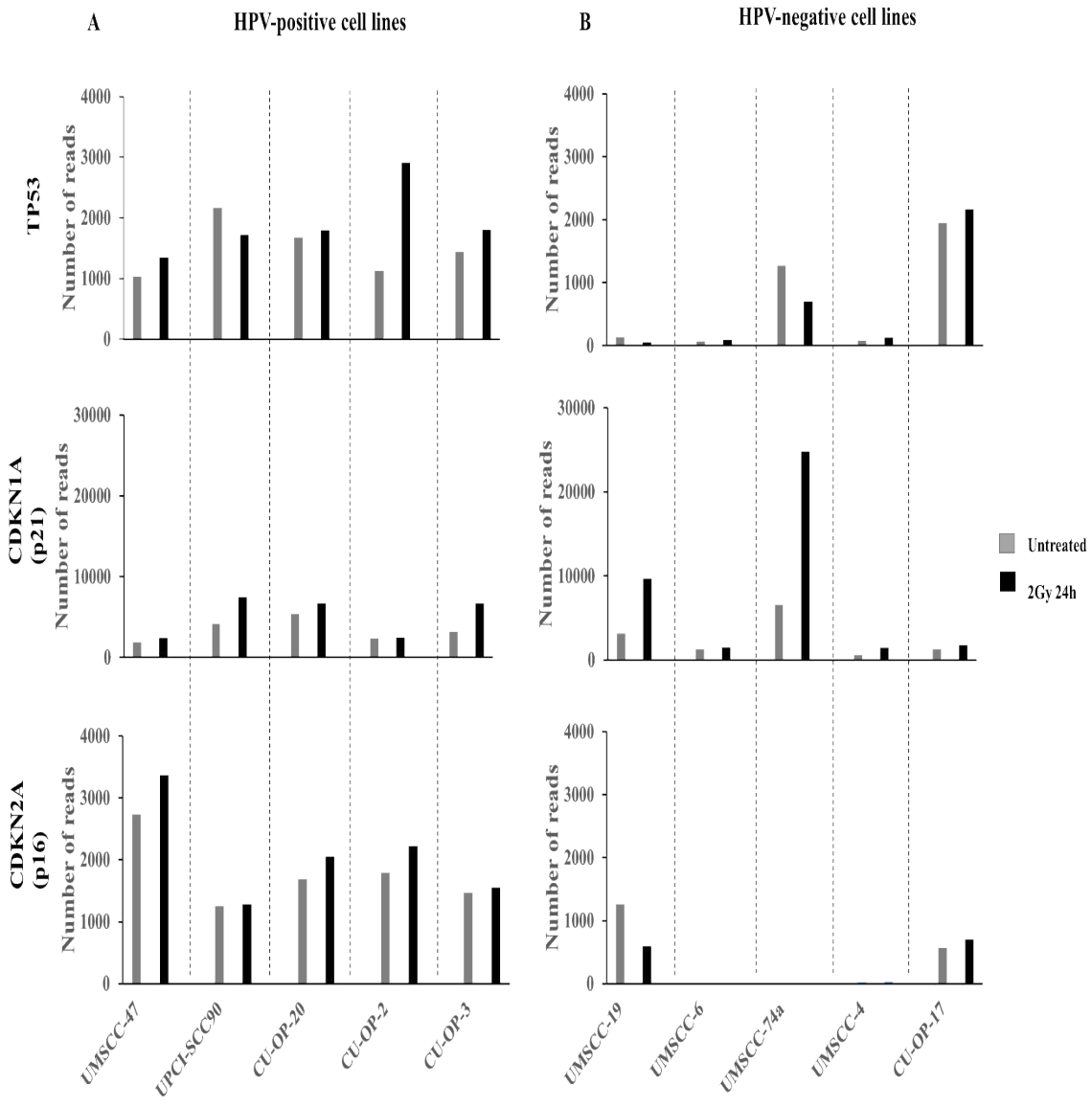



**Figure 5-21: TP53, CDKN1A (p21) and CDKN2A (p16) transcript levels for HPV-positive and HPV-negative OPSCC cell lines**

(A) HPV-positive (B) HPV-negative). HPV-positive paired comparison between untreated and treated groups showed no significant differences of TP53, p21 and p16. Histograms generated using multiple counts (RPKM) on Genview. For TP53, the greatest fold change was detected with CU-OP-2, with a 2.5-fold increase compared to its untreated sample. The remaining HPV-positive cell lines showed a fold change between 0.79 (UPCI-SCC-90) and 1.3 (UMSCC-47). The greatest fold change in the HPV-negative OPSCC was seen in UMSCC-4 (low read number) with 1.7 FC. For p21, the greatest fold change was seen in the HPV-positive OPSCC cell line CU-OP-3, with a 2-fold increase, and in the HPV-negative OPSCC cell line UMSCC-74a with a fold change of 3. For p16, no obvious changes were seen in in the HPV-positive cell lines. Two HPV-negative cell lines showed reads for p16 (UMSCC-19 and CU-OP-17). P16 levels for UMSCC-19have been reported in a previous thesis (not published in any study by the original source).

## 5.8. Summary of main findings

The main findings of the mRNA sequencing analysis were:

- Confirmed that all HPV-positive OPSCC cell lines had active expression of HPV encoded oncogenes.
- Showed that the expression of the HPV encoded oncogenes did not change after IR.
- Identified the HPV variants present in UPCI-SCC-90, UMSCC-47, CU-OP-2 and CU-OP-3 (consistent with previous studies). Identified the HPV-16 variant in CU-OP-20 as Asian-American.
- Showed that CU-OP-20 has HPV DNA integrated at multiple sites.
- Confirmed the main HPV-integration sites, as reported in previous studies, for UMSCC-47, UPCI-SCC-90, CU-OP-2 and CU-OP-3.
- Identified an increase in integration sites after IR for all HPV-positive OPSCC cell lines.
- Confirmed the p53 status for the established cell lines and for CU-OP-2 and CU-OP-3. CU-OP-20 was classified as wild-type p53. CU-OP-17 showed a likely pathogenic SNP at the UTR 5' region (non-coding) but was wild-type p53 in the coding region.
- Clustering of cell lines according to their IR sensitivity showed variation in expression levels of numerous genes. However, these genes were not obviously associated with DNA repair, DNA damage or apoptosis.
- Clustering of all OPSCC cell lines showed a general difference between HPV-positive and HPV-negative cell lines (untreated or 2Gy treated comparison).
- Different GO processes were identified between the comparisons, these were generally not directly related to DNA repair or DNA damage, except for identification of JUNB.
- Investigations of expression levels of individual genes showed:
    - More abundant transcripts of TP53 in HPV-positive cell lines than in HPV-negative lines. Consistent transcription of p21 is observed throughout the HPV-positive cell lines.
    - CDKN2A (p16) expression was confirmed for all HPV-positive cell lines. Two HPV-negative cell lines showed transcription of p16.

- In the majority of HPV-positive cell lines, JUNB showed increased expression after IR.
- NDRG1 showed increased expression in 3/5 HPV-positive cell lines.

## 5.9. Discussion

### 5.9.1. Confirmation of HPV gene expression

Investigations of HPV gene expression confirmed active expression of HPV genes in all HPV-positive OPSCC cell lines in the current study. The level of HPV gene expression did not change with a great fold-change in the majority of HPV-positive OPSCC cell lines (4 out of 5) after IR treatment (2Gy). There was a similar pattern of HPV gene expression before and after treatment of cell lines, with different strength of expression of HPV ORF's for individual cell lines. However, CU-OP-20 was the only cell line that showed an obvious change in HPV gene expression, with an approximately 50% reduction in transcripts following IR for E6, E7, E2, E5 and E4. CU-OP-20 was also the only HPV-positive cell line, which did not show a main integration site for HPV-16. Down-regulation of certain HPV ORF's after IR treatment could be a sign of an non or incompletely (episomal) integrated virus.

HPV gene expression assessed by mRNA level, HPV DNA typing and p16 status are factors used to define an HPV-positive tumour (Schache et al., 2011b), and CU-OP-20 fulfils this definition. Splice sites were also investigated in untreated and 2Gy treated samples. Patterns of splicing were previously reported (Pirotte, 2017) for several of the cell lines, and the current data are consistent with that report. Splice sites after 2Gy treatment showed a similar pattern compared to the untreated samples. There was no increase of E6*I or E6*II transcripts, but an increase of minor splicing sites at E2, E5 and E1^E4 in the CU-OP cell lines. The role of E6*I and E6*II and other E6 isoforms is directly connected to cancer development. The presence of E2 could be connected to active splicing within the CU-OP cell lines (CU-OP-3 and CU-OP-20), as E2 is an mRNA binding protein which regulates E6 splicing (Olmedo-Nieva et al., 2018). Expression of E6*I and E6*II confirmed the classification of HPV-driven tumours, including the novel cell lines of CU-OP-2, CU-OP-3 and CU-OP-20.

### 5.9.2. Fusion transcripts

The sites of HPV-16 integration were assessed and confirmed for UMSCC-47, UPCI-SCC-90, CU-OP-2 and CU-OP-3 (Pirotte, 2017). No main integration site was seen in CU-OP-20. The location of integration sites was determined, as it was previously reported 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 proportion of fully integrated HPV virus in oropharyngeal cancer is around 43% (Olthof et al., 2014), whereas 57% of HPV-positive samples showed extrachromosomal or mixed integration (Vojtechova et al., 2016). Interestingly the number of reads of mapped integration between the human genome and the HPV-16 genome decreased, even if the chromosomal sites are increasing, in UMSCC-47 and UPCI-SCC-90 (sensitive to IR). CU-OP-2 and CU-OP-3 showed an increase of mapped reads between untreated and treated samples. In CU-OP-20, HPV integration only showed a limited number of reads (mostly a single read per site) and showed an increase of mapped reads between HPV-16 and hg19. This is of interest, as CU-OP-20 is a radio-sensitive cell line. The greater variation in response to IR is seen as well within the integration sites (mRNA sequencing) of the HPV-positive cell lines, as two cell lines showed a decrease of reads compared to the remaining cell lines showing an increase of reads.

The main integration sites for HPV were not affected by IR. However, the number of integration sites increased (number of reads) in the majority of the OPSCC cell lines. Initial integration is more likely to occur in cells harbouring E1 and E2 replication proteins (episomal) (McBride and Warburton, 2017). Expression of E1 and E2 is seen in the majority of the cell lines included in this study. This could lead to increase of integration sites after IR, as HPV DNA hijacks the host DNA damage response to replicate its own DNA (McBride and Warburton, 2017). The increase of integration sites (excluding the number of reads) can be associated with the induced double strand breaks in the human genome. Breaks in the host DNA and the circular viral episome must occur for integration. Studies on oncogenic viruses have shown that integration is increased by the induction of DNA double strand breaks. These DSB might be enabled in a higher frequency and contribute to carcinogenesis (Williams et al., 2011). Studies in hepatitis B (HBV) have shown that integration frequency increases with DNA damage, and a similar mechanism could operate for HPV in the current study (Dandri et al., 2002). According to a study on integration of HPV in relation to patient prognosis, no statistically significant

difference was found in patient survival between those with HPV-positive integrated and those with extrachromosomal (episomal)/mixed forms of the virus (Vojtechova et al., 2016).

In general, HPV-positive patients have a better prognosis in survival compared to HPV-negative patients in Wales, however, the role of viral integration sites after IR was not investigated (Evans et al., 2013). Furthermore, during additional integration or integration of HPV into the human genome, epigenetic changes (DNA methylation) might occur, influencing whether there is active or silent integration (McBride and Warburton, 2017). Therefore, additional experiments have to be performed to determine the functionality of the new integration sites (methylation status) and confirming new integration sites e.g. using PCR as an additional control to mRNA sequencing results.

### 5.9.3. Confirming p53 status

The p53 status of the cell lines was also investigated using the mRNA sequence data. The wild-type status of the established cell lines UMSCC-47, UPCI-SCC-90, UMSCC-6 and UMSCC-74a was confirmed (Bradford et al., 2003, Ferris et al., 2005). UMSCC-4 and UMSCC-19 were confirmed to contain mutated p53 (Mandic et al., 2005, Somers et al., 1992). This p53 status was confirmed within this study as well. Additionally, the status of p53 for CU-OP-2 and CU-OP-3 was confirmed as consistent with a previous report (Pirotte, 2017). No mutation was found for CU-OP-20. In CU-OP-17 a SNP was identified in the UTR 5'prime region, but no mutations were identified. Therefore, CU-OP-17 can be considered as wild-type p53.

The majority of HPV-positive head and neck cancers in which there is degradation of p53 by HPV E6, have wild-type p53, as the selective pressure for p53 mutation is low in these tumours (Stransky et al., 2011). Conversely, mutations of p53 are mainly found in HPV-negative tumours, which are associated with shorter survival times and resistance to chemotherapy and radiotherapy (Evans et al., 2013).

### 5.9.4. GO analysis of OPSCC cell lines and investigation of differentially expressed DNA repair genes (including sensitivity analysis and effect of IR)

Analysis of transcription levels of significantly expressed genes, revealed a clear difference between HPV-positive and HPV-negative OPSCC cell lines, within the untreated and treated comparisons. Furthermore, GO analysis for the basic comparison of HPV-positive and HPV-negative OPSCC cell lines revealed various differentially regulated biological processes, including regulation of signalling or cell communication, which could play a role in how cell communicate differentially in response to DNA damage. Directly connected GO categories to DNA damage, DNA repair or cell cycle regulation (arrest) were not significantly overrepresented.

The GO analysis (after IR – paired tests) comparing the response to IR of HPV-positive sensitive and resistant cell lines, revealed differences in levels of transcripts ascribed to several ontologies. The only BP identified was the "homeostatic process", which is an important process in an organism to maintain a stable environment and plays a key role in cell proliferation and growth or involvement in regulatory functions of apoptosis in head and neck cancer (Ressnerova et al., 2016).

While comparing the GO analysis (untreated versus IR treated) in a paired comparison manner for all OPSCC cell lines the BP "positive regulation of G0 to G1 transition" was identified, which modulates the transition from the G0 to the G1 phase of the cell cycle. This could be connected to cell cycle arrest after IR treatment, as observed in Chapter 4.

Due to the fact that no DNA repair or DNA damage pathways were identified in the paired analysis including all OPSCC cell lines after IR, the focus was put on the HPV-positive cell lines. This was done because these lines showed the greatest variation in sensitivity in the radio-sensitivity assay. One of the goals of this chapter was to identify a gene, a pathway or a process explaining the variation in sensitivity to IR within the HPV-positive cell lines. By comparing the HPV-positive cell lines with each other before and after treatment, and including the genes identified with a greater fold-change and specific function, two genes were identified (JUNB and NDRG1).

JUNB is involved in regulation of cell death, apoptotic processes and response to radiation. It is known as a cell proliferation inhibitor, a senescence inducer, a tumour suppressor and is involved in different phases of the cell cycle. It not only has cell-division inhibiting but also cell-division promoting activities, depending on the cell cycle

stage (Piechaczyk and Farras, 2008). Another study reported normal expression of p53 and p21 in hepatocytes lacking JUNB, but both proteins were significantly elevated in Jun deficient fibroblasts (Hess et al., 2004). However, the protein levels of p53, p21 and JUNB observed in the current study do not support this direct association. Alternatively, JUNB be might be associated with differences in sensitivity due to its characteristics as a cell division inhibitor and senescence inducer.

NDRG1 has a connection to DNA damaging signalling (in connection with p53). NDRG1 is regulated by Myc and is induced under certain stress and pathological conditions. Furthermore, the encoded protein is important for p53-mediated apoptosis (Zhang et al., 2008).

During the paired comparisons no processes were identified in untreated or treated samples with a direct connection to DNA repair or DNA damage response pathways within the HPV-positive cell lines (sensitive versus resistant to IR). This was confirmed by the automatic generated heatmaps on Genview for BER, DSBR-HR and DSBR-NHEJ with no substantial differences between HPV-positive and HPV-negative OPSCC cell lines (untreated versus treated). However, such pathways have been previously reported in HPV-positive tumours in a larger dataset with a wider range of head and neck cancer cell lines (Zhang et al., 2016). This was not confirmed by the GO analysis in the current study, and is consistent with another study in our research group (Pirotte, 2017). These findings appear to contradict the reported postulated deficiencies in DSBR association with HPV-positive HNSCC cell lines (Kimple et al., 2013, Rieckmann et al., 2013).

### 5.9.5. Expression of genes of interest

Relative expression of several specific genes of interest were assessed, including p53 (TP53), p21 (CDKN1A), p16 (CDKN2A). None of these genes appeared to be consistently up or down regulated in response to IR.

The p53 protein is a critical tumour suppressor, and is held in check in unstressed cells by Mdm2 (E3 ubiquitin ligase), which binds to p53 and targets it for proteasomal degradation. In response to stimuli (e.g. stress, induction of DNA damage), inhibition is relieved and p53 target genes are trans-activated to cause e.g. apoptosis or cell cycle arrest (Beckerman and Prives, 2010). TP53 transcript levels were relatively abundant in all HPV-positive cell lines, and were present at equivalent levels in two of five HPV-negative cell lines (present in UMSCC-74a and CU-OP-17). Studies have shown that

overexpression of Myc and/or HPV E7 resulted in increased transcription of p53 in cell lines (Liu et al., 2007, Strickland and Vande Pol, 2016). Limited transcription of p53 in HPV-negative cell lines (except UMSCC-6) might be associated with wild-type or mutation status of p53, as a study has shown that mutated p53 could not trans-activate transcription from the p53 promotor (Deffie et al., 1993). This suggests that residual p53 activity in HPV-positive cell lines, could stimulate p53 transcription in the co-presence of p21 transcription. This process, would not be present in HPV-negative cell lines harbouring p53 mutations.

Transcription of CDKN1 (p21) is known to be directly regulated by p53 (Riley et al., 2008). However, p53-independent transcription of p21 is also possible (Galanos et al., 2016). Transcripts of p21 were present in all HPV-positive cell lines (untreated and treated). Less p21 transcription is seen in the majority of the HPV-negative cell lines. Functioning p53 could be present in UPCI-SCC-90 due the observed increase of (CDKN1) p21 transcription after IR. A similar pattern was seen for UMSCC-19 and UMSCC-74 after IR for p21 with fold change increase of 1.6 and 1.9, respectively.

The p16 (CDKN2A) gene is a tumour suppressor gene and is involved in the p16 – cyclin dependent kinase (CDK) – retinoblastoma (pRB) pathway of the cell cycle (Ai et al., 2003). It is known as a negative regulator of the cell cycle. The p16 gene encodes an inhibitor of CDK4 and CDK6, which regulates phosphorylation of pRB and the G1 to S phase transition. (Ai et al., 2003). Therefore, over-expression of p16 is directly connected to HPV E7 protein following down-regulation of pRB, through p16 connection in the pRB pathway. HPV DNA testing and p16 IHC are used as diagnostic tools to identify HPV driven tumours (Schache et al., 2011b). High levels of p16 were therefore expected in the panel of HPV-positive tumours. The p16 expression was confirmed through the provided patient data for the obtained biopsies and through CDKN2 (p16) expression after mRNA sequencing. The p16 IHC status is a known marker for the majority of HPV-positive tumours (Schache et al., 2011b).
However, transcription of p16 in the HPV-negative cell lines of UMSCC-19 and CU-OP-17 was found as well. For CU-OP-17 this was consistent with the p16 IHC status of the original tumour. This confirms previous reports of 14-17% p16 positive cases in HPV-negative tumours. A previous study of 187 OPSCC cases positive for p16 has identified that 26 cases were HPV-negative (14%) (Lewis et al., 2010). This was confirmed in

another study with a similar pattern where 17% of cases were identified as p16 positive and HPV-negative (Rietbergen et al., 2014).

Another possibility for no obvious up- or down-regulation of genes after IR treatment could be the heterogeneous population of cells within the cell lines used in the current study. The included CU-OP cell lines were used at as early as possible passage numbers, which should mimic the heterogeneous characteristics of the original tumour. Within the heterogeneous population, single sub-clones could be more or less sensitive to IR, which might lead to the data obtained after mRNA sequencing. To identify a possible different response of sub-clones to IR, the heterogeneous cell culture has to be sorted into single cells first. A single cell RNA sequencing on the single cells/sub-clones would identify differences in molecular mechanism, gene expression and mutation, as suggested by literature (Fan et al., 2018, Wang et al., 2018). The individual profile of single cell RNA sequencing could then be compared to the heterogeneous mRNA sequencing profile to identify similarities with a specific sub-clones. Furthermore, to identify outgrowth of a specific sub-clone, mRNA sequencing could be performed after each passage in parallel to single cell isolation.

Furthermore, a lack of DNA repair changes on mRNA level could be correlated to the lack of measurement of protein control level in activity. Indication of functioning or active protein could be identified through there phosphorylation status, such as p-p53 for p53), Identification of active protein level could lead to an active process in terms of DNA repair in gene expression. Parallel measurement of protein activity could have been used to correlate up- or downregulation of genes (e.g. DNA repair) to its functionality.

In summary, these data did not show a clear induction or down-regulation of genes after IR, connected to DNA repair or DNA damage. Activity in up- or downregulation of this genes (e.g. protein level in activity) need to be confirmed for the final conclusion, if these two mechanisms can explain the variation in radio-sensitivity within the HPV-positive cell lines. However, the data do show a response to IR across a wide range of genes, showing a clear effect of irradiation on the cell lines.

**Strengths and weaknesses**

The use of mRNA sequencing is a very robust and sensitive technique, which provided in depth analysis of the OPSCC cell lines panel. It was the first analysis by mRNA sequencing of the cell lines in the current study, including the novel cell lines CU-OP-20 and CU-OP-17, and assisted in their characterisation. Furthermore, it provided a deeper assessment of genes involved in different repair pathways.

A strength of the study is the sample size of 20 samples (10 cell lines in total), which included untreated samples and 2Gy IR treated samples. The main weakness, was that the mRNA sequencing data is based on one single sample for each cell line and condition (untreated and treated), which limits the statistical power. The study was limited to one replicate per sample by considering the size of the study cohort (number of extractions, cost and time factor). Additionally, the comparison between untreated to its treated counterpart was considered as sufficient during design of the experiment. To confirm the findings in this study, the experiment should be set-up and repeated a second time, within a larger data set, data points (treatment dose and time points) and replicates.

Furthermore, a direct comparison of the cell line model to the original tumours is difficult to make, because it is unclear how close gene expression in cell lines represents gene expression pattern in the original tumours. Hence, the original tumours consist of the entire heterozygous cell composition, whereas in cell lines model, a limited number of clones were able to grow on a plastic cell culture plate. However, it is important to point out, that growth conditions were standardized (e.g. identical FBS concentration and culture condition) and the differences observed on transcription (mRNA), molecular biology, mutation profile and HPV status reflected the characteristics of the original tumours.

## 5.10. Conclusion

The novel cell line CU-OP-20 derived in this study, showed all typical characteristics of a HPV driven tumour (expression of HPV genes, p53-wild-type, p16 positive and several integration sites). However, CU-OP-17 was defined as a HPV-negative OPSCC cell lines, with an unusual expression of p16 positivity (reflecting 14-17% of OPSCC).

IR seemed to have an effect on the number of HPV integration sites (fusion transcripts).

JUNB and NDRG1 were identified as genes, involved in cell death regulation and DNA damage signalling (connection to p53), which could play a part in the variation in sensitivity to IR within the HPV-positive cell lines.

The main conclusion is that there is no evidence for differential induction of DNA repair or DNA damage pathways in association with HPV-positive compared to HPV negative lines, or between radiation sensitive and resistant lines, which might suggest other factors involved in the different survival rate of HPV-positive and HPV-negative OPSCC cell lines. However, DNA repair and DNA damage cannot be excluded completely, as activation of these genes would need to get confirmed. This could be done via a protein based assay in controlling protein level changes in activity. Possible alternative factors might be involvement of the immune system or the tumour microenvironment, which was not possible to test in the cell line model used in the study.

# Chapter 6 : Final Discussion

The incidence of OPSCC is increasing in the UK and across the developed world, and an increasing number of cases are caused by HPV (Arenz et al., 2014, Attner et al., 2010, Evans et al., 2013, Gillison et al., 2012a, Kimple et al., 2013, Nasman et al., 2009, Nasman et al., 2015, Schache et al., 2016). However, a clear understanding of how this disease acts at the basic molecular level is lacking, and the reason patients with HPV-positive and HPV-negative OPSCC differ in terms of survival is still unclear. One reason for a lack of understanding of the disease could be the limitations of the OPSCC cell line models that were available when this project started (as discussed in Chapter 3), in particular HPV-positive OPSCC cell lines. A higher number of HPV-positive cell lines would provide more robust findings by providing a larger sample of a heterogeneous patient group. This would allow for individual patient variation and support investigation of the full range of responses to treatment (RT or CT).
Furthermore, new cell lines would provide a deeper understanding of molecular mechanisms, possibly through assessment of transcription levels of genes involved in the treatment response or DNA repair. This could assist the development of tailored treatment options for both OPSCC groups (HPV-positive and HPV-negative).

One of the aims of the current project was the development of new OPSCC cell lines for use in pre-clinical studies (described in Chapter 3). Two cell lines (CU-OP-17 and CU-OP-20) were successfully developed; CU-OP-20 was HPV-16 positive and p16 positive, and so was typical of most HPV-positive OPSCC; CU-OP-17 was HPV-negative and 16-positive, and this molecular profile is typical of 14-17% of all HPV-negative cases (Evans et al., 2013, Lewis et al., 2010, Rietbergen et al., 2014). These two cell lines were incorporated into the current study and their response to IR tested. Prior to this, both cell lines were fully characterised genetically, by comparing the STR profiles of original biopsy with the derived cell line, and through mRNA sequencing (Chapter 5). The mRNA-sequence analysis confirmed expression of HPV encoded oncogenes in CU-OP-20. However, the generation of novel cell lines was labour intensive and time consuming, and the process used had a success rate of just 18%. This may appear low, but is similar to rates observed in other studies (Kodack et al., 2017). This could possibly be improved by testing the behaviour of the biopsy growth with multiple types of media and different FBS concentrations. The media with the best growth characteristics, by increasing the total number of passages, would be considered as ideal and chosen for the individual biopsies. This could increase the success rate for generating cell lines and it might be

possible to generate multiple cell lines from each patient. The disadvantage of this method, is that the entire sample could be lost. The biopsy is small and may not be homogenous. Thus different parts of the biopsy might represent different tissue types and might not respond uniformly to changes in media composition. Another option could be growing the human tumour in mouse models, before developing a new cell line, by generating patient derived xenografts. This would provide advantage of consistent tissue sample for establishing cell lines. However, Kimple *et al.,* (2013) have investigated the treatment response and tumour growth in patient derived xenograft (PDX) models, which will be discussed in connection with immune targets later in the discussion (Kimple et al., 2013).

The initial plan is to share the newly derived cell lines through direct or indirect collaboration with other research groups. Availability of newly derived cell lines will be made public through publications and will be shared with other groups by contacting the principal investigator. Additionally, cell lines might be made available through a cell line bank at a later stage.

A major strength of the current study was that it was based on 10 cell lines (5 HPV-positive and 5 HPV-negative OPSCC cell lines). This was similar, or higher, than used in other studies of HNSCC including key publications in the same field (Arenz et al., 2014, Kimple et al., 2013, Rieckmann et al., 2013, Weaver et al., 2015, Ziemann et al., 2017). The advantage of the current panel of cell lines was that it constitutes a better model for studying OPSCC, because the lines were all derived from OPSCC (compared to other studies with a mix of HNSCC, including non-oropharyngeal) and included lines derived from HPV-positive non-smokers. This was not the case in previous studies (Kimple et al., 2013, Rieckmann et al., 2013).

One of the main findings during the radio-sensitivity assay was the variation in response within the HPV-positive OPSCC cell lines (Chapter 4). More cell lines were observed to be more sensitive to IR within the HPV-positive OPSCC cell lines, compared to the HPV-negative OPSCC cell lines, as stated in the hypothesis. Due to the variation in radio sensitivity within the HPV-positive cell lines, a clear distinction between the HPV-positive and negative cell lines could not be made. This finding was novel as in a previous study, a clear difference in radio sensitivity was determined between HPV-positive and HPV-negative HNSCC cell lines. A minor overlap between HPV-positive and HPV-negative cell line in terms of radio sensitivity towards IR was already reported in two

other HNSCC studies, with one HPV-negative HNSCC cell line lying in the middle of the HPV-positive cell lines (Rieckmann et al., 2013, Arenz et al., 2014). The differences in sensitivity to IR indicates a heterogeneous cell line panel. This may provide a model more accurately reflecting the heterogeneity and differential clinical response of OPSCC patients.

The role of p53 pathways in determining variation in response to IR within the HPV-positive OPSCC and in influencing differences in behaviour between HPV positive and negative cell lines was investigated by measuring expression of p53, p21 and p-p53. During the investigations no direct correlation between HPV status with different levels of expression of the three factors was observed. Similarly changes in transcription level after mRNA were not observed. This might lead to the conclusion that p53 is not directly involved in regulating the IR sensitivity within the HPV-positive cell lines, as p53 is still expressed and earlier cell cycle arrest was not seen, which might be a sign for inactive p53. By contrast, Kimple *et al.,* (2013) suggested that accumulation of p53 influenced IR response and was associated with greater sensitivity to IR (Kimple et al., 2013). Differing levels of p53, could contribute to the variation in surviving fraction within the HPV-positive OPSCC cell lines.

The dissimilar results in IR response of cell lines obtained in this study compared to previously published studies could be due to the use of different experimental conditions e.g. use of different cell lines for the study, inclusion and exclusion of feeder cells. No support of feeder cells was used for testing the effect of IR in Kimple *et al.,* (2013) and Arenz *et al.,* (2014). However, Rieckmann *et al.,* (2013) used feeder cells for only one of eight cell lines. The inconsistency between experiments could have influenced the results obtained for IR sensitivity or growth pattern of the cell lines. This was avoided for the radio-sensitivity assay in the current study, where 3T3 feeder cells were added for every cell line in this specific assay. 3T3 feeder cells were included in the radio-sensitivity assay to support cell growth at low cell density.

Clinically, radiotherapy as monotherapy could have the disadvantage of harming surrounding tissue during treatment. Future experiments might include more combinations of IR with chemotherapeutic agents. Studies for HNSCC in combination of IR with e.g. cisplatin or roscovitine have already been started, which might have a greater effect on survival or tumour treatment (Busch et al., 2016, Ziemann et al., 2017). One study showed a slight reduction in proliferation in HPV-positive cells, but only minor

effect of the combined treatment on colony formation (Busch et al., 2016). Another study on roscovitine did not show an increase in radio-sensitivity in HPV-positive HNSCC cell lines (Ziemann et al., 2017). However, a combined treatment with a chemotherapeutic agent or cell cycle inhibitor might lead to a reduction of the toxic effect of IR to the surrounding tissue and a possible increased survival of patients. *In-vitro* studies allow investigation of the different responses to varying treatment doses of IR with or without chemotherapeutic agents. The cell lines would provide the opportunity to test multiple combinations of treatments, such as combinations of several chemotherapeutic agents with IR, cell cycle inhibitors (e.g. Chk1/Chk2) or PARP inhibitors (e.g. Olaparib). Findings of these experiments could then be adapted for clinic, which could improve treatment of patient tumours without harming normal tissues.

Interestingly in previous studies, it was hypothesised that down regulation of p53 by E6 and treatment of IR would result in G2 arrest, within HPV-positive HNSCC cell lines. (Kimple et al., 2013, Rieckmann et al., 2013). According to the original hypothesis of the current study a greater G2 arrest after IR treatment, especially within the HPV-positive OPSCC cell lines was predicted. However, despite the differential accumulation of p53 within the HPV-positive OPSCC cell lines and the increased sensitivity of some cell lines to IR, a clear association between G2 arrest and HPV-status could not be made, as some G2 arrest was seen in some HPV-negative cell lines as well. Greater G2 arrest was observed as well in previous studies for HPV-positive HNSCC cell lines relative to HPV-negative lines (Rieckmann et al., 2013, Kimple et al., 2013). In general, more cell lines in the current study showed a greater G2 arrest as well compared to the HPV-negative cell lines. A possible explanation of the greater G2 arrest could perhaps be residual p53 activity due to incomplete degradation by E6. Cells arresting in G2 after IR could be observed independent of residual p53 activity (Bernhard et al., 1995, Branzei and Foiani, 2008), which questions the key involvement of p53 activity in G2 arrest.

In summary, more cells arresting in G2 could be an indication of preparing cells for repair, and the return to a more “normal” cell cycle distribution after 48 hours could be associated with DNA repair.

p53 plays a key role in regulation of the DNA damage response pathway, with roles including activating apoptosis, causing cell cycle arrest or inducing DNA repair pathways (HR, BER, NHEJ). The two main repair pathways of DSB are HR and NHEJ. Therefore, the lack of increase in p53 or p21 expression after IR might explain the lack of increased

transcript levels in genes associated with these DNA repair pathways noted in the mRNA-sequence analyses.

Defective repair of DSB after IR was previously reported as being a possible explanation for greater sensitivity of HPV-positive HNSCC cell lines (Rieckmann et al., 2013, Weaver et al., 2015). This lead to the hypothesis that genes within the DNA repair pathways are up or down-regulated after IR, and this results in variation of radio-sensitivity within the HPV-positive cell lines. No change in gene expression for major DNA repair pathways could be detected in the current study by mRNA sequencing of cells treated with IR. The GO analysis and general clustering of HPV-positive and HPV-negative cell lines revealed differences in gene expression, but these were not involved in viral oncogenesis or IR response. These results suggest other mechanisms could be involved in the differences in sensitivity to IR between cell lines. However, additional to the mRNA sequencing protein level in activity should have been measured to confirm functional up- or downregulation of certain genes.

It is possible that using a higher dose of IR or different time points might identify a significant change in gene expression that correlated to sensitivity to IR. This could be tested in the future but would require large scale experiments to properly assess the impact of both variables on DNA repair mechanisms.

One of the main novel findings of the mRNA sequencing was the increase in fusion transcripts originating from multiple sites of HPV integration after IR. The number of integration sites appeared to be increased by IR, most likely due to the induction of DSB breaks after 2 Gy IR. Initial integrations usually takes place in regions of a lesion in a proliferating cell, where host and viral DNA are close. Therefore, one explanation could result from the known characteristics of HPV-DNA, which hijacks the host DNA damage response to replicate its own DNA (McBride and Warburton, 2017). So the increase in integration sites could be associated with the increased number of DSB (increased level of lesions) induced by IR. The main DNA repair pathway after inducing DSB with IR are NHEJ and HR, with the main difference that for HR an intact template strand is required compared to NHEJ, which does not need an intact DNA strand, to start the repair of DSBs (Jackson and Bartek, 2009, Lieber, 2010). It was notable that increased numbers of integration events occurred mostly in the newly established cell lines, and this may indicate that the presence of episomal HPV might be associated with this effect.

It should be noted that no experiments were performed to assess levels of DNA damage (e.g. assessment of ɣH2AX), before looking at different gene expression after IR. Determination of the level of DNA damage, could help to adjust the level of IR used to dose cells before further analyses. This would need to get done in future, prior to mRNA sequencing experiments. ɣH2AX is the best known marker of DNA damage, and indicates levels of DSB. Expression of ɣH2AX could be tested prior to future experiments via double-immunofluorescence assays. Staining with specific antibodies for ɣH2AX, would provide the number of foci of ɣH2AX per cell lines. The total number of foci would then directly correlate with the level of DNA damage induced. Testing ɣH2AX at different treatment doses and time points could identify when the highest damage is induced and show the time course of its repair. This would indicate how long it would take for DNA repair in the specific panel of cell lines, and help guide future mRNA sequencing experiments.

Studies on cell line models (especially novel cell lines or early passages) are restricted in scale by the technical work necessary to culture the cell lines and by their relatively slow growth. These models also lack the tumour associated microenvironment and there is an absence of immune cells. One study has shown that HPV-positive tumours showed a better survival rate if there are a higher number of tumour infiltrating lymphocytes (TILs). There was an enrichment of $CD8^+$ T-cells in the tumours of OPSCC patients which associated with improved survival (King et al., 2014, Ward et al., 2014).This was confirmed by another study of HPV-positive OPSCC which additionally demonstrated that HPV-positive OPSCC patients showed an improved survival if expressing a high ratio of $CD8^+/FoxP3^+$ (Nasman et al., 2017).

Currently, immune compromised mice are widely used in HNSCC xenograft models, to avoid the risk of rejection by the host immune system. In a previous study, xenograft models (T-cell deficient) were used, by injecting HPV-positive and HPV-negative tumour cells and monitoring the growth of the tumours and the response to IR. In this study by Kimple *et al.,* (2013), no difference was seen between irradiated and non-irradiated tumours within the HPV-negative xenografts. However, it showed an increased prolongation in time to tumour quadrupling between irradiated and non-irradiated tumours in HPV-positive xenografts, which was concluded with a similar response seen in sensitivity of the HPV-positive *in-vitro* models after radiation (Kimple et al., 2013). Future experiments could also be done in transgenic mouse models. The advantages of transgenic mouse models, include potential use of gene-editing technology, to upregulate

or downregulate certain genes. HPV-associated HNSCC transgenic mouse models have already been established, and used to investigate a chemical carcinogen, 4-nitroquinoline-1-oxide (4-NQO) and to identify a synergetic effect of E6 and E7 in causing HNSCC (Jabbar et al., 2010). Designing a new transgenic HPV-positive mouse model with enriched expression of TILs, could mimic the high expression seen in human models. This would provide the advantage after treatment with IR or other chemotherapeutic agents, to include a possible influence of the immune system on the better treatment response.

In the current study, there was no indication that DNA repair mechanisms could contribute to greater sensitivity/resistance within HPV-positive cell lines. However, it has to be noted that this assay only showed changes in gene expression at a fixed time point. Additionally, protein levels should get investigated to confirm changes in activity of genes involved in the DNA repair mechanisms. Limited to the obtained that, it can suggest involvement of other mechanisms operating in these cell lines and the situation *in vivo* is likely to be even more complex. Other factors that could influence the response to radiotherapy could include the host immune response or the tissue microenvironment for cancer cells.

## 6.1. Final Conclusion

In conclusion, two novel cell lines were successfully derived and characterised. These novel cell lines, together with a panel of previously established cell lines allowed laboratory investigation of the role of HPV in OPSCC. The study demonstrated greater variation in response to IR within HPV-positive cell lines compared to HPV-negative OPSCC cell lines, and suggested that HPV positivity was not consistently associated with sensitivity to IR. In the IR sensitive lines there was no evidence for greater p53 activity or differences in expression of genes associated with DNA repair pathways. The results obtained from the *in-vitro* models, may reflect the variation in treatment response of HPV-positive OPSCC patients in terms of general survival, as not all HPV-positive patients respond better to treatment. This study highlights the need for pre-clinical laboratory studies to incorporate a wide range of well-characterised cell lines to accurately mimic differences in response in the patient population.

# Bibliography


ABBAS, T. & DUTTA, A. 2011. CRL4Cdt2: master coordinator of cell cycle progression and genome stability. *Cell Cycle,* 10**,** 241-9.

ABBAS, T., SIVAPRASAD, U., TERAI, K., AMADOR, V., PAGANO, M. & DUTTA, A. 2008. PCNA-dependent regulation of p21 ubiquitylation and degradation via the CRL4Cdt2 ubiquitin ligase complex. *Genes Dev,* 22**,** 2496-506.

AI, L., STEPHENSON, K. K., LING, W., ZUO, C., MUKUNYADZI, P., SUEN, J. Y., HANNA, E. & FAN, C. Y. 2003. The p16 (CDKN2a/INK4a) tumor-suppressor gene in head and neck squamous cell carcinoma: a promoter methylation and protein expression study in 100 cases. *Mod Pathol,* 16**,** 944-50.

AKAGI, K., LI, J., BROUTIAN, T. R., PADILLA-NASH, H., XIAO, W., JIANG, B., ROCCO, J. W., TEKNOS, T. N., KUMAR, B., WANGSA, D., HE, D., RIED, T., SYMER, D. E. & GILLISON, M. L. 2014. Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res,* 24**,** 185-99.

ALLEN, C. T., LEWIS, J. S., JR., EL-MOFTY, S. K., HAUGHEY, B. H. & NUSSENBAUM, B. 2010. Human papillomavirus and oropharynx cancer: biology, detection and clinical implications. *Laryngoscope,* 120**,** 1756-72.

ANDERSEN, M. P., NELSON, Z. W., HETRICK, E. D. & GOTTSCHLING, D. E. 2008. A genetic screen for increased loss of heterozygosity in Saccharomyces cerevisiae. *Genetics,* 179**,** 1179-95.

ANDROUTSOPOULOS, V. P., TSATSAKIS, A. M. & SPANDIDOS, D. A. 2009. Cytochrome P450 CYP1A1: wider roles in cancer progression and prevention. *BMC Cancer,* 9**,** 187.

ANG, K. K., HARRIS, J., WHEELER, R., WEBER, R., ROSENTHAL, D. I., NGUYEN-TAN, P. F., WESTRA, W. H., CHUNG, C. H., JORDAN, R. C., LU, C., KIM, H., AXELROD, R., SILVERMAN, C. C., REDMOND, K. P. & GILLISON, M. L. 2010. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med,* 363**,** 24-35.

ARENZ, A., ZIEMANN, F., MAYER, C., WITTIG, A., DREFFKE, K., PREISING, S., WAGNER, S., KLUSSMANN, J. P., ENGENHART-CABILLIC, R. & WITTEKINDT, C. 2014. Increased radiosensitivity of HPV-positive head and neck cancer cell lines due to cell cycle dysregulation and induction of apoptosis. *Strahlenther Onkol,* 190**,** 839-46.

ARGIRIS, A., KARAMOUZIS, M. V., RABEN, D. & FERRIS, R. L. 2008. Head and neck cancer. *Lancet,* 371**,** 1695-709.

ASHCROFT, M., KUBBUTAT, M. H. & VOUSDEN, K. H. 1999. Regulation of p53 function and stability by phosphorylation. *Mol Cell Biol,* 19**,** 1751-8.

ATTNER, P., DU, J., NASMAN, A., HAMMARSTEDT, L., RAMQVIST, T., LINDHOLM, J., MARKLUND, L., DALIANIS, T. & MUNCK-WIKLAND, E. 2010. The role of human papillomavirus in the increased incidence of base of tongue cancer. *Int J Cancer,* 126**,** 2879-84.

BAAN, R., STRAIF, K., GROSSE, Y., SECRETAN, B., EL GHISSASSI, F., BOUVARD, V., ALTIERI, A., COGLIANO, V. & GROUP, W. H. O. I. A. F. R. O. C. M. W. 2007. Carcinogenicity of alcoholic beverages. *Lancet Oncol,* 8**,** 292-3.

BAKER, S. R. 1985. An in vivo model for squamous cell carcinoma of the head and neck. *Laryngoscope,* 95**,** 43-56.

BANATH, J. P., MACPHAIL, S. H. & OLIVE, P. L. 2004. Radiation sensitivity, H2AX phosphorylation, and kinetics of repair of DNA strand breaks in irradiated cervical cancer cell lines. *Cancer Res,* 64**,** 7144-9.

BASKAR, R., LEE, K. A., YEO, R. & YEOH, K. W. 2012. Cancer and radiation therapy: current advances and future directions. *Int J Med Sci,* 9**,** 193-9.

BECKERMAN, R. & PRIVES, C. 2010. Transcriptional regulation by p53. *Cold Spring Harb Perspect Biol,* 2**,** a000935.

BEGUM, S., CAO, D., GILLISON, M., ZAHURAK, M. & WESTRA, W. H. 2005. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clin Cancer Res,* 11**,** 5694-9.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological),* 57**,** 289-300.

BERGHOLZ, J. & XIAO, Z. X. 2012. Role of p63 in Development, Tumorigenesis and Cancer Progression. *Cancer Microenviron,* 5**,** 311-22.

BERNHARD, E. J., MAITY, A., MUSCHEL, R. J. & MCKENNA, W. G. 1995. Effects of ionizing radiation on cell cycle progression. A review. *Radiat Environ Biophys,* 34**,** 79-83.

BLAGOSKLONNY, M. V. 2002. P53: an ubiquitous target of anticancer drugs. *Int J Cancer,* 98**,** 161-6.

BOGAARDS, J. A., WALLINGA, J., BRAKENHOFF, R. H., MEIJER, C. J. & BERKHOF, J. 2015. Direct benefit of vaccinating boys along with girls against oncogenic human papillomavirus: bayesian evidence synthesis. *BMJ,* 350**,** h2016.

BOREL, F., LACROIX, F. B. & MARGOLIS, R. L. 2002. Prolonged arrest of mammalian cells at the G1/S boundary results in permanent S phase stasis. *J Cell Sci,* 115**,** 2829-38.

BOULET, G., HORVATH, C., VANDEN BROECK, D., SAHEBALI, S. & BOGERS, J. 2007. Human papillomavirus: E6 and E7 oncogenes. *Int J Biochem Cell Biol,* 39**,** 2006-11.

BRADFORD, C. R., ZHU, S., OGAWA, H., OGAWA, T., UBELL, M., NARAYAN, A., JOHNSON, G., WOLF, G. T., FISHER, S. G. & CAREY, T. E. 2003. P53 mutation correlates with cisplatin sensitivity in head and neck squamous cell carcinoma lines. *Head Neck,* 25**,** 654-61.

BRANZEI, D. & FOIANI, M. 2008. Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol,* 9**,** 297-308.

BRAUNSTEIN, S., BADURA, M. L., XI, Q., FORMENTI, S. C. & SCHNEIDER, R. J. 2009. Regulation of protein synthesis by ionizing radiation. *Mol Cell Biol,* 29**,** 5645-56.

BRENNAN, J. A., BOYLE, J. O., KOCH, W. M., GOODMAN, S. N., HRUBAN, R. H., EBY, Y. J., COUCH, M. J., FORASTIERE, A. A. & SIDRANSKY, D. 1995. Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck. *N Engl J Med,* 332**,** 712-7.

BRENNER, J. C., GRAHAM, M. P., KUMAR, B., SAUNDERS, L. M., KUPFER, R., LYONS, R. H., BRADFORD, C. R. & CAREY, T. E. 2010. Genotyping of 73 UM-SCC head and neck squamous cell carcinoma cell lines. *Head Neck,* 32**,** 417-26.

BROUSTAS, C. G. & LIEBERMAN, H. B. 2014. DNA damage response genes and the development of cancer metastasis. *Radiat Res,* 181**,** 111-30.

BRYANT, D., ONIONS, T., RAYBOULD, R., FLYNN, A., TRISTRAM, A., MEYRICK, S., GILES, P., ASHELFORD, K., HIBBITTS, S., FIANDER, A. & POWELL, N. 2014a. mRNA sequencing of novel cell lines from human papillomavirus type-16 related vulval intraepithelial neoplasia: consequences of expression of HPV16 E4 and E5. *J Med Virol,* 86**,** 1534-41.

BRYANT, D., ONIONS, T., RAYBOULD, R., JONES, S., TRISTRAM, A., HIBBITTS, S., FIANDER, A. & POWELL, N. 2014b. Increased methylation of Human Papillomavirus type 16 DNA correlates with viral integration in Vulval Intraepithelial Neoplasia. *J Clin Virol,* 61**,** 393-9.

BUCK, C. B., CHENG, N., THOMPSON, C. D., LOWY, D. R., STEVEN, A. C., SCHILLER, J. T. & TRUS, B. L. 2008. Arrangement of L2 within the papillomavirus capsid. *J Virol,* 82**,** 5190-7.

BURK, R. D., HARARI, A. & CHEN, Z. 2013. Human papillomavirus genome variants. *Virology,* 445**,** 232-43.

BURKE, S. C., SMITH, K. V., SHARMIN, S. & WINKELMAN, C. 2014. Prevalence of risk factors related to head and neck squamous cell carcinoma (HNSCC) among college students. *Cancer and Oncology Research,* 2**,** 7-16.

BUSCH, C. J., BECKER, B., KRIEGS, M., GATZEMEIER, F., KRUGER, K., MOCKELMANN, N., FRITZ, G., PETERSEN, C., KNECHT, R., ROTHKAMM, K. & RIECKMANN, T. 2016. Similar cisplatin sensitivity of HPV-positive and -negative HNSCC cell lines. *Oncotarget,* 7**,** 35832-35842.

CARSON, A. & KHAN, S. A. 2006. Characterization of transcription factor binding to human papillomavirus type 16 DNA during cellular differentiation. *J Virol,* 80**,** 4356-62.

CHANG, Y., MOORE, P. S. & WEISS, R. A. 2017. Human oncogenic viruses: nature and discovery. *Philos Trans R Soc Lond B Biol Sci,* 372.

CHEN, A. Y., SCHRAG, N. M., HALPERN, M. T. & WARD, E. M. 2007. The impact of health insurance status on stage at diagnosis of oropharyngeal cancer. *Cancer,* 110**,** 395-402.

CHEN, Z., SCHIFFMAN, M., HERRERO, R., DESALLE, R., ANASTOS, K., SEGONDY, M., SAHASRABUDDHE, V. V., GRAVITT, P. E., HSING, A. W. & BURK, R. D. 2011. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One,* 6**,** e20183.

CHU, A., GENDEN, E., POSNER, M. & SIKORA, A. 2013. A patient-centered approach to counseling patients with head and neck cancer undergoing human papillomavirus testing: a clinician's guide. *Oncologist,* 18**,** 180-9.

CORNET, I., GHEIT, T., FRANCESCHI, S., VIGNAT, J., BURK, R. D., SYLLA, B. S., TOMMASINO, M., CLIFFORD, G. M. & GROUP, I. H. V. S. 2012. Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol,* 86**,** 6855-61.

CUBIE, H. A. 2013. Diseases associated with human papillomavirus infection. *Virology,* 445**,** 21-34.

D'SOUZA, G., CULLEN, K., BOWIE, J., THORPE, R. & FAKHRY, C. 2014. Differences in oral sexual behaviors by gender, age, and race explain observed differences in prevalence of oral human papillomavirus infection. *PLoS One,* 9**,** e86023.

D'SOUZA, G., KREIMER, A. R., VISCIDI, R., PAWLITA, M., FAKHRY, C., KOCH, W. M., WESTRA, W. H. & GILLISON, M. L. 2007. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med,* 356**,** 1944-56.

D'SOUZA, G., ZHANG, H. H., D'SOUZA, W. D., MEYER, R. R. & GILLISON, M. L. 2010. Moderate predictive value of demographic and behavioral characteristics for a diagnosis of HPV16-positive and HPV16-negative head and neck cancer. *Oral Oncol,* 46**,** 100-4.

DANDRI, M., BURDA, M. R., BURKLE, A., ZUCKERMAN, D. M., WILL, H., ROGLER, C. E., GRETEN, H. & PETERSEN, J. 2002. Increase in de novo HBV DNA integrations in response to oxidative DNA damage or inhibition of poly(ADP-ribosyl)ation. *Hepatology,* 35**,** 217-23.

DAVIS, A. J. & CHEN, D. J. 2013. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res,* 2**,** 130-143.

DAY, P. M., LOWY, D. R. & SCHILLER, J. T. 2008. Heparan sulfate-independent cell binding and infection with furin-precleaved papillomavirus capsids. *J Virol,* 82**,** 12565-8.

DAYYANI, F., ETZEL, C. J., LIU, M., HO, C. H., LIPPMAN, S. M. & TSAO, A. S. 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 Oncol,* 2**,** 15.

DE RODA HUSMAN, A. M., SNIJDERS, P. J., STEL, H. V., VAN DEN BRULE, A. J., MEIJER, C. J. & WALBOOMERS, J. M. 1995. Processing of long-stored archival cervical smears for human papillomavirus detection by the polymerase chain reaction. *Br J Cancer,* 72**,** 412-7.

DE VILLIERS, E. M. 2013. Cross-roads in the classification of papillomaviruses. *Virology,* 445**,** 2-10.

DE VILLIERS, E. M., FAUQUET, C., BROKER, T. R., BERNARD, H. U. & ZUR HAUSEN, H. 2004. Classification of papillomaviruses. *Virology,* 324**,** 17-27.

DE VUYST, H., CLIFFORD, G. M., NASCIMENTO, M. C., MADELEINE, M. M. & FRANCESCHI, S. 2009. Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: a meta-analysis. *Int J Cancer,* 124**,** 1626-36.

DEFFIE, A., WU, H., REINKE, V. & LOZANO, G. 1993. The tumor suppressor p53 regulates its own transcription. *Mol Cell Biol,* 13**,** 3415-23.

DESCAMPS, G., WATTIEZ, R. & SAUSSEZ, S. 2014. Proteomic study of HPV-positive head and neck cancers: preliminary results. *Biomed Res Int,* 2014**,** 430906.

DESCHLER, D. G., RICHMON, J. D., KHARIWALA, S. S., FERRIS, R. L. & WANG, M. B. 2014. The "new" head and neck cancer patient-young, nonsmoker, nondrinker, and HPV positive: evaluation. *Otolaryngol Head Neck Surg,* 151**,** 375-80.

DESOUKY, O., DING, N. & ZHOU, G. 2015. Targeted and non-targeted effects of ionizing radiation. *Journal of Radiation Research and Applied Sciences,* 8**,** 247-254.

DI NICOLANTONIO, F., MERCER, S. J., KNIGHT, L. A., GABRIEL, F. G., WHITEHOUSE, P. A., SHARMA, S., FERNANDO, A., GLAYSHER, S., DI PALMA, S., JOHNSON, P., SOMERS, S. S., TOH, S., HIGGINS, B., LAMONT, A., GULLIFORD, T., HURREN, J., YIANGOU, C. & CREE, I. A. 2005. Cancer cell adaptation to chemotherapy. *BMC Cancer,* 5**,** 78.

DONA, M. G., PAOLINI, F., BENEVOLO, M., VOCATURO, A., LATINI, A., GIGLIO, A., VENUTI, A. & GIULIANI, M. 2013. Identification of episomal human papillomavirus and other DNA viruses in cytological anal samples of HIV-uninfected men who have sex with men. *PLoS One,* 8**,** e72228.

DOORBAR, J. 2005. The papillomavirus life cycle. *J Clin Virol,* 32 Suppl 1**,** S7-15.

DOORBAR, J., EGAWA, N., GRIFFIN, H., KRANJEC, C. & MURAKAMI, I. 2015. Human papillomavirus molecular biology and disease association. *Rev Med Virol,* 25 Suppl 1**,** 2-23.

DOORBAR, J., QUINT, W., BANKS, L., BRAVO, I. G., STOLER, M., BROKER, T. R. & STANLEY, M. A. 2012. The biology and life-cycle of human papillomaviruses. *Vaccine,* 30 Suppl 5**,** F55-70.

DRACOPOLI, N. C. & FOGH, J. 1983. Loss of heterozygosity in cultured human tumor cell lines. *J Natl Cancer Inst,* 70**,** 83-7.

DREXLER, H. G. & UPHOFF, C. C. 2002. Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention. *Cytotechnology,* 39**,** 75-90.

DROST, J. & CLEVERS, H. 2018. Organoids in cancer research. *Nat Rev Cancer,* 18**,** 407-418.

DUENSING, S., DUENSING, A., CRUM, C. P. & MUNGER, K. 2001. Human papillomavirus type 16 E7 oncoprotein-induced abnormal centrosome synthesis is an early event in the evolving malignant phenotype. *Cancer Res,* 61**,** 2356-60.

DUENSING, S., LEE, L. Y., DUENSING, A., BASILE, J., PIBOONNIYOM, S., GONZALEZ, S., CRUM, C. P. & MUNGER, K. 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. *Proc Natl Acad Sci U S A,* 97**,** 10002-7.

DUENSING, S. & MUNGER, K. 2004. Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *Int J Cancer,* 109**,** 157-62.

DURST, M., GISSMANN, L., IKENBERG, H. & ZUR HAUSEN, H. 1983. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc Natl Acad Sci U S A,* 80**,** 3812-5.

EGAWA, N., EGAWA, K., GRIFFIN, H. & DOORBAR, J. 2015. Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia. *Viruses,* 7**,** 3863-90.

EL-DEIRY, W. S., HARPER, J. W., O'CONNOR, P. M., VELCULESCU, V. E., CANMAN, C. E., JACKMAN, J., PIETENPOL, J. A., BURRELL, M., HILL, D. E., WANG, Y. & ET AL. 1994. WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer Res,* 54**,** 1169-74.

ELFSTROM, K. M., LAZZARATO, F., FRANCESCHI, S., DILLNER, J. & BAUSSANO, I. 2016. Human Papillomavirus Vaccination of Boys and Extended Catch-up Vaccination: Effects on the Resilience of Programs. *J Infect Dis,* 213**,** 199-205.

EVANS, M., NEWCOMBE, R., FIANDER, A., POWELL, J., ROLLES, M., THAVARAJ, S., ROBINSON, M. & POWELL, N. 2013. Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population. *BMC Cancer,* 13**,** 220.

EWA, B. & DANUTA, M. S. 2017. Polycyclic aromatic hydrocarbons and PAH-related DNA adducts. *J Appl Genet,* 58**,** 321-330.

FAKHRY, C. & GILLISON, M. L. 2006. Clinical implications of human papillomavirus in head and neck cancers. *J Clin Oncol,* 24**,** 2606-11.

FALCON, S. & GENTLEMAN, R. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics,* 23**,** 257-8.

FAN, J., LEE, H. O., LEE, S., RYU, D. E., LEE, S., XUE, C., KIM, S. J., KIM, K., BARKAS, N., PARK, P. J., PARK, W. Y. & KHARCHENKO, P. V. 2018. Linking transcriptional and genetic tumor heterogeneity through allele analysis of single-cell RNA-seq data. *Genome Res,* 28**,** 1217-1227.

FERRIS, R. L., MARTINEZ, I., SIRIANNI, N., WANG, J., LOPEZ-ALBAITERO, A., GOLLIN, S. M., JOHNSON, J. T. & KHAN, S. 2005. Human papillomavirus-16 associated squamous cell carcinoma of the head and neck (SCCHN): a natural disease model provides insights into viral carcinogenesis. *Eur J Cancer,* 41**,** 807-15.

FLORES, E. R., ALLEN-HOFFMANN, B. L., LEE, D., SATTLER, C. A. & LAMBERT, P. F. 1999. Establishment of the human papillomavirus type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line. *Virology,* 262**,** 344-54.

FONTECHA, N., BASARAS, M., ARRESE, E., HERNAEZ, S., ANDIA, D. & CISTERNA, R. 2015. Human Papillomavirus 16 Variants May Be Identified by E6 Gene Analysis. *Intervirology,* 58**,** 143-8.

FRANCESCHI, S., COMBES, J. D., DALSTEIN, V., CAUDROY, S., CLIFFORD, G., GHEIT, T., TOMMASINO, M., CLAVEL, C., LACAU ST GUILY, J., BIREMBAUT, P., STUDY OF NATURAL HISTORY OF HUMAN PAPILLOMAVIRUS, I. & PRECANCEROUS LESIONS IN THE, T. 2015. Deep brush-based cytology in tonsils resected for benign diseases. *Int J Cancer,* 137**,** 2994-9.

FRANCESCHI, S., TALAMINI, R., BARRA, S., BARON, A. E., NEGRI, E., BIDOLI, E., SERRAINO, D. & LA VECCHIA, C. 1990. Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy. *Cancer Res,* 50**,** 6502-7.

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

FREEMAN, J. A. & ESPINOSA, J. M. 2013. The impact of post-transcriptional regulation in the p53 network. *Brief Funct Genomics,* 12**,** 46-57.

FRIEDBERG, E. C., MCDANIEL, L. D. & SCHULTZ, R. A. 2004. The role of endogenous and exogenous DNA damage and mutagenesis. *Curr Opin Genet Dev,* 14**,** 5-10.

FRIEDL, F., KIMURA, I., OSATO, T. & ITO, Y. 1970. Studies on a new human cell line (SiHa) derived from carcinoma of uterus. I. Its establishment and morphology. *Proc Soc Exp Biol Med,* 135**,** 543-5.

GALANOS, P., VOUGAS, K., WALTER, D., POLYZOS, A., MAYA-MENDOZA, A., HAAGENSEN, E. J., KOKKALIS, A., ROUMELIOTI, F. M., GAGOS, S., TZETIS, M., CANOVAS, B., IGEA, A., AHUJA, A. K., ZELLWEGER, R., HAVAKI, S., KANAVAKIS, E., KLETSAS, D., RONINSON, I. B., GARBIS, S. D., LOPES, M., NEBREDA, A., THANOS, D., BLOW, J. J., TOWNSEND, P., SORENSEN, C. S., BARTEK, J. & GORGOULIS, V. G. 2016. Chronic p53-independent p21 expression causes genomic instability by deregulating replication licensing. *Nat Cell Biol,* 18**,** 777-89.

GALLAGHER, K. E., LAMONTAGNE, D. S. & WATSON-JONES, D. 2018. Status of HPV vaccine introduction and barriers to country uptake. *Vaccine,* 36**,** 4761-4767.

GELBOIN, H. V. 1980. Benzo[alpha]pyrene metabolism, activation and carcinogenesis: role and regulation of mixed-function oxidases and related enzymes. *Physiol Rev,* 60**,** 1107-66.

GIARRE, M., CALDEIRA, S., MALANCHI, I., CICCOLINI, F., LEAO, M. J. & TOMMASINO, M. 2001. Induction of pRb degradation by the human papillomavirus type 16 E7 protein is essential to efficiently overcome p16INK4a-imposed G1 cell cycle Arrest. *J Virol,* 75**,** 4705-12.

GILLESPIE, K. A., MEHTA, K. P., LAIMINS, L. A. & MOODY, C. A. 2012. Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. *J Virol,* 86**,** 9520-6.

GILLET, J. P., VARMA, S. & GOTTESMAN, M. M. 2013. The clinical relevance of cancer cell lines. *J Natl Cancer Inst,* 105**,** 452-8.

GILLISON, M. L., BROUTIAN, T., PICKARD, R. K., TONG, Z. Y., XIAO, W., KAHLE, L., GRAUBARD, B. I. & CHATURVEDI, A. K. 2012a. Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA,* 307**,** 693-703.

GILLISON, M. L., CHATURVEDI, A. K., ANDERSON, W. F. & FAKHRY, C. 2015. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J Clin Oncol,* 33**,** 3235-42.

GILLISON, M. L. & 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. *Curr Opin Oncol,* 13**,** 183-8.

GILLISON, M. L., ZHANG, Q., JORDAN, R., XIAO, W., WESTRA, W. H., TROTTI, A., SPENCER, S., HARRIS, J., CHUNG, C. H. & ANG, K. K. 2012b. Tobacco smoking and increased risk of death and progression for patients with p16-positive and p16-negative oropharyngeal cancer. *J Clin Oncol,* 30**,** 2102-11.

GISSMANN, L., DEVILLIERS, E. M. & ZUR HAUSEN, H. 1982a. Analysis of human genital warts (condylomata acuminata) and other genital tumors for human papillomavirus type 6 DNA. *Int J Cancer,* 29**,** 143-6.

GISSMANN, L., DIEHL, V., SCHULTZ-COULON, H. J. & ZUR HAUSEN, H. 1982b. Molecular cloning and characterization of human papilloma virus DNA derived from a laryngeal papilloma. *J Virol,* 44**,** 393-400.

GOLDENBERG, D., BEGUM, S., WESTRA, W. H., KHAN, Z., SCIUBBA, J., PAI, S. I., CALIFANO, J. A., TUFANO, R. P. & KOCH, W. M. 2008. Cystic lymph node metastasis in patients with head and neck cancer: An HPV-associated phenomenon. *Head Neck,* 30**,** 898-903.

GUERRERO-PRESTON, R., MICHAILIDI, C., MARCHIONNI, L., PICKERING, C. R., FREDERICK, M. J., MYERS, J. N., YEGNASUBRAMANIAN, S., HADAR, T., NOORDHUIS, M. G., ZIZKOVA, V., FERTIG, E., AGRAWAL, N., WESTRA, W., KOCH, W., CALIFANO, J., VELCULESCU, V. E. & SIDRANSKY, D. 2014. Key tumor suppressor genes inactivated by "greater promoter" methylation and somatic mutations in head and neck cancer. *Epigenetics,* 9**,** 1031-46.

GUMUS, Z. H., DU, B., KACKER, A., BOYLE, J. O., BOCKER, J. M., MUKHERJEE, P., SUBBARAMAIAH, K., DANNENBERG, A. J. & WEINSTEIN, H. 2008. Effects of tobacco smoke on gene expression and cellular pathways in a cellular model of oral leukoplakia. *Cancer Prev Res (Phila),* 1**,** 100-11.

HAFKAMP, H. C., MOOREN, J. J., CLAESSEN, S. M., KLINGENBERG, B., VOOGD, A. C., BOT, F. J., KLUSSMANN, J. P., HOPMAN, A. H., MANNI, J. J., KREMER, B., RAMAEKERS, F. C. & SPEEL, E. J. 2009. P21 Cip1/WAF1 expression is strongly associated with HPV-positive tonsillar carcinoma and a favorable prognosis. *Mod Pathol,* 22**,** 686-98.

HAKENEWERTH, A. M., MILLIKAN, R. C., RUSYN, I., HERRING, A. H., NORTH, K. E., BARNHOLTZ-SLOAN, J. S., FUNKHOUSER, W. F., WEISSLER, M. C. & OLSHAN, A. F. 2011. Joint effects of alcohol consumption and polymorphisms in alcohol and oxidative stress metabolism genes on risk of head and neck cancer. *Cancer Epidemiol Biomarkers Prev,* 20**,** 2438-49.

HALBEISEN, R. E., GALGANO, A., SCHERRER, T. & GERBER, A. P. 2008. Post-transcriptional gene regulation: from genome-wide studies to principles. *Cell Mol Life Sci,* 65**,** 798-813.

HAN, J. Y., LEE, G. K., JANG, D. H., LEE, S. Y. & LEE, J. S. 2008. Association of p53 codon 72 polymorphism and MDM2 SNP309 with clinical outcome of advanced nonsmall cell lung cancer. *Cancer,* 113**,** 799-807.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HARPER, D. M., FRANCO, E. L., WHEELER, C. M., MOSCICKI, A. B., ROMANOWSKI, B., ROTELI-MARTINS, C. M., JENKINS, D., SCHUIND, A., COSTA CLEMENS, S. A., DUBIN, G. & GROUP, H. P. V. V. S. 2006. Sustained efficacy up to 4.5 years of a bivalent L1 virus-like particle vaccine against human papillomavirus types 16 and 18: follow-up from a randomised control trial. *Lancet,* 367**,** 1247-55.

HARRISON, L. B., SESSIONS, R. B. & HONG, W. K. 2009. *Head and neck cancer : a multidisciplinary approach,* Philadelphia, Pa. ; London, Lippincott Williams & Wilkins.

HASHIM, D., SARTORI, S., BRENNAN, P., CURADO, M. P., WUNSCH-FILHO, V., DIVARIS, K., OLSHAN, A. F., ZEVALLOS, J. P., WINN, D. M., FRANCESCHI, S., CASTELLSAGUE, X., LISSOWSKA, J., RUDNAI, P., MATSUO, K., MORGENSTERN, H., CHEN, C., VAUGHAN, T. L., HOFMANN, J. N., D'SOUZA, G., HADDAD, R. I., WU, H., LEE, Y. C., HASHIBE, M., VECCHIA, C. L. & BOFFETTA, P. 2016. The role of oral hygiene in head and neck cancer: results from International Head and Neck Cancer Epidemiology (INHANCE) consortium. *Ann Oncol,* 27**,** 1619-25.

HAUPT, S., BERGER, M., GOLDBERG, Z. & HAUPT, Y. 2003. Apoptosis - the p53 network. *J Cell Sci,* 116**,** 4077-85.

HERMANS, R. & LENZ, M. 1996. Imaging of the oropharynx and oral cavity. Part I: Normal anatomy. *Eur Radiol,* 6**,** 362-8.

HESS, J., ANGEL, P. & SCHORPP-KISTNER, M. 2004. AP-1 subunits: quarrel and harmony among siblings. *J Cell Sci,* 117**,** 5965-73.

HO, K. F., FOWLER, J. F., SYKES, A. J., YAP, B. K., LEE, L. W. & SLEVIN, N. J. 2009. IMRT dose fractionation for head and neck cancer: variation in current approaches will make standardisation difficult. *Acta Oncol,* 48**,** 431-9.

HORVATH, C. A., BOULET, G. A., RENOUX, V. M., DELVENNE, P. O. & BOGERS, J. P. 2010. Mechanisms of cell entry by human papillomaviruses: an overview. *Virol J,* 7**,** 11.

JABBAR, S., STRATI, K., SHIN, M. K., PITOT, H. C. & LAMBERT, P. F. 2010. Human papillomavirus type 16 E6 and E7 oncoproteins act synergistically to cause head and neck cancer in mice. *Virology,* 407**,** 60-7.

JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature,* 461**,** 1071-8.

JANG, N. Y., KIM, D. H., CHO, B. J., CHOI, E. J., LEE, J. S., WU, H. G., CHIE, E. K. & KIM, I. A. 2015. Radiosensitization with combined use of olaparib and PI-103 in triple-negative breast cancer. *BMC Cancer,* 15**,** 89.

JEON, S., ALLEN-HOFFMANN, B. L. & LAMBERT, P. F. 1995. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J Virol,* 69**,** 2989-97.

KADAJA, M., SILLA, T., USTAV, E. & USTAV, M. 2009. Papillomavirus DNA replication - from initiation to genomic instability. *Virology,* 384**,** 360-8.

KAJITANI, N., SATSUKA, A., KAWATE, A. & SAKAI, H. 2012. Productive Lifecycle of Human Papillomaviruses that Depends Upon Squamous Epithelial Differentiation. *Front Microbiol,* 3**,** 152.

KAMMER, C., WARTHORST, U., TORREZ-MARTINEZ, N., WHEELER, C. M. & PFISTER, H. 2000. Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *J Gen Virol,* 81**,** 1975-81.

KASAI, F., HIRAYAMA, N., OZAWA, M., IEMURA, M. & KOHARA, A. 2016. Changes of heterogeneous cell populations in the Ishikawa cell line during long-term culture: Proposal for an in vitro clonal evolution model of tumor cells. *Genomics,* 107**,** 259-66.

KATZ, Y., WANG, E. T., SILTERRA, J., SCHWARTZ, S., WONG, B., THORVALDSDOTTIR, H., ROBINSON, J. T., MESIROV, J. P., AIROLDI, E. M. & BURGE, C. B. 2015. Quantitative visualization of alternative exon expression from RNA-seq data. *Bioinformatics,* 31**,** 2400-2.

KIMPLE, R. J., SMITH, M. A., BLITZER, G. C., TORRES, A. D., MARTIN, J. A., YANG, R. Z., PEET, C. R., LORENZ, L. D., NICKEL, K. P., KLINGELHUTZ, A. J., LAMBERT, P. F. & HARARI, P. M. 2013. Enhanced radiation sensitivity in HPV-positive head and neck cancer. *Cancer Res,* 73**,** 4791-800.

KING, E. V., OTTENSMEIER, C. H. & THOMAS, G. J. 2014. The immune response in HPV(+) oropharyngeal cancer. *Oncoimmunology,* 3**,** e27254.

KLINGELHUTZ, A. J., FOSTER, S. A. & MCDOUGALL, J. K. 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature,* 380**,** 79-82.

KMIETOWICZ, Z. 2018. Boys in England to get HPV vaccine from next year. *BMJ,* 362**,** k3237.

KODACK, D. P., FARAGO, A. F., DASTUR, A., HELD, M. A., DARDAEI, L., FRIBOULET, L., VON FLOTOW, F., DAMON, L. J., LEE, D., PARKS, M., DICECCA, R., GREENBERG, M., KATTERMANN, K. E., RILEY, A. K., FINTELMANN, F. J., RIZZO, C., PIOTROWSKA, Z., SHAW, A. T., GAINOR, J. F., SEQUIST, L. V., NIEDERST, M. J., ENGELMAN, J. A. & BENES, C. H. 2017. Primary Patient-Derived Cancer Cells and Their Potential for Personalized Cancer Patient Care. *Cell Rep,* 21**,** 3298-3309.

KOSHIOL, J., HILDESHEIM, A., GONZALEZ, P., BRATTI, M. C., PORRAS, C., SCHIFFMAN, M., HERRERO, R., RODRIGUEZ, A. C., WACHOLDER, S., YEAGER, M., CHANOCK, S. J., BURK, R. D. & WANG, S. S. 2009. Common genetic variation in TP53 and risk of human papillomavirus persistence and progression to CIN3/cancer revisited. *Cancer Epidemiol Biomarkers Prev,* 18**,** 1631-7.

KRZYWINSKI, M., SCHEIN, J., BIROL, I., CONNORS, J., GASCOYNE, R., HORSMAN, D., JONES, S. J. & MARRA, M. A. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res,* 19**,** 1639-45.

KUBBUTAT, M. H., JONES, S. N. & VOUSDEN, K. H. 1997. Regulation of p53 stability by Mdm2. *Nature,* 387**,** 299-303.

KUKURBA, K. R. & MONTGOMERY, S. B. 2015. RNA Sequencing and Analysis. *Cold Spring Harb Protoc,* 2015**,** 951-69.

LAHTZ, C., BATES, S. E., JIANG, Y., LI, A. X., WU, X., HAHN, M. A. & PFEIFER, G. P. 2012. Gamma irradiation does not induce detectable changes in DNA methylation directly following exposure of human cells. *PLoS One,* 7**,** e44858.

LEE, Y. C., BOFFETTA, P., STURGIS, E. M., WEI, Q., ZHANG, Z. F., MUSCAT, J., LAZARUS, P., MATOS, E., HAYES, R. B., WINN, D. M., ZARIDZE, D., WUNSCH-FILHO, V., ELUF-NETO, J., KOIFMAN, S., MATES, D., CURADO, M. P., MENEZES, A., FERNANDEZ, L., DAUDT, A. W., SZESZENIA-DABROWSKA, N., FABIANOVA, E., RUDNAI, P., FERRO, G., BERTHILLER, J., BRENNAN, P. & HASHIBE, M. 2008. Involuntary smoking and head and neck cancer risk: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiol Biomarkers Prev,* 17**,** 1974-81.

LEWIS, J. S., JR., THORSTAD, W. L., CHERNOCK, R. D., HAUGHEY, B. H., YIP, J. H., ZHANG, Q. & EL-MOFTY, S. K. 2010. p16 positive oropharyngeal squamous cell carcinoma:an entity with a favorable prognosis regardless of tumor HPV status. *Am J Surg Pathol,* 34**,** 1088-96.

LI, G. M. 2008. Mechanisms and functions of DNA mismatch repair. *Cell Res,* 18**,** 85-98.

LIAO, Y., SMYTH, G. K. & SHI, W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics,* 30**,** 923-30.

LIEBER, M. R. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem,* 79**,** 181-211.

LILL, C., BACHTIARY, B., SELZER, E., MITTLBOECK, M. & THURNHER, D. 2017. A 5year update of patients with HPV positive versus negative oropharyngeal cancer after radiochemotherapy in Austria. *Wien Klin Wochenschr,* 129**,** 398-403.

LIN, B. Y., MAKHOV, A. M., GRIFFITH, J. D., BROKER, T. R. & CHOW, L. T. 2002. Chaperone proteins abrogate inhibition of the human papillomavirus (HPV) E1 replicative helicase by the HPV E2 protein. *Mol Cell Biol,* 22**,** 6592-604.

LIN, C. J., GRANDIS, J. R., CAREY, T. E., GOLLIN, S. M., WHITESIDE, T. L., KOCH, W. M., FERRIS, R. L. & LAI, S. Y. 2007. Head and neck squamous cell carcinoma cell lines: established models and rationale for selection. *Head Neck,* 29**,** 163-88.

LINDENBERGH-VAN DER PLAS, M., BRAKENHOFF, R. H., KUIK, D. J., BUIJZE, M., BLOEMENA, E., SNIJDERS, P. J., LEEMANS, C. R. & BRAAKHUIS, B. J. 2011. Prognostic significance of truncating TP53 mutations in head and neck squamous cell carcinoma. *Clin Cancer Res,* 17**,** 3733-41.

LISSOWSKA, J., PILARSKA, A., PILARSKI, P., SAMOLCZYK-WANYURA, D., PIEKARCZYK, J., BARDIN-MIKOLLAJCZAK, A., ZATONSKI, W., HERRERO, R., MUNOZ, N. & FRANCESCHI, S. 2003. Smoking, alcohol, diet, dentition and sexual practices in the epidemiology of oral cancer in Poland. *Eur J Cancer Prev,* 12**,** 25-33.

LIU, X., DAKIC, A., CHEN, R., DISBROW, G. L., ZHANG, Y., DAI, Y. & SCHLEGEL, R. 2008. Cell-restricted immortalization by human papillomavirus correlates with telomerase activation and engagement of the hTERT promoter by Myc. *J Virol,* 82**,** 11568-76.

LIU, X., DISBROW, G. L., YUAN, H., TOMAIC, V. & SCHLEGEL, R. 2007. Myc and human papillomavirus type 16 E7 genes cooperate to immortalize human keratinocytes. *J Virol,* 81**,** 12689-95.

LIU, X., YUAN, H., FU, B., DISBROW, G. L., APOLINARIO, T., TOMAIC, V., KELLEY, M. L., BAKER, C. C., HUIBREGTSE, J. & SCHLEGEL, R. 2005. The E6AP ubiquitin ligase is required for transactivation of the hTERT promoter by the human papillomavirus E6 oncoprotein. *J Biol Chem,* 280**,** 10807-16.

LOMAX, M. E., FOLKES, L. K. & O'NEILL, P. 2013. Biological consequences of radiation-induced DNA damage: relevance to radiotherapy. *Clin Oncol (R Coll Radiol),* 25**,** 578-85.

LONGWORTH, M. S. & LAIMINS, L. A. 2004. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev,* 68**,** 362-72.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

MADISON, K. C. 2003. Barrier function of the skin: "la raison d'etre" of the epidermis. *J Invest Dermatol,* 121**,** 231-41.

MANDIC, R., SCHAMBERGER, C. J., MULLER, J. F., GEYER, M., ZHU, L., CAREY, T. E., GRENMAN, R., DUNNE, A. A. & WERNER, J. A. 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. *Clin Cancer Res,* 11**,** 6845-52.

MATA, J., MARGUERAT, S. & BAHLER, J. 2005. Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci,* 30**,** 506-14.

MAXWELL, J. H., KUMAR, B., FENG, F. Y., WORDEN, F. P., LEE, J. S., EISBRUCH, A., WOLF, G. T., PRINCE, M. E., MOYER, J. S., TEKNOS, T. N., CHEPEHA, D. B., MCHUGH, J. B., URBA, S. G., STOERKER, J., WALLINE, H. M., KURNIT, D. M., CORDELL, K. G., DAVIS, S. J., WARD, P. D., BRADFORD, C. R. & CAREY, T. E. 2010. Tobacco use in human papillomavirus-positive advanced oropharynx cancer patients related to increased risk of distant metastases and tumor recurrence. *Clin Cancer Res,* 16**,** 1226-35.

MCBRIDE, A. A. & WARBURTON, A. 2017. The role of integration in oncogenic progression of HPV-associated cancers. *PLoS Pathog,* 13**,** e1006211.

MEHANNA, H., PALERI, V., WEST, C. M. & NUTTING, C. 2010. Head and neck cancer--Part 1: Epidemiology, presentation, and prevention. *BMJ,* 341**,** c4684.

MIDDLETON, K., PEH, W., SOUTHERN, S., GRIFFIN, H., SOTLAR, K., NAKAHARA, T., EL-SHERIF, A., MORRIS, L., SETH, R., HIBMA, M., JENKINS, D., LAMBERT, P., COLEMAN, N. & DOORBAR, J. 2003. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *J Virol,* 77**,** 10186-201.

MILLER, C. J., KASSEM, H. S., PEPPER, S. D., HEY, Y., WARD, T. H. & MARGISON, G. P. 2003. Mycoplasma infection significantly alters microarray gene expression profiles. *Biotechniques,* 35**,** 812-4.

MITRA, A., MISHRA, L. & LI, S. 2013. Technologies for deriving primary tumor cells for use in personalized cancer therapy. *Trends Biotechnol,* 31**,** 347-54.

MOLIJN, A., KLETER, B., QUINT, W. & VAN DOORN, L. J. 2005. Molecular diagnosis of human papillomavirus (HPV) infections. *J Clin Virol,* 32 Suppl 1**,** S43-51.

MOMAND, J., ZAMBETTI, G. P., OLSON, D. C., GEORGE, D. & LEVINE, A. J. 1992. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell,* 69**,** 1237-45.

MOORE, P. S. & CHANG, Y. 2010. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nat Rev Cancer,* 10**,** 878-89.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods,* 5**,** 621-8.

MOSCICKI, A. B., MA, Y., WIBBELSMAN, C., DARRAGH, T. M., POWERS, A., FARHAT, S. & SHIBOSKI, S. 2010. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. *Obstet Gynecol,* 116**,** 1373-80.

MUNOZ, N., BOSCH, F. X., DE SANJOSE, S., HERRERO, R., CASTELLSAGUE, X., SHAH, K. V., SNIJDERS, P. J., MEIJER, C. J. & INTERNATIONAL AGENCY FOR RESEARCH ON CANCER MULTICENTER CERVICAL CANCER STUDY, G. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med,* 348**,** 518-27.

MUNRO, J., STOTT, F. J., VOUSDEN, K. H., PETERS, G. & PARKINSON, E. K. 1999. Role of the alternative INK4A proteins in human keratinocyte senescence: evidence for the specific inactivation of p16INK4A upon immortalization. *Cancer Res,* 59**,** 2516-21.

MUNSHI, A., HOBBS, M. & MEYN, R. E. 2005. Clonogenic cell survival assay. *Methods Mol Med,* 110**,** 21-8.

NARISAWA-SAITO, M. & KIYONO, T. 2007. Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: roles of E6 and E7 proteins. *Cancer Sci,* 98**,** 1505-11.

NASMAN, A., ATTNER, P., HAMMARSTEDT, L., DU, J., ERIKSSON, M., GIRAUD, G., AHRLUND-RICHTER, S., MARKLUND, L., ROMANITAN, M., LINDQUIST, D., RAMQVIST, T., LINDHOLM, J., SPAREN, P., YE, W., DAHLSTRAND, H., MUNCK-WIKLAND, E. & DALIANIS, T. 2009. Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma? *Int J Cancer,* 125**,** 362-6.

NASMAN, A., BERSANI, C., LINDQUIST, D., DU, J., RAMQVIST, T. & DALIANIS, T. 2017. Human Papillomavirus and Potentially Relevant Biomarkers in Tonsillar and Base of Tongue Squamous Cell Carcinoma. *Anticancer Res,* 37**,** 5319-5328.

NASMAN, A., NORDFORS, C., HOLZHAUSER, S., VLASTOS, A., TERTIPIS, N., HAMMAR, U., HAMMARSTEDT-NORDENVALL, L., MARKLUND, L., MUNCK-WIKLAND, E., RAMQVIST, T., BOTTAI, M. & DALIANIS, T. 2015. Incidence of human papillomavirus positive tonsillar and base of tongue carcinoma: a stabilisation of an epidemic of viral induced carcinoma? *Eur J Cancer,* 51**,** 55-61.

NAUTA, I. H., RIETBERGEN, M. M., VAN BOKHOVEN, A., BLOEMENA, E., LISSENBERG-WITTE, B. I., HEIDEMAN, D. A. M., BAATENBURG DE JONG, R. J., BRAKENHOFF, R. H. & LEEMANS, C. R. 2018. Evaluation of the eighth TNM classification on p16-positive oropharyngeal squamous cell carcinomas in the Netherlands and the importance of additional HPV DNA testing. *Ann Oncol,* 29**,** 1273-1279.

NDIAYE, C., MENA, M., ALEMANY, L., ARBYN, M., CASTELLSAGUE, X., LAPORTE, L., BOSCH, F. X., DE SANJOSE, S. & TROTTIER, H. 2014. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol,* 15**,** 1319-31.

NICHOLS, A. C., DHALIWAL, S. S., PALMA, D. A., BASMAJI, J., CHAPESKIE, C., DOWTHWAITE, S., FRANKLIN, J. H., FUNG, K., KWAN, K., WEHRLI, B., HOWLETT, C., SIDDIQUI, I., SALVADORI, M. I., WINQUIST, E., ERNST, S., KURUVILLA, S., READ, N., VENKATESAN, V., TODOROVIC, B., HAMMOND, J. A., KOROPATNICK, J., MYMRYK, J. S., YOO, J. & BARRETT, J. W. 2013a. Does HPV type affect outcome in oropharyngeal cancer? *J Otolaryngol Head Neck Surg,* 42**,** 9.

NICHOLS, A. C., YOO, J., HAMMOND, J. A., FUNG, K., WINQUIST, E., READ, N., VENKATESAN, V., MACNEIL, S. D., ERNST, D. S., KURUVILLA, S., CHEN, J., CORSTEN, M., ODELL, M., EAPEN, L., THEURER, J., DOYLE, P. C., WEHRLI, B., KWAN, K. & PALMA, D. A. 2013b. Early-stage squamous cell carcinoma of the oropharynx: radiotherapy vs. trans-oral robotic surgery (ORATOR)--study protocol for a randomized phase II trial. *BMC Cancer,* 13**,** 133.

NUTTING, C. M., MORDEN, J. P., HARRINGTON, K. J., URBANO, T. G., BHIDE, S. A., CLARK, C., MILES, E. A., MIAH, A. B., NEWBOLD, K., TANAY, M., ADAB, F., JEFFERIES, S. J., SCRASE,

C., YAP, B. K., A'HERN, R. P., SYDENHAM, M. A., EMSON, M., HALL, E. & GROUP, P. T. M. 2011. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. *Lancet Oncol,* 12**,** 127-36.

OH, S. T., KYO, S. & LAIMINS, L. A. 2001. Telomerase activation by human papillomavirus type 16 E6 protein: induction of human telomerase reverse transcriptase expression through Myc and GC-rich Sp1 binding sites. *J Virol,* 75**,** 5559-66.

OLINER, J. D., PIETENPOL, J. A., THIAGALINGAM, S., GYURIS, J., KINZLER, K. W. & VOGELSTEIN, B. 1993. Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature,* 362**,** 857-60.

OLMEDO-NIEVA, L., MUNOZ-BELLO, J. O., CONTRERAS-PAREDES, A. & LIZANO, M. 2018. The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis. *Viruses,* 10.

OLTHOF, N. C., HUEBBERS, C. U., KOLLIGS, J., HENFLING, M., RAMAEKERS, F. C., CORNET, I., VAN LENT-ALBRECHTS, J. A., STEGMANN, A. P., SILLING, S., WIELAND, U., CAREY, T. E., WALLINE, H. M., GOLLIN, S. M., HOFFMANN, T. K., DE WINTER, J., KREMER, B., KLUSSMANN, J. P. & SPEEL, E. J. 2015. Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. *Int J Cancer,* 136**,** E207-18.

OLTHOF, N. C., SPEEL, E. J., KOLLIGS, J., HAESEVOETS, A., HENFLING, M., RAMAEKERS, F. C., PREUSS, S. F., DREBBER, U., WIELAND, U., SILLING, S., LAM, W. L., VUCIC, E. A., KREMER, B., KLUSSMANN, J. P. & HUEBBERS, C. U. 2014. Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. *PLoS One,* 9**,** e88718.

ORAZIZADEH, M., HASHEMITABAR, M., BAHRAMZADEH, S., DEHBASHI, F. N. & SAREMY, S. 2015. Comparison of the enzymatic and explant methods for the culture of keratinocytes isolated from human foreskin. *Biomed Rep,* 3**,** 304-308.

OWADALLY, W., HURT, C., TIMMINS, H., PARSONS, E., TOWNSEND, S., PATTERSON, J., HUTCHESON, K., POWELL, N., BEASLEY, M., PALANIAPPAN, N., ROBINSON, M., JONES, T. M. & EVANS, M. 2015. PATHOS: a phase II/III trial of risk-stratified, reduced intensity adjuvant treatment in patients undergoing transoral surgery for Human papillomavirus (HPV) positive oropharyngeal cancer. *BMC Cancer,* 15**,** 602.

PAI, S. I. & WESTRA, W. H. 2009. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu Rev Pathol,* 4**,** 49-70.

PANTELEEVA, A., SLONINA, D., BRANKOVIC, K., SPEKL, K., PAWELKE, J., HOINKIS, C. & DORR, W. 2003. Clonogenic survival of human keratinocytes and rodent fibroblasts after irradiation with 25 kV x-rays. *Radiat Environ Biophys,* 42**,** 95-100.

PARFENOV, M., PEDAMALLU, C. S., GEHLENBORG, N., FREEMAN, S. S., DANILOVA, L., BRISTOW, C. A., LEE, S., HADJIPANAYIS, A. G., IVANOVA, E. V., WILKERSON, M. D., PROTOPOPOV, A., YANG, L., SETH, S., SONG, X., TANG, J., REN, X., ZHANG, J., PANTAZI, A., SANTOSO, N., XU, A. W., MAHADESHWAR, H., WHEELER, D. A., HADDAD, R. I., JUNG, J., OJESINA, A. I., ISSAEVA, N., YARBROUGH, W. G., HAYES, D. N., GRANDIS, J. R., EL-NAGGAR, A. K., MEYERSON, M., PARK, P. J., CHIN, L., SEIDMAN, J. G., HAMMERMAN, P. S., KUCHERLAPATI, R. & CANCER GENOME ATLAS, N. 2014. Characterization of HPV and host genome interactions in primary head and neck cancers. *Proc Natl Acad Sci U S A,* 111**,** 15544-9.

PARSONS, J. T., MENDENHALL, W. M., STRINGER, S. P., AMDUR, R. J., HINERMAN, R. W., VILLARET, D. B., MOORE-HIGGS, G. J., GREENE, B. D., SPEER, T. W., CASSISI, N. J. & MILLION, R. R. 2002. Squamous cell carcinoma of the oropharynx: surgery, radiation therapy, or both. *Cancer,* 94**,** 2967-80.

PATEL, S. G. & SHAH, J. P. 2005. TNM staging of cancers of the head and neck: striving for uniformity among diversity. *CA Cancer J Clin,* 55**,** 242-58; quiz 261-2, 264.

PATTILLO, R. A., HUSSA, R. O., STORY, M. T., RUCKERT, A. C., SHALABY, M. R. & MATTINGLY, R. F. 1977. Tumor antigen and human chorionic gonadotropin in CaSki cells: a new epidermoid cervical cancer cell line. *Science,* 196**,** 1456-8.

PEARSON, K. 1901. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* 2**,** 559-572.

PETT, M. & COLEMAN, N. 2007. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *J Pathol,* 212**,** 356-67.

PIECHACZYK, M. & FARRAS, R. 2008. Regulation and function of JunB in cell proliferation. *Biochem Soc Trans,* 36**,** 864-7.

PIGNON, J. P., LE MAITRE, A., MAILLARD, E., BOURHIS, J. & GROUP, M.-N. C. 2009. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol,* 92**,** 4-14.

PIM, D., BERGANT, M., BOON, S. S., GANTI, K., KRANJEC, C., MASSIMI, P., SUBBAIAH, V. K., THOMAS, M., TOMAIC, V. & BANKS, L. 2012. Human papillomaviruses and the specificity of PDZ domain targeting. *FEBS J,* 279**,** 3530-3537.

PIROTTE, E. 2017. PARP inhibition in novel oropharyngeal cancer cell lines *Ph.D. thesis, Division of Cancer and Genetics, School of Medicine, Cardiff University*

PRUE, G. 2014. Vaccinate boys as well as girls against HPV: it works, and it may be cost effective. *BMJ,* 349**,** g4834.

RAFEHI, H., ORLOWSKI, C., GEORGIADIS, G. T., VERVERIS, K., EL-OSTA, A. & KARAGIANNIS, T. C. 2011. Clonogenic assay: adherent cells. *J Vis Exp*.

RAFF, A. B., WOODHAM, A. W., RAFF, L. M., SKEATE, J. G., YAN, L., DA SILVA, D. M., SCHELHAAS, M. & KAST, W. M. 2013. The evolving field of human papillomavirus receptor research: a review of binding and entry. *J Virol,* 87**,** 6062-72.

RAMPIAS, T., SASAKI, C. & PSYRRI, A. 2014. Molecular mechanisms of HPV induced carcinogenesis in head and neck. *Oral Oncol,* 50**,** 356-63.

RAMQVIST, T. & DALIANIS, T. 2010. Oropharyngeal cancer epidemic and human papillomavirus. *Emerg Infect Dis,* 16**,** 1671-7.

RAMQVIST, T. & DALIANIS, T. 2011. An epidemic of oropharyngeal squamous cell carcinoma (OSCC) due to human papillomavirus (HPV) infection and aspects of treatment and prevention. *Anticancer Res,* 31**,** 1515-9.

RESSNEROVA, A., RAUDENSKA, M., HOLUBOVA, M., SVOBODOVA, M., POLANSKA, H., BABULA, P., MASARIK, M. & GUMULEC, J. 2016. Zinc and Copper Homeostasis in Head and Neck Cancer: Review and Meta-Analysis. *Curr Med Chem,* 23**,** 1304-30.

RHEINWALD, J. G. & GREEN, H. 1975a. Formation of a keratinizing epithelium in culture by a cloned cell line derived from a teratoma. *Cell,* 6**,** 317-30.

RHEINWALD, J. G. & GREEN, H. 1975b. Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. *Cell,* 6**,** 331-43.

RIECKMANN, T., TRIBIUS, S., GROB, T. J., MEYER, F., BUSCH, C. J., PETERSEN, C., DIKOMEY, E. & KRIEGS, M. 2013. HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. *Radiother Oncol,* 107**,** 242-6.

RIETBERGEN, M. M., SNIJDERS, P. J., BEEKZADA, D., BRAAKHUIS, B. J., BRINK, A., HEIDEMAN, D. A., HESSELINK, A. T., WITTE, B. I., BLOEMENA, E., BAATENBURG-DE JONG, R. J., LEEMANS, C. R. & BRAKENHOFF, R. H. 2014. Molecular characterization of p16-immunopositive but HPV DNA-negative oropharyngeal carcinomas. *Int J Cancer,* 134**,** 2366-72.

RILEY, T., SONTAG, E., CHEN, P. & LEVINE, A. 2008. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol,* 9**,** 402-12.

ROBINSON, J. T., THORVALDSDÓTTIR, H., WINCKLER, W., GUTTMAN, M., LANDER, E. S., GETZ, G. & MESIROV, J. P. 2011. Integrative Genomics Viewer. *Nature biotechnology,* 29**,** 24-26.

ROBINSON, M., SCHACHE, A., SLOAN, P. & THAVARAJ, S. 2012. HPV specific testing: a requirement for oropharyngeal squamous cell carcinoma patients. *Head Neck Pathol,* 6 Suppl 1**,** S83-90.

ROMAGOSA, C., SIMONETTI, S., LOPEZ-VICENTE, L., MAZO, A., LLEONART, M. E., CASTELLVI, J. & RAMON Y CAJAL, S. 2011. p16(Ink4a) overexpression in cancer: a tumor suppressor gene associated with senescence and high-grade tumors. *Oncogene,* 30**,** 2087-97.

ROSENBERGER, S., DE-CASTRO ARCE, J., LANGBEIN, L., STEENBERGEN, R. D. & ROSL, F. 2010. Alternative splicing of human papillomavirus type-16 E6/E6* early mRNA is coupled to EGF signaling via Erk1/2 activation. *Proc Natl Acad Sci U S A,* 107**,** 7006-11.

ROSENTHAL, A. N., RYAN, A., AL-JEHANI, R. M., STOREY, A., HARWOOD, C. A. & JACOBS, I. J. 1998. p53 codon 72 polymorphism and risk of cervical cancer in UK. *Lancet,* 352**,** 871-2.

ROUS, P. 1910. A Transmissible Avian Neoplasm. (Sarcoma of the Common Fowl.). *J Exp Med,* 12**,** 696-705.

ROUS, P. 1911. A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. *J Exp Med,* 13**,** 397-411.

RUITBERG, C. M., REEDER, D. J. & BUTLER, J. M. 2001. STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res,* 29**,** 320-2.

RYNDOCK, E. J. & MEYERS, C. 2014. A risk for non-sexual transmission of human papillomavirus? *Expert Rev Anti Infect Ther,* 12**,** 1165-70.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science,* 239**,** 487-91.

SAPP, M. & BIENKOWSKA-HABA, M. 2009. Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. *FEBS J,* 276**,** 7206-16.

SCHACHE, A. G., LILOGLOU, T., RISK, J. M., FILIA, A., JONES, T. M., SHEARD, J., WOOLGAR, J. A., HELLIWELL, T. R., TRIANTAFYLLOU, A., ROBINSON, M., SLOAN, P., HARVEY-WOODWORTH, C., SISSON, D. & SHAW, R. J. 2011a. Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous cell carcinoma: sensitivity, specificity, and prognostic discrimination. *Clin Cancer Res,* 17**,** 6262-71.

SCHACHE, A. G., POWELL, N. G., CUSCHIERI, K. S., ROBINSON, M., LEARY, S., MEHANNA, H., RAPOZO, D., LONG, A., CUBIE, H., JUNOR, E., MONAGHAN, H., HARRINGTON, K. J., NUTTING, C. M., SCHICK, U., LAU, A. S., UPILE, N., SHEARD, J., BROUGHAM, K., WEST, C. M., OGUEJIOFOR, K., THOMAS, S., NESS, A. R., PRING, M., THOMAS, G. J., KING, E. V., MCCANCE, D. J., JAMES, J. A., MORAN, M., SLOAN, P., SHAW, R. J., EVANS, M. & JONES, T. M. 2016. HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology. *Cancer Res,* 76**,** 6598-6606.

SCHACHE, A. G., SIMCOCK, R., GILBERT, D. C. & SHAW, R. J. 2011b. Changing face of HPV related cancer in the UK. *BMJ,* 343**,** d6675.

SCHELHAAS, M., SHAH, B., HOLZER, M., BLATTMANN, P., KUHLING, L., DAY, P. M., SCHILLER, J. T. & HELENIUS, A. 2012. Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis. *PLoS Pathog,* 8**,** e1002657.

SCHERER, W. F., SYVERTON, J. T. & GEY, G. O. 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. *J Exp Med,* 97**,** 695-710.

SCHLECHT, N. F., FRANCO, E. L., PINTOS, J. & KOWALSKI, L. P. 1999. Effect of smoking cessation and tobacco type on the risk of cancers of the upper aero-digestive tract in Brazil. *Epidemiology,* 10**,** 412-8.

SCHMITZ, M., DRIESCH, C., BEER-GRONDKE, K., JANSEN, L., RUNNEBAUM, I. B. & DURST, M. 2012. Loss of gene function as a consequence of human papillomavirus DNA integration. *Int J Cancer,* 131**,** E593-602.

SCOUMANNE, A., CHO, S. J., ZHANG, J. & CHEN, X. 2011. The cyclin-dependent kinase inhibitor p21 is regulated by RNA-binding protein PCBP4 via mRNA stability. *Nucleic Acids Res,* 39**,** 213-24.

SEDMAN, J. & STENLUND, A. 1998. The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *J Virol,* 72**,** 6893-7.

SHAIKH, M. H., KHAN, A. I., SADAT, A., CHOWDHURY, A. H., JINNAH, S. A., GOPALAN, V., LAM, A. K., CLARKE, D. T. W., MCMILLAN, N. A. J. & JOHNSON, N. W. 2017. Prevalence and types of high-risk human papillomaviruses in head and neck cancers from Bangladesh. *BMC Cancer,* 17**,** 792.

SHIEH, S. Y., IKEDA, M., TAYA, Y. & PRIVES, C. 1997. DNA damage-induced phosphorylation of p53 alleviates inhibition by MDM2. *Cell,* 91**,** 325-34.

SMEETS, S. J., HESSELINK, A. T., SPEEL, E. J., HAESEVOETS, A., SNIJDERS, P. J., PAWLITA, M., MEIJER, C. J., BRAAKHUIS, B. J., LEEMANS, C. R. & BRAKENHOFF, R. H. 2007. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer,* 121**,** 2465-72.

SOMERS, K. D., MERRICK, M. A., LOPEZ, M. E., INCOGNITO, L. S., SCHECHTER, G. L. & CASEY, G. 1992. Frequent p53 mutations in head and neck cancer. *Cancer Res,* 52**,** 5997-6000.

ST CLAIR, S., GIONO, L., VARMEH-ZIAIE, S., RESNICK-SILVERMAN, L., LIU, W. J., PADI, A., DASTIDAR, J., DACOSTA, A., MATTIA, M. & MANFREDI, J. J. 2004. DNA damage-induced downregulation of Cdc25C is mediated by p53 via two independent mechanisms: one involves direct binding to the cdc25C promoter. *Mol Cell,* 16**,** 725-36.

STANLEY, M. A. 2002. *CHAPTER 5 Culture of Human Cervical Epithelial Cells* New York; Glasgow, Wiley-Liss, Inc.

STEENBERGEN, R. D., HERMSEN, M. A., WALBOOMERS, J. M., JOENJE, H., ARWERT, F., MEIJER, C. J. & SNIJDERS, P. J. 1995. Integrated human papillomavirus type 16 and loss of heterozygosity at 11q22 and 18q21 in an oral carcinoma and its derivative cell line. *Cancer Res,* 55**,** 5465-71.

STEGER, G. & CORBACH, S. 1997. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J Virol,* 71**,** 50-8.

STEIN, A. P., SAHA, S., KRANINGER, J. L., SWICK, A. D., YU, M., LAMBERT, P. F. & KIMPLE, R. J. 2015. Prevalence of Human Papillomavirus in Oropharyngeal Cancer: A Systematic Review. *Cancer J,* 21**,** 138-46.

STOREY, A., THOMAS, M., KALITA, A., HARWOOD, C., GARDIOL, D., MANTOVANI, F., BREUER, J., LEIGH, I. M., MATLASHEWSKI, G. & BANKS, L. 1998. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature,* 393**,** 229-34.

STRANSKY, N., EGLOFF, A. M., TWARD, A. D., KOSTIC, A. D., CIBULSKIS, K., SIVACHENKO, A., KRYUKOV, G. V., LAWRENCE, M. S., SOUGNEZ, C., MCKENNA, A., SHEFLER, E., RAMOS, A. H., STOJANOV, P., CARTER, S. L., VOET, D., CORTES, M. L., AUCLAIR, D., BERGER, M. F., SAKSENA, G., GUIDUCCI, C., ONOFRIO, R. C., PARKIN, M., ROMKES, M., WEISSFELD, J. L., SEETHALA, R. R., WANG, L., RANGEL-ESCARENO, C., FERNANDEZ-LOPEZ, J. C., HIDALGO-MIRANDA, A., MELENDEZ-ZAJGLA, J., WINCKLER, W., ARDLIE, K., GABRIEL, S. B., MEYERSON, M., LANDER, E. S., GETZ, G., GOLUB, T. R., GARRAWAY, L. A. & GRANDIS, J. R. 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science,* 333**,** 1157-60.

STRICKLAND, S. W. & 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. *J Virol,* 90**,** 5611-5621.

STUART, S. A. & WANG, J. Y. 2009. Ionizing radiation induces ATM-independent degradation of p21Cip1 in transformed cells. *J Biol Chem,* 284**,** 15061-70.

STUBENRAUCH, F. & LAIMINS, L. A. 1999. Human papillomavirus life cycle: active and latent phases. *Semin Cancer Biol,* 9**,** 379-86.

SYRJANEN, K., SYRJANEN, S., LAMBERG, M., PYRHONEN, S. & NUUTINEN, J. 1983. Morphological and immunohistochemical evidence suggesting human papillomavirus (HPV) involvement in oral squamous cell carcinogenesis. *Int J Oral Surg,* 12**,** 418-24.

SYRJANEN, K., SYRJANEN, S. & PYRHONEN, S. 1982. Human papilloma virus (HPV) antigens in lesions of laryngeal squamous cell carcinomas. *ORL J Otorhinolaryngol Relat Spec,* 44**,** 323-34.

SYRJANEN, S. 2010a. Current concepts on human papillomavirus infections in children. *APMIS,* 118**,** 494-509.

SYRJANEN, S. 2010b. The role of human papillomavirus infection in head and neck cancers. *Ann Oncol,* 21 Suppl 7**,** vii243-5.

TAKEBAYASHI, S., HICKSON, A., OGAWA, T., JUNG, K. Y., MINETA, H., UEDA, Y., GRENMAN, R., FISHER, S. G. & CAREY, T. E. 2004. Loss of chromosome arm 18q with tumor progression in head and neck squamous cancer. *Genes Chromosomes Cancer,* 41**,** 145-54.

TANG, A. L., HAUFF, S. J., OWEN, J. H., GRAHAM, M. P., CZERWINSKI, M. J., PARK, J. J., WALLINE, H., PAPAGERAKIS, S., STOERKER, J., MCHUGH, J. B., CHEPEHA, D. B., BRADFORD, C. R., CAREY, T. E. & PRINCE, M. E. 2012. UM-SCC-104: a new human papillomavirus-16-positive cancer stem cell-containing head and neck squamous cell carcinoma cell line. *Head Neck,* 34**,** 1480-91.

TAYLOR, A. & POWELL, M. E. 2004. Intensity-modulated radiotherapy--what is it? *Cancer Imaging,* 4**,** 68-73.

THERMOFISHERSCIENTIFIC 2009. Human Epidermal Keratinocytes, neoatal (HEKn); Cat.no. C-001-5C. *Manual & Protocol*.

THOMAS, M., DASGUPTA, J., ZHANG, Y., CHEN, X. & BANKS, L. 2008. Analysis of specificity determinants in the interactions of different HPV E6 proteins with their PDZ domain-containing substrates. *Virology,* 376**,** 371-8.

THORLAND, E. C., MYERS, S. L., GOSTOUT, B. S. & SMITH, D. I. 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene,* 22**,** 1225-37.

TIBBETTS, R. S., BRUMBAUGH, K. M., WILLIAMS, J. M., SARKARIA, J. N., CLIBY, W. A., SHIEH, S. Y., TAYA, Y., PRIVES, C. & ABRAHAM, R. T. 1999. A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes Dev,* 13**,** 152-7.

TILANUS, M. G. 2006. Short tandem repeat markers in diagnostics: what's in a repeat? *Leukemia,* 20**,** 1353-5.

TODARO, G. J. & GREEN, H. 1963. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J Cell Biol,* 17**,** 299-313.

TRIKALINOS, T. A., TERASAWA, T., IP, S., RAMAN, G. & LAU, J. 2009. *Particle Beam Radiation Therapies for Cancer.* Rockville (MD).

TSAI, S. T., WONG, T. Y., OU, C. Y., FANG, S. Y., CHEN, K. C., HSIAO, J. R., HUANG, C. C., LEE, W. T., LO, H. I., HUANG, J. S., WU, J. L., YEN, C. J., HSUEH, W. T., WU, Y. H., YANG, M. W., LIN, F. C., CHANG, J. Y., CHANG, K. Y., WU, S. Y., LIAO, H. C., LIN, C. L., WANG, Y. H., WENG, Y. L., YANG, H. C. & CHANG, J. S. 2014. The interplay between alcohol consumption, oral hygiene, ALDH2 and ADH1B in the risk of head and neck cancer. *Int J Cancer,* 135**,** 2424-36.

VAN DOORSLAER, K., CHEN, Z., BERNARD, H. U., CHAN, P. K. S., DESALLE, R., DILLNER, J., FORSLUND, O., HAGA, T., MCBRIDE, A. A., VILLA, L. L., BURK, R. D. & ICTV REPORT, C. 2018. ICTV Virus Taxonomy Profile: Papillomaviridae. *J Gen Virol,* 99**,** 989-990.

VAN DOORSLAER, K., LI, Z., XIRASAGAR, S., MAES, P., KAMINSKY, D., LIOU, D., SUN, Q., KAUR, R., HUYEN, Y. & MCBRIDE, A. A. 2017. The Papillomavirus Episteme: a major update to the papillomavirus sequence database. *Nucleic Acids Res,* 45**,** D499-D506.

VAN TINE, B. A., DAO, L. D., WU, S. Y., SONBUCHNER, T. M., LIN, B. Y., ZOU, N., CHIANG, C. M., BROKER, T. R. & CHOW, L. T. 2004. Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning. *Proc Natl Acad Sci U S A,* 101**,** 4030-5.

VIGNARD, J., MIREY, G. & SALLES, B. 2013. Ionizing-radiation induced DNA double-strand breaks: a direct and indirect lighting up. *Radiother Oncol,* 108**,** 362-9.

VILLA, L. L., COSTA, R. L., PETTA, C. A., ANDRADE, R. P., PAAVONEN, J., IVERSEN, O. E., OLSSON, S. E., HOYE, J., STEINWALL, M., RIIS-JOHANNESSEN, G., ANDERSSON-ELLSTROM, A., ELFGREN, K., KROGH, G., LEHTINEN, M., MALM, C., TAMMS, G. M., GIACOLETTI, K., LUPINACCI, L., RAILKAR, R., TADDEO, F. J., BRYAN, J., ESSER, M. T., SINGS, H. L., SAAH, A. J. & BARR, E. 2006. High sustained efficacy of a prophylactic quadrivalent human papillomavirus types 6/11/16/18 L1 virus-like particle vaccine through 5 years of follow-up. *Br J Cancer,* 95**,** 1459-66.

VILLA, L. L. & DENNY, L. 2006. CHAPTER 7 Methods for detection of HPV infection and its clinical utility. *Int J Gynaecol Obstet,* 94 Suppl 1**,** S71-S80.

VOJTECHOVA, Z., SABOL, I., SALAKOVA, M., TUREK, L., GREGA, M., SMAHELOVA, J., VENCALEK, O., LUKESOVA, E., KLOZAR, J. & TACHEZY, R. 2016. Analysis of the integration of human papillomaviruses in head and neck tumours in relation to patients' prognosis. *Int J Cancer,* 138**,** 386-95.

VON AHLFEN, S., MISSEL, A., BENDRAT, K. & SCHLUMPBERGER, M. 2007. Determinants of RNA quality from FFPE samples. *PLoS One,* 2**,** e1261.

WALBOOMERS, J. M., JACOBS, M. V., MANOS, M. M., BOSCH, F. X., KUMMER, J. A., SHAH, K. V., SNIJDERS, P. J., PETO, J., MEIJER, C. J. & MUNOZ, N. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol,* 189**,** 12-9.

WANG, W., GAO, D. & WANG, X. 2018. Can single-cell RNA sequencing crack the mystery of cells? *Cell Biol Toxicol,* 34**,** 1-6.

WANG, Y. & PRIVES, C. 1995. Increased and altered DNA binding of human p53 by S and G2/M but not G1 cyclin-dependent kinases. *Nature,* 376**,** 88-91.

WARD, M. J., THIRDBOROUGH, S. M., MELLOWS, T., RILEY, C., HARRIS, S., SUCHAK, K., WEBB, A., HAMPTON, C., PATEL, N. N., RANDALL, C. J., COX, H. J., JOGAI, S., PRIMROSE, J., PIPER, K., OTTENSMEIER, C. H., KING, E. V. & THOMAS, G. J. 2014. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. *Br J Cancer,* 110**,** 489-500.

WATSON, J. V., CHAMBERS, S. H. & SMITH, P. J. 1987. A pragmatic approach to the analysis of DNA histograms with a definable G1 peak. *Cytometry,* 8**,** 1-8.

WEAVER, A. N., COOPER, T. S., RODRIGUEZ, M., TRUMMELL, H. Q., BONNER, J. A., ROSENTHAL, E. L. & YANG, E. S. 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**,** 26995-7007.

WEINBERG, R. A. 2007. *The biology of cancer,* New York ; London, Garland Science.

WEST, C. M., DAVIDSON, S. E., ROBERTS, S. A. & HUNTER, R. D. 1993. Intrinsic radiosensitivity and prediction of patient response to radiotherapy for carcinoma of the cervix. *Br J Cancer,* 68**,** 819-23.

WHO 2014. Human papillomavirus vaccines: WHO position paper, October 2014. *Wkly Epidemiol Rec,* 89**,** 465-91.

WHO 2015. Human papillomavirus vaccines: WHO position paper, October 2014-Recommendations. *Vaccine,* 33**,** 4383-4.

WIEST, T., SCHWARZ, E., ENDERS, C., FLECHTENMACHER, C. & BOSCH, F. X. 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**,** 1510-7.

WILLERS, H., XIA, F. & POWELL, S. N. 2002. Recombinational DNA Repair in Cancer and Normal Cells: The Challenge of Functional Analysis. *J Biomed Biotechnol,* 2**,** 86-93.

WILLIAMS, V. M., FILIPPOVA, M., SOTO, U. & DUERKSEN-HUGHES, P. J. 2011. HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future Virol,* 6**,** 45-57.

WILSON, R., FEHRMANN, F. & LAIMINS, L. A. 2005. Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31. *J Virol,* 79**,** 6732-40.

YOUNG, M. D., WAKEFIELD, M. J., SMYTH, G. K. & OSHLACK, A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol,* 11**,** R14.

ZHANG, J., CHEN, S., ZHANG, W., ZHANG, J., LIU, X., SHI, H., CHE, H., WANG, W., LI, F. & YAO, L. 2008. Human differentiation-related gene NDRG1 is a Myc downstream-regulated gene that is repressed by Myc on the core promoter region. *Gene,* 417**,** 5-12.

ZHANG, W., EDWARDS, A., FANG, Z., FLEMINGTON, E. K. & ZHANG, K. 2016. Integrative Genomics and Transcriptomics Analysis Reveals Potential Mechanisms for Favorable Prognosis of Patients with HPV-Positive Head and Neck Carcinomas. *Sci Rep,* 6**,** 24927.

ZHANG, Z. F., MORGENSTERN, H., SPITZ, M. R., TASHKIN, D. P., YU, G. P., HSU, T. C. & SCHANTZ, S. P. 2000. Environmental tobacco smoking, mutagen sensitivity, and head and neck squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev,* 9**,** 1043-9.

ZHAO, M., SANO, D., PICKERING, C. R., JASSER, S. A., HENDERSON, Y. C., CLAYMAN, G. L., STURGIS, E. M., OW, T. J., LOTAN, R., CAREY, T. E., SACKS, P. G., GRANDIS, J. R., SIDRANSKY, D., HELDIN, N. E. & MYERS, J. N. 2011. Assembly and initial characterization of a panel of 85 genomically validated cell lines from diverse head and neck tumor sites. *Clin Cancer Res,* 17**,** 7248-64.

ZIEMANN, F., SELTZSAM, S., DREFFKE, K., PREISING, S., ARENZ, A., SUBTIL, F. S. B., RIECKMANN, T., ENGENHART-CABILLIC, R., DIKOMEY, E. & WITTIG, A. 2017. Roscovitine strongly enhances the effect of olaparib on radiosensitivity for HPV neg. but not for HPV pos. HNSCC cell lines. *Oncotarget,* 8**,** 105170-105183.

ZUR HAUSEN, H. 1975. Oncogenic Herpes viruses. *Biochim Biophys Acta,* 417**,** 25-53.

ZUR HAUSEN, H. 1977. Human papillomaviruses and their possible role in squamous cell carcinomas. *Curr Top Microbiol Immunol,* 78**,** 1-30.

ZUR HAUSEN, H. 2002. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer,* 2**,** 342-50.

ZUR HAUSEN, H. 2009. Papillomaviruses in the causation of human cancers - a brief historical account. *Virology,* 384**,** 260-5.

# Appendix 1: PCOC study protocol

# 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: powelln@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.<br>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.

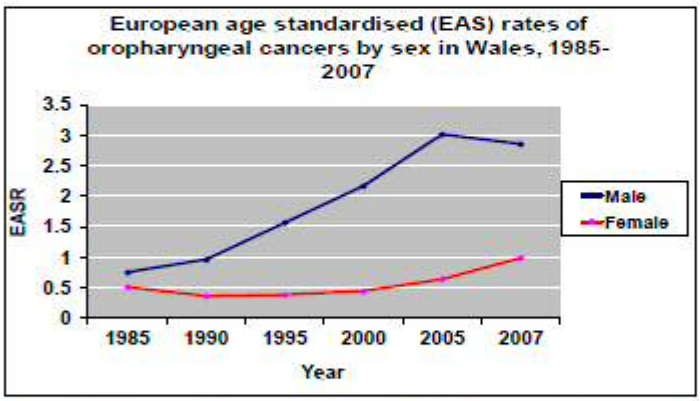



Fig 1: Incidence of oropharyngeal cancer in Wales in men and women 1987-2007 (WICSU data).

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.

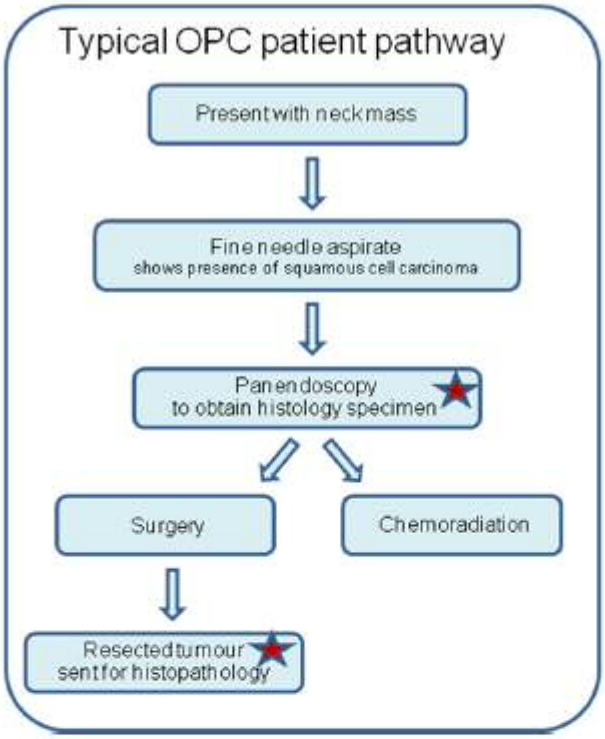
## 3. STUDY DESIGN

### 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.

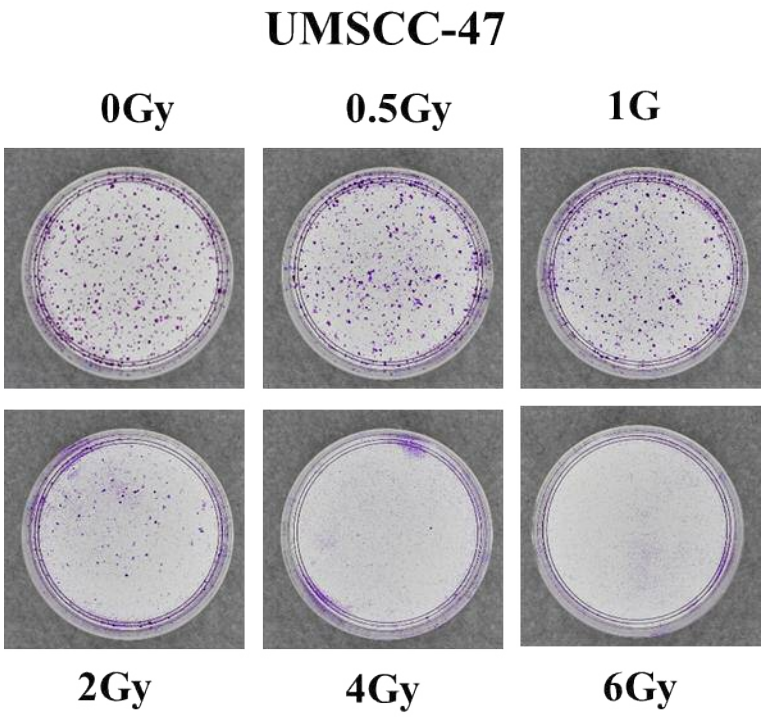
## 10. PUBLICATION POLICY

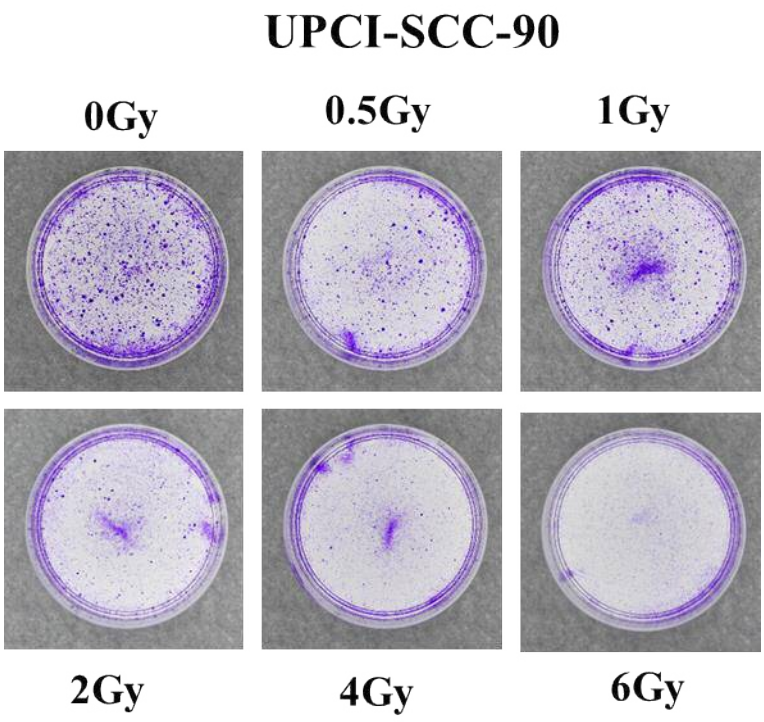
The results of these investigations will be published in the peer-reviewed scientific literature. No patient identifiable data will be published, but the contribution of the patients and of the funding body will be acknowledged.

## 11. REFERENCES

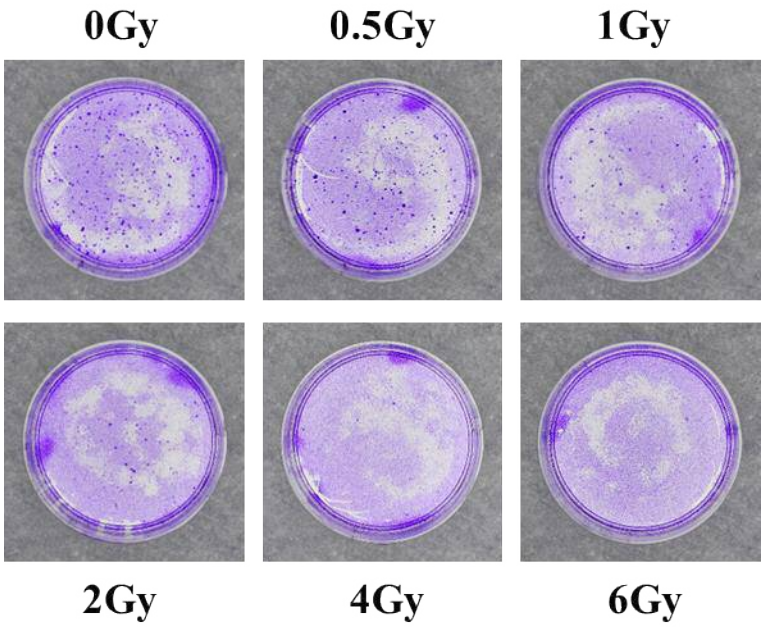

1. Nasman, A., et al., *Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma?* Int J Cancer, 2009. **125**(2): p. 362-6.
2. Mellin, H., et al., *Human papillomavirus type 16 is episomal and a high viral load may be correlated 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

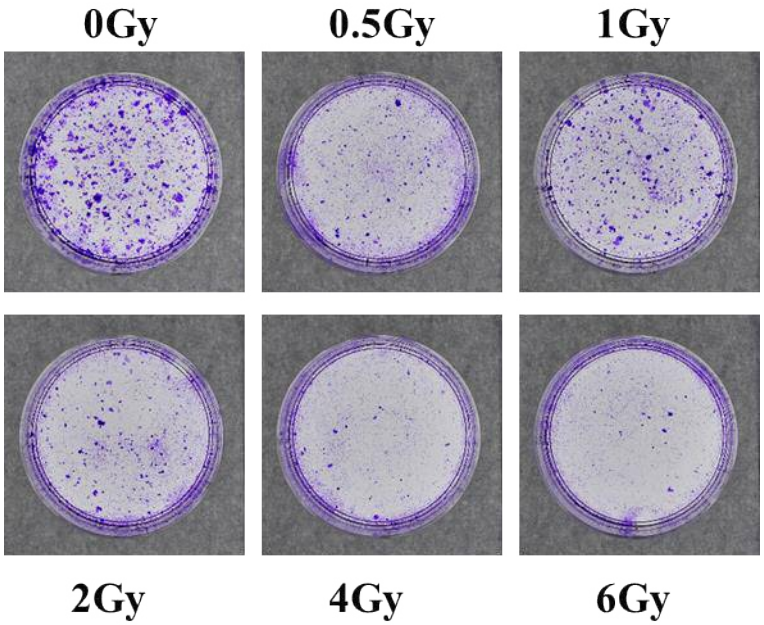


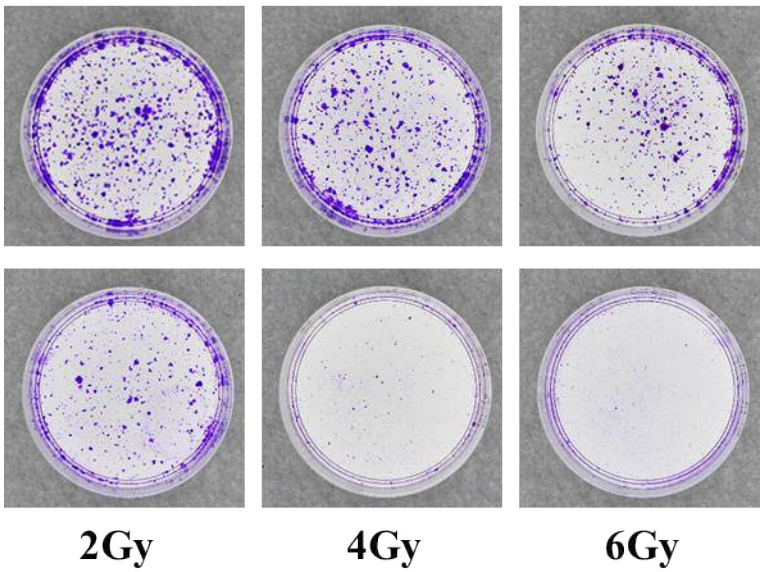
## CU-OP-2



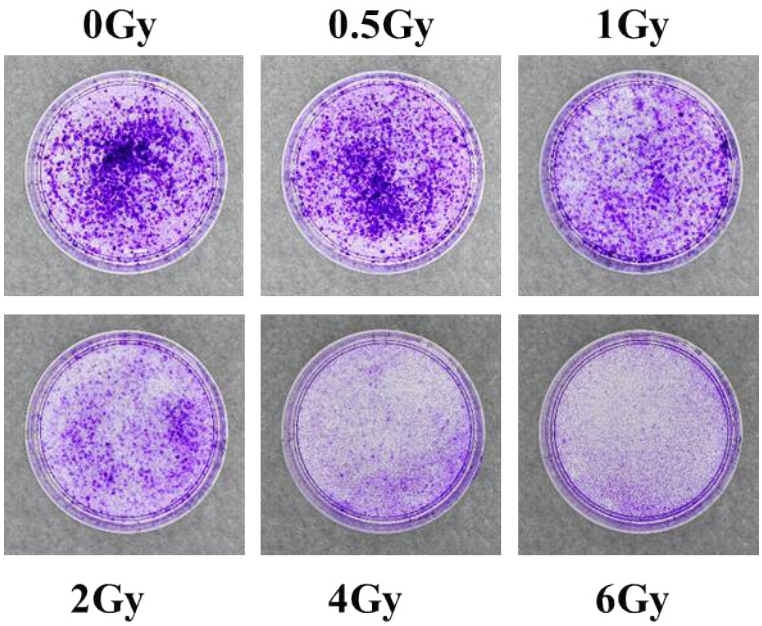

## UMSCC-19

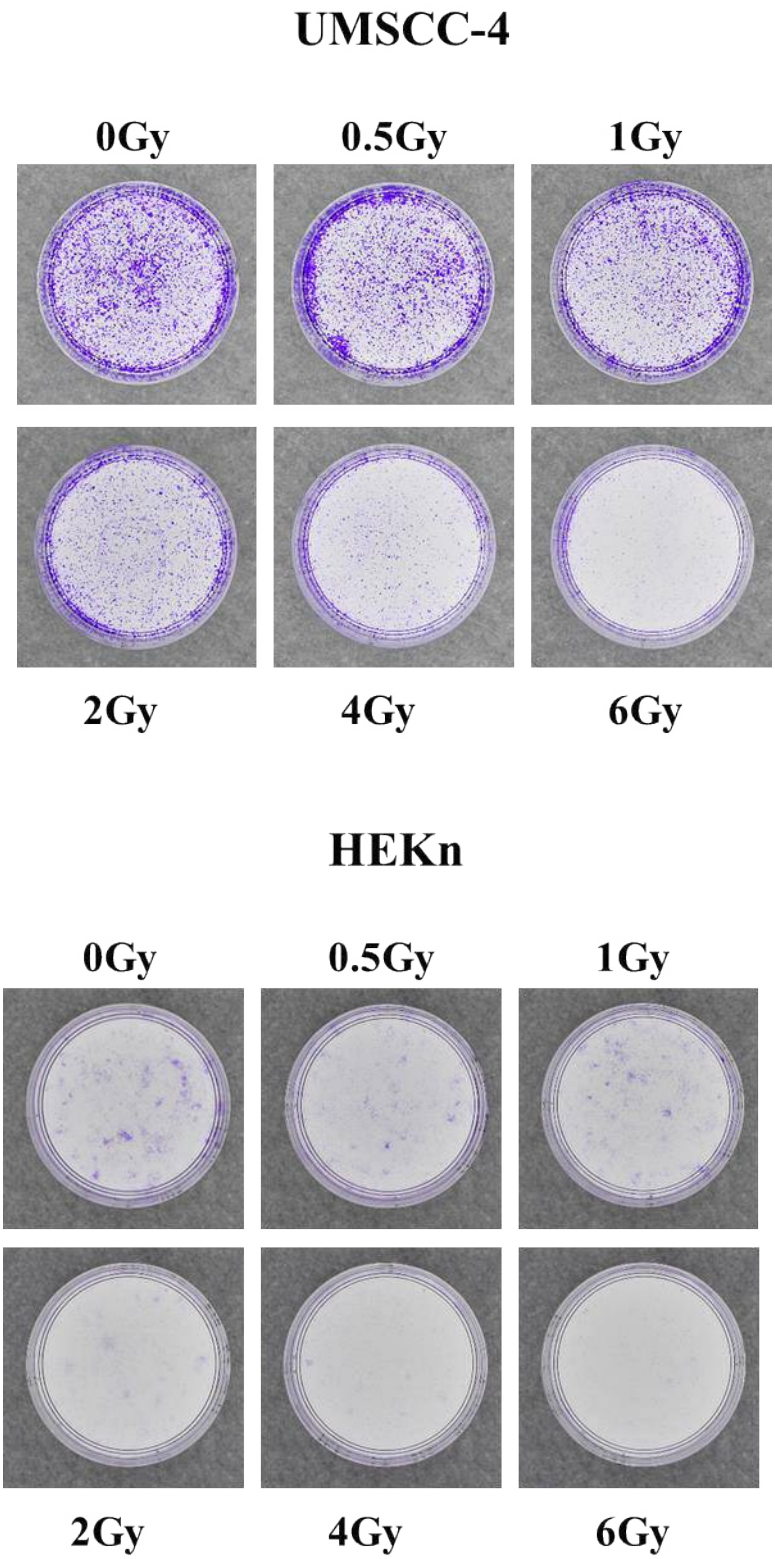


## UMSCC-6




## UMSCC-74a




UMSCC-4
0Gy
0.5Gy
1Gy
2Gy
4Gy
6Gy
HEKn
0Gy
0.5Gy
1Gy
2Gy
4Gy
6Gy